INTERNATIONAL JOURNAL OF AUDIOLOGY
2020, VOL. 59, NO. 7, 524–533
https://doi.org/10.1080/14992027.2020.1761035

ORIGINAL ARTICLE

Effects of nonlinear frequency compression on Mandarin speech and sound-quality perception in hearing-aid users

Xueqing Chen[a,b], Yanyan You[a,b], Jing Yang[c], Jinyu Qian[d,e], Qian Lu[f], Volker Kuehnel[g], Julia Rehmann[g], Bo Liu[a,b] and Li Xu[h]

[a]Beijing Tongren Hospital, Capital Medical University, Beijing, China; [b]Key Laboratory of Otolaryngology - Head and Neck Surgery, Beijing Institute of Otolaryngology, Capital Medical University, Ministry of Education, Beijing, China; [c]Department of Communication Sciences and Disorders, University of Wisconsin - Milwaukee, Milwaukee, WI, USA; [d]Innovation Center Toronto, Sonova Canada, Mississauga, ON, Canada; [e]Department of Communicative Disorders and Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA; [f]Sonova China, Shanghai, China; [g]Sonova AG, Stäfa, Switzerland; [h]Division of Communication Sciences and Disorders, Ohio University, Athens, OH, USA

ABSTRACT

Objective: The purpose of the present study was to examine the effects of NLFC fitting in hearing aids and auditory acclimatisation on speech perception and sound-quality rating in hearing-impaired, native Mandarin-speaking adult listeners.

Design: Mandarin consonant, vowel and tone recognition were tested in quiet and sentence recognition in noise (speech-shaped noise at a +5 dB signal-to-noise ratio) with NLFC-on and NLFC-off. Sound-quality ratings were collected on a 0–10 scale at each test session. A generalised linear model and correlational analyses were performed.

Study sample: Thirty native Mandarin-speaking adults with moderate-to-severe sensorineural hearing loss were recruited.

Results: The hearing-impaired listeners showed significantly higher accuracy with NLFC-on than with NLFC-off for consonant and sentence recognition and the recognition performance improved with both NLFC-on and off as a function of increased length of use. The satisfaction score of sound-quality ratings for different types of sounds significantly increased with NLFC-on than with NLFC-off. The speech recognition results showed moderate to strong correlation with the unaided hearing thresholds.

Conclusion: For native Mandarin-speaking listeners with hearing loss, the NLFC technology provided modest but significant improvement in Mandarin fricative and sentence recognition. Subjectively, the naturalness and overall preference of sound-quality satisfaction judgement also improved with NLFC.

ARTICLE HISTORY
Received 29 October 2019
Revised 20 April 2020
Accepted 21 April 2020

KEYWORDS
Hearing aids; speech perception; nonlinear frequency compression; sound quality

Introduction

Most listeners with sensorineural hearing loss have limited or no access to high-frequency information (Agrawal, Platz, and Niparko 2008). The high frequencies (i.e. 3 kHz and above) contain a substantial amount of speech/linguistic information (e.g. Stelmachowicz et al. 2004). The inability to access a full range of high-frequency information causes great challenges for hearing-impaired listeners to understand everyday conversations, particularly in noisy environments (Healy and Yoho 2016). Hearing aids have been the most commonly adopted method of intervention for sensorineural hearing loss. However, conventional hearing aids provide limited benefit for several reasons. Firstly, the available frequency bandwidth of hearing aids is limited. Conventional processing technologies used in hearing aids amplify the sound energy but fail to provide sufficient high-frequency information beyond 6 kHz due to either a limited gain or detrimental acoustical feedbacks (Pittman et al. 2003). Thus, a number of consonant sounds such as /s/ and /z/ are filtered out. Second, the physiological audible bandwidth of the cochlea can be limited due to the existence of dead regions within the cochlea. When the hearing threshold is > 90 dB HL, the hair cells in the cochlear region are likely to be damaged (Moore and Alcantara 2001). There might be little benefit in amplifying frequencies in the dead region (Vickers, Moore, and Baer 2001; Salorio-Corbetto, Baer, and Moore 2017).

Given the crucial role of high frequencies in speech perception, frequency-lowering techniques have been recently proposed and implemented. It aims to shift the higher-frequency components from the incoming signals to a lower-frequency region where hearing is relatively less impaired. To date, various frequency-lowering schemes have been developed and adopted in commercial hearing aids (see Simpson 2009; Alexander 2013; Mao et al. 2017; Akinseye, Dickinson, and Munro 2018 for reviews). One earlier effort to lower frequency is linear frequency compression in which all frequencies of the signal are subject to the same amount of compression. In contrast, to preserve the acoustic structure of the lower frequency in a signal, a more recent approach is nonlinear frequency compression (NLFC) in which only the higher frequency region of the signal is compressed and the lower frequency region is unchanged. Phonak has implemented an NLFC algorithm, named SoundRecover (SR), in its hearing aids. The SR algorithm involves two adjustable parameters: cut-off frequency (CT) and compression ratio (CR). CT divides the input signal into two parts and determines the start point of compression. The frequencies below the CT remain uncompressed while the frequencies above the CT are

CONTACT Li Xu xul@ohio.edu Communication Sciences and Disorders, Ohio University, Athens OH 45701, USA

nonlinearly compressed. The higher frequencies are compressed to a greater extent than the lower frequencies. The strength of compression is specified by the CR (see also Yang et al. 2018). In the SR algorithm, the compressed and non-compressed portions of the signal have no spectral overlap, which preserves the first-formant (F1) and most of the second-formant (F2) frequency (Yang et al. 2018). On the other hand, a potential disadvantage of the NLFC is that the compressed high-frequency could lead to changes in harmonic spacing, spectral peak levels and shapes, which might affect the sound quality (McDermott 2011).

So far, a handful of studies have been conducted to compare the perceptual performance on consonants, vowels and sound quality with NLFC and with conventional processing in English-speaking hearing-impaired listeners (Glista et al. 2009; McCreery et al. 2014; Hopkins et al. 2014; Alexander, Kopun, and Stelmachowicz 2014; Alexander 2016; Alexander and Rallapalli 2017; Wolfe et al. 2010, 2011, 2017; Parsa et al. 2013; Picou, Marcrum, and Ricketts 2015; Brennan et al. 2014). Little research has been done to evaluate the efficacy of NLFC in hearing-impaired listeners from other language backgrounds. To fill this gap, the present study focussed on the application of NLFC on a Mandarin-speaking population. China has an estimated 28.7 million people with a "hearing-disability" [defined as having pure-tone average hearing threshold of 0.5, 1, 2 and 4 kHz (PTA) > 40 dB HL], many of whom can potentially benefit from hearing aids. Research findings on hearing aid techniques based on the English language may not be readily applicable to Mandarin-speaking listeners because Mandarin has its own distinct language system and phonetic structure. Mandarin Chinese is a tonal language in which four types of pitch contours are used to differentiate lexical meanings. In addition, Mandarin differs from English in the segmental level. Mandarin has a simple syllable structure in the form of (C)(V)V(V)(N). Mandarin has a smaller inventory of monophthong vowels (i.e. /a, i, u, y, ɤ/) but a larger number of diphthongs and triphthongs than English. The corner vowels (i.e. /a, i, u/) represent the most peripheral articulatory positions that also define the lowest and highest formant values of the vowels. The F1 and F2 ranges of Mandarin vowels are 0.3–1.1 kHz and 0.7–3 kHz, respectively, which are similar to those of English vowels (Zee and Lee 2001; Zee 2001; Yang et al. 2018). More importantly, Mandarin has three-way contrasts of sibilant fricatives /s, ɕ, ʂ/ and affricates /ts, tɕ, tʂ/ and /tsʰ, tɕʰ, tʂʰ/. For listeners with high-frequency hearing loss, the three-way-contrast high-frequency sounds likely cause more difficulties than do the two-way-contrast sibilant fricatives in English. Tseng et al. (2018) tested the recognition of monosyllables, sentences, and sound quality preferences in 14 native Mandarin-speaking listeners with moderate to severe hearing loss under unaided, NLFC aided, and extended-bandwidth NLFC aided conditions. The extended-bandwidth NLFC expanded the conventional NLFC input frequency of 2–6 kHz to 2–10 kHz. The authors found that all participants showed greatly improved perceptual outcomes with hearing aids than without the aids. Moreover, the extended-bandwidth NLFC provided a greater improvement in word and consonant recognition. Extending from Tseng et al. (2018) that had a relatively small number of hearing-impaired patients, the present study aimed to test the efficacy of NLFC on various aspects of auditory perception in a relatively larger group of hearing-impaired, native Mandarin-speaking listeners. Previous studies reported the beneficial role of auditory acclimatisation to NLFC processing (Glista, Scollie, and Sulkers 2012; Wolfe et al. 2011). In the present study, we were also interested in understanding whether auditory acclimatisation with NLFC benefits speech recognition and sound quality rating in hearing-impaired native Mandarin-speaking listeners. In addition, listener characteristics, such as hearing threshold, slope of hearing loss, age of hearing loss onset, chronological age and cognitive performance play important roles in determining the speech recognition outcomes and magnitude of benefit of NLFC in hearing aids (Ellis and Munro 2015; Shehorn, Marrone, and Muller 2018). In the present study, correlational analyses were performed between the unaided hearing thresholds and speech recognition performance with or without NLFC in the hearing aids.

