Individual Variability in Recognition of Frequency-Lowered Speech

Joshua M. Alexander, Ph.D.[1]

ABSTRACT

Frequency lowering in hearing aids is not a new concept, but modern advances in technology have allowed it to be performed more efficiently and on select portions of the spectrum. Nonlinear frequency compression reduces the frequency spacing in a band of high-frequency energy so that more information is carried in the audible bandwidth. Frequency transposition and translation techniques lower only the part of the high-frequency spectrum that likely contains important speech information. These advances may help overcome the limited bandwidth in conventional hearing aids, which restrict access to high-frequency information even for those with mild to moderate hearing loss. This is especially important for young children learning speech and language. A framework is advanced in which factors that influence individual differences in speech recognition can be divided into extrinsic factors that affect the representation of the frequency-lowered speech at the auditory periphery, including the specific technique and the settings chosen for it, and intrinsic factors that contribute to an individual's ability to learn and benefit from this signal. Finally, the importance of electroacoustically verifying the output to avoid too little or too much lowering and the importance of validating effectiveness of outcomes in individual users of the technology are emphasized.

KEYWORDS: Hearing aids, frequency lowering, frequency compression, frequency transposition

Learning Outcomes: As a result of this activity, the participant will describe how different frequency-lowering technologies alter the speech signal, how characteristics of the hearing aid wearer might influence the ability to benefit from this signal, and what to watch for when verifying output.

---

[1] Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, Indiana.

Address for correspondence: Joshua M. Alexander, Ph.D., Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN 47907 (e-mail: alexan14@purdue.edu).

Individual Variability in Aided Outcomes; Guest Editor, Jason A. Galster, Ph.D.

Semin Hear 2013;34:86–109. 333 Seventh Avenue, New York, NY 10001, USA. Tel: +1(212) 584-4662.
DOI: http://dx.doi.org/10.1055/s-0033-1341346.
ISSN 0734-0451.

Individuals with high-frequency sensorineural hearing loss (SNHL) are denied access to potentially important speech information. For milder losses, this can occur if the miniature electronics in hearing aids are unable to provide sufficient high-frequency amplification or cannot do so without audible whistling and overtones caused by feedback. For more severe losses, the inner hairs cells that code these frequencies may simply be "dead," possibly rendering amplification in this region less useful.[1–5] As the severity of loss increases to include more low-frequency content, the amount of "lost" speech information increases along with the challenges for the listener. Most of the sounds we will be concerned about in this discussion involve the fricatives, affricates, and the initial segments of the stop consonants, which are primarily characterized by aperiodic mid- to high-frequency spectral information. Fig. 1 shows a spectrogram of the sentence "children like strawberries" with the previously mentioned sound classes denoted by arrows. As can be seen, the acoustic energy in these sounds can be quite high in frequency, with peak energy sometimes around 9,000 Hz, especially for women and children talkers.[6–8] In contrast, the vowels and other consonants are primarily characterized by bands of energy, formants, at relatively lower frequencies. Unlike the diffuse spread of energy that is characteristic of the fricatives, affricates, and initial stop consonant segments, changing the frequency relationship of the formants can have serious consequences for the identity of the corresponding speech sound.

Speech is linguistically and acoustically redundant and, with varying degrees of success, listeners can identify high-frequency phonemes using only the transitions from the lower-frequency formants of the coarticulated phonemes that precede and follow them.[9,10] Despite this, there is evidence that speech perception improves for both adults and children when an effort is made to preserve the high-frequency noise energy associated with frication, especially when identifying /s/.[7,11–21] These findings and the fact that hearing aids have limited usable bandwidth have been used to explain the continued difficulty experienced by young children using hearing aids when perceiving and producing these sounds compared with vowels and other consonant sound classes.[22–25]

The gravity of this problem is compounded by the regularity with which /s/ and its voiced cognate, /z/, occur in the English language (~ 8% of all spoken consonants) and by their linguistic importance.[26] Rudmin identifies over 20 linguistic uses for /s/ and /z/, including plurality, third-person present tense, past versus present tense, to show possession, possessive pronouns, contractions, and so on.[26] Developmentally, inconsistent exposure to these phonemes for a child with SNHL may have long-term consequences for morphosyntactic development.[24,27] Findings like these have been the

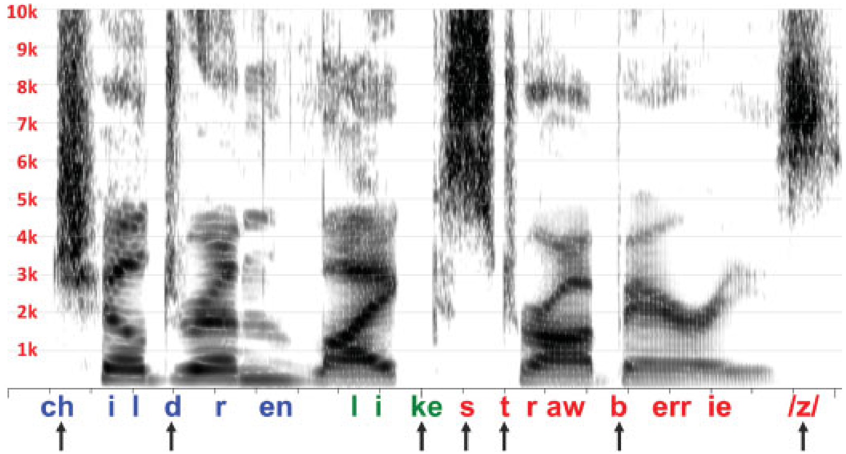


Figure 1 Spectrogram of the sentence "children like strawberries" before processing as spoken by a female talker. Arrows denote the fricative, affricate, and stop sound classes that contain significant high-frequency energy and are therefore the subject of frequency-lowering techniques.

inspiration for a variety of frequency-lowering (FL) techniques (i.e., methods of moving high-frequency speech information into lower-frequency regions) in commercially available hearing aids. In contrast to a few years ago when FL seemed like just a signal processing novelty, now more than half of the world's major hearing aid manufacturers include FL as an optional feature.

Table 1 provides a summary of peer-reviewed research on modern, commercially available FL technologies along with a few studies under review in which the author has personal involvement. A supplementary description of early FL techniques and a summary of related research findings can be found in Simpson.[28] Specific studies in Table 1 are discussed where appropriate and specific technologies are discussed in the next section.

At the outset, it is important to make note that the data indicate that we, as a field of clinicians and researchers, do not yet know enough to predict who will and who will not benefit from FL technology. First, because modern FL techniques have only been around for a short period compared with other hearing aid processing strategies, there just simply is not a lot of data by independent researchers as can be inferred from Table 1. Furthermore, we have the same difficulties as we do when trying to understand individual variability with conventional hearing aids in addition to nuances associated with FL technology. Specifically, we now have the added difficulty of understanding (1) how the different technologies alter the speech signal that is transduced by the impaired auditory periphery and (2) how this interacts with individual characteristics that influence the ability to learn it.

EXTRINSIC FACTORS RELATED TO THE SIGNAL PROCESSING AND IMPAIRMENT

Fig. 2 provides a framework for understanding the different factors that likely influence individual variability in recognition of frequency-lowered speech. The first set of factors to be discussed are extrinsic to the listener and relates to the representation of the signal at the auditory periphery. The specific speech sound and its environmental context serve as input to the digital signal processor that alters the acoustic representation in a way that is specific to the signal and to the FL technology and its settings. How this newly coded input is then "seen" by the central auditory system, and beyond, depends on the integrity of peripheral processing (e.g., threshold elevation and broadened auditory filters associated with varying degrees of outer and inner hair cell impairment). It is argued here that these factors contribute the most to differences in individual outcomes. Because there is nothing one can do to control the input signal or the peripheral processing, it is incumbent upon the hearing aid dispenser to choose the appropriate technology and settings that match the speech perception deficit with the impairment. As will be discussed, with a few exceptions, there are no firm guidelines on how this should be done during fitting, which puts the onus on the dispenser to use appropriate measures to validate individual outcomes.

Signal Processing Details

To make informed choices about when, what, and how to implement FL in a hearing aid fitting, it is critical that the dispenser understands the technology and how the handles in the fitting software manipulate the signal. Unfortunately, these details are sometimes hard to come by. Presented below is a summary of information gathered from detailed acoustic analyses by the author for all but the most recent FL techniques. Industry interest in FL had a slow start, but has recently surged. Currently, five FL techniques are implemented in commercially available hearing aids:

- Linear frequency compression by AVR Sonovation of Israel (introduced 1991)
- Linear frequency transposition (LFT) by Widex of Lynge, Denmark (introduced 2006)
- Nonlinear frequency compression (NFC) by Phonak of Stäfa, Switzerland (introduced 2008). Also used by Unitron of Kitchener, Ontario, Canada (starting 2012)
- Spectral envelope warping by Starkey Hearing Technologies of Minnesota (Eden Prarie, MN), USA (introduced 2011)

