

Speech Perception and Sound-Quality Rating with an Adaptive Nonlinear Frequency Compression Algorithm in Mandarin-Speaking Hearing Aid Users

Li Xu¹ Solveig C. Voss² Jing Yang³ Xianhui Wang¹ Qian Lu⁴ Julia Rehmann⁵ Volker Kuehnel⁵
Jinyu Qian²

¹Communication Sciences and Disorders, Ohio University, Athens, Ohio

Address for correspondence Li Xu, MD, PhD, xul@ohio.edu.

²Innovation Centre Toronto, Sonova Canada, Inc., Mississauga, Ontario, Canada

³Communication Sciences and Disorders, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin

⁴Innovation Center Shanghai, Sonova China, Shanghai, China

⁵R&D, Sonova AG, Stäfa, Switzerland

J Am Acad Audiol 2020;31:590–598.

Abstract

Background Mandarin Chinese has a rich repertoire of high-frequency speech sounds. This may pose a remarkable challenge to hearing-impaired listeners who speak Mandarin Chinese because of their high-frequency sloping hearing loss. An adaptive nonlinear frequency compression (adaptive NLFC) algorithm has been implemented in contemporary hearing aids to alleviate the problem.

Purpose The present study examined the performance of speech perception and sound-quality rating in Mandarin-speaking hearing-impaired listeners using hearing aids fitted with adaptive NLFC (i.e., SoundRecover2 or SR2) at different parameter settings.

Research Design Hearing-impaired listeners' phoneme detection thresholds, speech reception thresholds, and sound-quality ratings were collected with various SR2 settings.

Study Sample The participants included 15 Mandarin-speaking adults aged 32 to 84 years old who had symmetric sloping severe-to-profound sensorineural hearing loss.

Intervention The participants were fitted bilaterally with Phonak Naida V90-SP hearing aids.

Data Collection and Analysis The outcome measures included phoneme detection threshold using the Mandarin Phonak Phoneme Perception test, speech reception threshold using the Mandarin hearing in noise test (M-HINT), and sound-quality ratings on human speech in quiet and noise, bird chirps, and music in quiet. For each test, five experimental settings were applied and compared: SR2-off, SR2-weak, SR2-default, SR2-strong 1, and SR2-strong 2.

Results The results showed that listeners performed significantly better with SR2-strong 1 and SR2-strong 2 settings than with SR2-off or SR2-weak settings for speech reception threshold and phoneme detection threshold. However, no significant improvement was observed in sound-quality ratings among different settings.

Keywords

- ▶ hearing aids
- ▶ Mandarin Chinese
- ▶ nonlinear frequency compression
- ▶ speech perception
- ▶ sound quality

received

March 22, 2019

accepted after revision

January 25, 2020

published online

April 27, 2020

Copyright © 2020 by the American Academy of Audiology. All rights reserved. Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. Tel: +1(212) 760-0888.

DOI <https://doi.org/10.1055/s-0040-1709450>.
ISSN 1050-0545.

Conclusions These preliminary findings suggested that the adaptive NLFC algorithm provides perceptual benefit to Mandarin-speaking people with severe-to-profound hearing loss.

Frequency lowering, a signal-processing technique that essentially shifts inaudible high-frequency information to the audible lower-frequency region, has been widely used in hearing aids to help hearing-impaired listeners gain accessibility to the linguistic information contained in the high-frequency region (see Simpson,¹ Mao et al,² and Simpson et al³ for reviews). Among the various types of frequency-lowering techniques, nonlinear frequency compression (NLFC) has drawn a great deal of attention from hearing aid research community.^{1,2}

In contrast to linear frequency compression, where all frequency components are lowered by a constant factor, the NLFC algorithm uses a predefined cut-off frequency (CT) and compression ratio (CR) such that all input signals above the CT are compressed into the output frequency area that is between the CT and the maximum output frequency of the NLFC algorithm. Higher frequencies above the CT are compressed in greater amounts than lower frequencies above the CT. Through disproportionately compressing high-frequency components to the low-frequency regions, the inaudible high frequencies become audible to the hearing aid users.⁴⁻⁷ SoundRecover (SR) by Phonak Naída is an NLFC algorithm commercially used in modern hearing aids.⁸ While NLFC enables hearing-impaired listeners to gain access to inaudible high frequencies, to those listeners with severe hearing loss who have very limited output bandwidth and require a more aggressive parameter setting (e.g., a lower CT and a higher CR), low-frequency sounds (e.g., vowels and sonorant consonants) might be adversely affected due to potential distortions of the spectral structure.⁷ To further improve the audibility for listeners with severe hearing impairment and maintain the spectral shape of vowels and low-frequency speech signals, an adaptive NLFC algorithm (known as SoundRecover2 or SR2) was developed.⁹ The algorithm uses two cut-offs, one lower cut-off (CT1) and one higher cut-off (CT2). The selection of CT1 or CT2 is automatically determined by the short-term energy distribution of the input signal. When the energy of the incoming signal mainly resides in low-frequency regions, CT2 is applied. On the other hand, when the incoming acoustic energy is located in high-frequency regions, CT1 is used. This adaptive selection of the CTs maintains the vowel formant structures and maximizes the access to high-frequency sounds.^{7,9} Clinically, the settings of CT1, CT2, and CR depend on the severity of the hearing loss. The range of CT1 is between 800 and 7,000 Hz and that of CT2 is between 1,600 and 7,700 Hz. The CR is set between 1.1 and 1.5.

As reviewed in Mao et al,² Simpson et al,³ and Akinseye et al,¹⁰ there have been several studies evaluating the efficacy of NLFC with different parameter settings on speech perception in hearing-impaired listeners (e.g., Alexander and Rallapalli,⁶ Brennan et al,¹¹ Ching et al,¹² Glista et al,¹³ Wolfe

et al,¹⁴⁻¹⁶ Wright et al,¹⁷ McCreery et al,³¹ Hopkins et al³²). Only a few studies provided data regarding speech perception of English-speaking hearing-impaired people using adaptive NLFC-fitted hearing devices with different parameter settings, which might be because the adaptive NLFC is a newly developed algorithm and has not been widely adopted yet. Wolfe et al¹⁶ compared the performance of speech perception with NLFC and adaptive NLFC in 14 hearing-impaired English-speaking children with mild-to-moderate low-frequency hearing loss and severe-to-profound high-frequency hearing loss. The perceptual ability was evaluated through several tests including the University of Western Ontario (UWO) plurals test,¹⁸ Consonant-nucleus-consonant (CNC) monosyllabic word recognition test,¹⁹ and Phonak Phoneme Perception test (PPT).²⁰ The authors found that these children demonstrated a higher recognition accuracy with the adaptive NLFC-fitted device than with the static NLFC-fitted device for the UWO plural test and CNC word recognition test. Moreover, many children showed lower recognition thresholds in PPT with adaptive NLFC than with static NLFC.