Materials and methods

Participants

The participants included 30 native-Mandarin-speaking adults (23 males and 7 females) all recruited from Beijing, China. The age of the participants ranged from 43 to 83 years (mean ± SD: 67.93 ± 8.17 years). Ninety percent (27/30) of the participants were between 61 and 78 years of age. Figure 1 shows the individual and group mean pure-tone air-conduction thresholds. Bone-conduction thresholds (not shown in Figure 1) were all within 10 dB of those of the air-conduction thresholds. All participants had sloping sensorineural hearing loss. Of the 30 participants, the mean PTA of both ears were between 40 and 70 dB HL (i.e. moderate-to-severe sensorineural hearing loss) except one participant who showed a mean PTA of 73 dB HL (i.e. severe sensorineural hearing loss). The hearing loss of both ears was symmetrical (i.e. interaural difference ≤ 15 dB at all octave frequencies between 0.25 and 8 kHz) in 23 of the 30 participants. The other seven participants had interaural differences of 20 or 25 dB at one or two audiometric frequencies. The duration of hearing loss ranged from 1 to 12 years, with a mean of 6.30 years. No participants had experience with hearing aids prior to participating in the study. All participants spoke Mandarin Chinese in their daily communications. The use of human subjects was reviewed and approved by the Institutional Review Boards of Ohio University and Beijing Tongren Hospital.

Hearing-aid fitting

All participants were fitted bilaterally with the Phonak Bolero Q50-P BTE or Bolero Q50-SP BTE hearing aids according to their required amplification needs. Occluding ear-moulds were used to ensure that the participants obtained as much access to acoustic speech information as possible within the limits of clinical prescriptive gain targets. The hearing aids were programmed using Phonak's Target (v. 4.1) programming software. The NLFC (i.e. SR) parameters of the hearing aids were prescribed by the fitting software based on the participants' hearing loss configurations. The settings in hearing aids were programmed with the APDT (Adaptive Phonak Digital Tonal). The NLFC parameters (CT and CR) as well as the advanced features (digital noise reduction, reverberation tail suppression, impulse noise reduction, directionality, as well as automatic programme selectors) of the device were set at default (CT = 4.3 to 6.0 kHz and CR = 1.5 to 3.2). A trained audiologist performed all the fitting procedures.

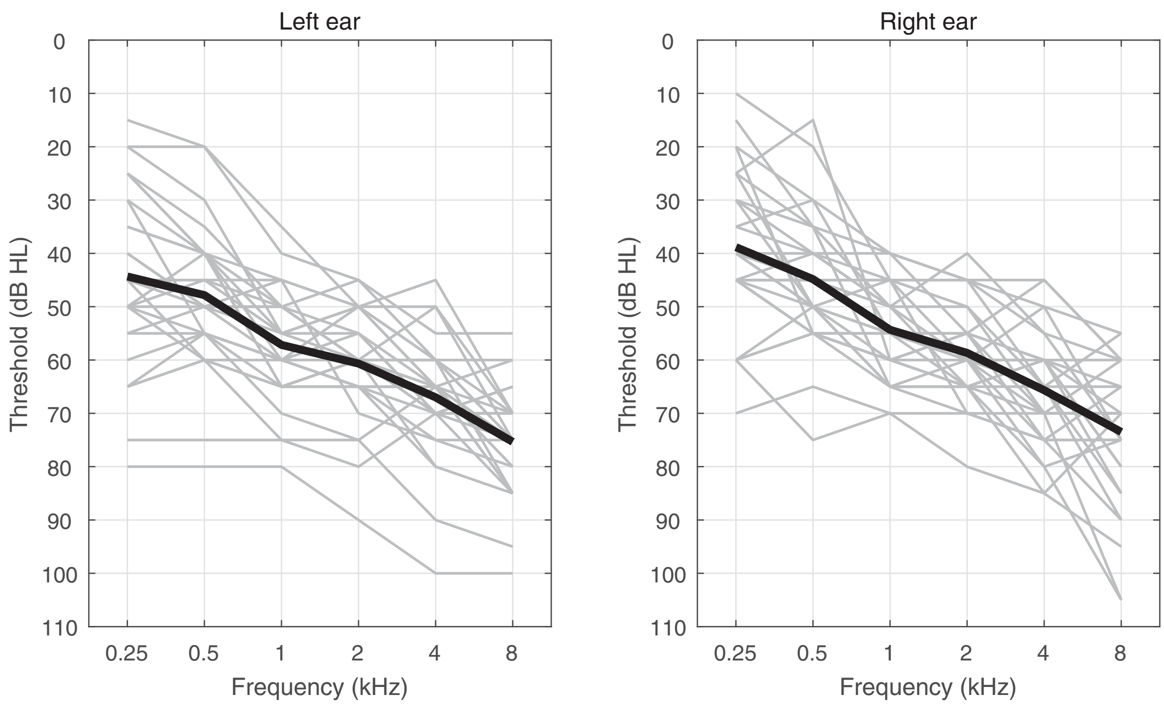

Figure 1. Individual and group mean pure-tone threshold for the 30 participants. Each thin grey line represents one participant. The thick black line denotes the group mean threshold ($n = 30$).

Consonant recognition

The consonant recognition test included five Mandarin fricatives /f/, /x/, /s/, /ɕ/ and /ʂ/ embedded in a/Ca/syllable in tone 1 (i.e. "fā" 发, "hā" 哈, "sā" 撒, "xiā" 虾, and "shā" 沙). The tokens were recorded from 10 native adult Mandarin speakers (5 males and 5 females). The mean F0 for the five male and five female speakers were 0.158 kHz (ranging from 0.137 to 0.184 kHz) and 0.266 kHz (ranging from 0.252 to 0.290 kHz), respectively. The test included two presentations of the 50 tokens, for a total of 100 tokens (5 words × 10 speakers× 2 presentations). A graphical user interface (GUI) was built in MATLAB programming environment to present the consonant recognition test in a 5-alternative forced-choice paradigm. In the GUI, five buttons labelled with the Chinese characters and the pinyin (i.e. phonemic spellings) were shown on a computer screen. After listening each consonant stimulus, the subjects were required to indicate what they had heard by pointing and clicking on one of the five buttons using a computer mouse. The order of the presentations was randomised. The intensity of the presentation was set at 65 dB SPL.