Table 1 Summary of Studies Using Modern Frequency-Lowering Technology

Reference/Technique	Participant Details	Methods/Outcome Measures	Results
Kuk et al (2009)[29]/LFT	8 adults with severe to profound SNHL > 2 kHz	● Daily training for first mo with LFT; no control for training without LFT ● Consonant recog (nonsense CVCVCs) in quiet (50- & 68-dB SPL) & babble noise, assessed after initial fit, 1 mo postfit + training, 2 mo postfit	● At 50-dB SPL, fricatives improved 5–10% initially & an additional 10% after 2 mo ● At 68-dB SPL, fricatives improved 10% after 2 mo; an initial decrease in stops was back to baseline after 2 mo ● No difference for speech in noise, except fricatives (> 15% improvement)
Auriemmo et al (2009)[41]/LFT	10 children with severe to profound SNHL > 3 kHz	● Phoneme recog (nonsense CVCVs) in quiet & fricative production ● Conditions: participant's own aid, new device without LFT (3 wk), & with LFT (6 wk) ● Weekly auditory training without LFT & with LFT	● After 6 wk, consonant recog with LFT at 30-dB HL was > 20% better than without LFT; no significant difference at 50-dB HL ● Worst performers without LFT had greatest improvement when activated ● No vowel recog differences ● Production accuracy of /s/ & /z/ improved by > 10% when activated
Smith et al (2009)[96]/LFT	6 children with severe to profound SNHL > 1 kHz	● Audio & audiovisual recog of live voice monosyllables (CVCs) ● Consonant production in words & sentences ● Parental & teacher reports ● No randomization or control for maturation effects	● Improvements in CVC recog at 3 & 6 mo for audio only (not audiovisual) & production at 6 mo
Alexander et al (2008)[42]/LFT & NFC	24 adults with mild to moderate SNHL & 24 normal-hearing controls	● Recog of fricatives & affricates (VCs) in noise recorded through aids with LFT & NFC ● Against manufacturers' candidacy ● Conditions: no LFT/no NFC (conventional), LFT, NFC, wideband (by adding high-pass filtered speech to the recordings)	● Wideband improved over conventional ● NFC with input bandwidth ~ 9 kHz improved over conventional ● LFT significantly degraded recog, especially /s/ & /z/, for both groups
Simpson et al (2005)[34]/NFC	17 adults (experienced HA users) with moderate to profound SNHL	● Recog of CVC monosyllabic words ● Conditions: conventional aid, experimental NFC (4–6 wk) ● No evidence that audibility for the lowered speech was measured	● 8 had improvement in high-freq. word recog with NFC, 8 showed no difference, & 1 performed worse
Simpson et al (2006)[35]/NFC	7 adults with precipitous SNHL	● Recog in quiet (open-set monosyllabic words, closed-set VCVs) & noise (open-set sentences) ● Subjective measure: APHAB ● 2 conditions: conventional aid, experimental NFC (4–6 wk) ● Start freq. ranged from 1.0 to 1.6 kHz ● No evidence that audibility for the lowered speech was measured	● No significant difference in monosyllabic word & consonant recog in quiet ● Only 1 listener showed improvement in sentence recog in noise with NFC ● APHAB: higher global scores for 4 listeners with conventional aid

(*Continued*)

Table 1 (*Continued*)

Reference/Technique	Participant Details	Methods/Outcome Measures	Results
Glista et al (2009)[40]/NFC	13 adults & 11 children with sloping high-freq. SNHL ranging from moderately severe to profound	● Aided speech detection, speech recog for consonants, plurals, & vowels ● 3 phases: acclimatization to conventional processing, treatment with prototype NFC, treatment withdrawal (conventional processing)	● 5 adults had significant improvement, mostly for plurals ● 7 children had significant improvement, mostly for plurals ● Benefit for plurals depended on age group & high-freq. loss ● Children > adults preferred NFC ● 1 adult had worse vowel perception
Bohnert et al (2010)[47]/NFC	11 adults with severe to profound SNHL	● NFC compared with listeners' own device after 2 & 4 mo ● Start frequencies ≤ 2 kHz for 7 listeners ● SRT: nonsense sentences in noise ● Subjective questionnaires ● Comparison between devices not valid; also, no blinding to the treatment	● Reports of fricatives as sounding unnatural with NFC, despite better sound-quality ratings
Wolfe et al (2010)[78]/NFC	15 children with moderate to moderately severe SNHL	● NFC enabled versus NFC disabled (6-wk acclimatization period each) ● Aided speech detection, plural recog in quiet, sentence recog in multitalker babble, high-freq. consonant discrimination in quiet	● Plural recog significantly better with NFC (84% versus 99%) ● Improved consonant discrimination (/s/ & /d/) with NFC ● No significant difference for sentence recog in noise
Wolfe et al (2011)[25]/NFC	See Wolfe et al (2010)[78]	● 6-mo follow-up on Wolfe et al 2010 using the same measures ● No control or withdrawal condition, thus, cannot rule out maturation/learning effects	● Improved consonant discrimination with NFC after 6 mo versus 6 wk ● Improved sentence recog in noise only after 6 mo
McCreery et al (2012)[45]/NFC	20 adults with normal hearing	● Processing for 3 audiograms with varying degrees of high-freq. hearing loss ● Recog of fricatives & affricates (nonsense CVCs) ● 3 conditions (using the Purdue hearing aid simulator): conventional processing, manufacturer's default NFC settings, NFC settings with optimized bandwidth	● Speech recog improved across conditions as estimated audibility & bandwidth increased ● Optimized fitting method resulted in significantly higher speech recog compared with other conditions
Glista et al (2012)[51]/NFC	6 children with at least moderately severe high-freq. SNHL	● Aided speech detection, plural recog, /s/-/ʃ/ discrimination (CV pairs), consonant recog ● 3 phases: conventional processing (baseline), NFC (treatment ~ 16 wk), conventional processing (withdrawal) ● Single-subject design; cannot discount maturation effects	● Significant benefit (up to 20%) from NFC following acclimatization for 5 listeners on at least 1 test, especially fricatives; the other listener was near ceiling with & without NFC for most tests ● Improved plural recog (3 listeners) at varying times during treatment phase ● Improved /s/-/ʃ/ discrimination (3 listeners) over time ● Acclimatization trends were variable

Table 1 (*Continued*)

Reference/Technique	Participant Details	Methods/Outcome Measures	Results
Ellis & Munro (2013)[95]/NFC	15 adults with normal hearing	• Sentence recog in noise • Cognitive measures: working memory span (reading span test), executive control function (trail marking test) • 3 conditions: no NFC & 2 NFC with a 1.6 kHz start freq. (2:1 or 3:1 compression ration)	• NFC significantly decreased speech recog • Significant correlation between recog of unprocessed speech & 3 of 4 cognitive measures • No significant relationship between recog of NFC speech & cognitive measures
Kopun et al (2012)[97]/NFC	12 children & 24 adults with mild to severe SNHL	• Recog of monosyllabic words with & without NFC using Purdue hearing aid simulator, assessed before & after short-term audiovisual exposure with NFC	• NFC significantly improved word recog • No preferential effect for exposure or for age group • Only 2 adults performed worse with NFC
Brennan et al (2012)[48]/NFC	19 children & 24 adults with mild to severe SNHL	• Sound-quality ratings for speech (sentences with 3 fricatives) & for a variety of music genres • 3 conditions processed with Purdue hearing aid simulator: restricted bandwidth at 5 kHz, extended bandwidth at 10 kHz, & NFC	• Improved sound quality with extended bandwidth/NFC compared with restricted bandwidth for speech; no differences for music • Listeners' preference for extended bandwidth/NFC positively correlated with amplification experience, but no relationship with audiogram
Alexander (2012)[46]/NFC	28 adults with mild to moderately severe SNHL	• Recog of consonants & vowels from nonsense syllables in noise • Purdue hearing aid simulator used to create a control & 6 NFC conditions with fixed output bandwidth (2 groups at 3.3 & 5.0 kHz) but varying start freq. & compression ratio	• Improvement in fricative recog for most NFC settings relative to control • 1.6 kHz start freq. decreased vowel & nonfricative consonant recog (increasing start freq. was more effective at restoring recog than decreasing compression ratio)

Abbreviations: APHAB, Abbreviated Profile of Hearing Aid Benefit; C, consonant; Freq., frequency; HA, hearing aid; HL, hearing level; LFT, linear frequency transposition; NFC, nonlinear frequency compression; recog., recognition; SNHL, sensorineural hearing loss; SPL, sound pressure level; SRT, speech recognition threshold; V, vowel.

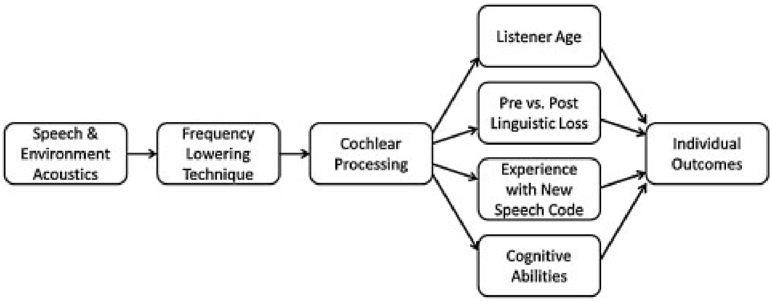


Figure 2 A framework for understanding factors that likely influence individual differences in speech recognition. Extrinsic factors affect the representation of the frequency-lowered speech at the auditory periphery, including the specific technique and the settings chosen for it, and intrinsic factors contribute to an individual's ability to learn and benefit from this signal.

- Frequency compression by Siemens of Erlangen, Germany (introduced 2012)

Linear Frequency Compression

One of the earliest commercially available FL techniques was linear frequency compression introduced by AVR Sonovation in analog devices in 1991 and later in digital devices in 2004. A key element of this technique is its switching behavior. A spectral balance detector is used at the front end of the processing to determine when FL should occur. Specifically, if the energy above 2,500 Hz is greater than that below 2,500 Hz, then FL occurs, otherwise amplification is provided as usual. Hence, FL is dynamic because of its all or none behavior. At least for speech in quiet, this is a good method for identifying the speech sounds characterized by high-frequency aperiodic energy. FL is implemented quite simply using a "slow play" effect; in fact, the early analog version used a magnetic tape that was played at a slow rate of speed. Digitally, this is done by using two different analog-to-digital converters (ADCs) at the frontend. The first, ADC1, has a sampling rate equal to the sampling rate of the digital-to-analog converter (DAC). The other, ADC2, has a programmable sampling rate that is an integer multiple (2, 3, 4, or 5) of the DAC sampling rate, called the dynamic frequency compression coefficient. The former, ADC1, is used when no lowering occurs (i.e., when the signal is low-frequency dominated, as with vowels) and the other is used for lowering (i.e., when the signal is high-frequency dominated, as with fricatives, etc.). Because the sampling rate of the DAC (output) is slower than the ADC2 (input), the entire frequency range is shifted proportionally lower by a factor equal to the dynamic frequency compression coefficient (Figs. 3 and 4).