Glista et al²¹ tested phoneme perception with adaptive NLFC, static NLFC, and conventional processing (i.e., no frequency compression) in eight hearing-impaired listeners, including both children and young adults. The phoneme perception tasks included detection, distinction, and recognition of selected English consonant sounds. The test materials for phoneme detection were English high-frequency sounds /j/ centered at 3,000 and 5,000 Hz and /s/ centered at 6,000 and 9,000 Hz. For phoneme distinction, /j/ centered at 5,000 Hz and /s/ centered at 6,000 Hz were used. For phoneme recognition, seven English consonants, /d/, /f/, /h/, /k/, /m/, /s/, and /j/, were used. The results suggested that while both types of NLFC algorithms outperformed the conventional processing, the adaptive NLFC did not provide additional perception benefit compared with the static NLFC in phoneme perception tests. In a more recent study, Glista and colleagues²² assessed sound-quality ratings in normal-hearing adults and children as well as hearing-impaired adults and children with static NLFC or adaptive NLFC. The sound-quality rating was tested in individually fine-tuned and intentionally modified conditions. The results suggested that while both normal-hearing and hearing-impaired listeners rated the stimuli that were intentionally modified to have poor sound quality as having below-average scores, the hearing-impaired listeners provided similar above-average ratings for sound quality with all static and adaptive NLFC settings.

So far, little data have been reported on the usage of the adaptive NLFC on non-English-speaking hearing-impaired listeners. To fill this knowledge gap, the present study was conducted to (1) evaluate the effect of adaptive NLFC algorithm (i.e., SR2) in participants who speak and listen to

Mandarin Chinese and (2) evaluate the effect of different adaptive NLFC parameter settings on speech perception. Mandarin Chinese has systematic differences from English in the phonetic system. Particularly, in terms of high-frequency speech sounds, Mandarin has three sibilant fricatives (/s, ʃ, ʂ/) and six affricates (/ts, tsʰ, tɕ, tɕʰ, tʂ, tʂʰ/) that present a three-way contrast of alveolar–alveopalatal–retroflex postalveolar for the place of articulation. The three-way contrasts in the high-frequency speech signal present particular challenges for Mandarin-speaking hearing-impaired listeners. The present study was designed to examine whether the adaptive NLFC improves speech perception for Mandarin-speaking hearing aid users. In addition, this study aimed at evaluating whether the current parameter settings yield perceptual benefits for Mandarin-speaking hearing-impaired listeners or not.

Methods

Participants

The participants included 15 Mandarin-speaking adults (2 females and 13 males) with symmetric, sloping severe-to-profound sensorineural hearing loss. The participants were between 32 and 84 years old with an average age of 67.5 years (standard deviation = 15.3 years). The adaptive NLFC has been recommended for hearing-impaired listeners with high-frequency loss, including those with severe-to-profound hearing loss, left corner audiograms, and ski-slope losses.⁹ For the present study, all participants met the candidacy for adaptive NLFC. The exclusion criteria were as follows: unilateral, mild or moderate, flat or reversed shaped or conductive hearing loss, not being a native speaker of Mandarin, or having a language disability that could interfere with testing. The individual and

group mean unaided audiometric thresholds of the 15 participants are plotted in ►Fig. 1. The air conduction thresholds were measured with an Aurical audiometer and circumaural TDH 39 headphones. A B transducer was used to measure bone conduction thresholds. Tympanometry was conducted with a Madsen OtoFlex 100 tympanometer to exclude conductive hearing loss. Each participant had at least 3 months of hearing aid experience, but none of them had experience using any kind of devices fitted with frequency-lowering algorithms. No participant reported to have cognitive or speech–language impairments. The use of human subjects was reviewed and approved by the Institutional Review Board of Ohio University and participants signed a consent form prior to participating in the study.

Hearing Aid Fitting

The participants were fitted bilaterally with Phonak Naida V90-SP behind-the-ear hearing aids. The participants could use their own earmolds as long as the condition and vent of the earmolds were deemed good. Otherwise, newly made earmolds were provided with vents according to the recommendation in Target fitting software version 5.2. The hearing aids were fitted with Adaptive Phonak Digital Tonal gain prescription developed for tonal languages. A 100% gain level was applied to match estimated targets as estimated by the fitting software. The RECD (real-ear-to-coupler difference) was estimated based on the feedback and real-ear test conducted in the fitting software. The feedback and real-ear test measures of the feedback path provided an estimate of the vent loss and corresponding compensation. A manual listening program for speech in quiet and another for speech in noise were created and used in the corresponding speech test conditions. Adaptive parameters, such as noise