Vowel recognition

The Mandarin vowel list consisted of 20 Mandarin monosyllabic words in a /dV/ syllable structure in tone 1. The 20 words included the following vowels or vowels followed by nasal endings: /a/, /ɤ/, /i/, /u/, /aɪ/, /ɑʊ/, /oʊ/, /iɛ/, /iɑʊ/, /ioʊ/, /uo/, /ueɪ/, /an/, /ɑŋ/, /əŋ/, /iɛn/, /iŋ/, /uan/, /un/, and /uŋ/ (i.e. "dā" 搭, "dē" 嘚, "dī" 低, "dū" 督, "dāi" 呆, "dāo" 刀, "dōu" 兜, "diē" 跌, "diāo" 雕, "diū" 丢, "duō" 多, "dūi" 堆, "dān" 丹, "dāng" 当, "dēng" 灯, "diān" 颠, "dīng" 丁, "duān" 端, "dūn" 吨, and "dōng" 东). Sixteen of these vowels were selected from the vowel perception test in Li et al. (2014) and the remaining four vowels (/ɤ, iɛ, an, iɛn/) were added to make the vowel repertoire more complete. The tokens were recorded from the same 10 native adult Mandarin speakers. Considering the large number of words for the vowel recognition test, tokens from randomly selected three male speakers and three female speakers were used for the vowel recognition test. Therefore, there were a total number of 120 tokens for the vowel recognition test (20 words × 6 speakers). A different GUI was built to present the vowel recognition test. Except that vowel recognition test was in a 20-alternative forced-choice paradigm, the design of the GUI and the procedures were similar to those of the consonant recognition test. The order of the presentations was randomised and the intensity of the presentation was set at 65 dB SPL.

Tone recognition

The stimuli for the tone recognition test consisted of 10 syllables (/fu/ "fu", /tɕi/ "ji", /ma/ "ma", /tɕʰi/ "qi", /uan/ "wan", /ɕi/ "xi", /ɕiɛn/ "xian", /iɛn/ "yan", /iaŋ/ "yang", and /i/ "yi") each in four tones. The tokens were produced by one male and one female native Mandarin speaker selected from the above 10 speakers. The mean F0 for the male and female speakers were 0.137 and 0.257 kHz, respectively. The durations of the four tones of each syllable were equalised to the mean duration of the four tones of each syllable using the method of PSOLA (pitch synchronous overlap and add, Charpentier and Stella 1986). The tone recognition test contained 80 tokens (i.e. 2 speakers × 10 syllables × 4 tones). A different GUI was built to present the tone recognition test that used a 4-alternative forced-choice paradigm. Because 10 different syllables were used in tone recognition test, after each response, the GUI would refresh the screen, display a new list of four choices, and present the next stimulus. The order of the presentations was randomised and the intensity of the presentation was set at 65 dB SPL.

Sentence recognition

The stimuli used for sentences recognition were Mandarin Hearing in Noise Test (M-HINT; Wong et al. 2007) presented with speech-shaped noise at a +5-dB signal-to-noise ratio. M-HINT contains 12 lists. Each list consists of 20 sentences and each sentence is made of 10 Chinese characters. The presentation level was set at 65 dB SPL. The speech-shaped noise, generated by filtering a white noise to the long-term average speech spectrum of the M-HINT sentences (Soli and Wong 2008), was added to achieve a +5-dB signal-to-noise ratio in order to avoid the ceiling effect for speech recognition. The masking noise started 500 ms before the sentence and ended 500 ms after the sentence. During the test, for each participant, one sentence list was randomly selected and the order of the sentences in each list was randomly presented. The participants were required to verbally repeat the sentence that they had heard.

Sound-quality rating

The stimuli used for sound-quality rating included own voice, male voice, female voice, bird chirps and music. Own voice is a familiar stimulus to everyone including those with postlingual hearing loss. Therefore, any alterations and unfamiliar perception are more easily identified compared to other stimuli. Maintaining good perception of own voice is also a goal for any hearing aid signal processing (e.g. Bohnert, Nyffeler, and Keilmann 2010). The male and female voices were from recorded text by native Mandarin speakers. The text was a paragraph of 127 Chinese characters and the lengths of the recordings were 34 s and 36 s for the male and female voices, respectively. The bird chirps were originally provided in MATLAB which contained 8 chirps that were typically downward sweeps with frequency dropping from approximately 4 to 2 kHz in roughly 90 ms. We duplicated the bird chirps three times so that the final stimuli contained 24 chirps and lasted 5.5 s. The music stimuli were a recorded piano performance of a classic piece of Chinese folk music entitled “Liang Zhu” (The Butterfly Lovers) that was known to all participants. The duration of the music stimuli was 105 s. The five types of sound samples were played in a random order to each listener. No repetition of the sound stimuli was allowed. Listeners were requested to rate the loudness, clearness, naturalness, as well as overall sound quality of each stimulus after they finished listening to each individual sample on a 0 to 10 scale with 0 and 10 being “extremely poor” and “perfect”, respectively.

Procedures

Evaluation was performed at three time intervals for each participant. The first evaluation was performed immediately after the hearing aids were fitted. Then, follow-up evaluations were conducted 2–3 weeks and 12 weeks post fitting, respectively. The participants had worn the NLFC-equipped hearing aids bilaterally for the entire 12-week period. On the evaluation day, the participants were tested in two conditions: NLFC-on and NLFC-off. The order of the test conditions (NLFC-on or off) was randomised across subjects. All tests were administered through a custom MATLAB programmed GUI and were conducted in a sound booth. The test stimuli were presented through a loudspeaker mounted 1 m in front of the participant at 0° azimuth.

For each participant at each test session, a speech recognition test was conducted first and was followed by the sound-quality rating. The test order of speech recognition for consonant, vowel, tone and sentence stimuli was randomised. Before the real test, a short practice session was provided in order to familiarise the participants with the test procedures. For consonant, vowel and tone recognition, the practice session used randomly selected 20, 20 and 16 tokens from the test stimuli, respectively. For sentence recognition, five M-HINT sentences (different from those used in the real test) were used for practice. Feedback was provided during the practice session.

Data analysis

Data analysis was performed in MATLAB with the Statistics Toolbox. The percent-correct scores of the speech recognition test were treated as binomial data (Thornton and Raffin 1978). Following a logit transformation of the percent-correct data, a generalised linear model (GLM; Warton and Hui 2011) was used to examine the effects of (1) NLFC condition (NLFC-on or NLFC-off) and (2) NLFC acclimatisation on the percent-correct scores. For the sound-quality rating, the data were treated as normal distribution and no transformations were applied. A GLM model for each category of percept (loudness, clarity, naturalness and overall preference) was then used to examine the effects of (1) NLFC condition, (2) NLFC acclimatisation and (3) type of sound involved (i.e. own voice, male voice, female voice, bird chirp and music) on the sound-quality rating scores.

Results

Figure 2 shows the performance of consonant, vowel, tone and sentence recognition tests with NLFC-off and NLFC-on conditions. The individual data are shown in the upper panels and the group mean data are shown in the bottom panels (Figure 2). The statistical results are summarised in Table 1. Large variability was evident in the speech recognition performance. Correlational analyses indicated that such variability was not related to chronological age of the participants. Nonetheless, we could not rule out that that cognitive decline due to aging might have affected the results.

There was a significant improvement in consonant, vowel and sentence recognition, with test session for both NLFC-on and NLFC-off, but not in tone recognition. Meanwhile, both consonant and sentence recognition improved significantly with NLFC on compared to that with NLFC off in all three sessions. Further, there was not a significant interaction between test sessions and NLFC on/off conditions for any recognition test. The lack of interaction effects suggested that the improvement as a function of test session might reflect listeners’ adaptation to the test stimuli rather than the NLFC itself.

