Original Article

Can Dual Compression Offer Better Mandarin Speech Intelligibility and Sound Quality Than Fast-Acting Compression?

Yuan Chen[1] ⬤, Lena L. N. Wong[2], Volker Kuehnel[3], Jinyu Qian[4,5], Solveig Christina Voss[4] and Wang Shangqiguo[2]

Abstract

The aim of this study was to evaluate the efficacy of dual compression for Mandarin-speaking hearing aid users. Dual compression combines fast and slow compressors operating simultaneously across all frequency channels. The study participants were 31 hearing aid users with symmetrical moderate-to-severe hearing loss, with a mean age of 67 years. A new pair of 20-channel behind-the-ear hearing aids (i.e., Phonak Bolero B90-P) was used during the testing. The results revealed a significant improvement in speech reception thresholds in noise when switching from fast-acting compression to dual compression. The sound quality ratings revealed that most listeners preferred dual compression to fast-acting compression for listening effort, listening comfort, speech clarity, and overall sound quality at +4 dB signal-to-noise ratio. These results are consistent with predictions based on the theoretical understanding of dual and fast-acting compression. However, whether these results can be generalized to other languages or other dual compression systems should be verified by future studies.

Compression is typically implemented in hearing aids (HAs) to address the loudness recruitment and reduced dynamic range associated with sensorineural hearing loss by reducing gain for high-level sounds while increasing gain for low-level sounds (Moore, 2008). However, the effects of compression systems on the amplified signal depend on the time constants, namely, the attack time (AT) and the release time (RT). Compression can be categorized as fast (AT: 0.5–20 ms; RT: 10–100 ms) and slow (RT > 500 ms) (Cox & Xu, 2010; Kuk et al., 2019; Moore, 2008). Fast gain recovery as determined by the RT improves the audibility of soft sounds and thus may improve speech intelligibility. However, fast-acting compression leads to amplification of low-level noise, resulting in a reduced signal-to-noise ratio (S/N) (Bor et al., 2008). In addition, fast-acting compression tends to reduce intensity contrasts and the modulation depth of speech, introducing distortion of temporal cues (Moore, 2008). Furthermore, many commercial HAs, like the ones used in this study, have 20 or more channels. The number of compression channels may influence the effect of compression speed on speech intelligibility. Multiple channels provide better

[1]Department of Special Education and Counselling, The Education University of Hong Kong, Hong Kong SAR, China
[2]Clinical Hearing Sciences (CHearS) Laboratory, Faculty of Education, The University of Hong Kong, Hong Kong SAR, China
[3]Sonova AG, Stäfa, Switzerland
[4]Innovation Center Toronto, Sonova Canada, Mississauga, Ontario, Canada
[5]Department of Communicative Disorders and Sciences, University at Buffalo, State University of New York, Buffalo, New York, United States

The first two authors contributed equally to this work.

Corresponding author:
Yuan Chen, Department of Special Education and Counselling, The Education University of Hong Kong, Hong Kong SAR, China.
Email: cheny@eduhk.hk

frequency shaping to match the prescribed targets. However, as multichannel compression can change gain across frequency as well as time, both spectral and temporal contrast may be reduced when fast-acting compression is applied with multiple channels (Plomp, 1988; Stone & Moore, 2008).

Although slow-acting compression may provide lower audibility and less information from temporal dips in the background, it is better at preserving original temporal cues and the S/N than fast-acting compression (Moore, 2008). Therefore, slow-acting compression is often preferred over fast-acting compression for reduced distortion, naturalness of sounds, greater comfort, and better sound quality structure (Gatehouse et al., 2006; Hansen, 2002; Moore, 2008; Neuman et al., 1995, 1998). The application of slow-acting compressions is preferred in listeners with poorer sensitivity to temporal fine structure since this leads to greater reliance on the temporal envelop which is better preserved by the slow compression (Moore & Sęk, 2016).

Contrasting results regarding the effect of compression time constants on speech intelligibility exist in literature. Although Davies-Venn et al. (2009), Gatehouse et al. (2006) and Jenstad and Souza (2005) reported better speech intelligibility with fast-acting compression, Reinhart and Souza (2016) and Stone and Moore (2004) reported better speech intelligibility using slow-acting compression. In addition, Salorio-Corbetto et al. (2020) and Novick et al. (2001) did not find a significant difference in the intelligibility of speech in noise between slow- and fast-acting compression. These varied results may be due to the trade-off effects among the distortion introduced by compression, the audibility produced by compression, and the working memory (WM) (Kuk, 2016). According to the Ease of Language Understanding (ELU) model (Rönnberg et al., 2019), any mismatch between the perceptual input and phonological representation stored in long-term memory (LTM) disrupts automatic lexical retrieval, leading to explicit, effortful processing mechanisms based on WM (Füllgrabe & Rosen, 2016). As mentioned, fast-acting compression introduces more distortion of temporal information in speech but is likely to increase audibility for HA users. HA users with better WM may be more likely to tolerate these distortions and benefit from the improved audibility, leading to improved speech intelligibility (Kuk, 2016). In contrast, the distortion introduced by fast-acting compression could be more detrimental for those with poor WM. These listeners are more likely to benefit from slow-acting compression (Kuk, 2016). At present, no protocols have been made for determining the amount of distortion someone with poor WM may tolerate, or for identifying HA users who have difficulty understanding speech processed with fast-acting compression. Thus, clinicians often make arbitrary adjustments to HA compression parameters based on client feedback.