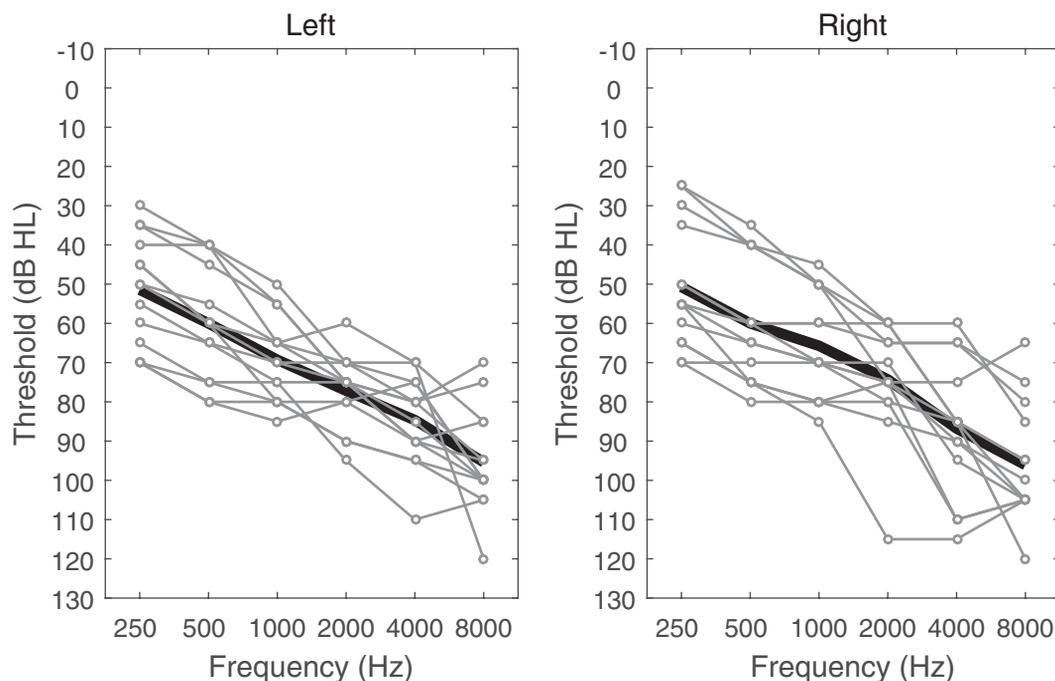


Fig. 1 Audiometric results in the left and right ears for all participants. The thick black line shows the average of the hearing thresholds across all participants.

cancellation, transient noise reduction, feedback cancellation, beamformer, and binaural coordination, were set to default values as recommended by the fitting software to create a realistic hearing aid fitting and support speech understanding in the speech in noise test. Fine tuning was allowed if the participants reported that the hearing aids were too soft or too loud. For some participants, the overall loudness was reduced using the overall gain level, which can be set to 70, 80, 90, 100, or 110%.

The order of the five experimental settings (i.e., SR2-off, SR2-default, SR2-weak, SR2-strong 1, and SR2-strong 2) was randomized for each participant. The CTs and CR were changed by adjusting two sliders in the software. In the first, a 20-step slider, moving the slider to the left (Audibility) decreased the CT1 and CT2 and moving the slider to the right (Distinction) increased the CT1 and CT2. The CR changed relative to the position of CT1 and CT2. A second, 4-point slider allowed the change of the position of CT1 and CT2 without changing the CR. CT1 and CT2 could be adjusted in a limited range that was determined by the position of the first slider. Moving the slider to the left (position “a,” more Clarity) could decrease CT1 and CT2 and moving the slider to the right (position “d,” Comfort) could increase CT1 and CT2. However, in the present study, the second slider remained at the default position “a” for all fittings. All SR2 adjustments in this study were applied using only the first slider. Therefore, the five experimental settings were (1) SR2-off, (2) SR2-default (parameters calculated based on audiogram), (3) SR2-weak (three steps toward Distinction relative to default on the first slider), (4) SR2-strong 1 (three steps toward Audibility relative to default on the first slider), and (5) SR2-strong 2 (six steps toward Audibility relative to default on the first slider). ►Table 1 shows the parameters of the four SR2 settings based on a fitting with the average hearing loss of all subjects.

For each individual, the testbox measurements of hearing aid outputs in response to the phonemes “s,” “sh,” and the ISTS (International Speech Test Signal)^{23,24} presented at 65 dB (sound pressure level, SPL) were performed in the Audioscan Verifit with a 2-cc coupler. The hearing aid was set to a listening program for quiet environment with adaptive parameters disabled and the SR2 was set as described above. ►Fig. 2 shows such measurements based on the average hearing loss of all subjects. The testbox measurements performed based on individuals’ hearing thresholds showed similar results as represented in ►Fig. 2, that is, the

Table 1 The SR2 parameters based on a fitting with the average hearing loss of all subjects

	CT1 (kHz)	CT2 (kHz)	CR
Weak	3.0	4.7	1.1
Default	2.0	3.5	1.2
Strong 1	1.5	2.8	1.3
Strong 2	1.1	2.5	1.4

Abbreviations: CR, compression ratio; CT1, cut-off frequency 1; CT2, cut-off frequency 2.

ISTS and phoneme “sh” were audible in all SR2 settings including the SR2-off condition for all subjects. Phoneme “s” was audible in the SR2-default setting in seven of the 15 subjects and in SR2-strong 1 and SR2-strong 2 settings in all 15 subjects.

Outcome Measures

The outcome measures included speech perception and sound-quality ratings. Speech perception included the phoneme detection test and speech reception threshold (SRT) test in quiet and in noise. The phoneme detection was tested using the Mandarin version of PPT (M-PPT). M-PPT consists of three subtests that assess three aspects of perceptual ability for the Mandarin-speaking population: detection, distinction, and recognition. In the present study, the detection subtest was used. The five high-frequency stimuli for the detection subtest included /s/ centered at 3,000 Hz (“sh3”)