When comparing the recognition performance of individual consonants, it is worth noting that not all fricatives were recognised with the same accuracy rate or improved to the same extent. Figure 3 presents the confusion matrices of individual fricatives in each session with NLFC-off and NLFC-on. Among the five fricatives, /x/ in “ha” had the highest recognition accuracy for all three sessions while /s/ in “sa” had the lowest recognition accuracy and showed substantial confusion with /f/ in “fa” and /ʂ/ in “sha”. With NLFC on relative to NLFC off, the two fricatives /f/ and /s/ showed relatively greater improvement in recognition accuracy from the first test session to the third test session than the other fricatives. It is noteworthy that when NLFC was off, the participants also showed observable

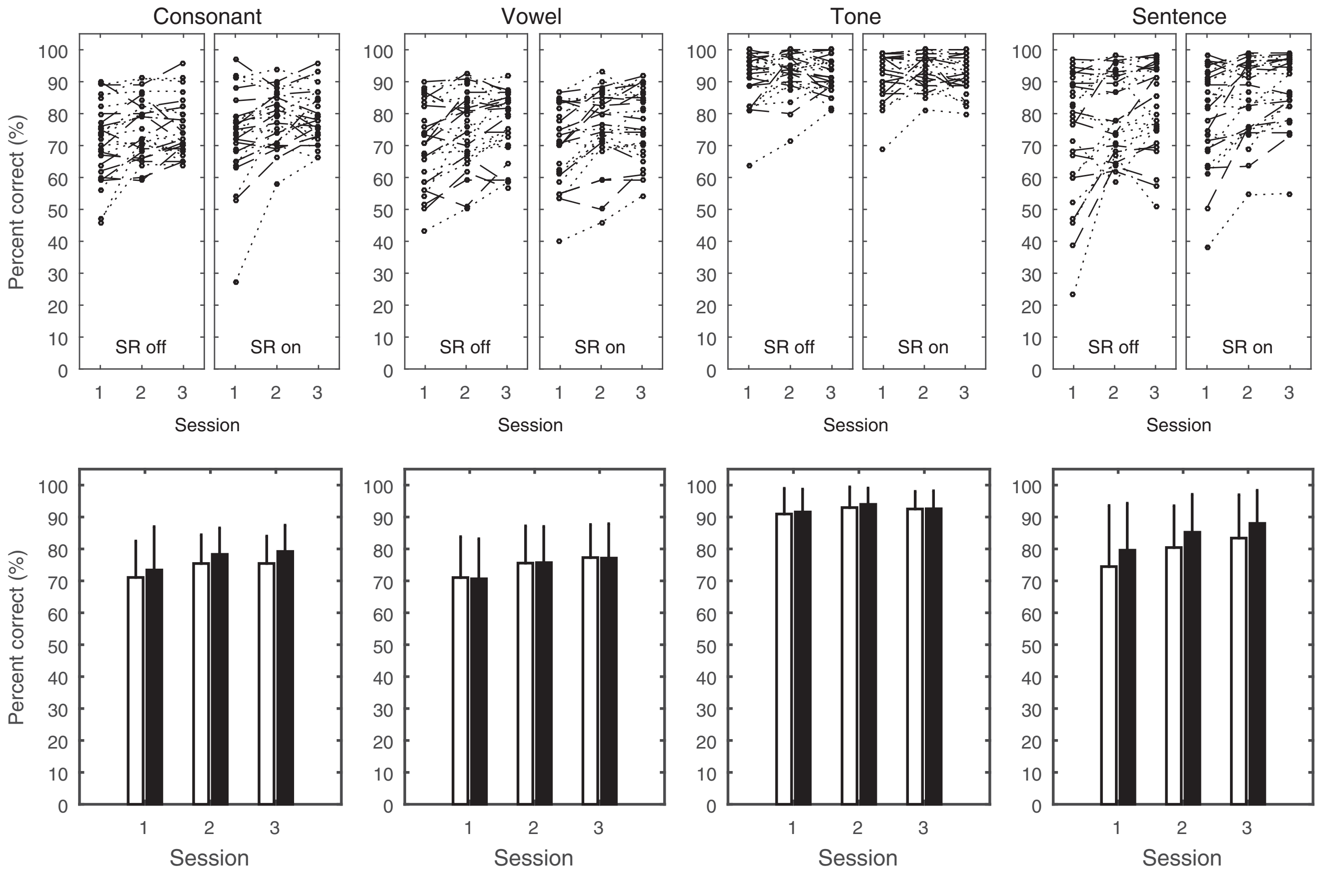

Figure 2. Speech recognition performance. Upper panels show the individual recognition scores. The four sets of two panels represent consonant, vowel, tone, sentence recognition performance in NLFC-on and NLFC-off conditions, respectively. In each panel, data from each individual for sessions 1, 2 and 3 (i.e. 0, 2–3 weeks and 12 weeks post-fitting) are plotted with a line. Lower panels show the group mean and standard deviation of speech recognition performance. The four panels represent data from consonant, vowel, tone and sentence recognition tests, respectively. The white and black bars represent scores under the NLFC-off and NLFC-on conditions, respectively. The error bars represent 1 SD. The statistical results related to speech recognition performance can be found in Table 1.

Table 1. Summary of GLM results for the factors of session and NLFC-on/off conditions on the performance of each speech perception task.

	Session			NLFC on/off			Interaction		
	β	*t*	*p* Value	*β*	*t*	*p* Value	*β*	*t*	*p* Value
Consonant	0.127	2.81	**0.005**	−0.120	−2.40	**0.016**	−0.034	−1.21	0.228
Vowel	0.100	2.50	**0.013**	0.009	2.50	0.843	0.0004	0.02	0.989
Tone	−0.004	−0.05	0.960	−0.116	−1.28	0.200	−0.116	0.45	0.651
Sentence	0.230	6.11	**<0.0001**	−0.295	−7.70	**<0.0001**	−0.032	−1.38	0.168

Bold font represents $p < 0.05$.

improvement in the recognition of all five fricatives from the first session to the second session.

Given the fairly large individual variability, it is interesting to compare NLFC-on and NLFC-off conditions at an individual level. For consonant recognition, 9 out of the 30 participants had an average increase of consonant recognition scores across all three sessions by more than five percentage points. One of the 30 participants had a mean score decreased by more than five percentage points. The remaining 20 participants had a change of less than five percentage points in consonant recognition between NLFC-on and NLFC-off conditions. For sentence recognition, 14 out of the 30 participants had an average increase of sentence recognition scores across all three sessions for more than five percentage points. The remaining 16 participants had a change of less than five percentage points in sentence recognition between NLFC-on and NLFC-off conditions.

Correlation analyses were performed to examine the relationship between speech-recognition scores and the hearing thresholds. We used the mean PTA of the two ears to represent the hearing threshold of an individual. Speech-recognition scores were averaged across the three sessions for the NLFC-off condition and NLFC-on condition, respectively. Figure 4 shows the mean speech-recognition scores as a function of the mean PTA. There were significant correlations of consonant, vowel and sentence recognition with PTA for both NLFC on and NLFC off, but not for tone recognition. Following this analysis, correlational tests between the differences in speech recognition scores of the NLFC-on and NLFC-off conditions and the hearing thresholds were conducted. We found no significant correlation between the potential benefit of NLFC and PTA for the four types of speech recognition tasks (all $p > 0.05$).