The limitations of fast- and slow-acting compression were the starting point for the development of a dual compression system that combines both these types of compression. Moore & Glasberg (1988) and Moore et al. (1991) described a dual front-end automatic gain control system that included a fast and a slow control voltage generator. The system was determined by the slow-acting control voltage, while the fast-acting control voltage came into operation when intense sound occurred to protect the user from sudden transient loud sounds without affecting the long-term gain. The HAs used in this study blended fast and slow dynamic range compression (DRC) simultaneously across all frequency channels at an adjustable ratio (Figure 1). Compression was not entirely independent in the different channels but a coupling across neighboring bands that was motivated by cochlea Bark filter resolution was applied (Phonak, 2000). The overall DRC was applied as a weighted average of gain that was smoothed with either a fast or a slow set of time constants. The dynamic behavior for fast and dual compression was similar for all frequency channels, and the compression time constants were fixed. More specifically, the fast-acting compressor operated with an AT of 10 ms and an RT of 60 ms (Figure 2). The dual compressor (see Figure 2A and B) combined the fast-acting compressor described earlier with a slow-acting compressor that had an AT of 1 s and an RT of 8 s. In this implementation, the slow-acting compressor dominated the overall system dynamics. The resulting AT and RT of the dual compressor were 1.2 s and 7.1 s, respectively (Figure 2A). Figure 2B, which is an enlarged section of Figure 2A, depicts the recovery process of the output signal. At $t = 40$ s, there was a fast recovery of the output from 44 dB sound pressure level (SPL) at RT of 60 ms for the fast compression settings. The slow recovery of the output persisted until $t = 47$ s where output returned to 4 dB below the stationary gain (ANSI/ASA S3.22-2014), but a partial recovery with a fast RT of 60 ms can be observed for the dual compression system as well. The lower compression paths used identical compression knee points. The theoretical advantage of a dual compressor (such as this one) over a compression system with a fast AT and slow RT is the increased audibility of weaker sounds after strong sounds that trigger the fast attack. This can help avoid the wearers' perception of devices being "dead" after sudden strong sounds.

The dual compression system was matched in root mean square (RMS) output sound pressure level to the fast compression system. This was achieved by choosing a similar ratio between the RT and AT constants for the two compressors (Figure 1). Choosing a similar ratio of AT and RT for the fast and slow compressor results in
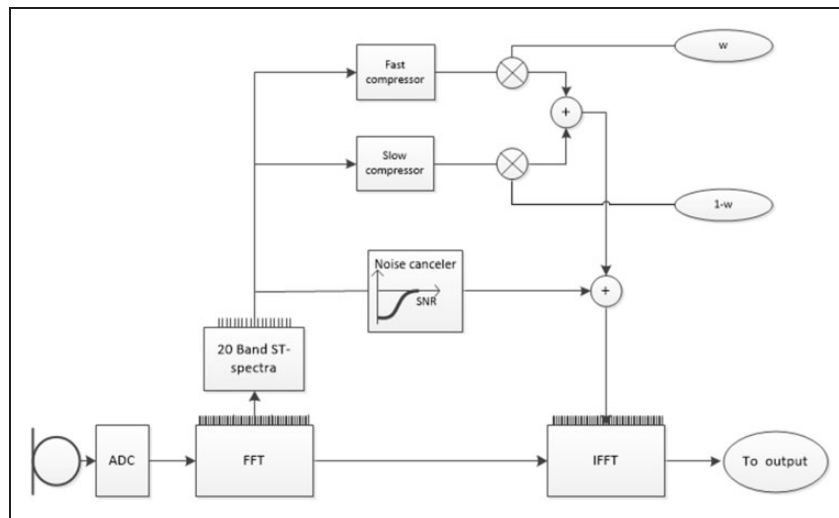
Figure 1. Schematic Diagram of the Signal Processing Applied in This Study. The processing of the gain with fast-acting and slow-acting gain regulation and a noise canceller used the same input signal of a 20-frequency-band short-term spectrum derived from a FFT filter bank. Fast and slow compression and the noise canceller delivered a gain that was combined additively in the logarithmic domain. These gains were applied as filter weights to the filter bank signal before it was converted back to a time-domain signal with an IFFT. The dual compressor had a control parameter, w, which determined the amount of fast-acting and slow-acting compression. A value of $w = 1$ resulted in fast-acting compression only, whereas $w = 0$ resulted in slow-acting compression only. In the study, $w = 1$ was used for fast-acting compression and $w = 0.4$ for dual compression. In each of the 20 bands, the noise canceller derived a signal-to-noise ratio (S/N) estimate and calculated a gain depending on this S/N. For low S/N values, negative gain was applied, whereas for large S/N values, the gain approached 0 dB. ADC = Analog-to-digital conversion; FFT = fast Fourier transform; IFFT = inverse fast Fourier transform; SNR = signal-to-noise ratio; ST = short-term.