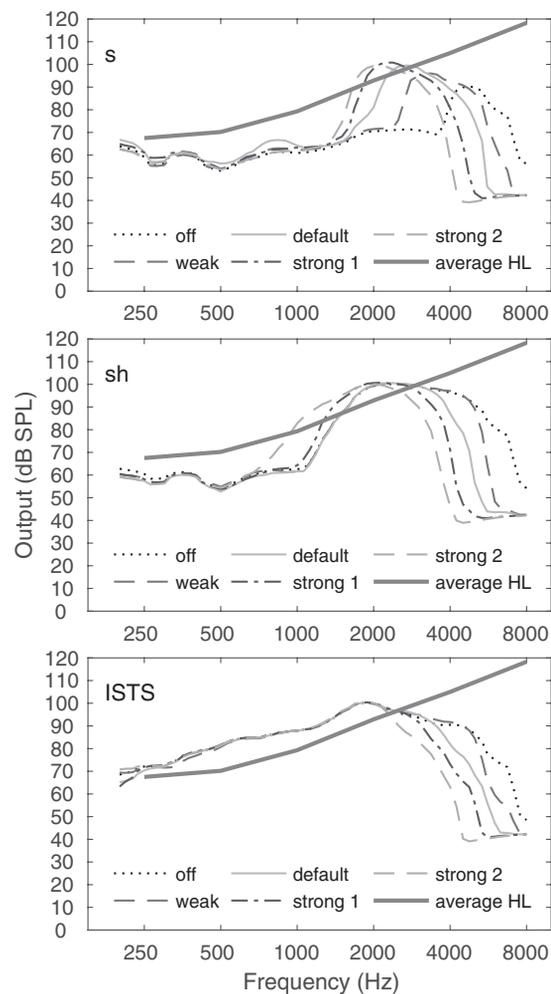


Fig. 2 The testbox measurements of hearing aid outputs in response to speech signals “s” (top panel), “sh” (middle panel), and ISTS (bottom panel) measured with a fitting based on the average hearing loss of all subjects. Different curves represent different SR2 settings. The thick blue line represents the average audiograms of all subjects. ISTS, International Speech Test Signal.

and 5,000 Hz (“sh5”), /s/ centered at 6,000 Hz (“s6”) and 9,000 Hz (“s9”), and /ç/ (“x”).

The SRT was tested using the extended version of Mandarin-hearing in noise test (M-HINT). Developed from the original M-HINT,²⁵ the extended M-HINT consists of 24 lists. Each list consists of 10 sentences and each sentence is made of 10 Chinese characters. The participants were seated in a sound booth. Their heads were situated at 0° azimuth, 1.45 m from the loudspeaker. One randomly selected and randomly ordered sentence list was presented to each participant. The participants were required to repeat what they had heard, and their responses were counted in the speech test software by an experimenter sitting outside of the booth. The speech signals were initially set at 65 dB(A) for testing in quiet. The output level was adaptively changed to yield the SRT, i.e., 50% sentence recognition accuracy. When testing in noise, a speech-spectrum-shaped noise was set at 60 dB(A) and the speech signals were adaptively varied to reach the SRT in noise, i.e., the signal-to-noise ratio that yielded 50% sentence recognition accuracy.

In addition to the speech perception abilities, participants' subjective ratings for different types of sounds were collected. The stimuli for the sound-quality rating included male speech in quiet, male speech in noise, female speech in quiet, female speech in noise, bird chirps, and music in quiet. The male and female voices were recordings of a text read aloud by native Mandarin speakers. The text was a paragraph of 127 Chinese characters and the lengths of the recordings were 34 and 36 seconds for the male and female voices, respectively. The bird chirps were originally provided in MATLAB which contained eight chirps. We duplicated the bird chirps three times so that the final stimuli contained 24 chirps and lasted 5.5 seconds. The music stimulus was a recorded piano performance of a classic Chinese folk music that was known to all participants. The duration of the music stimulus was 105 seconds. All stimuli were presented from a loudspeaker situated at 0° azimuth and 1.45 m away from the participant. The stimulus levels were set to match natural sound levels. The human voices were presented at 65 dB(A) in quiet and 60 dB(A) in noise. The noise was the speech-spectrum-shaped noise presented at a level of 65 dB(A). The bird chirps and music were presented at 65 and 70 dB (A), respectively. The stimuli were recorded prior to the data collection using hearing aids that were programmed with the participants' individual fitting. The hearing aids were placed on a head and torso simulator (KEMAR) and the electrical output signal was recorded through a soundcard. For the rating, participants wore another pair of hearing aids with their individual earmolds. The prerecorded stimuli were electrically routed to those hearing aids and played back upon pressing a button on a screen. This allowed for a calibrated presentation of the sounds while maintaining the natural acoustics of an individual earmold. For each type of tested sounds (i.e., male voice in quiet and in noise, female male voice in quiet and in noise, bird chirps, and music in quiet), the participants were asked to rate the loudness, familiarity, clarity, unusualness, and overall sound quality on a 5-point categorical rating scale.

Procedures

Each participant attended three appointments. During the first appointment, a hearing test was conducted to determine the type and severity of hearing loss. Then, each participant was tested with one list of M-HINT sentences in a quiet listening condition with the participant's own hearing aids. Only participants who scored $\geq 60\%$ correct in the M-HINT sentence recognition test were enrolled for the subsequent perceptual tasks in the present study. Those who scored $< 60\%$ correct for M-HINT sentence recognition in quiet would not be able to measure a reliable SRT and were thus excluded from the study. Due to time constraints, the participants were asked to come back in two additional appointments to complete the hearing aid fitting, speech recognition testing, and sound-quality rating. The second appointment occurred 1 or 2 weeks after the first appointment. Each participant was fitted bilaterally with the Phonak hearing aids with SR2 activated. Then speech tests in quiet and noise were administered, followed by the detection subtest of M-PPT. In the third appointment, which occurred 1 or 2 weeks after the second appointment, the sound-quality rating task was performed. For each test, the five different settings of SR2 were applied and compared. The testing order of the settings was randomized.

Results

► **Fig. 3** presents the M-PPT detection threshold for the five speech stimuli. Among the five stimuli, “s9” had the highest detection threshold while “sh3” had the lowest detection threshold. The stimulus “x” showed a similar threshold to “s6.” Among the five experimental settings, all five stimuli showed lower detection thresholds with the two stronger settings (i.e., SR2-strong 1 and SR2-strong 2) than with the weak setting (i.e., SR2-weak). Since the sample size was small, Friedman rank sum tests were used to examine the differences among the five SR2 settings for each speech stimulus. The results showed significant differences for “s6,” “s9,” “sh5,” and “x” (all $p < 0.05$) but not for “sh3” ($p > 0.05$). The Wilcoxon signed-rank test was conducted for subsequent pairwise comparisons. Bonferroni corrections were applied for multiple comparisons. The results revealed no significant results between any two settings for “sh5.” For the other three stimuli “s6,” “s9,” and “x,” the detection threshold with SR2-off was significantly higher than those with SR2-default, SR2-strong 1, and SR2-strong 2 (all $p < 0.005$). The detection threshold of SR2-weak was significantly higher than those of SR2-strong 1 and SR2-strong 2 (all $p < 0.005$). In addition, the detection threshold with SR2-default was significantly lower than that of SR2-weak for “s6” ($p = 0.003$). The detection threshold with SR2-strong 2 was also significantly lower than that of SR2-default for “s9” ($p = 0.001$). Finally, the detection threshold with SR2-strong 1 was significantly lower than that of SR2-default for “x” ($p = 0.002$).