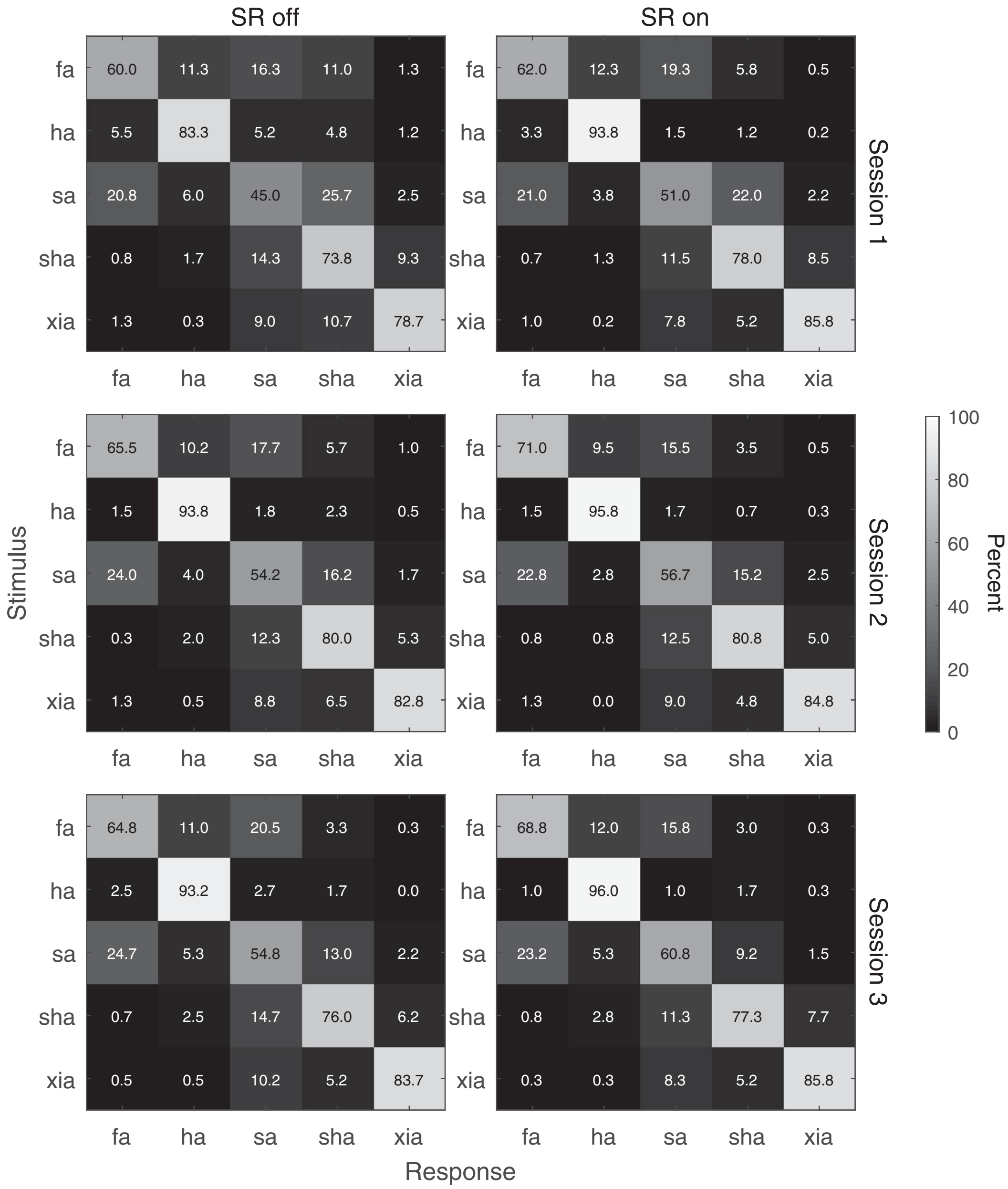

Figure 3. Confusion matrices of the group mean recognition of the five Mandarin fricatives. The two columns represent the NFLC-off and NLFC-on conditions and the three rows represent the three test sessions. In each panel, the stimulus is represented by the ordinate and the response by the abscissa. The value in each cell of the matrix indicates the percent of the stimuli being identified as a particular fricative.

Figure 5 plots the group mean data of sound-quality rating of each percept (abscissa) for all tested sound types (ordinate) with NLFC-off and on (left and right panels) in all three test sessions (rows of panels). The mean sound-quality rating ranged from 7.5 to 9.1 on a scale that ranged from 0 to 10. The GLM analyses revealed a significant improvement in quality rating for loudness with test session. Meanwhile, the sound quality ratings of naturalness and overall preference improved with NLFC on than NLFC off. Further, the rating scores were significantly different across different types of tested sounds. But no significant interaction effect between test sessions and NLFC on/off conditions was yielded. The detailed statistical results are summarised in Table 2.

Discussion

In the present study, we evaluated the efficacy of NLFC on Mandarin speech perception. A series of speech perception tests including consonant, vowel, tone recognition in quiet and sentence recognition in noise were conducted and subjective sound-quality ratings were collected in a group of 30 adult hearing-impaired participants. The different perceptual tasks produced varying outcomes.

Tone recognition in the group of listeners with moderate-to-severe hearing loss was on average >90% correct with NLFC-on or NLFC-off (Figure 2). These results were consistent with previous findings of tone perception in listeners with sensorineural hearing loss (Wang et al. 2016). Acoustically, Mandarin tones are represented in the F0 contours. The average F0 of adult speakers is normally below 0.300 kHz, which is much lower than the start frequency of NLFC. Compared to consonant and vowel perception, tone perception is fairly robust for listeners with sensorineural hearing loss. It is reassuring to see that the NLFC processing, as implemented in SR, exerted no detrimental effects on tone recognition in the hearing aid users.

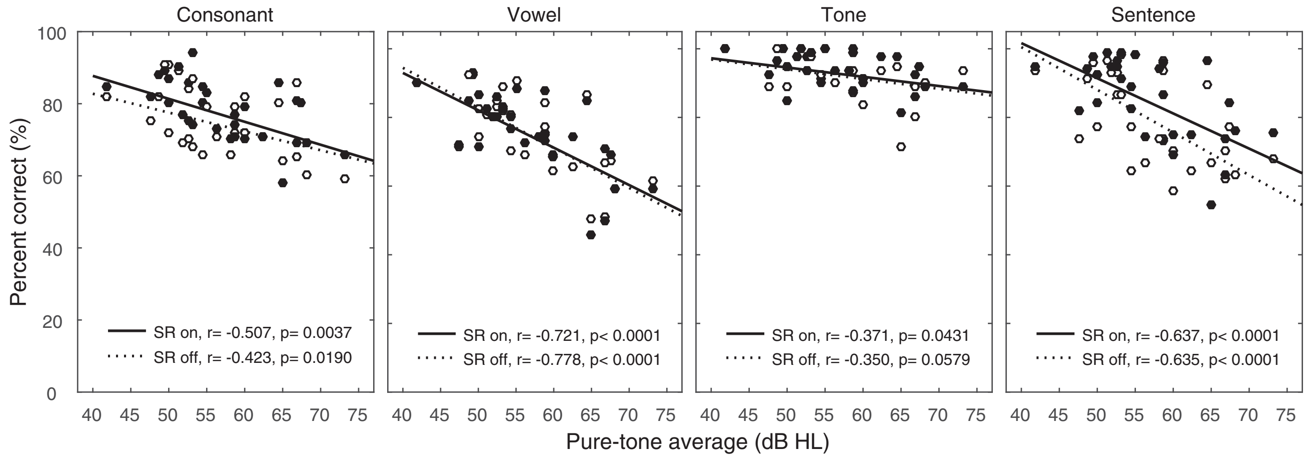

Figure 4. Relationship between speech-recognition scores and the hearing thresholds. The four panels from left to right represent data from consonant, vowel, tone and sentence recognition, respectively. Each symbol represents data from one participant. Hearing threshold (the abscissa) is the mean pure-tone average thresholds of the two ears between 0.5 and 4 kHz. The ordinate represents the mean speech-recognition score across the three test sessions. The open and filled symbols represent NLFC-off and NLFC-on conditions, respectively. The dotted and solid lines represent linear fit of the data for the NLFC-off and NLFC-on conditions, respectively.