similar output for soft to loud speech signal without the need to choose different knee points or compression ratios for the two compressors. Table 1 shows the 2 cc coupler gains for the International Speech Test Signal (ISTS) for 50, 65, and 80 dB input levels. The gains differed by less than 2 dB between fast and dual compression. Figure 3 shows a separation between the 99th and 30th percentile output levels for the dual compressor than for the fast-acting compressor. This indicates that less distortion is introduced as the modulation depth of speech is reduced less by the dual compressor. We expected the dual compressor to outperform the fast-acting compressor for the intelligibility of speech in noise as well as for sound quality rating.

In summary, this study compared speech intelligibility and sound-quality preference between fast-acting compression and dual compression using commercially available HAs. According to previous studies, better hearing thresholds (Souza & Sirow, 2014), higher education level (Chen et al., 2020), previous experience with a given HA signal processing scheme (Ng & Rönnberg, 2020) and better WM (Cox & Xu, 2010) were significantly correlated with speech intelligibility. The effects of these variables on speech intelligibility and sound-quality preference with different compression schemes were also examined.

Methods

Participants

HA users from the Shenkang Hearing Center (Beijing, China) were recruited via telephone. Patient files were reviewed, and all patients who met the following inclusion criteria were contacted: (a) symmetrical moderate to severe hearing loss, defined as an interaural difference of less than 10 dB HL at octave frequencies from 250 to 8000 Hz; (b) native monolingual standard Mandarin speakers living in Beijing; (c) normal cognitive function and a passing score on the Montreal Cognitive Assessment (MoCA, Chinese version; Yu et al., 2012); and (d) bilateral HA use for at least 3 months with a daily wearing time of at least 5 hr.

A total of 31 HA users agreed to participate (23 males and 8 females). Their age was between 33 and 87 years (mean = 67, standard deviation [SD] = 14, median = 69). The mean pure-tone hearing thresholds are shown in Figure 4. The mean education level was 6.1 years (SD = 1.7, median = 6.0). Participants had used HAs for an average of 10.5 years (SD = ±8.6, median = 8.0).

Seventeen participants had used HAs equipped with the same dual compression scheme as in this study for more than 6 months. The other 14 participants had no experience with dual compression HAs and were
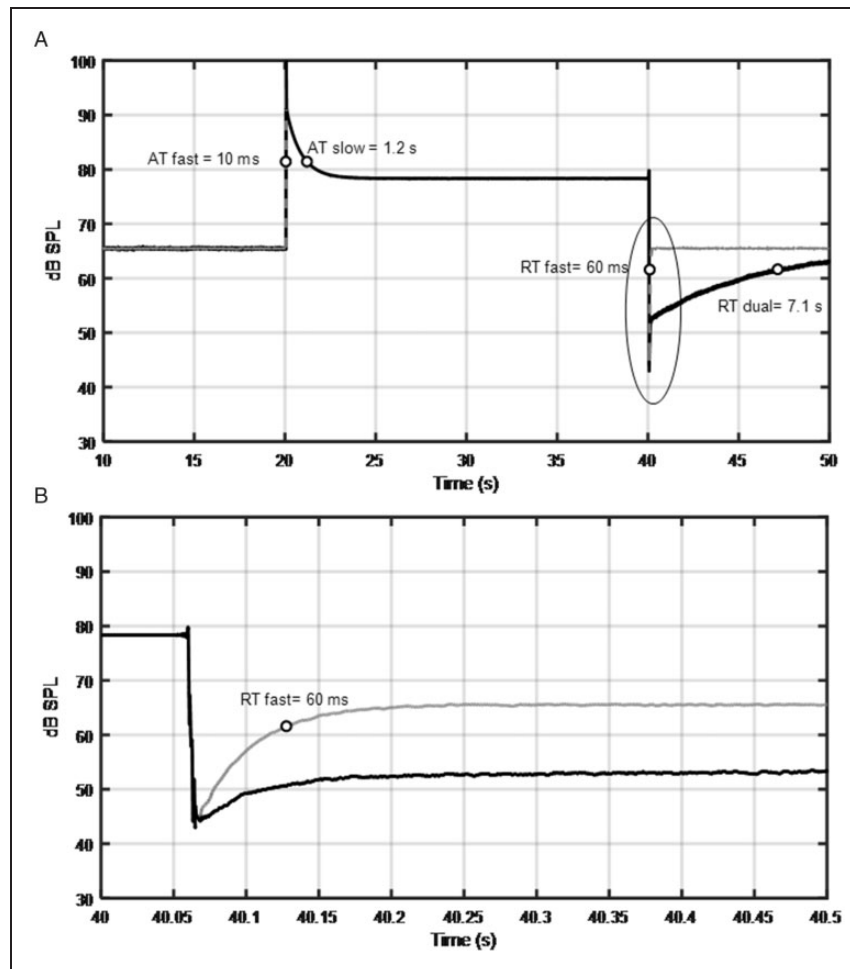
Figure 2. AT and RT for fast compression and dual compression. A: American National Standards Institute (ANSI) S3.22-2014 step response and derived AT and RT for fast compression (gray line) and dual compression (solid black curve). The fast compression had an AT of 10 ms and RT of 60 ms, whereas the dual compression had an AT of 1.2 s and RT of 7.1 s. The static compression ratio was set to CR = 2.7. The section marked with an ellipse is shown magnified in Panel B. B: An enlarged section of Panel A at *t* = 40 s showing the release behavior for dual compression (solid line) and fast compression (gray line). The gain for dual compression recovered to 4 dB below the stationary gain at *t* = 47.1 s (see Panel A), which is outside the range visible in this detailed view. AT = attack time; RT = release time; SPL = sound pressure level.