► **Fig. 4** shows the SRTs in quiet and noise with the five experimental settings. There were little differences between the SR2-off and SR2-weak settings. All three other settings

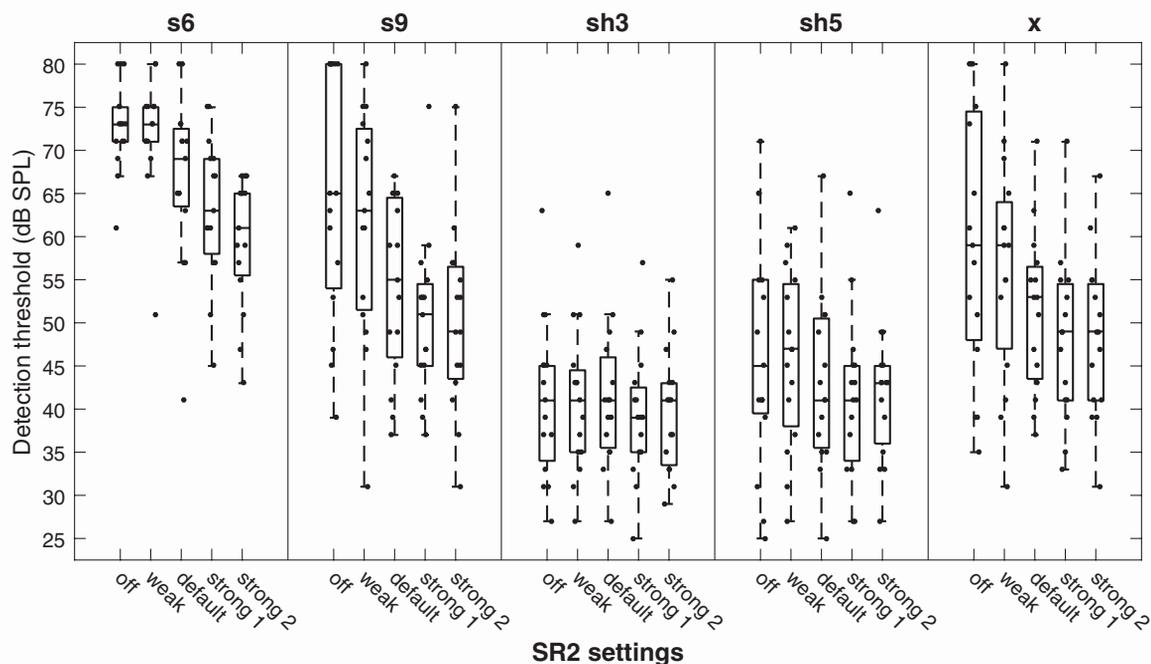


Fig. 3 Boxplot of M-PPT detection threshold results. Each panel shows the result for one of the five different stimuli. The five stimuli were /s/ centered at 6,000 Hz (“s6”) and 9,000 Hz (“s9”), /ʃ/ centered at 3,000 Hz (“sh3”) and 5,000 Hz (“sh5”), and /ç/ (“x”). Each dot indicates data from one participant. The horizontal line in the box is the average of the group scores. The bottom and the top lines of the box represent the 25th and 75th percentiles of the data. The whiskers represent the ranges. Any outliers are identified outside of the whiskers.

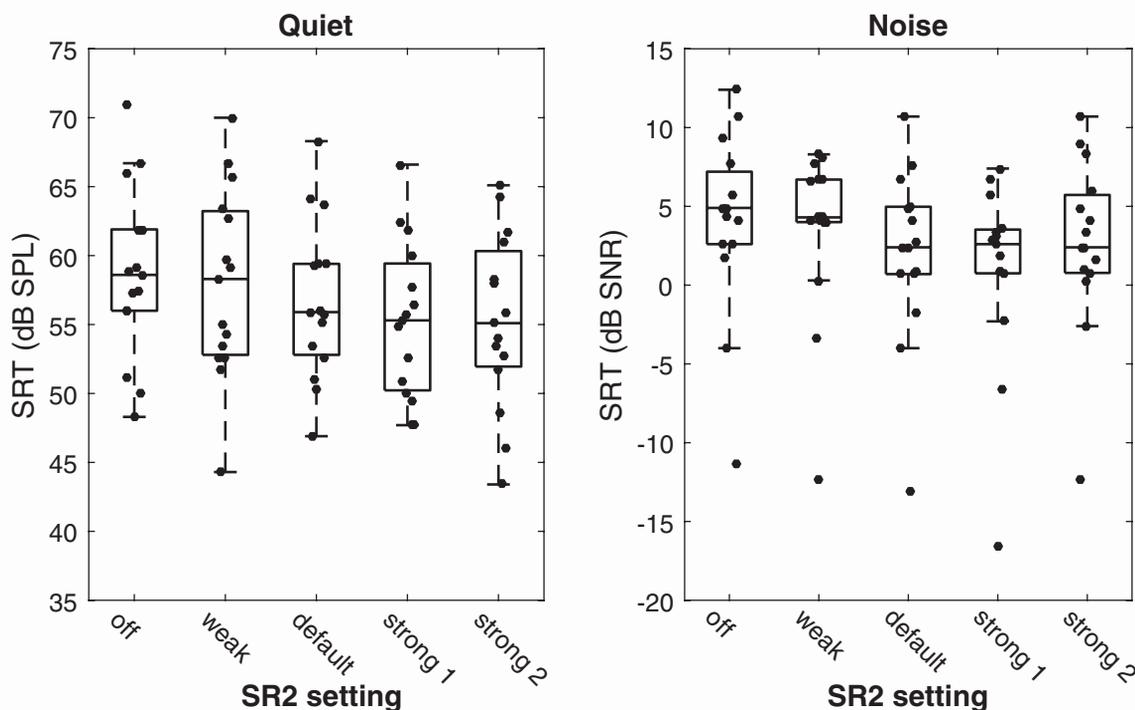


Fig. 4 Boxplot showing speech reception threshold (SRT) in quiet (left panel) and in noise (right panel) using the five SR2 settings. The horizontal line inside the box indicates the average of the group. The bottom and top of the box indicate the lower 25th and top 75th percentiles of the data. The whiskers show the range of the data. Individual data points are plotted with circles. Data points outside of the whiskers are outliers.