Linear Frequency Transposition

In 2006, Widex introduced LFT to the market as the "Audibility Extender." Whereas AVR Sonovation hearing aids were primarily niche products, this was the first time that the concept of FL went mainstream. LFT was simply an optional add-on feature of an already fully developed line of products. When the LFT feature is activated, the algorithm continually searches for the most intense spectral peak in a limited frequency range known as the "source

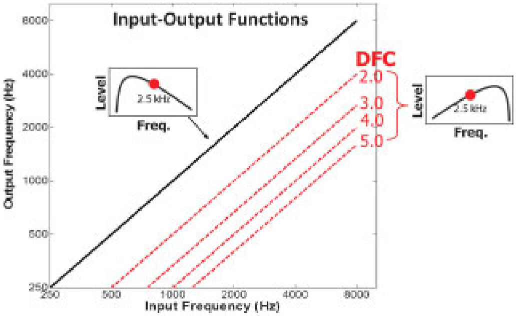


Figure 3 Relationship between input and output frequencies for linear frequency compression. When the input is dominated by energy below 2,500 Hz, no lowering occurs. When it is dominated by energy above 2,500 Hz, the entire frequency range is compressed by a factor determined by the dynamic frequency compression coefficient. Abbreviation: DFC, dynamic frequency compression.

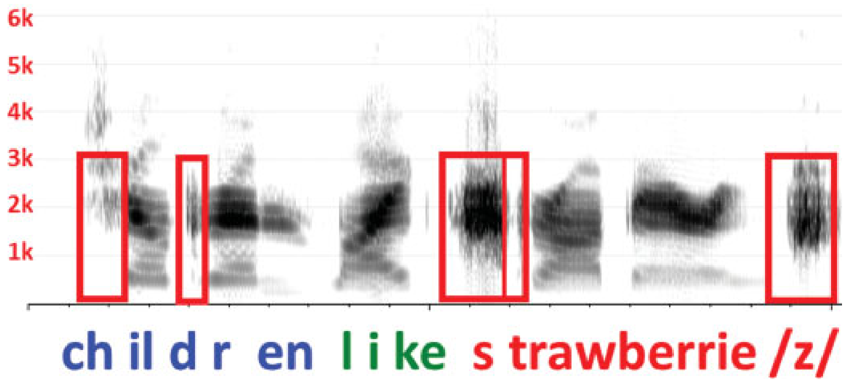


Figure 4 A spectrogram of the sentence "children like strawberries" from Fig. 1 after processing with linear frequency compression and a dynamic frequency compression coefficient of three. Energy from 0 to 8 kHz is compressed down to 0 to 2.67 kHz when the spectral balance of the input segment is high-frequency dominated. Boxes highlight the visually identifiable energy altered by the processing.

region." The source region is determined by a programmable "start frequency" that includes the one-third-octave band frequencies from 630 to 6,000 Hz. The source region begins a half octave below the start frequency and extends one octave above it or to the limit of the input bandwidth of the microphone and/or ADC. The frequency region to which the input is transposed is called the "target region" and is one octave below the source region (Table 2). An octave-wide band (relative to the target destination) is filtered around the dominant spectral peak in the source region and is then resynthesized one octave down (a factor of two), thus mixing with any low-frequency energy that might be present.[29] An optional expanded mode exists for start frequencies $\leq$ 2,500 Hz where the source region begins a half octave *above* the start frequency and extends for an additional octave (Table 3). As with the basic mode, an octave-wide band (relative to the target destination) is filtered around the dominant spectral peak in the source region, but is then resynthesized down by a factor of three instead of two. Hence, below the nominal start frequency there is a potential mixing of energy from the original input signal, the transposed signal from the basic mode, and the transposed signal from the expanded mode (Fig. 5A).

Table 2 The Approximate Source and Target Regions of the Basic Transposed Signals for Different LFT Start Frequencies (in Hz)

Start Freq.	Source Region	Target Region
630	445–1260	223–630
800	566–1600	283–800
1,000	707–2000	354–1000
1,250	884–2500	442–1250
1,600	1131–3200	566–1600
2,000	1414–4000	707–2000
2,500	1768–5000	884–2500
3,200	2263–6400	1131–3200
4,000	2828–8000	1414–4000
6,000	4242–(max)	2121–(max/2)

Abbreviations: Freq, frequency; LFT, linear frequency; Note: Actual values will depend on the audibility of the lowered signal. The value for "max" corresponds to the maximum frequency represented by the analog-to-digital converter.

Table 3 The Approximate Source and Target Regions of the Expanded Transposed Signals for Different LFT Start Frequencies (in Hz)

Start Freq.	Source Region	Target Region
630	891–1782	297–594
800	1131–2263	377–754
1,000	1414–2828	471–943
1,250	1768–3536	589–1179
1,600	2263–4525	754–1508
2,000	2828–5657	943–1886
2,500	3536–(max)	1179–(max/3)

Abbreviations: Freq, frequency; LFT, linear frequency; Note: Actual values will depend on the audibility of the lowered signal. The value for "max" corresponds to the maximum frequency represented by the analog-to-digital converter.

Unlike the previous technique, when activated, FL is continuous, although its behavior is dynamic because what is lowered and where depends on the spectral content of the input signal (Fig. 6).

Nonlinear Frequency Compression

In 2008, Phonak introduced their first hearing aids with NFC, known as "SoundRecover" to the market. In some respects, frequency compression as implemented by Phonak is opposite of frequency compression as implemented by AVR Sonovation. Whereas the latter is selective in time and is implemented linearly in the temporal domain (i.e., by exploiting sampling rate) across the entire frequency range, the former is frequency selective and is implemented nonlinearly in the spectral domain across time. That is, when activated, NFC is always operating, but it only does so over a limited analysis band that is determined by a programmable start frequency that ranges from 1,500 to 6,000 Hz. Frequencies below the start frequency do *not* undergo FL (Fig. 7). This is a key difference between how LFT and NFC are controlled by the handles within the Widex and Phonak programming software. With Widex, all the FL occurs *below* the start frequency whereas with Phonak, all the FL occurs *above* the start frequency.

With the first generation product (Naída UP), only frequencies up to 6.3 kHz were subject to frequency compression. With the second generation of products, a band about 4.5 to 4.8 kHz wide beginning just below the start frequency was subject to frequency compression, with an upper limit of 10 kHz. With the latest generation of products introduced in 2011, all frequencies beginning with the start frequency and continuing through 10 kHz undergo frequency compression. The relationship between input and output frequencies is determined by the compression ratio, which can vary from 1.5:1 to 4.0:1. As confirmed by calculations performed by the author, because FL occurs on a log scale, the compression ratio corresponds to the psychophysical reduction in spectral resolution in terms of auditory filters (e.g., 2.0:1 means that information that would normally span two auditory filters in the

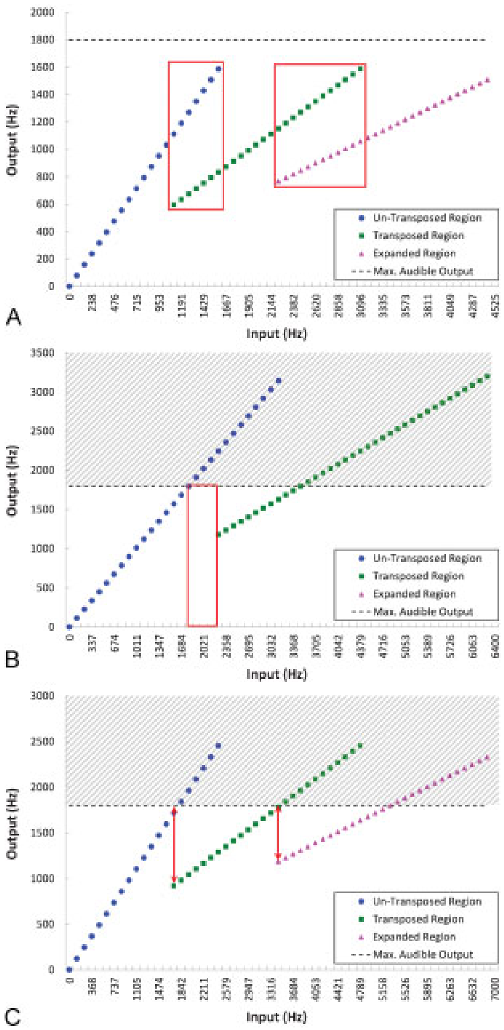


Figure 5 Relationship between input and output frequencies for three different start frequencies (different panels) for linear frequency transposition. The dotted line in each panel corresponds to a hypothetical loss in which the maximum audible output frequency with amplification is 1,800 Hz. The untransposed signal is represented by circles, the transposed signal by squares, and the expanded transposed signal by triangles. (A) The boxes indicate frequencies of overlap between the untransposed, transposed, and expanded transposed signals when the start frequency is 1,600 Hz. (B) The box indicates an island of intermediate input frequencies that are inaudible when the start frequency is 3,200 Hz. (C) The arrows indicate an ideal scenario when the start frequency is 2,500 Hz in which the input frequency corresponding to where audibility for the untransposed signal ends is close to the start frequency for transposition and where audibility for this signal ends is close to the start frequency for expanded transposition.

Figure 6 A spectrogram of the carrier sentence "children like strawberries" from Fig. 1 and the stimulus "eeSH" after processing with linear frequency transposition in which peak energy from ~ 4,242 to 7,000 Hz is transposed down to 2,121 to 3,500 Hz. Boxes highlight the visually identifiable energy altered by the processing.

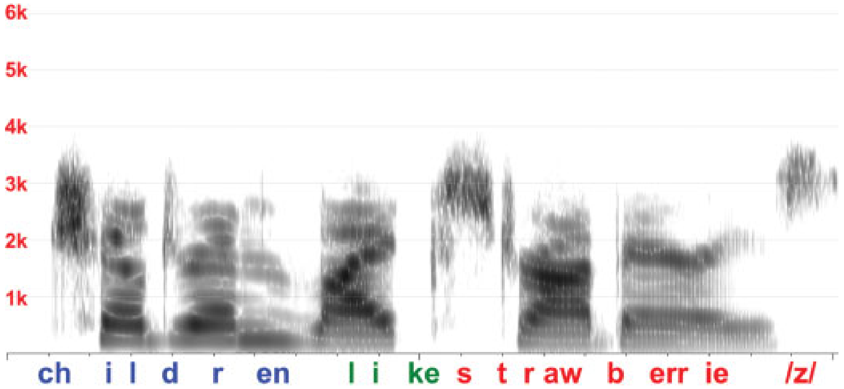


Figure 7 A spectrogram of the sentence "children like strawberries" from Fig. 1 after processing with nonlinear frequency compression in which the entire band from 1.5 to 6.0 kHz is compressed to 1.5 to 3.5 kHz. One consequence of a low start frequency that can be seen from the figure is that formant transitions become flattened compared with the input in Fig. 1.

unimpaired ear before processing will only span one auditory filter after processing). To prevent dispensers from being overwhelmed, options in the programming software are limited to 12 to 15 preset combinations of start frequency and compression ratio. The exact combinations depend on the hearing loss entered into the software. Shown in Fig. 8 is a family of frequency input-output curves that were measured empirically using a hearing aid with the second generation algorithm.