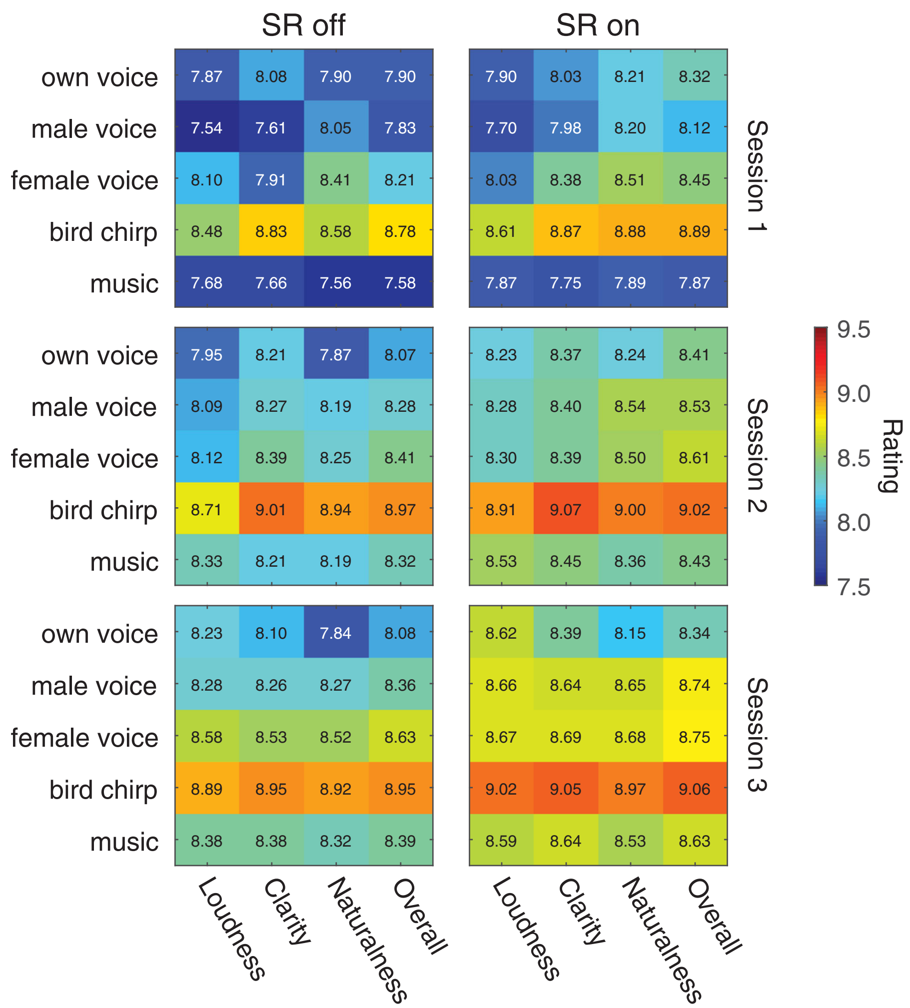

Figure 5. Group mean sound-quality rating. The two columns of panels represent the NLFC-off and NLFC-on condition. The three rows of panels represent the three test sessions. In each panel, the colour represents the mean sound-quality rating for a particular category of percepts (abscissa) and type of sounds (ordinate). The statistical results related to sound-quality rating can be found in Table 2.

Table 2. Summary of GLM results for the factors of session, NLFC-on/off conditions and sound type on the rating scores of each percept.

	Session			NLFC on/off			Session × NLFC interaction			Sound type		
	β	t	p Value	β	t	p Value	β	t	p Value	β	t	p Value
Loudness	0.26	2.48	**0.013**	−0.12	−1.00	0.319	−0.04	−0.67	0.500	0.09	2.98	**0.003**
Clarity	0.17	1.62	0.106	−0.15	−1.20	0.229	−0.02	−0.36	0.715	0.07	2.50	**0.013**
Naturalness	0.07	0.71	0.479	−0.24	−2.00	**0.046**	0.01	0.08	0.934	0.08	2.69	**0.007**
Overall	0.10	1.02	0.309	−0.24	−2.06	**0.039**	0.01	0.17	0.865	0.06	2.36	**0.019**

Bold font represents $p < 0.05$.

Vowel recognition ranged from 70% to 77% correct on average (Figure 2). NLFC exerted no effects on vowel recognition in these listeners. This result was consistent with previous findings on vowel recognition with NLFC in native English-speaking listeners (Glista et al. 2009). Compared to high-frequency consonants, vowels produced by adult speakers are normally characterised by the first three formants located below 3 kHz. Such acoustic characteristics and phonetic correlates for the recognition of vowel identities are not affected by the NLFC processing (Yang et al. 2018). In the present study, because almost all participants had moderate-to-severe hearing loss, the compression parameters (CT and CR) were set at the software defaults to match with the participants' hearing loss. The relatively high CT value (i.e. CT = 4.3 to 6.0 kHz) did not disrupt the vowel formant structures. Therefore, no observable change was found for vowel recognition.

Since the majority of listeners with sensorineural hearing loss have difficulties perceiving high-frequency sounds, improving the recognition of these sounds has been an important goal of modern hearing-aid techniques including the NLFC processing. Several previous studies reported improved consonant recognition in native English-speaking listeners using NLFC in comparison to listeners using conventional processing (Glista et al. 2009; Wolfe et al. 2010, 2011; Hopkins et al. 2014; McCreery et al. 2014). In the present study, the Mandarin-speaking listeners with hearing loss experienced a small but significant improvement in consonant recognition with NLFC-on as compared with NLFC-off. In particular, the hearing-impaired listeners exhibited improved recognition accuracy for /s/ and less confusion between /s/ and /ʂ/ with NLFC-on than with NLFC-off. As stated earlier, Mandarin has five fricatives with three of them being high-frequency sibilants. Such a complicated fricative system might bring challenges to the NLFC technology. In a previous study, we examined the changes of acoustic features in Mandarin speech segments induced by NLFC and the effects of these acoustic changes on phoneme recognition in normal-hearing Mandarin-speaking listeners (Yang et al. 2018). We noticed that the spectral features of NLFC-processed fricatives changed more as the CT was decreased. The three sibilant fricatives (/s/, /ɕ/, /ʂ/) showed similar spectral patterns as a result of downward-shifted high-frequency prominences, which caused a significant decrease in recognition of /s/ sound in the normal-hearing listeners. In particular, the recognition accuracy of /s/ was 96% correct for the unprocessed speech signal but approximately 60% correct for signals processed with low-CT settings in the NLFC algorithm. In the present study for listeners with hearing loss, /s/ was also the most confusing sound with less than 60% accuracy regardless of NLFC-on or off at all three test sessions. This finding indicated that even though the overall consonant recognition accuracy was improved with NLFC, the confusion of the three-way contrast of Mandarin sibilant fricatives caused by the frequency-lowering technique was still evident in hearing-impaired listeners. Note also that consonant recognition in both NLFC-on and off conditions was correlated with the PTA of the hearing-impaired listeners. These results suggest that the presentation level of 65 dB SPL might have not reached audibility for some of the high-frequency consonants in listeners with more severe high-frequency hearing loss. These results also highlight the importance of verification and real ear measurement as well as fine tuning at the individual level of the hearing-impaired listeners.

Previous research on sentence perception with NLFC revealed mixed findings. Some studies found that a reduced speech perception threshold only occurred in certain participants (e.g. Glista et al. 2009; Glista, Scollie, and Sulkers 2012). A few studies found no significant change in speech perception threshold for sentence recognition in competing noise with NLFC as compared to conventional processing (e.g. John et al. 2014). In the present study, The NLFC processing improved the group performance in Mandarin sentence recognition in noise to a greater extent by approximately five percentage points (Figure 2) than that of any phonemes per se. This result suggests that the slightly improved consonant recognition and better audibility of high-frequency information facilitates the top-down recognition processing in addition to providing clearer acoustic inputs. The sentence materials used in the present study was M-HINT sentences that were recorded using a male voice. Female voice tends to produce a higher spectral component for certain phonemes (e.g. fricative /s/; Boothroyd and Medwetsky 2012). We might speculate that the NLFC would produce greater benefits for female voices than for male voices in sentence recognition. However, such speculation remains to be tested. While we observed improved group mean recognition accuracy, we also noticed the large variability among the hearing-impaired listeners for sentence perception performance. Some participants still showed low recognition accuracy with NLFC-on for 12 weeks. This observation suggests that listeners' higher-level linguistic knowledge and cognitive ability might affect the efficacy of NLFC on sentence perception.

Hearing threshold might be another factor for the benefit of NLFC processing or the lack of it in sentence recognition. For listeners with relatively good hearing thresholds, sentence recognition might be at the ceiling. The gap between NLFC on and off was increasingly larger as hearing loss became more severe (Figure 4, right panel). Thus, there was a trend of greater improvement using NLFC processing in participants with more severe hearing loss. These results were consistent with the Ellis and Munro (2015) study in which English speech recognition, with or without NLFC, was correlated with the degree of high-frequency hearing loss.