Table 1. Measured 2 cc Coupler Gain and MPO Recorded for the Mean Hearing Loss Configuration of the Participants of the Study (Average for Left and Right Ears).

Frequency (Hz)	250	500	750	1000	1500	2000	3000	4000	6000	8000
MPO (dB SPL 2 cc)	95	97	100	102	104	103	106	106	83	67
2cc gain at 80 dB (dB)	1	2	7	14	23	24	31	28	5	−9
2cc gain at 65 dB (dB)	12	15	19	25	33	35	42	39	15	2
2cc gain at 50 dB (dB)	18	22	26	31	37	39	47	45	20	7
Static CR	1.7	1.7	2	2	2.2	2.5	2.7	2.8	3	2.9
Compression knee point (dB SPL in 1/3 octave band)	47	48	46	43	39	38	34	34	36	38

Note. The gains are shown as 2 cc coupler gains for ISTS signals (IEC 60118-15, 2015) at 80, 65, and 50 dB SPL input levels. The static CRs are indicated as well as the compression knee points expressed as third octave band levels. The measured 2 cc response curves differed by less than 2 dB between fast and dual compression and hence the average across fast and dual compression is presented here. MPO = maximum power output; CR = compression ratio; SPL = sound pressure level.
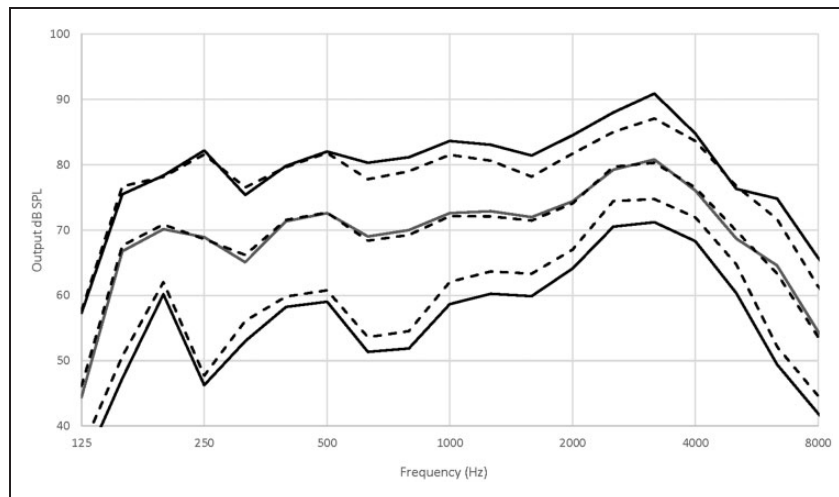
Figure 3. Third Octave Output Spectra Recorded on KEMAR (Burkhard & Sachs, 1975) for 20 s of the International Speech Test Signal (ISTS; Holube et al., 2010; International Electrotechnical Commission, 2012) Played at 65 dB SPL. The signals were recorded with a prescribed gain for a flat 60 dB HL hearing loss according to the Adaptive Phonak Digital Tonal (APDT) gain prescription rule. The three solid lines show the 30th percentile, the root mean square (RMS), and the 99th percentile of the recorded third octave spectrum for the dual compression setting. Level percentiles were calculated using short-term spectra of 125 ms duration having 50% overlap. Dashed lines show the same values for fast-acting compression. The RMS levels of the two compression strategies were matched within ±2 dB. The dynamic range for the dual compressor was larger than that for the fast-acting compressor and closer to the dynamic range of the uncompressed ISTS (not shown), which is about 30 dB between the 30th and 99th percentiles (Holube et al., 2010). SPL = sound pressure level.
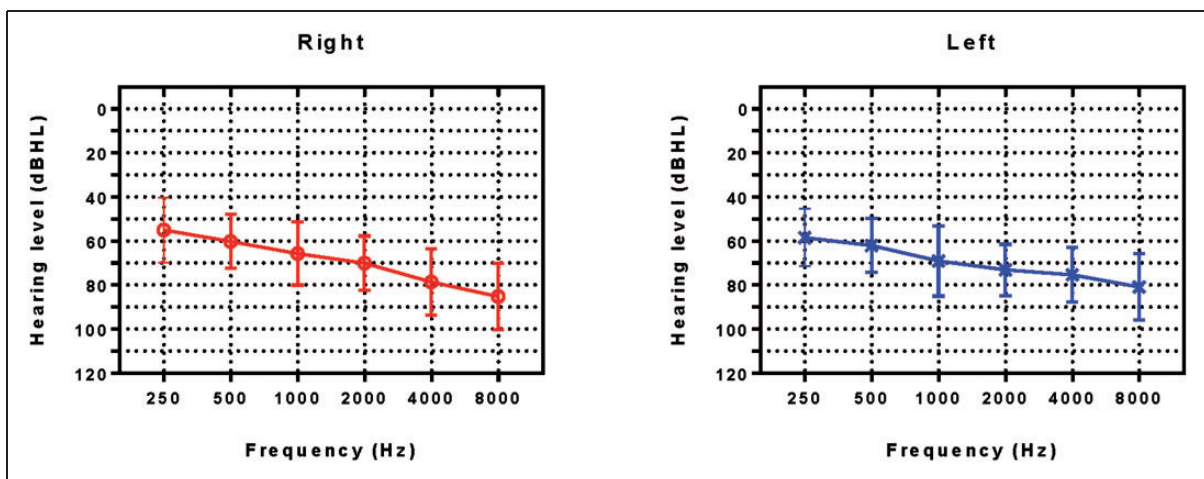
Figure 4. Mean and Standard Deviation of the Pure-Tone Hearing Thresholds of Participants' Right and Left Ears.