demonstrated observable lower SRTs than the SR2-off and SR2-weak settings. Meanwhile, the SRTs with the two stronger settings did not differ much from that with the default setting. Friedman rank sum tests were used to compare the SRT results in quiet and in noise conditions, respectively, among the five SR2 settings. The results showed significant differences in both quiet and noise ($p < 0.01$). Pair-wise comparison was conducted using the Wilcoxon signed-rank test with Bonferroni correction applied for multiple comparisons. The results showed significantly lower SRTs for SR2-strong 1 and SR2-strong 2 than for SR2-off in quiet ($p < 0.005$). For speech test in noise, the SR2-strong 1 showed significantly lower SRTs than SR2-off and SR2-weak conditions ($p < 0.005$).

As a group, no significant differences were found in sound-quality ratings regarding the loudness, familiarity, clarity, unusualness, and overall sound quality among the five SR2 settings for all types of tested sounds (Friedman rank sum test, all $p > 0.05$). All settings were rated as “just right” in terms of loudness. ▶**Fig. 5** shows the overall sound-quality rating scores for the four types of sounds in each SR2 setting. To simplify the data presentation, the rating scores of male voice in quiet and noise conditions were pooled together. Likewise, the rating scores of female voice in quiet and noise conditions were pooled together. Among the four different types of sounds, bird chirps were rated with the lowest overall sound-quality scores while music and male speech were rated with higher overall sound-quality scores. In addition, for both music and female voice, the overall sound quality with SR2-strong 2 was rated with lower score than with other settings. Note that no significant results were obtained for all sound-quality ratings. The variation in sound-quality ratings did not suggest the superiority of one setting to the other.

Discussion

This preliminary study reported the performance of speech perception and sound-quality rating through hearing aids fitted with various settings of an adaptive NLFC algorithm in Mandarin-speaking hearing-impaired adults. As Mandarin is characterized by a richer number of high-frequency sounds in comparison to English,²⁶ the present study provides valuable information regarding the application of adaptive NLFC to a non-English-speaking population. Such information is valuable for future research in NLFC. Our results demonstrated that the speech detection and recognition performance in the Mandarin-speaking hearing-impaired listeners improved as a result of the stronger SR2 settings, when compared with the SR2-off or SR2-weak settings. However, no significant improvement was observed in different aspects of sound-quality rating among those SR2 settings.

PPT was designed to evaluate the efficacy of frequency-lowering algorithms. In M-PPT, speech signals unique in Mandarin Chinese were used to evaluate detection and recognition ability. With the processed high-frequency stimuli centered at different frequencies, M-PPT is sensitive to different SR2 settings for high-frequency phonemes in Mandarin Chinese. Our results showed that for the high-frequency sounds /s/ (both “s6” and “s9”) and /ç/ (“x”), the participants demonstrated improved detection ability as reflected by a significantly reduced threshold with SR2 activated in comparison to SR2-off (▶**Fig. 3**). Moreover, the detection ability improved further from the SR2-default setting to stronger settings. However, for the sound /ʃ/, especially “sh3” that is characterized by frequency components in a lower frequency range, the detection thresholds did not show significant changes across different settings. Note that the detection

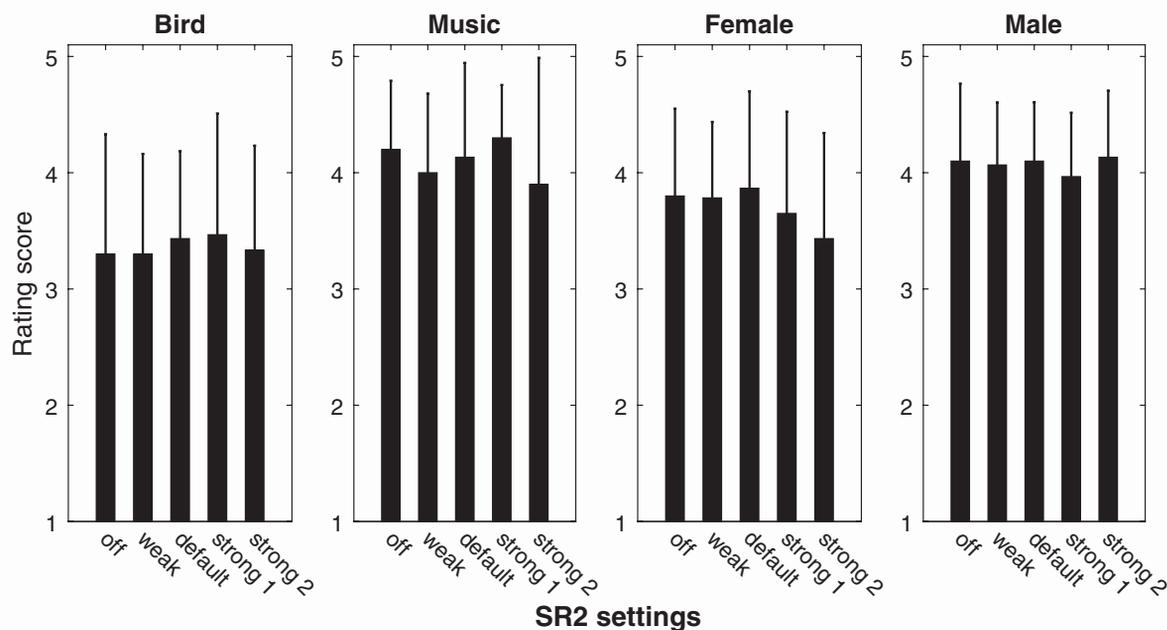


Fig. 5 Overall sound-quality rating for the four different test stimuli in each SR2 setting. The ratings in both noise and quiet conditions were pooled together for the female and male voices. The rating scale was from 1 (very poor) to 5 (excellent). Error bars represent 1 standard deviation.

threshold for the sound /s/, especially “sh3,” was already low enough with SR2-off. The textbox measurements also confirmed that phoneme “sh” was audible in all SR2 settings including the SR2-off condition (►Fig. 2).