In addition to speech recognition, the present study also revealed the benefit of NLFC on satisfaction with certain aspects of sound quality. Tseng et al. (2018) and some other studies (Simpson, Hersbach, and McDermott 2006; Picou, Marcrum, and Ricketts 2015) found no significant change in the sound-quality ratings using NLFC relative to conventional devices. However, Brennan et al. (2014) indicated that hearing-impaired listeners with a greater degree of hearing loss might be more likely to prefer NLFC. Note that the 17 participants in the Picou, Marcrum, and Ricketts (2015) study all had flat, mild to moderate hearing loss. On the other hand, the seven participants in Simpson, Hersbach, and McDermott (2006) study had essentially normal hearing in the low frequency but dropped steeply to more than 100 dB HL at and above 2 kHz. Our participants had sloping moderate-to-severe sensorineural hearing loss (Figure 1). In the present study, the participants showed slightly improved subjective experience in loudness after a certain period of NLFC adaptation as well as naturalness and overall sound quality with NLFC-on than with NLFC-off (Figure 5). Two factors might account for these results: (1) NLFC might provide immediate improved sound quality to the hearing aid users and (2) the participants became increasingly adapted to the sound quality delivered through the device after a certain period of continuous usage of NLFC-fitted hearing aids. Nonetheless, the observed benefits in subjective preference might be associated with the particular configurations of hearing loss.

In addition to the efficacy of NLFC on various perception tests, we examined the role of auditory acclimatisation in speech perception with hearing aids. The listeners had worn hearing aids with NLFC enabled for the entire 12 weeks during the study period. We found that consonant, vowel and sentence recognition performance as well as participants' subjective experience with different types of sounds such as human voice, music and bird chirps all improved with an increased length of device use with both NLFC-on and NLFC-off conditions. However, as the same test materials were used in all three sessions in the present study, the fact that the amount of improvement with NLFC-on relative to NLFC-off did not change as a function of test sessions, i.e. no interaction of session versus NLFC condition, suggested that the improvement might be due to perceptual learning or training effects rather than auditory acclimatisation *per se*. Meanwhile, for the three perceptual tasks that showed improved recognition accuracy as a function of length of NLFC use, the listeners appeared to experience greater improvement from session one to session two than from session two to session three. There was only a two-to-three-week time interval between session 1 and session 2 but a nine-to-ten-week interval between session 2 and session 3. This result suggested that the hearing-impaired listeners showed a rapid adaptation to the speech stimuli processed with NLFC. The listeners might still show continuing improvement in speech recognition in a longer period of time, but the magnitude of improvement was not as great as that in the first three weeks.

Conclusions

In summary, the NLFC technology provides modest but significant improvement in Mandarin fricative and sentence recognition as well as the naturalness and overall preference of sound-quality satisfaction judgement for native Mandarin-speaking listeners with hearing loss. The NLFC processing does not alter the performance of vowel and tone recognition as compared to the conventional processing. In addition, a certain period of auditory adaptation ensures better recognition performance and the loudness of sound-quality satisfaction judgement. Such a period is likely to be fairly short (e.g. two to three weeks) for Mandarin-speaking listeners although smaller improvement in performance may still occur in a longer period.

Acknowledgements

Solveig Voss, Yitao Mao, Beier Qi, Yu Zhang, Lexi Neltner and Lauren Muscari provided technique assistance in various stages of the research.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

This research was funded in part by a research contract from Sonova AG.

ORCID

Jing Yang http://orcid.org/0000-0001-6702-5104
Li Xu http://orcid.org/0000-0002-0988-7934

References

Akinseye, G. A., A. Dickinson, and K. J. Munro. 2018. “Is Non-Linear Frequency Compression Amplification Beneficial to Adults and Children with Hearing Loss? A Systematic Review.” *International Journal of Audiology* 57 (4): 262–273. doi:10.1080/14992027.2017.1420255.

Alexander, J. M., J. G. Kopun, and P. G. Stelmachowicz. 2014. “Effects of Frequency Compression and Frequency Transposition on Fricative and Affricate Perception in Listeners with Normal Hearing and Mild to Moderate Hearing Loss.” *Ear and Hearing* 35 (5): 519–532. doi:10.1097/AUD.0000000000000040.

Alexander, J. M., and V. Rallapalli. 2017. “Acoustic and Perceptual Effects of Amplitude and Frequency Compression on High-Frequency Speech.” *Journal of the Acoustical Society of America* 142 (2): 908–923. doi:10.1121/1.4997938.

Alexander, J. M. 2013. “Individual Variability in Recognition of Frequency-Lowered Speech.” *Seminars in Hearing* 34 (2): 086–109. doi:10.1055/s-0033-1341346.

Alexander, J. M. 2016. “Nonlinear Frequency Compression: Influence of Start Frequency and Input Bandwidth on Consonant and Vowel Recognition.” *The Journal of the Acoustical Society of America* 139 (2): 938–957. doi:10.1121/1.4941916.

Agrawal, Y., E. A. Platz, and J. K. Niparko. 2008. “Prevalence of Hearing Loss and Differences by Demographic Characteristics among US Adults: Data from the National Health and Nutrition Examination Survey, 1999-2004.” *Archives of Internal Medicine* 168 (14): 1522–1530. doi:10.1001/archinte.168.14.1522.

Bohnert, A., M. Nyffeler, and A. Keilmann. 2010. “Advantages of a Non-Linear Frequency Compression Algorithm in Noise.” *European Archives of Oto-Rhino-Laryngology : official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 267 (7): 1045–1053. doi:10.1007/s00405-009-1170-x.

Boothroyd, A., and L. Medwetsky. 2012. “Spectral Distribution of/s/and the Frequency Response of Hearing Aids.” *Ear Hear* 13: 150–157. doi:10.1097/00003446-199206000-00003.

Brennan, M. A., R. McCreery, J. Kopun, B. Hoover, J. Alexander, D. Lewis, P. G. Stelmachowicz, et al. 2014. “Paired Comparisons of Nonlinear Frequency Compression, Extended Bandwidth, and Restricted Bandwidth Hearing Aid Processing for Children and Adults with Hearing Loss.” *Journal of the American Academy of Audiology* 25 (10): 983–998. doi:10.3766/jaaa.25.10.7.

Charpentier, F., and M. Stella, 1986. “Diphone Synthesis Using an Overlap-Add Technique for Speech Waveforms Concatenation.” Paper presented at the Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'86, 2015–2018, Tokyo, Japan, April 7–11. doi:10.1109/ICASSP.1986.1168657.

Ellis, R. J., and K. J. Munro. 2015. “Predictors of Aided Speech Recognition, with and without Frequency Compression, in Older Adults.” *International Journal of Audiology* 54 (7): 467–475. doi:10.3109/14992027.2014.996825.

Glista, D., S. Scollie, M. Bagatto, R. Seewald, V. Parsa, and A. Johnson. 2009. “Evaluation of Nonlinear Frequency Compression: Clinical Outcomes.” *International Journal of Audiology* 48 (9): 632–644. doi:10.1080/14992020902971349.

Glista, D., S. Scollie, and J. Sulkers. 2012. “Perceptual Acclimatization Post Nonlinear Frequency Compression Hearing Aid Fitting in Older

Children." *Journal of Speech, Language, and Hearing Research* 55 (6): 1765–1787. doi:10.1044/1092-4388(2012/11-0163).

Healy, E. W., and S. E. Yoho. 2016. "Difficulty Understanding Speech in Noise by the Hearing Impaired: Underlying Causes and Technological Solutions." *Conference Proceedings IEEE Engineering in Medicine and Biology Society* 2016: 89–92. doi:10.1109/EMBC.2016.7590647.

Hopkins, K., M. Khanom, A. M. Dickinson, and K. J. Munro. 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." *International Journal of Audiology* 53 (4): 219–228. doi:10.3109/14992027.2013.873956.