Siemens has recently implemented a form of FL simply known as "frequency compression."[30] Details about its signal processing are unknown at this time. Dispenser control over FL is different from SoundRecover in that two handles, f_{min} and f_{max}, determine the start and end frequencies of the target region, respectively. The compression ratio is then determined by the settings for the two handles, which have lower and upper limits of 1.5 and 8.0 kHz, respectively.

Spectral Envelope Warping

Starkey is also one of the latest companies to implement FL, known as "Spectral iQ," in its hearing aids. The algorithm is described as "spectral envelope warping." The term *spectral feature detection* is used in the algorithm description because a classifier looks for spectral features in the high-frequency spectrum that are characteristic of speech. The term *translation* has been used to describe the algorithm

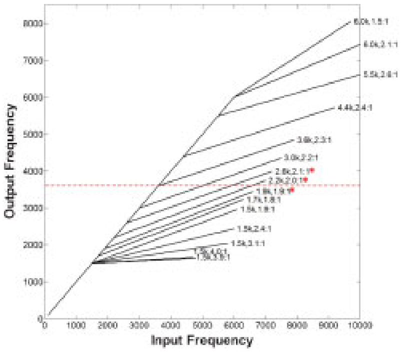


Figure 8 A family of frequency input-output curves from the second generation implementation of nonlinear frequency compression for one particular hearing loss. For each setting, the first number is the start frequency, followed by the compression ratio. The dotted line represents the maximum audible output frequency achievable with the hearing aid for a hypothetical hearing loss. Asterisks indicate the only appropriate settings for this hypothetical loss based on criteria outlined in the text (i.e., audibility for the lowered signal in a way that does not restrict the audible bandwidth of the processed signal).

behavior as these features are added to the low-frequency signal (Fig. 9) in a way that preserves their natural harmonic structure (Fig. 10). Another key feature is that, unlike other FL techniques, this technique does not roll off the high frequencies beyond the two upper channels (> 5.7 kHz). This is done to minimize the risk of the dispenser unintentionally limiting audible bandwidth by choosing too aggressive a FL setting, the importance of which will be discussed later. Dispenser control of Spectral iQ takes the form of a seven-point scale that corresponds to the bandwidth of the source region, with higher settings reserved for more severe hearing losses. A gain control is also provided; this control independently adjusts the level of the translated spectral feature.

Summary of Techniques

Fig. 11 provides a visual schematic for comparing the commercially available FL techniques. Frequency is represented by the color map—the lowest frequencies are red (bottom) and the highest frequencies are violet (top). The first bar shows the input band and indicates that the source region is the frequency range for which the hearing aid cannot provide sufficient audibility. The following four plots show how each technique approaches the goal of bringing down this information into a region of aided audibility.

Fig. 12 shows a classification of these techniques along two dimensions: algorithm activation (input dependent versus always active) and technique that serves as the basis for FL (compression versus transposition). Key characteristics of frequency compression are (1) the target region (the frequency range where information is moved to) is contained within the source region (the frequency range or analysis band that is subject to lowering), (2) the bandwidth of the source region is reduced, and (3) the start frequency (which can be 0 Hz) is like an anchor that does not move. Key characteristics of frequency transposition/translation are (1) there is less overlap between target and source regions, (2) the bandwidth of the source region is not reduced, and (3) the start frequency is moved to a lower frequency (i.e., there is a mixing of lowered and unlowered signals).

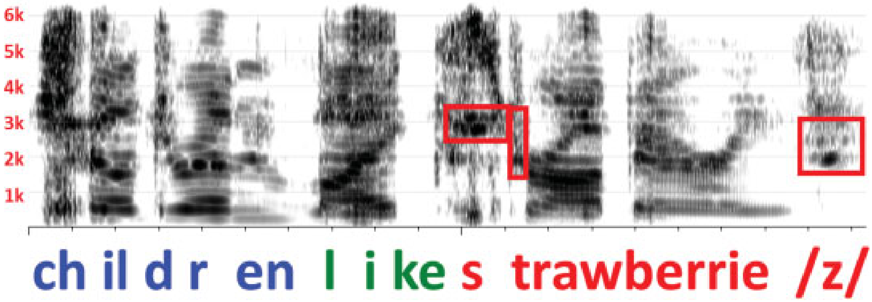


Figure 9 A spectrogram of the sentence "children like strawberries" from Fig. 1 after processing with spectral envelope warping. Boxes highlight the visually identifiable energy altered by the processing.

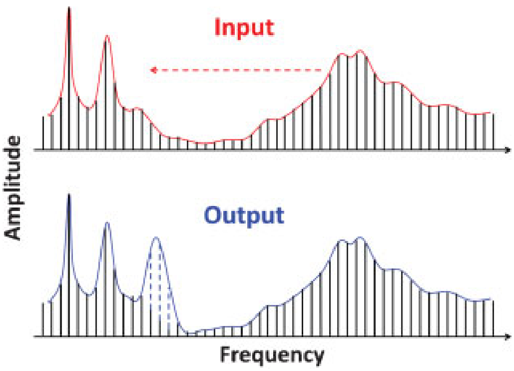


Figure 10 Author's rendition of how spectral envelope warping adds information from the high-frequency spectrum to the low-frequency spectrum while maintaining the harmonic structure of the source signal and the original high-frequency spectral content.

Modeling the Interaction between the Signal Processing, Hearing Loss, and Speech

Historically, attempts to implement FL have been limited to individuals with severe to profound SNHL.[29,31–35] For these individuals, arguments for the use of FL are relatively easy to make because deficits in speech recognition increase as high-frequency audibility decreases.[2,36–38] However, as the cutoff frequency of audibility decreases, the challenges involved with FL increase. Reasons for this include the fact that there is more *lost information* that needs to be recovered and a correspondingly smaller region for recoding it. In addition, the recoded information must be moved to regions where critical low-frequency speech information (i.e. formants) might already exist.

On the other hand, arguments justifying the use of FL for individuals with mild to moderate SNHL are more difficult to make because the overall deficit attributed to bandwidth reduction for these losses is less, which means that potential benefit is also less. In addition, early FL technology involved drastic alterations of the signal and possible audible artifacts.[33,39] As such, the likelihood of "doing more harm than good" was relatively high. Modern techniques and advances in signal processing have reduced some of these risks associated with FL and, in general, may be more amenable to listeners with milder degrees of SNHL, especially children who depend on the full bandwidth of speech for normal speech and language development (see following discussion).

As just highlighted, the relationship between SNHL and FL seems to be give and take. As the severity of loss increases, the deficit and corresponding potential benefit increase along with the risks associated with using FL.[35,40,41] With less severe loss, there is less deficit and potential benefit, but also less risk involved with FL.[42] This relationship can be better understood using the schematic in Fig. 13. On the abscissa is the amount of FL. Because the acoustic fidelity of the signal is radically altered by FL, the abscissa also is labeled "distortion."

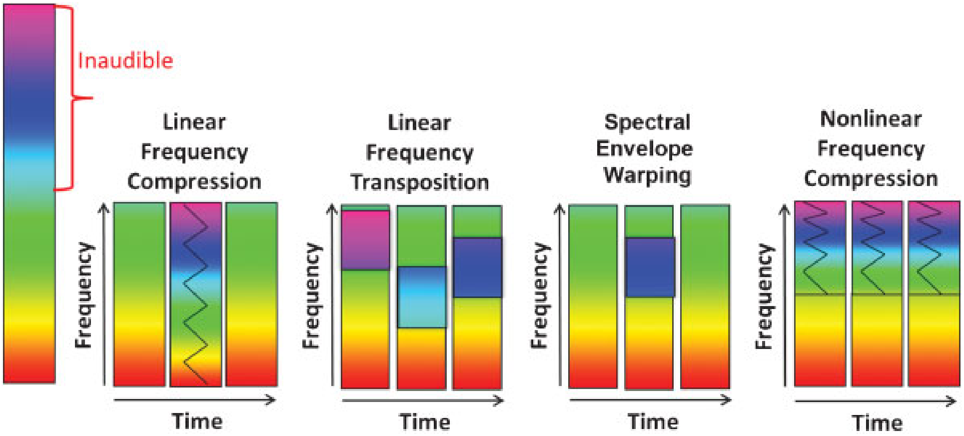
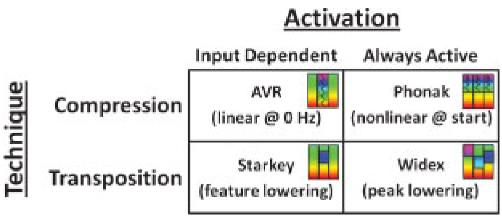


Figure 11 Visual schematic for comparing the four frequency-lowering techniques. See text for details.



Activation

Technique	Input Dependent	Always Active
Compression	AVR (linear @ 0 Hz)	Phonak (nonlinear @ start)
Transposition	Starkey (feature lowering)	Widex (peak lowering)

Figure 12 Classification of the frequency-lowering techniques. Inserts correspond to the plots shown in Fig. 11.