The improved perceptual ability was also reflected by decreased speech recognition thresholds measured with M-HINT sentences in both quiet and noise conditions when SR2 was activated. An improved speech recognition performance was found with stronger SR2 settings. The improvement was even more evident for the SR2-strong 1 setting as compared to the SR2-off setting (►Fig. 4). As shown in ►Fig. 2, the SR2-weak setting provided limited benefit in audibility. The SR2-default setting provided a marginal benefit in audibility, whereas SR2-strong 1 and particularly SR2-strong 2 settings provided greater benefit in audibility to high-frequency speech signals. Note that SR2 settings were always calculated based on the better ear if side-independent calculation was deactivated. However, highly asymmetrical audiograms were necessary to measure a benefit of side-independent calculation. Since all participants in the present study had a symmetrical hearing loss, this calculation approach is not expected to have great influence on the study outcomes. For the two SR2-strong settings, SR2-strong 2 had a lower cut-off but a higher CR in comparison to SR2-strong 1, which resulted in a visible change of output for higher amplitude for frequency at 2.5 kHz. However, the recognition performance of SR2-strong 2 was not better than that of SR2-strong 1. This result indicated that while SR2 with a stronger setting provided perceptual benefit to listeners with severe hearing loss in comparison to SR2-off or SR2-default, the recognition performance did not improve further as the parameter setting became more aggressive.

Previous studies pointed out that while more aggressive settings of NLFC increased the accessibility to high-frequency cues for certain phonemes such as sibilant fricatives or affricates, the recognition performance of low-frequency sounds decreased as a tradeoff effect due to the altered formant relationship.²⁷⁻²⁹ In the present study, most participants had steeply sloping hearing loss with very high thresholds above 4 kHz. The audiometric results of these participants suggested that these participants need strong settings with low CT and high CR to achieve “optimal” outcomes for high-frequency sound recognition. With the stronger settings, especially SR2-strong 2, the spectral features of the speech signal were substantially distorted as a result of the low CT1 and CT2, which would have negated potential improvement in speech recognition. Therefore, for those patients with profound hearing loss, the benefit of hearing aids might be limited, and cochlear implantation should be considered as an alternative intervention.

Although the participants showed improved detection and recognition performance with SR2-default and with stronger SR2 settings, the sound-quality rating did not show significant differences between SR2-off and SR2 activated with different settings (►Fig. 5). Nonetheless, sounds with high-frequency components appeared to be rated lower. This might be due to a lack of acclimatization to the high-frequency sounds with SR2 in our participants. Previous research suggested that sound-quality rating was negatively

impacted with more aggressive SR2 settings.³⁰ The stronger SR2 setting (e.g., SR2-strong 1 and SR2-strong 2) alters the speech signals with high-frequency components more aggressively. This could cause unwanted distortion of spectral representations that might in turn lead to unpleasant perception of sound quality. In the present study, although no improved sound-quality rating was found between SR2-off and SR2 activated with different settings, there was no deterioration of sound quality with various SR2 settings as compared with the SR2-off condition. We found no significant differences in the rating scores with the two strong settings or with the weak or default settings. These results suggest that stronger settings of SR2 were tolerated, as the perceived sound quality was not reduced significantly. Glista and colleagues²² reported that while both normal-hearing and hearing-impaired listeners rated the stimuli that were intentionally modified to have poorer sound quality as having below-average scores, the hearing-impaired listeners provided similar above-average rating for sound quality with all static and adaptive NLFC settings. Our findings, together with the results reported in Glista et al,²² suggest that the adaptive NLFC with strong settings may not negatively affect sound-quality rating in hearing-impaired listeners.

The sample size in the present study was relatively small. In addition, all participants had hearing aid experience with Phonak products. Although none of them had used devices fitted with frequency-lowering algorithms before the study was conducted, the subjective sound-quality rating could be biased. Moreover, the present study only tested the thresholds for speech recognition and phoneme detection after the initial fitting process. The limited set of outcome measures used in the present study constrained the generalization of the current findings to overall perception performance with the adaptive NLFC algorithms. It is noteworthy that even though the participants could detect different phonemes after compression, some phonemes such as /s/ and /ç/ become similar to /ʃ/ due to the lowered spectral prominences and become difficult to discriminate.⁷ Therefore, in a future study, more participants should be tested on different aspects of speech perception including detection, distinction, and recognition. Also, the perceptual outcome with different types of NLFC algorithms in comparison to NLFC-off should be evaluated at different times to examine the factor of acclimatization in speech perception with NLFC-fitted hearing aids.

Conclusion

The preliminary study suggested that the adaptive NLFC algorithm (SR2) provides perceptual benefit to Mandarin-speaking people with severe-to-profound hearing loss. Specifically, a significantly better performance in phoneme detection and speech recognition was found in SR2 with a more aggressive setting (i.e., a higher CR and a lower CT) than with the default setting. No major negative effects were observed on sound-quality ratings among different SR2 settings. Given that we only tested the phoneme detection threshold and sentence reception threshold, it remains unclear whether the distortion of spectral features as a result of stronger SR2 settings will lead

to any negative effects on other aspects of speech perception. Therefore, SR2 fittings should be applied with caution and should be accompanied by real ear or electroacoustic verification measurements. Future studies should be conducted to elucidate the relationship between audibility and distortion in NLFC applied in hearing-impaired listeners.

Notes

This study was presented at the 46th Annual Scientific and Technology Conference of the American Auditory Society, Scottsdale, Arizona.

Conflict of Interest

Dr. Qian and Dr. Voss report being employees of Sonova Canada, Inc., outside the submitted work. Dr. Xu reports a grant from Sonova AG, outside the submitted work. All other authors report no conflict of interest.

Acknowledgments

We are grateful to Dr. Fei Chen who provided the recordings of the extended M-HINT sentences. Lexi Neltner provided editorial assistance in the preparation of the manuscript.