wearing HAs from Phonak, Oticon, or Resound that did not implement the dual compression scheme used in this study.

Materials and Test Equipment

A sentence recognition test and sound quality paired comparison measures were conducted. In addition, WM was measured.

Sentence Recognition. The Mandarin Hearing in Noise Test (MHINT) was used to measure sentence reception thresholds (SRTs) in quiet and noisy conditions. The MHINT is an adaptive test that measures the presentation level or S/N ratio at which half the keywords are correctly recognized in quiet or in noise and was developed using the same rationale as the English Hearing in Noise Test (Wong, Ho, et al., 2007). The test corpus contains 12 sentence lists, and each list consists of 20 sentences. Distributions of phonemes and lexical tones are balanced between the lists. Each sentence consists of 10 characters, and the background noise is a speech-spectrum-shaped noise initiated 1 s before the start of

each sentence (Wong, Ho, et al., 2007) and ending 1 s after the sentence.

Four test conditions were used (two compression strategies and two listening conditions). Each condition was evaluated using one MHINT list. In the noisy condition, speech and noise were presented from a loudspeaker situated 1 m away from the participants at 0° azimuth, with the noise level fixed at 65 dBA. The level of the speech was adjusted to determine the SRT. Before testing, calibration was conducted using a sound level meter placed at the center of the head position of participant. The order of the test conditions and sentence lists was randomized for each participant. Participants were encouraged to make a guess, even if they did not hear the whole sentence. A practice list was presented for participants to get used to the test procedure. A 5- to 10-min break was offered upon request or when the participant seemed tired.

Sound Quality Paired Comparison. Paired comparisons were conducted to evaluate preferences. Participants listened to the same recording under two compression settings. The target speech was presented in noise at +1 dB and +4 dB S/N, and participants were asked to indicate their preference between dual and fast-acting compression in terms of the listening effort, listening comfort, speech clarity, and overall sound quality, all in one block. First, participants were asked to indicate their preferred compression setting and then to indicate how much better the preferred compression setting was compared to the other. The order of the two compression settings and test conditions (i.e., +1 dB and +4 dB S/N) was randomized for each participant.