FL might be considered "constructive distortion" when it aids in speech recognition and "destructive distortion" when it does not,[43] so the term is appropriate either way. On the ordinate is information, which can be quantified using information-theoretic terms like *bits*, or units from the Speech Intelligibility Index,[44] and so on. The dashed line represents the potential information gained by moving otherwise inaudible high-frequency content to lower-frequency regions of audibility. Up to a certain point, the amount of high-frequency information increases with increases in FL, beyond which, information decreases as cochlear limitations take over (e.g., severe amounts of frequency compression). The line does not go below zero because it represents only that information that is inaccessible with conventional amplification, so presumably FL cannot make this information worse than not having it at all. The dotted line represents the potential decrease in information for that part of the spectrum that can be amplified normally. It never goes above zero because it is hard to imagine that FL can actually make this information better. For small degrees of FL, this information may be undisturbed; for example, when NFC has a high start frequency.[42,46] However, as FL increases, the information contained in the low-frequency spectrum becomes degraded. The magnitude of the line for the high-frequency content is intentionally less than that for the low-frequency content to respect the differences in the amount each contributes to overall speech recognition.[44] The absolute magnitude and the underlying form of each function will then shift depending on the severity of loss, the specific FL technology and settings, and the speech sounds involved.

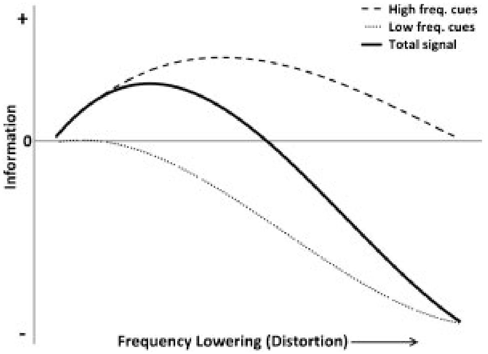


Figure 13 Schematic of how information in the speech signal is affected by the distortion introduced by frequency lowering. The dashed line represents the potential increase of information gained by moving otherwise inaudible high-frequency content to lower-frequency regions of audibility. The dotted line represents the potential decrease in information for that part of the spectrum that can be amplified normally. The solid black line represents the summation of information gained and information lost attributable to frequency lowering. It is hypothesized that the underlying forms of the functions will depend on the factors depicted in Fig. 2.

The thick solid line in Fig. 13 represents the summation of information gained and information lost attributable to FL. The goal of the hearing aid dispenser is to choose the FL setting that maximizes this function for the individual. Although one cannot know what the underlying function is, or where each setting is at on the function, probe microphone measurements or subjective listening tests can help guard against the two extremes.[40] The first extreme occurs when the FL information is moved to a region that is still inaudible for the listener. If no additional information is made audible, then no benefit should be expected. Although this seems obvious, confusion about the technology or a failure to verify aided audibility can result in this scenario. For example, as already mentioned, with Widex all the FL occurs below the start frequency whereas with Phonak all the FL occurs above the start frequency. If the understanding for the former were confused for the latter, it is likely that little to none of the FL information would be audible. Verification of audibility for frequency-lowered speech should also be a consideration when evaluating research outcomes as in Table 1. For example, the earliest reported investigations of NFC with adults did not report if or how audibility for the frequency-lowered speech was obtained, thereby limiting inferences or comparisons one might draw.[34,35]

The other extreme that must be guarded against is unintentionally limiting the audible bandwidth by choosing an overly aggressive FL setting. After all, the treatment should follow the principle of "do no harm." Fig. 13 illustrates the possible negative effects of FL. Shown by the solid black line, as FL is made more aggressive and a broader input bandwidth is affected, a patient's ability to extract information from the lowered signal decreases. Not only does the "information value" of the recoded high-frequency content decrease, but the information from low-frequency content is progressively degraded as well. The exact details will, of course, depend on the specific technology. Adults being fit with this FL will verbally object if they feel that they are losing too much information, but young children cannot. This risk of impeding information extraction has lead manufacturers to intentionally limit the adjustment parameters associated with FL and/or provide the dispenser with recommended candidacy guidelines. Regardless of who is being fit with FL hearing aids, a guideline the author recommends is to use probe microphone measurements to obtain the maximum audible frequency after fine-tuning the hearing aid with FL *deactivated*, and then do the same with FL activated using the settings under consideration to ensure that audible bandwidth is not limited by FL.

To assist the dispenser in choosing FL settings for NFC and LFT, the author has developed online tools that visually plot how frequencies are altered by the different settings (available at www.tinyURL.com). For NFC, the basic principles can be visualized in Fig. 8. For a given audiogram, only a limited number of combinations of start frequency and compression ratio are available to the dispenser. In this example, we will assume that the maximum audible frequency in the output that can be obtained with NFC deactivated is 3.6 kHz, as indicated by the dashed line. From the figure, it should be clear that only a select number of the available options, as indicated by the asterisks, avoid the two extremes just described. On the one extreme, none of the FL signal will be audible for start frequencies ≥ 3.6 kHz. On the other extreme, settings with start frequencies ≤ 1.7 kHz begin to restrict the audible bandwidth of the output. Of the remaining settings, the dispenser might reasonably choose the one that objectively maximizes the bandwidth of the input signal that is made available in the output after FL or the one that is subjectively most pleasing to the listener.[40,45] For LFT, in addition to these two extremes, one might want to consider how the information in the input is repackaged in the output. For example, Fig. 5A shows how a low start frequency might lead to less than optimal outcomes because too many frequencies in the input are over represented in the output due to the overlap between the untransposed, transposed, and expanded transposed signals. Fig. 5B on the other hand, shows the opposite in which a high start frequency can lead to an island of intermediate input frequencies that are still inaudible after FL. Fig. 5C shows an ideal scenario in which the input frequency corresponding to where audibility for an untransposed signal ends is close to the start frequency for transposition and where audibility for this signal ends is close to the start frequency for expanded transposition.

Some Data

Two studies by the author highlight how the choice of technology and its settings can influence outcomes for different hearing losses. Alexander et al investigated the efficacy of LFT and NFC using 24 adults with mild to moderate SNHL and 24 normal-hearing controls.[42] Participants listened monaurally through headphones to a series of nine fricatives and affricates spoken by three women in a vowel-consonant context that had been mixed with speech-shaped noise at 10-dB signal-to-noise ratio and recorded through hearing aids with LFT or NFC that were programmed for a mild to moderate loss such that FL occurred only for input frequencies ≥ 4 kHz. It should be noted that the uses of LFT and NFC for these losses were outside the recommended candidacy guidelines established by the manufacturers. Control stimuli included recordings made while FL was deactivated in each hearing aid (restricted bandwidth condition) and these same recordings mixed with high-pass filtered versions of the input stimuli, so that average audibility extended out to 9 kHz (wide bandwidth condition). Consistent with previous findings that demonstrated the importance of high-frequency information for fricative identification,[7,19] performance for the hearing-impaired listeners in the wide bandwidth conditions was significantly better than the restricted bandwidth conditions. Individual performance for conditions where NFC provided audibility for input frequencies up to 8 to 9 kHz was similar to the wide bandwidth conditions, indicating that the benefit observed with increasing bandwidth also can be obtained using NFC. In contrast, performance with LFT for both hearing-impaired and normal-hearing listeners was significantly worse compared with the restricted bandwidth condition by about 10 and 20%, respectively. Significant differences between most conditions could be largely attributed to an increase or decrease in confusions for the phonemes /s/ and /z/. Differences in outcomes between the two FL techniques might be attributed to the degree to which LFT altered the low-frequency spectrum compared with NFC in this particular population of listeners. As discussed later, it is possible that this difference would have become less following extended experience with the technology.

Using simulated hearing aid processing in MATLAB (i.e., "the Purdue hearing aid simulator") with NFC modeled after Simpson et al and flexible multichannel wide dynamic range compression, [34]Alexander investigated the effect of varying the start frequency and compression ratio for two fixed output bandwidths.[46] Twenty-eight listeners with mild to moderately severe SNHL identified consonants and vowels from nonsense syllables in noise. All speech output was low-pass filtered at 3.3 or 5.0 kHz across two groups of listeners to control for high-frequency thresholds when simulating two clinical scenarios whereby the dispenser has a variety of NFC options for repackaging different amounts of high-frequency information in a limited band of audibility. For both groups there was significant improvement in fricative/affricate identification for most NFC settings relative to the low-pass control conditions. However, when start frequency was low (1.6 kHz), there was a decrease in vowel and nonfricative consonant identification. Recognition of these sounds improved when the start frequency was increased ($\geq$ 2.2 kHz) even though the compression ratio also had to be increased to provide audibility to the same band of input frequencies, a process that reduced spectral resolution within the FL signal. Alternatively, when less compression was used with the low start frequency, vowel and nonfricative consonant identification was closer to that for the low-pass control, although this came at the expense of bringing less high-frequency fricative information down into the range of audibility (cf. Fig. 13). Overall, the results of this study indicate that many factors likely determine how much information individual listeners can extract from frequency-lowered speech, including the frequency regions altered by FL and the severity of loss in the regions to where information is moved.

INTRINSIC FACTORS RELATED TO THE ABILITY TO LEARN AND BENEFIT FROM FREQUENCY-LOWERED SPEECH

The second set of factors in Fig. 2 to be discussed are highly interconnected and relate to intrinsic characteristics of the individual listeners and their abilities to make use of the new and altered speech cues associated with FL. These factors include age of fitting, age of hearing loss onset, listening experience, and cognitive factors. The preceding section indicates that each FL strategy is a form of speech recoding that uniquely alters the information contained in individual speech sounds at the auditory periphery in a complex manner. For listeners to benefit from the recoded speech information, they must first learn how to interpret the new signal. For some FL settings, little to no learning may be necessary and immediate benefit can observed, even in the laboratory.[29,45,46] For example, if NFC has a high start frequency (e.g., > 3,500 Hz), FL will primarily affect information contained in the aperiodic high-frequency spectrum by making it narrower in bandwidth. Because the processed signal can have at least the same sound quality as conventional amplification,[47,48] listeners with SNHL may not even notice the alteration for most speech sounds. The potential for benefit is related to the extent that the new information made available with FL reduces uncertainty about the spoken message. The degree to which listeners have uncertainty in the first place will depend on the linguistic context of the message and the listeners' abilities to use this context. The role of linguistic context was discussed earlier. Listeners' abilities to use context to bootstrap understanding of a message fragmented by "information dropouts" (inaudibility of critical acoustic cues) depends on their knowledge of grammar, semantics, and pragmatics. This is where we might expect to see differences between children and adults. It also introduces a further distinction between whether the loss is pre- or postlinguistic, because listeners who have auditory experience with the natural productions of the recoded sounds might adapt more quickly to them.