References

- Simpson A. Frequency-lowering devices for managing high-frequency hearing loss: a review. *Trends Amplif* 2009;13(02):87–106
- Mao Y, Yang J, Hahn E, Xu L. Auditory perceptual efficacy of nonlinear frequency compression used in hearing aids: a review. *J Otol* 2017;12(03):97–111
- Simpson A, Bond A, Loeliger M, Clarke S. Speech intelligibility benefits of frequency-lowering algorithms in adult hearing aid users: a systematic review and meta-analysis. *Int J Audiol* 2018; 57(04):249–261
- McDermott HJ. A technical comparison of digital frequency-lowering algorithms available in two current hearing aids. *PLoS One* 2011;6(07):e22358
- 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* 2013;34(02):e24–e27
- Alexander JM, Rallapalli V. Acoustic and perceptual effects of amplitude and frequency compression on high-frequency speech. *J Acoust Soc Am* 2017;142(02):908–923
- Yang J, Qian J, Chen X, et al. Effects of nonlinear frequency compression on the acoustic properties and recognition of speech sounds in Mandarin Chinese. *J Acoust Soc Am* 2018;143(03):1578–1590
- McDermott H. Background story: SoundRecover—the importance of wide perceptual bandwidth. Phonak AG March 2010. Available at https://www.phonakpro.com/content/dam/phonakpro/gc_hq/en/resources/evidence/white_paper/documents/technical_paper/bgs_SoundRecover_Bandwidth.pdf. Accessed March 16, 2020
- Rehmann J, Jha S, Allegro Baumann S. SoundRecover2—The first adaptive frequency compression algorithm. Phonak Insight Paper. Stäfa, Switzerland: Phonak; 2016. Available at: https://www.phonaknhs.co.uk/wp-content/uploads/2018/11/Insight_SoundRecover2.pdf. Accessed March 16, 2020
- Akinseye GA, Dickinson AM, Munro KJ. Is non-linear frequency compression amplification beneficial to adults and children with hearing loss? A systematic review. *Int J Audiol* 2018;57(04):262–273
- Brennan MA, McCreery R, Kopun J, et al. Paired comparisons of nonlinear frequency compression, extended bandwidth, and restricted bandwidth hearing aid processing for children and adults with hearing loss. *J Am Acad Audiol* 2014;25(10):983–998
- Ching TY, Day J, Zhang V, et al. A randomized controlled trial of nonlinear frequency compression versus conventional processing in hearing aids: speech and language of children at three years of age. *Int J Audiol* 2013;52(Suppl 2):S46–S54
- Glista D, Scollie S, Bagatto M, Seewald R, Parsa V, Johnson A. Evaluation of nonlinear frequency compression: clinical outcomes. *Int J Audiol* 2009;48(09):632–644
- Wolfe J, John A, Schafer E, Nyffeler M, Boretzki M, Caraway T. Evaluation of nonlinear frequency compression for school-age children with moderate to moderately severe hearing loss. *J Am Acad Audiol* 2010;21(10):618–628
- Wolfe J, John A, Schafer E, et al. Long-term effects of non-linear frequency compression for children with moderate hearing loss. *Int J Audiol* 2011;50(06):396–404
- Wolfe J, Duke M, Schafer EC, et al. Preliminary evaluation of a novel non-linear frequency compression scheme for use in children. *Int J Audiol* 2017;56(12):976–988
- Wright SC, Joshi AP, Schafer EC, et al. A pilot study on the effects of nonlinear frequency compression on performance of individuals who speak Mandarin Chinese. *J Acad Rehabilitative Audiol* 2017; 50:1–12
- 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(06):1765–1787
- Peterson GE, Lehiste I. Revised CNC lists for auditory tests. *J Speech Hear Disord* 1962;27:62–70
- Schmitt N, Winkler A, Boretzki M, Holube I. A phoneme perception test method for high-frequency hearing aid fitting. *J Am Acad Audiol* 2016;27(05):367–379
- Glista D, Hawkins M, Bohnert A, Rehmann J, Wolfe J, Scollie S. The effect of adaptive nonlinear frequency compression on phoneme perception. *Am J Audiol* 2017;26(04):531–542
- Glista D, Hawkins M, Vaisberg JM, Pourmand N, Parsa V, Scollie S. Sound quality effects of an adaptive nonlinear frequency compression processor with normal-hearing and hearing-impaired listeners. *J Am Acad Audiol* 2019;30(07):552–563
- Holube I, Fredelake S, Vlaming M, Kollmeier B. Development and analysis of an International Speech Test Signal (ISTS). *Int J Audiol* 2010;49(12):891–903
- Scollie S, Glista D, Seto J, et al. Fitting frequency-lowering signal processing applying the American Academy of Audiology Pediatric Amplification Guideline: Updates and protocols. *J Am Acad Audiol* 2016;27(03):219–236
- Wong LLN, Soli SD, Liu S, Han N, Huang MW. Development of the Mandarin hearing in noise test (MHINT). *Ear Hear* 2007;28(02): 70S–74S
- Duanmu S, ed. Consonants. In: *The Phonology of Standard Chinese*. 2nd ed. New York, NY: Oxford University Press; 2007:23–25
- Alexander JM. Individual variability in recognition of frequency-lowered speech. *Semin Hear* 2013;34:86–109
- Alexander JM. Nonlinear frequency compression: influence of start frequency and input bandwidth on consonant and vowel recognition. *J Acoust Soc Am* 2016;139(02):938–957
- Alexander JM, Kopun JG, Stelmachowicz PG. Effects of frequency compression and frequency transposition on fricative and affricate perception in listeners with normal hearing and mild to moderate hearing loss. *Ear Hear* 2014;35(05):519–532
- Parsa V, Scollie S, Glista D, Seelisch A. Nonlinear frequency compression: effects on sound quality ratings of speech and music. *Trends Amplif* 2013;17(01):54–68
- McCreery R, Alexander J, Brennan M, Hoover B, Kopun J, Stelmachowicz P. The influence of audibility on speech recognition with nonlinear frequency compression for children and adults with hearing loss. *Ear Hear* 2014;35(04):440–447
- Hopkins K, Khanom M, Dickinson AM, Munro KJ. Benefit from nonlinear frequency compression hearing aids in a clinical setting: the effects of duration of experience and severity of high-frequency hearing loss. *Int J Audiol* 2014;53(04):219–228