A recording of restaurant noise played at 66 dBA and a recording of a Mandarin translation of "The North Wind and the Sun" (Holube et al., 2010) played at 70 dBA and at 67 dBA were used as target speech material to yield S/Ns of +4 dB and +1 dB. Before the paired comparison, oral and written instructions were given to ensure that the participants fully understood the five rating questions. The paired comparison rating scale ranged from 1 to 7 for listening effort, listening comfort, speech clarity, and overall quality: 1—dual compression was much better, 2—dual compression was better, 3—dual compression was slightly better, 4—no difference, 5—fast-acting compression was slightly better, 6—fast-acting was compression better, 7—fast-acting compression was much better. For loudness judgments, although the rating scale still ranged from 1 to 7, participants were asked to judge which compression scheme sounded louder. Letters (A or B) were randomly assigned to the stimuli presented using the dual or fast-acting compression in order not to reveal the order of the compression conditions to participants. The speech and noise were replayed as many times as needed.

Working Memory. WM was measured using the One Back Test (OBK) from the CogState Battery, which has been adapted for use among Chinese speakers and has good reliability and validity (Zhong et al., 2013). In the OBK, a playing card is presented at the center of the screen, and the participant is asked to indicate whether the card is the same as the previous card by pressing keys on a keyboard. The results are used to analyze the speed of performance (mean of the $\log^{10}$ transformed reaction times for correct responses). Lower scores indicate better WM.

Procedures

A pair of Phonak Bolero B90-P, 20-channel behind-the-ear HAs were used. Participants' own custom ear molds were used during testing. If participants wore in-the-ear HAs, disposable universal ear tips were used. The vent of the earmolds was prescribed by the fitting software, typically with vents smaller than 1 mm. Phonak Target fitting software (version 5.1) was used to fit and fine-tune the HAs. The Adaptive Phonal Digital Tonal (APDT; Wong et al., 2018) prescription rule was used to prescribe HA gain parameters. The APDT rule prescribed more gain for soft low-frequency inputs than an older prescription rule, the Adaptive Phonak Digital (APD; Latzel, 2013), to enhance the audibility of tonal information in speech, as low frequencies carry more information for Mandarin than English speech intelligibility. Dual compression, as described earlier, is implemented in the APDT as the default compression method. A research version of the fitting software was used to fit participants with fast-acting or dual compression based on APDT and to allow manual switching between the two types of compression. The electroacoustic properties of the HAs were checked both for proper functioning and for matching to the prescribed settings prior to and after the study. Table 1 shows measured 2 cc coupler gains for soft, moderate, and loud speech for the average hearing loss of the participants.

During initial fitting, a real-ear loop-gain measurement was conducted to determine the maximum stable gain before feedback. The fitted gain was limited to this maximum stable gain. All functions were turned off except for the feedback management and noise reduction (NR). NR was set at 14, which is a moderate setting activated with an AT of several seconds and RT of several milliseconds. This NR function employs a Wiener filter-type algorithm in all the HA channels. The maximum attenuation applied by the NR was 7 dB for an unmodulated broadband noise. This setting was used since Wong et al. (2018) found it to be the preferred setting among a range of mild to more aggressive NR settings. Thus, it was expected that most HA users in this
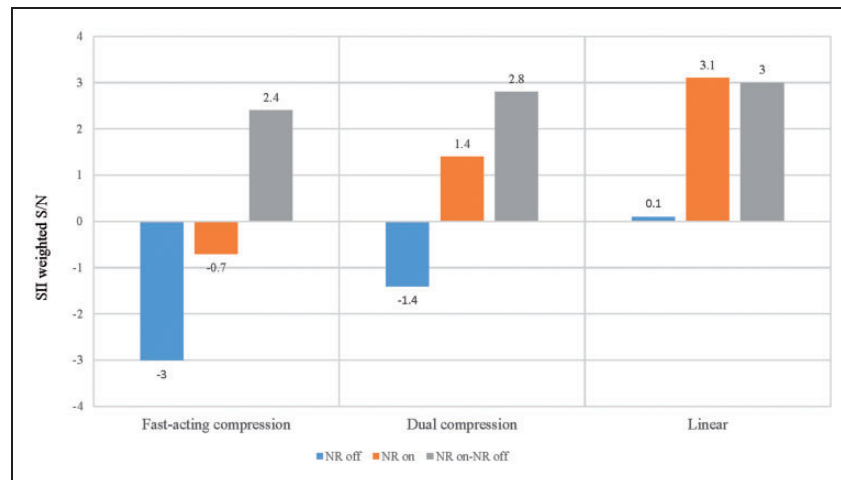
Figure 5. SII Weighted Signal to Noise Ratio (S/N) in dB of the Hearing Aid Output for the Average Amplification Setting Used in the Study. NR = noise reduction.

study would prefer this setting. More information about this NR setting can be found in Wong et al. (2018).