For other FL settings, more implicit and/or explicit learning may be necessary before full benefit can be realized. These settings likely involve a manipulation of the primary formants of speech, which generally reside in the part of the spectrum below 3,500 Hz or so.[49] For these settings, altered sound quality might be observed along with new perceptual confusions.[29,40,48,51] Whether listeners can adapt to the altered sound quality is an important consideration because the technology will likely be rejected if they cannot. There is some indication that sound quality or preference ratings can improve for both adults and children over time with modern FL strategies.[50,52] Assuming that sound quality is not an issue, other factors that might determine benefit are listeners' experience with the technology, which can be gained implicitly through acclimatization and/or explicitly through training, and the listeners' underlying cognitive ability to apply the necessary effort involved with the perceptual learning process.

Children versus Adults

Children might experience greater benefit from FL compared to adults simply because they appear to have a greater "deficit" when identifying speech under identical conditions. In other words, in conditions where adults are performing near the ceiling of their abilities using conventional amplification, children might still be able to benefit from additional information gained via frequency-lowered speech. As mentioned, when linguistic knowledge confers an advantage, adult–child differences in speech recognition might be expected. For example, when identifying words or sentences in noise, children require more favorable signal-to-noise ratios for similar performance.[53–57] However, even when linguistic knowledge is of little apparent value, as when identifying nonsense syllables, children still underperform adults and require higher levels of audibility.[18,58–61] Reasons for this are beyond the scope of this article, but likely include differences in phonological development and phonotactic knowledge.[62] That is, adults have relatively stable categories for the various acoustic forms sounds in a language can take and have an implicit knowledge about the allowable sound sequences. This helps to restrict the range of possible response options even when nonwords are tested.[63] More relevant to this discussion are findings that suggest that children benefit more than adults when the availability of speech information is augmented by increases in signal bandwidth.[7,64] Whether this finding holds when the same information is made available via FL has yet to be demonstrated.

Only a small number of studies have tested adults and children using identical methodologies, which is an important evaluation criterion because different procedures, especially different methods for selecting individual FL settings, can significantly affect outcomes as noted previously. Glista et al reported outcomes for both children and adults with sloping high-frequency SNHL ranging from moderately severe to profound.[40] A withdrawal design was used in which performance following exposure to NFC for a minimum of 3 to 4 weeks was compared with terminal performance following exposure to conventional processing for another 4 weeks or so. Age group was a significant predictor of performance on a test of plural recognition, with 4 of 13 adults and 7 of 11 children showing significant improvement with FL. However, it is unknown whether this result is due to developmental differences or to the fact that children were given 5- to 10-dB more gain than adults per the Desired Sensation Level v5.0 prescriptive guidelines.[65]

Using the Purdue hearing aid simulator, Kopun et al tested adults and children with mild to severe SNHL on monosyllabic word recognition before and after 22 minutes of audio-visual exposure to two children's stories that had been processed with customized amplification and NFC processing.[97] They found that both children and adults performed significantly better before and after exposure with NFC compared to without. Exposure did not provide additional benefit for NFC because performance also improved for the processing condition without NFC. Importantly, they did not find a significant difference in benefit for children and adults.

In a study that exclusively examined sound-quality judgments for speech and music following processing using the Purdue hearing aid simulator with NFC and without NFC at two bandwidths, Brennan et al did not find preference differences between adults and children with mild to severe SNHL.[48] For speech stimuli, both age groups preferred wide bandwidth processing (10,000 Hz) or NFC more often than restricted bandwidth processing (5,000 Hz). This preference was positively correlated with amplification experience, but there was no relationship with the audiogram. For music stimuli, no significant differences were found between the three conditions. It should be noted that the start frequency of NFC was $> 3{,}000$ Hz for most listeners, which limited the processing to frequencies that do not contribute much to pitch. These results are somewhat in contrast to those of Glista et al,[40] who found that children had a greater preference for NFC than adults, and to those of Auriemmo et al,[41] who noted that children more often preferred LFT than adults after initial fitting, but not after 2 weeks of usage due to increases in preference by the adults.[29]

In summary, despite several valid reasons to expect adult–child differences in benefit and/or preference for FL, supporting evidence is not available within studies that used a matched experimental design across age groups.

Age of Hearing Loss Onset

Discussion of child–adult differences implies more than auditory and linguistic development. Most of the adults in the studies described in the preceding section had acquired losses (namely presbycusis and noise exposure), and most of the children had congenital losses (e.g., genetic causes, meningitis, hyperbilirubinemia, anoxia at birth, etc.). Differences in etiology have implications for the degree of involvement of the various cochlear structures (e.g., outer versus inner hair cell loss) and the audiometric configuration—factors discussed in the first part of this article. For example, Pittman and Stelmachowicz analyzed almost 500 audiograms from a clinical population and found that children compared with adults had a greater variety of audiometric configurations other than the classic sloping loss, had greater variability in thresholds across frequency, and had a greater prevalence of asymmetrical hearing losses.[66] As discussed, audiometric differences influence the amount of unaided deficit for individual speech sounds and the choice of FL technology and its settings.

Another factor to consider in this discussion is the role of prelinguistic versus postlinguistic hearing loss. It has been documented that young children who are deprived of rich auditory input because of untreated SNHL or unaidable high-frequency SNHL often have delayed phonological development compared with their normal-hearing peers.[22–25,67,68] In other words, the acoustic properties that help to categorize one sound as being different from another are less defined (acoustic-phonetic boundaries are more variable), resulting in perceptual confusions and expressive deficits. It is unknown what effect this could have on perception of frequency-lowered speech relative to conventional hearing aid processing, but the amount of deficit and the ability to learn the new speech code are likely mediating factors. For example, listeners with better phonological representations of the phonemes /s/ and /z/ due to prior auditory experience might have fewer deficits with conventional hearing aids, hence less room for improvement as performance nears the ceiling. These listeners might also be able to learn to use the new and altered speech cues more quickly, whereas those with poorer phonological representations might require an extended period to acclimatize and/or require explicit training.

Acclimatization and Training

Perceptual acclimatization refers to the *process* by which individuals adapt to altered sensory input to maintain optimum performance in their environment. For speech perception, the effects of acclimatization are often gauged by improvement in recognition scores over time,[69–76] but they also could include gradual improvements in rated sound quality, decreased listening effort, and so on.[77] The latter are often overlooked, but are important to consider because initial reactions to frequency-lowered speech, which contribute to the overall experience, might exhibit even greater change over time compared with conventional amplification because of the increased amount of signal alteration. These considerations are mentioned briefly elsewhere in this article.

Two studies[25,51] that have explicitly examined acclimatization effects for frequency-lowered speech without training have been conducted on children. Both studies reported significant improvements in speech recognition following extended experience listening through hearing aids with NFC. Wolfe et al presented results from a 6-month follow-up on 15 young children (ages 5 to 13 years) with moderate to moderately severe SNHL who were initially tested after two counterbalanced 6-week intervals in which they listened through the same hearing aids with NFC activated or deactivated during their daily routines.[25,78] Discrimination of singular/plural contrasts in quiet, which tests perception of /s/ and /z/, and discrimination of high-frequency consonants embedded in nonsense vowel-consonant-vowel utterances in quiet were significantly improved by NFC after 6 weeks. Sentence recognition in multitalker babble did not improve during this time, however. When tested again at 6 months, plural discrimination maintained at ceiling performance levels and there were continued improvements in the discrimination of high-frequency consonants relative to testing at 6 weeks. Interestingly, there was significant improvement the signal-to-noise ratio threshold for sentence recognition that was not present after 6 weeks. These results indicate that some children may undergo an acclimatization period lasting at least 6 months for NFC. However, maturation and/or learning effects cannot be ruled out as contributing factors in the Wolfe et al study because there was no control condition or withdrawal condition whereby listeners were tested following an acclimatization period with NFC deactivated.[25]

To minimize the likelihood of maturation effects, Glista et al tested 6 older children (ages 11 to 18 years) with at least moderately severe high-frequency SNHL in a single-subject design on phoneme detection, plural recognition, discrimination between /s/ and /ʃ/, and consonant recognition at regular intervals over a 4-month period.[51] One listener did not show improvement over time due to ceiling effects. For the other five listeners, they found significant improvement (up to 20%) with NFC on at least one test, especially with fricatives, although the pattern of improvement over time was quite variable within and between listeners. Therefore, the authors caution about extrapolating from individual results and warn that maturation effects are still possible, although unlikely.

Explicit auditory training is another way, and in some cases may be the only way, that listeners can get the experience they need to learn the new and altered speech cues introduced by FL. The concept of training on frequency-lowered speech goes back to the earliest techniques described in the literature, where it was more commonplace than modern techniques (see Simpson[28] for a review). One reason for this shift in approaches may be that the modern techniques described in the recent literature are usually implemented in wearable devices, which provides listeners with an opportunity to acclimatize to the processing. The only exceptions that incorporated formal training procedures utilized LFT.[29,41]

Auriemmo et al tested 10 children (ages 6 to 13 years) with severe to profound SNHL on consonant and vowel recognition in quiet before and after a 3-week control period during which they wore a hearing aid with LFT deactivated during their daily routines.[41] Then they repeated the same testing after two additional 3-week treatment periods during which LFT in the hearing aid was activated. Throughout the control and treatment phases of the experiment, listeners attended weekly auditory training sessions. The main findings were that during the 6-week period when LFT was activated, consonant recognition at a low presentation level (30-dB hearing level) significantly improved by more than 20% and the production accuracy of /s/ and /z/ improved by more than 10%. There was no significant difference in vowel recognition when LFT was activated versus deactivated because performance was near ceiling levels during initial testing for both conditions. Because listeners received training along with their daily exposure to LFT, it cannot be determined if the observed effects were due to acclimatization and/or to training.