John, A., J. Wolfe, S. Scollie, E. Schafer, M. Hudson, W. Woods, J. Wheeler, K. Hudgens, and S. Neumann. 2014. "Evaluation of Wideband Frequency Responses and Nonlinear Frequency Compression for Children with Cookie-Bite Audiometric Configurations." *Journal of the American Academy of Audiology* 25 (10): 1022–1033. doi:10.3766/jaaa.25.10.10.

Li, Y., G. Zhang, J. J. Galvin, III, and Q. J. Fu. 2014. "Mandarin Speech Perception in Combined Electric and Acoustic Stimulation." *PLoS One.* 9 (11): e112471. doi:10.1371/journal.pone.0112471..

Mao, Y., J. Yang, E. Hahn, and L. Xu. 2017. "Auditory Perceptual Efficacy of Nonlinear Frequency Compression Used in Hearing Aids: A Review." *Journal of Otology* 12 (3): 97–111. doi:10.1016/j.joto.2017.06.003.

McCreery, R. W., J. Alexander, M. A. Brennan, B. Hoover, J. Kopun, and P. G. Stelmachowicz. 2014. "The Influence of Audibility on Speech Recognition with Nonlinear Frequency Compression for Children and Adults with Hearing Loss." *Ear and Hearing* 35 (4): 440–447. doi:10.1097/AUD.0000000000000027.

McDermott, H. J. 2011. "A Technical Comparison of Digital Frequency-Lowering Algorithms Available in Two Current Hearing Aids." *PLoS One.* 6 (7): e22358. doi:10.1371/journal.pone.0022358.

Moore, B. C., and J. I. Alcantara. 2001. "The Use of Psychophysical Tuning Curves to Explore Dead Regions in the Cochlea." *Ear and Hearing* 22 (4): 268–278. doi:10.1097/00003446-200108000-00002.

Parsa, V., S. Scollie, D. Glista, and A. Seelisch. 2013. "Nonlinear Frequency Compression: Effects on Sound Quality Ratings of Speech and Music." *Trends in Amplification* 17 (1): 54–68. doi:10.1177/1084713813480856.

Picou, E. M., S. C. Marcrum, and T. A. Ricketts. 2015. "Evaluation of the Effects of Nonlinear Frequency Compression on Speech Recognition and Sound Quality for Adults with Mild to Moderate Hearing Loss." *International Journal of Audiology* 54 (3): 162–169. doi:10.3109/14992027.2014.961662.

Pittman, A. L., P. G. Stelmachowicz, D. E. Lewis, and B. M. Hoover. 2003. "Spectral Characteristics of Speech at the Ear: Implications for Amplification in Children." *Journal of Speech, Language, and Hearing Research* 46 (3): 649–657. doi:10.1044/1092-4388(2003/051)..

Shehorn, J., N. Marrone, and T. Muller. 2018. "Speech Perception in Noise and Listening Effort of Older Adults with Nonlinear Frequency Compression Hearing Aids." *Ear and Hearing* 39 (2): 215–225. doi:10.1097/AUD.0000000000000481.

Simpson, A., A. A. Hersbach, and H. J. McDermott. 2006. "Frequency-Compression Outcomes in Listeners with Steeply Sloping Audiograms." *International Journal of Audiology* 45 (11): 619–629. doi:10.1080/14992020600825508.

Simpson, A. 2009. "Frequency-Lowering Devices for Managing High-Frequency Hearing Loss: A Review." *Trends in Amplification* 13 (2): 87–106. doi:10.1177/1084713809336421..

Salorio-Corbetto, M., T. Baer, and B. C. J. Moore. 2017. "Evaluation of Frequency-Lowering Algorithm for Adults with High-Frequency Hearing Loss." *Trends in Hearing* 21: 1–23. doi:10.1177/2331216517734455.

Soli, S. D., and L. L. N. Wong. 2008. "Assessment of Speech Intelligibility in Noise with the Hearing in Noise Test." *International Journal of Audiology* 47 (6): 356–361. doi:10.1080/14992020801895136.

Stelmachowicz, P. G., A. L. Pittman, B. M. Hoover, D. E. Lewis, and M. P. Moeller. 2004. "The Importance of High-Frequency Audibility in the Speech and Language Development of Children with Hearing Loss." *Archives of Otolaryngology-Head & Neck Surgery* 130 (5): 556–562. doi:10.1001/archotol.130.5.556.

Tseng, W. H., D. L. Hsieh, W. T. Shih, and T. C. Liu. 2018. "Extended Bandwidth Nonlinear Frequency Compression in Mandarin-Speaking Hearing-Aid Users." *Journal of the Formosan Medical Association* 117 (2): 109–116. doi:10.1016/j.jfma.2017.01.013.

Thornton, A. R., and M. J. Raffin. 1978. "Speech-Discrimination Scores Modeled as a Binomial Variable." *Journal of Speech and Hearing Research* 21 (3): 507–518. doi:10.1044/jshr.2103.507.

Vickers, D. A., B. C. Moore, and T. Baer. 2001. "Effects of Low-Pass Filtering on the Intelligibility of Speech in Quiet for People with and without Dead Regions at High Frequencies." *The Journal of the Acoustical Society of America* 110 (2): 1164–1175. doi:10.1121/1.1381534.

Wang, S., R. Dong, D. Liu, Y. Wang, Y. Mao, H. Zhang, L. Zhang, et al. 2016. "Perceptual Separation of Sensorineural Hearing Loss and Auditory Neuropathy Spectrum Disorder." *The Laryngoscope* 126 (6): 1420–1425. doi:10.1002/lary.25595.

Warton, D. I., and F. K. Hui. 2011. "The Arcsine is Asinine: The Analysis of Proportions in Ecology." *Ecology* 92 (1): 3–10. doi:10.1890/10-0340.1.

Wolfe, J., A. John, E. Schafer, M. Nyffeler, M. Boretzki, T. Caraway, M. Hudson, et al. 2011. "Long-Term Effects of Non-Linear Frequency Compression for Children with Moderate Hearing Loss." *International Journal of Audiology* 50 (6): 396–404. doi:10.3109/14992027.2010.551788.

Wolfe, J., A. John, E. Schafer, M. Nyffeler, M. Boretzki, and T. Caraway. 2010. "Evaluation of Nonlinear Frequency Compression for School-Age Children with Moderate to Moderately Severe Hearing Loss." *Journal of the American Academy of Audiology* 21 (10): 618–628. doi:10.3109/14992027.2010.551788.

Wolfe, J., M. Duke, E. C. Schafer, J. Rehmann, S. Jha, S. Allegro Baumann, A. John, et al. 2017. "Preliminary Evaluation of a Novel Non-Linear Frequency Compression Scheme for Use in Children." *International Journal of Audiology* 56 (12): 976–988. doi:10.1080/14992027.2017.1358467.

Wong, L. L., S. D. Soli, S. Liu, N. Han, and M. W. Huang. 2007. "Development of the Mandarin Hearing in Noise Test (MHINT)." *Ear and Hearing* 28 (2 Suppl): 70S–74S. doi:10.1097/AUD.0b013e31803154d0.

Yang, J., J. Qian, X. Chen, V. Kuehnel, J. Rehmann, A. von Buol, Y. Li, et al. 2018. "Effects of Nonlinear Frequency Compression on the Acoustic Properties and Recognition of Speech Sounds in Mandarin Chinese." *The Journal of the Acoustical Society of America* 143 (3): 1578–1590. doi:10.1121/1.5027404.

Zee, E., and W. S. Lee. 2001. "An Acoustical Analysis of the Vowels in Beijing Mandarin." Paper presented at the Seventh European Conference on Speech Communication and Technology, Aalborg, Denmark, September 3–7.

Zee, E. 2001. "The Phonetic Value of the Vowels, Diphthongs and Triphthongs in Beijing Mandarin." Paper presented at the 5th National Conference on Modern Phonetics, Beijing, August.