To verify that there was only little interaction between the compression and NR, additional technical measurements were performed using the phase inversion method described by Hagerman and Olofsson (2004). The International Speech Test Signal (ISTS) signal and the spectrally matched IFnoise (European Hearing Instrument Manufacturers Association, 2016; Holube et al., 2010) were played back simultaneously at 65 dB SPL. An HA was programmed to the average hearing loss of the participants with the gain according to Table 1. The output of the HA was recorded using a 2 cc coupler. To ensure that the phase inversion method was valid, we confirmed that the signal and noise extracted with the phase inversion method had only minimal contributions from the other signal (below −15 dB). The SII (ANSI/ASA S3.5-1997 (R2017)) weighted S/N of the output signal was determined for the conditions NR off or NR (set to 14), both combined with fully linear gain, fast-acting compression or dual compression setting. Figure 5 shows the SII-weighted S/N. With the fast compression, the overall S/N was 3 dB lower than for the linear setting and the S/N improvement with the NR was reduced to 2.4 dB compared to 2.8 dB with the slow compression. The change of 0.4 dB was markedly smaller than the effect of compression on the S/N, suggesting little interaction between NR and compression system.

For fine-tuning, the choice of compression type was randomized. The settings for 15/31 randomly chosen participants were fine-tuned using dual compression, while the settings for the remaining participants were fine-tuned using fast-acting compression. A recording of "The North Wind and the Sun" (Holube et al., 2010) was presented at 65 dBA via a loudspeaker 1

meter in front of participants (0° azimuth). Participants were asked to comment on the overall loudness and the loudness balance between ears. The broadband gain for 65 dB input (G65) was adjusted in 1 dB steps until participants were satisfied with the balance and loudness comfort in both ears. Next, to ensure that the loudness of the music was comfortable, a short, lively orchestral piece, which provided greater variations in sound level than average speech, was presented at 70 dBA. Participants were asked to judge the sound quality of the music. To adjust the tonal balance, the gain in the frequency regions above and below 1.5 kHz was increased or decreased in 1 dB steps for all input levels until the participants indicated that the music was neither boomy nor tinny. The two ears were fitted together. Nearly 3/4 of the participants preferred fine-tuning of 1 dB while no changes were made for the others. After the HA fitting and fine-tuning, sentence recognition and sound quality comparisons were conducted. The order of compression settings and measures for sentence recognition was randomized for each participant. All testing was conducted in a sound booth in the Shenkang Hearing Center (Beijing, China) with background noise lower than 35 dBA. All participants received an honorarium of 100 RMB (equal to about 13 USD). Ethical approval was obtained from the University of Hong Kong. Written informed consent was obtained from all participants prior to the study.

Results

Sentence Recognition in Quiet and in Noise

Pearson product–moment correlation coefficients with Bonferroni corrections showed that age, hearing
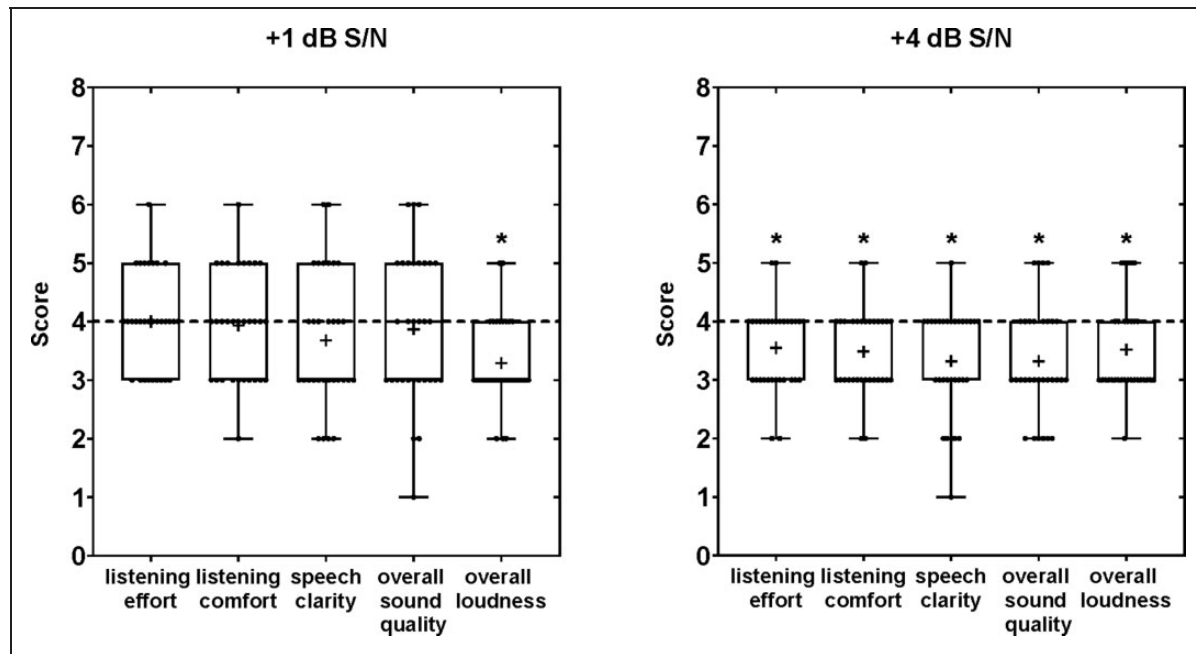
Figure 6. Box-and-Whisker Plots of Sound Quality Paired Comparisons for the +1 dB and +4 dB S/N Conditions. The "+" indicates mean score, the box represents the quartiles, and the whiskers indicate the range of ratings. The "*" indicates a score significantly lower than 4, suggesting a preference for dual compression.