Kuk et al tested eight adults with severe to profound SNHL on consonant recognition in quiet and in babble noise, first with LFT deactivated and activated after the initial fit, then after 1 month of daily exposure and training ($\sim$ 20 to 30 minutes per a day), and then after another month of daily exposure but without training.[29] A feature analysis of the confusions revealed that the main source of improvement with LFT was fricatives. At a presentation level of 50-dB sound pressure level (SPL), fricative recognition improved 5 to 10% initially and then an additional 10% after 2 months. At a presentation level of 68-dB SPL, a significant difference was only seen for fricatives, which improved by 10% after 2 months. At this level, there was also a decrease in recognition of stop consonants after the initial fit, which significantly improved after 2 months to performance levels slightly better than baseline. For speech in noise, the only significant difference was about a 15% improvement for fricatives after 2 months of training. Like the Auriemmo et al study, one cannot determine if the observed effects were due to acclimatization and/or to training, because listeners received training along with their daily exposure to LFT.[41] Furthermore, because there was no control group that trained on the same hearing aid with LFT deactivated, one cannot determine if the observed main effects are due to the additional information introduced by LFT or to simple practice effects. To address this concern, a brief follow-up report by Kuk and Keenan indicates that there were no training effects per se when using the hearing aids with LFT deactivated.[79]

In summary, the few reports that have examined the effects of extended listening experience with frequency-lowered speech support the notion that at least some listeners require a period of acclimatization and/or training with the new speech code before full benefit can be realized. One can speculate that the more drastic the alteration of speech cues following FL, namely involvement of low-frequency speech cues, the greater the need for extended or explicit listening experience for the listener to learn how to most effectively process the newly introduced and altered information.

Cognitive Factors

There has been recent interest in understanding how cognitive processes influence outcomes for hearing aid algorithms other than FL (see Lunner et al[80] for a review). Cognitive processing has been implicated in at least two ways. One way is as a moderating variable (e.g., verbal working memory) that influences who benefits and under what circumstances, as with wide dynamic range compression.[81–85] Another way is as a dependent variable; that is, the construct by which benefit is measured (e.g., in terms of listening effort, recall, reaction time), as with digital noise reduction.[86–88] As noted in the previous paragraphs, FL techniques and settings that alter more of the natural speech code may require a learning period consisting of exposure with or without training to achieve maximum potential. On the one hand, if learning puts a high demand on cognitive resources, then FL might not be beneficial or might even be detrimental to those listeners who have fewer resources available, such as the elderly.[89–93] On the other hand, if FL can provide additional information that helps reduce uncertainty about the spoken message without the need for learning, then it might be able to reduce the cognitive load associated with listening in real-world scenarios, thereby benefiting these same listeners in terms of listening effort, recall, comprehension, multitasking, and so on.[92]

Two recent studies have explored the relationship between cognitive processing and recognition of speech with NFC.[94,95] Using older adults (ages 62 to 92 years) with mild to severe sloping SNHL, Arehart et al[94] measured working memory and recognition of sentences in noise that had been processed with a hearing aid simulator that implemented a form of NFC similar to that described by Simpson et al.[34] The signal processing differed in that only the most intense peaks in the high-frequency spectrum were compressed, instead of the entire high-frequency spectrum. NFC settings were intentionally aggressive with start frequencies as low as 1,000 Hz because the primary purpose was to obtain correlates for speech recognition of distorted speech, not FL per se. As the amount of distortion increased, speech recognition was more adversely affected in listeners with the poorest working memory abilities compared with those with better working memory abilities. In contrast, Ellis and Munro[95] used similar test materials and the same cognitive test as Arehart et al[94] but did not find a significant relationship between the working memory of young normal-hearing adults (ages 18 to 50 years) and recognition of speech with an NFC start frequency of 1,600 Hz. Interestingly, they found a significant relationship between working memory and recognition of the unprocessed speech, which indicates that the variation in working memory abilities for this nonclinical population should have been sufficiently large to capture an effect for NFC processing had there truly been one. Further research is clearly needed to better understand the role that cognition plays in processing frequency-lowered speech and how this influences candidacy for this technology.

CONCLUSION

After reviewing all the studies mentioned in this article, it should be clear that differences in outcomes with FL vary along with the dependency of the test materials and individual listeners on information in the high-frequency spectrum. Accordingly, when FL shows a benefit for speech recognition it is primarily for fricative consonants with restricted linguistic context. Not all of the individuals in the studies reviewed showed improvement in speech recognition with FL. For most of these, speech recognition with FL was not statistically different from speech recognition without FL, which might offer some assurance from a "do no harm" perspective. However, for some listeners, FL negatively affected speech recognition. In light of the model presented in Fig. 13, one might question whether the settings in these cases actually optimized the total information in the signal. In addition, full benefit might not be realized until after several months of experience with the technology, especially for speech in noise.

Although the focus of measuring benefit has been on speech recognition, one should be open to the possibility that benefit from FL might be less evident in terms of decreased listening effort, improved speech production and vocal quality, improved localization and spatial unmasking, and so on, even when there are no observable improvements in speech recognition as measured in the laboratory.

Finally, the reader is reminded of the need to (1) understand the specific FL technology before implementing it (i.e., what frequencies are lowered, where are they lowered to, when are they lowered, and how are they lowered), (2) verify electroacoustically that FL moves additional speech information to a region of audibility, but does not unduly limit the audible bandwidth of the entire signal, and (3) validate outcomes to ensure effectiveness for the individual.

ACKNOWLEDGMENTS
Varsha Hariram helped prepare the references and the information in Table 1, although any errors contained therein are solely the responsibility of the author.

REFERENCES

1. Hogan CA, Turner CW. High-frequency audibility: benefits for hearing-impaired listeners. J Acoust Soc Am 1998;104:432–441
2. Ching TY, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. J Acoust Soc Am 1998; 103:1128–1140

3. Turner CW, Cummings KJ. Speech audibility for listeners with high-frequency hearing loss. Am J Audiol 1999;8:47–56
4. Baer T, Moore BCJ, Kluk K. Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. J Acoust Soc Am 2002;112(3 Pt 1): 1133–1144
5. Vickers DA, Moore BCJ, Baer T. Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. J Acoust Soc Am 2001;110:1164–1175
6. Boothroyd A, Medwetsky L. Spectral distribution of /s/ and the frequency response of hearing aids. Ear Hear 1992;13:150–157
7. 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
8. Fox RA, Nissen SL. Sex-related acoustic changes in voiceless English fricatives. J Speech Lang Hear Res 2005;48:753–765
9. Mann VA, Repp BH. Influence of vocalic context on perception of the /š/-/s/ distinction: spectral factors. Percept Psychophys 1980;28:213–228
10. Whalen DH. Effects of vocalic formant transitions and vowel quality on the English [s]-[ŝ] boundary. J Acoust Soc Am 1981;69:275–282
11. Heinz JM, Stevens KN. On the properties of voiceless fricative consonants. J Acoust Soc Am 1961;33:589–596
12. Nittrouer S, Studdert-Kennedy M. The role of coarticulatory effects in the perception of fricatives by children and adults. J Speech Hear Res 1987;30:319–329
13. Zeng FG, Turner CW. Recognition of voiceless fricatives by normal and hearing-impaired subjects. J Speech Hear Res 1990;33:440–449
14. Nittrouer S. Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. J Phonetics 1992;20: 351–382
15. Nittrouer S, Miller ME. Predicting developmental shifts in perceptual weighting schemes. J Acoust Soc Am 1997;101:2253–2266
16. Nittrouer S, Miller ME. Developmental weighting shifts for noise components of fricative-vowel syllables. J Acoust Soc Am 1997;102:572–580
17. Hedrick MS. Effect of acoustic cues on labeling fricatives and affricates. J Speech Lang Hear Res 1997;40:925–938
18. Pittman AL, Stelmachowicz PG. Perception of voiceless fricatives by normal-hearing and hearing-impaired children and adults. J Speech Lang Hear Res 2000;43:1389–1401
19. Stelmachowicz PG, Lewis DE, Choi S, Hoover B. Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. Ear Hear 2007;28:483–494
20. Pittman AL. Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. J Speech Lang Hear Res 2008;51: 785–797
21. Hornsby BWY, Johnson EE, 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 Hear 2011;32:543–555
22. Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE. Aided perception of /s/ and /z/ by hearing-impaired children. Ear Hear 2002;23: 316–324
23. Moeller MP, Hoover BM, Putman CA, et al. Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I—phonetic development. Ear Hear 2007;28:605–627
24. Moeller MP, McCleary E, Putman C, Tyler-Krings A, Hoover B, Stelmachowicz P. Longitudinal development of phonology and morphology in children with late-identified mild-moderate sensorineural hearing loss. Ear Hear 2010;31: 625–635
25. Wolfe J, John A, Schafer EC, et al. Long-term effects of non-linear frequency compression for children with moderate hearing loss. Int J Audiol 2011;50:396–404
26. Rudmin F. The why and how of hearing /s/. Volta Review 1983;85:263–269
27. Elfenbein JL, Hardin-Jones MA, Davis JM. Oral communication skills of children who are hard of hearing. J Speech Hear Res 1994;37: 216–226
28. Simpson A. Frequency-lowering devices for managing high-frequency hearing loss: a review. Trends Amplif 2009;13:87–106
29. Kuk F, Keenan D, Korhonen P, Lau CC. Efficacy of linear frequency transposition on consonant identification in quiet and in noise. J Am Acad Audiol 2009;20:465–479
30. Serman M, Hanneman R, Kornagel U. White paper: Micon frequency compression. Siemens AG, 2012. Available at: http://hearing/siemens.com/Resources/literatures/Global/publications/2012%20-%20white-paper%20mocon%20frequcny%20compression.pdf?_blob=publicationfile. Accessed on March 19, 2013
31. Braida LD, Durlach NI, Lippmann RP, Hicks BL, Rabinowitz WM, Reed CM. Hearing aids—a review of past research on linear amplification, amplitude compression, and frequency lowering. ASHA Monogr 1979;19:1–114
32. Parent TC, Chmiel R, Jerger J. Comparison of performance with frequency transposition hearing aids and conventional hearing aids. J Am Acad Audiol 1998;9:67–77

33. McDermott HJ, Dorkos VP, Dean MR, Ching TYC. Improvements in speech perception with use of the AVR TranSonic frequency-transposing hearing aid. J Speech Lang Hear Res 1999;42: 1323–1335
34. Simpson A, Hersbach AA, McDermott HJ. Improvements in speech perception with an experimental nonlinear frequency compression hearing device. Int J Audiol 2005;44:281–292
35. Simpson A, Hersbach AA, McDermott HJ. Frequency-compression outcomes in listeners with steeply sloping audiograms. Int J Audiol 2006;45: 619–629
36. Turner CW, Robb MP. Audibility and recognition of stop consonants in normal and hearing-impaired subjects. J Acoust Soc Am 1987;81:1566–1573
37. Dubno JR, Dirks DD, Ellison DE. Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. I: The contribution of selected frequency regions. J Acoust Soc Am 1989;85:347–354
38. Horwitz AR, Dubno JR, Ahlstrom JB. Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing. J Acoust Soc Am 2002;111(1 Pt 1):409–416
39. Kuk F, Korhonen P, Peeters H, Keenan D, Jessen A, Andersen H. Linear frequency transposition: extending the audibility of high frequency information. Hearing Review 2006;13:42–48
40. Glista D, Scollie S, Bagatto M, Seewald R, Parsa V, Johnson A. Evaluation of nonlinear frequency compression: clinical outcomes. Int J Audiol 2009; 48:632–644
41. Auriemmo J, Kuk F, Lau C, et al. Effect of linear frequency transposition on speech recognition and production of school-age children. J Am Acad Audiol 2009;20:289–305
42. Alexander JM, Lewis DE, Kopun JG, McCreery RW, Stelmachowicz PG. Effects of frequency lowering in wearable devices on fricative and affricate peroeption, 2008 International Hearing Aid Conference, August 13–16, 2008; Lake Tahoe, CA
43. Kuk F, Jessen A, Baekgaard L. Ensuring high-fidelity in hearing aid sound processing. Hearing Review 2009;16:34–43
44. ANSI. ANSI S3.5–1997 (R2007), Methods for Calculation of the Speech Intelligibility Index. New York, NY: American National Standards Institute
45. McCreery RW, Brennan MA, Hoover B, Kopun J, Stelmachowicz PG. Maximizing audibility and speech recognition with nonlinear frequency compression by estimating audible bandwidth. Ear Hear 2012;20:1–4
46. Alexander JM. Nonlinear frequency compression: balancing start frequency and compression ratio. Paper presented at: 39th Annual meeting of the American Auditory Society; March 8–10, 2012; Scottsdale, AZ
47. Bohnert A, Nyffeler M, Keilmann A. Advantages of a non-linear frequency compression algorithm in noise. Eur Arch Otorhinolaryngol 2010;267:1045–1053
48. Brennan M, McCreery R, Kopun J, Hoover B, Stelmachowicz PG. Signal processing preference for music and speech in adults and children with hearing loss, 39th Annual meeting of the American Auditory Society, March 8–10, 2012; Scottsdale, AZ
49. Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. J Acoust Soc Am 1995;97(5 Pt 1):3099–3111
50. Kuk F, Peeters H, Keenan D, Lau C. Use of frequency transposition on in thin-tube, open-ear fittings. Hearing Journal 2007;60:59–63
51. Glista D, Scollie S, Sulkers J. Perceptual acclimatization post nonlinear frequency compression hearing aid fitting in older children. J Speech Lang Hear Res 2012;55:1765–1787
52. Nyffeler M. Study finds that non-linear frequency compression boosts speech intelligibility. Hearing Journal 2008;61:22–26
53. Nittrouer S, Boothroyd A. Context effects in phoneme and word recognition by young children and older adults. J Acoust Soc Am 1990;87:2705–2715
54. Gravel JS, Fausel N, Liskow C, Chobot J. Children's speech recognition in noise using omni-directional and dual-microphone hearing aid technology. Ear Hear 1999;20:1–11
55. Fallon M, Trehub SE, Schneider BA. Children's use of semantic cues in degraded listening environments. J Acoust Soc Am 2002;111(5 Pt 1):2242–2249
56. Hall JW III, Grose JH, Buss E, Dev MB. Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. Ear Hear 2002;23:159–165
57. Bonino AY, Leibold LJ, Buss E. Release from perceptual masking for children and adults: benefit of a carrier phrase. Ear Hear 2013;34:3–14
58. Stelmachowicz PG, Hoover BM, Lewis DE, Kortekaas RWL, Pittman AL. The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. J Speech Lang Hear Res 2000;43:902–914
59. Scollie SD. Children's speech recognition scores: the Speech Intelligibility Index and proficiency factors for age and hearing level. Ear Hear 2008; 29:543–556
60. Nishi K, Lewis DE, Hoover BM, Choi S, Stelmachowicz PG. Children's recognition of American English consonants in noise. J Acoust Soc Am 2010;127:3177–3188

61. McCreery RW, Stelmachowicz PG. Audibility-based predictions of speech recognition for children and adults with normal hearing. J Acoust Soc Am 2011;130:4070–4081
62. Hnath-Chisolm TE, Laipply E, Boothroyd A. Age-related changes on a children's test of sensory-level speech perception capacity. J Speech Lang Hear Res 1998;41:94–106
63. Edwards J, Beckman ME, Munson B. The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. J Speech Lang Hear Res 2004;47:421–436
64. Kortekaas RWL, Stelmachowicz PG. Bandwidth effects on children's perception of the inflectional morpheme /s/: acoustical measurements, auditory detection, and clarity rating. J Speech Lang Hear Res 2000;43:645–660
65. Scollie S, Seewald R, Cornelisse L, et al. The Desired Sensation Level multistage input/output algorithm. Trends Amplif 2005;9:159–197
66. Pittman AL, Stelmachowicz PG. Hearing loss in children and adults: audiometric configuration, asymmetry, and progression. Ear Hear 2003;24: 198–205
67. Jerger S, Lai L, Marchman VA. Picture naming by children with hearing loss: I. Effect of semantically related auditory distractors. J Am Acad Audiol 2002;13:463–477
68. Eisenberg LS. Current state of knowledge: speech recognition and production in children with hearing impairment. Ear Hear 2007;28:766–772
69. Cox RM, Alexander GC. Maturation of hearing aid benefit: objective and subjective measurements. Ear Hear 1992;13:131–141
70. Gatehouse S. The time course and magnitude of perceptual acclimatization to frequency responses: evidence from monaural fitting of hearing aids. J Acoust Soc Am 1992;92:1258–1268
71. Gatehouse S. Role of perceptual acclimatization in the selection of frequency responses for hearing aids. J Am Acad Audiol 1993;4:296–306
72. Bentler RA, Niebuhr DP, Getta JP, Anderson CV. Longitudinal study of hearing aid effectiveness. I: Objective measures. J Speech Hear Res 1993;36: 808–819
73. Turner CW, Humes LE, Bentler RA, Cox RM. A review of past research on changes in hearing aid benefit over time. Ear Hear 1996;17(3, Suppl): 14S–25S
74. Cox RM, Alexander GC, Taylor IM, Gray GA. Benefit acclimatization in elderly hearing aid users. J Am Acad Audiol 1996;7:428–441
75. Horwitz AR, Turner CW. The time course of hearing aid benefit. Ear Hear 1997;18:1–11
76. Saunders GH, Cienkowski KM. Acclimatization to hearing aids. Ear Hear 1997;18:129–139
77. Bentler RA, Niebuhr DP, Getta JP, Anderson CV. Longitudinal study of hearing aid effectiveness. II: Subjective measures. J Speech Hear Res 1993;36: 820–831
78. Wolfe J, John A, Schafer EC, Nyffeler M, Boretzki M, Caraway T. Evaluation of nonlinear frequency compression for school-age children with moderate to moderately severe hearing loss. J Am Acad Audiol 2010;21:618–628
79. Kuk F, Keenan D. Frequency transposition: training is only half the story. Hearing Review 2010; 17:38–46
80. Lunner T, Rudner M, Rönnberg J. Cognition and hearing aids. Scand J Psychol 2009;50:395–403
81. Gatehouse S, Naylor G, Elberling C. Linear and nonlinear hearing aid fittings—1. Patterns of benefit. Int J Audiol 2006;45:130–152
82. Lunner T, Sundewall-Thorén E. Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. J Am Acad Audiol 2007; 18:604–617
83. Cox RM, Xu J. Short and long compression release times: speech understanding, real-world preferences, and association with cognitive ability. J Am Acad Audiol 2010;21:121–138
84. Foo C, Rudner M, Rönnberg J, Lunner T. Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. J Am Acad Audiol 2007;18:618–631
85. Masterson KM, Alexander JM. Factors influencing release from masking with fast vs. slow compression. 38th Annual meeting of the American Auditory Society, March 10–12, 2011; Scottsdale, AZ
86. Sarampalis A, Kalluri S, Edwards B, Hafter E. Objective measures of listening effort: effects of background noise and noise reduction. J Speech Lang Hear Res 2009;52:1230–1240
87. Pittman A. Children's performance in complex listening conditions: effects of hearing loss and digital noise reduction. J Speech Lang Hear Res 2011;54:1224–1239
88. Pittman A. Age-related benefits of digital noise reduction for short-term word learning in children with hearing loss. J Speech Lang Hear Res 2011; 54:1448–1463
89. Humes LE. Factors underlying the speech-recognition performance of elderly hearing-aid wearers. J Acoust Soc Am 2002;112(3 Pt 1):1112–1132
90. Humes LE. The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. J Am Acad Audiol 2007;18:590–603
91. Humes LE, Floyd SS. Measures of working memory, sequence learning, and speech recognition in the elderly. J Speech Lang Hear Res 2005;48:224–235

92. Schneider BA. How age affects auditory-cognitive interactions in speech comprehension. Audiol Res 2011;1(e10):34–39

93. Gosselin PA, Gagné J-P. Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. Int J Audiol 2011; 50:786–792

94. Arehart KH, Souza P, Baca R, Kates JM. Working Memory, Age, and Hearing Loss: Susceptibility to Hearing Aid Distortion. Ear Hear 2013;(Jan):3

95. Ellis RJ, Munro KJ. Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? Int J Audiol 2013;52:14–22

96. Smith J, Dann M, Brown M. An evaluation of frequency transposition for hearing-impaired school-age children. Deafness Educ Int 2009; 11:62–82

97. Kopun J, McCreery R, Hoover B, Spalding J, Brennan M, Stelmachowicz P. Effects of Exposure on Speech Recognition with Nonlinear Frequency Compression, 39th Annual meeting of the American Auditory Society, March 8–10, 2012; Scottsdale, AZ