thresholds, education level, previous experience with dual compression, and WM were not significantly related to SRTs for dual compression and fast-acting compression in quiet or in noise. Therefore, data from all participants were combined for further analysis.

Paired samples *t*-tests showed no significant difference between SRTs for dual-compression (Mean = 49.1, SD = 3.7) and fast-acting compression (Mean = 48.7, SD = 4.1) in quiet (*p* > .05) but a significant difference was found between SRTs for dual compression (Mean = 0.9, SD = 2.2), and for fast-acting compression (Mean = 1.8, SD = 2.1) in noise; $t(30) = -4.0$, $p < .001$, Cohen's $d = 0.72$.

Sound-Quality Preferences. Spearman's rank-order correlation coefficients showed that hearing thresholds, age, education level, WM, and previous experience with dual compression were not significantly related to compression setting preferences. Therefore, data from participants with and without prior experience with dual compression were combined for further analysis.

Figure 6 shows sound quality paired comparison results at +1 dB and +4 dB S/N. A rating of 4 indicated no difference between the two compression strategies. A score above 4 indicated a preference for fast-acting compression or that fast-acting compression sounded louder than dual compression, while a score below 4 indicated a preference for dual compression or that dual compression sounded louder than fast compression. Mean

ratings at +4 dB S/N were 3.5 (SD = 0.7) for listening effort, 3.5 (SD = 0.7) for listening comfort, 3.3 (SD = 0.9) for speech clarity, 3.3 (SD = 0.9) for overall sound quality, and 3.5 (SD = 0.8) for loudness. Wilcoxon signed-rank tests with Bonferroni correction indicated that paired comparison ratings for all five sound qualities were significantly lower than 4, suggesting slight preferences for dual compression (*p*s < .05) at +4 dB S/N with effect sizes (*r*) ranging from −0.52 to −0.60. In other words, participants found that listening to speech with dual compression at +4 dB S/N required less effort and was more comfortable, clearer, louder, and of better overall quality.

At +1 dB S/N, only the loudness rating was significantly lower than 4. In other words, participants found that dual compression was louder than fast-acting compression, but there was no overall preference for dual compression regarding listening effort, listening comfort, speech clarity, or overall sound quality. Correlation analyses among preference ratings obtained at +1 dB S/N were not conducted since participants reported extremely low speech intelligibility in this condition.

At +4 dB S/N, positive correlations (with Bonferroni correction) were found among four of the five sound quality measures: listening effort, listening comfort, speech clarity, and overall quality (*r*s = 0.52–0.60). Loudness judgments were not correlated with any sound quality judgments (Table 2). This was confirmed by a principal component analysis with orthogonal

Table 2. Spearman's Rank Correlation Coefficients Between Paired Comparison Sound Quality Ratings at +4 dB S/N.

	Comfort	Clarity	Overall	Loudness
Effort	0.66**	0.56*	0.52*	0.07
Comfort		0.73**	0.85**	0.15
Clarity			0.78**	0.15
Overall				0.27

*p < .05. ** p < .01 after Bonferroni correction.

rotation (varimax). Only one component was found, which yielded an eigenvalue value greater than 1 and explained 62.5% of the variance. Factor loadings for listening effort, listening comfort, speech clarity, overall quality, and loudness were 0.77, 0.92, 0.88, 0.91, and 0.30, respectively. Thus, listening effort, listening comfort, speech clarity, and overall quality represented a single dimension, with loudness being different from the other four.

Discussion

In this study, we evaluated the effects of dual and fast-acting compression on SRTs and sound-quality preferences. In quiet, there was no difference in SRTs for the two compression settings. However, in noise, dual compression yielded better SRTs and higher sound-quality preferences than fast-acting compression.

Sentence Recognition

Slow-acting compression provides less amplification for short-term low-level inputs than fast-acting compression (Moore, 2008). The dual compression investigated in this study was designed to give an RMS output SPL matched with that of the fast-acting compression, while offering larger signal dynamics for speech signals (Figure 3). However, there was no significant difference in SRTs in quiet between dual and fast-acting compression settings, suggesting that the slightly lower low-level gain of the dual-compression system had no material effects in quiet.

In noise, SRTs with dual compression were significantly lower (better) than those with fast-acting compression. As mentioned previously, this may have happened because dual compression is better at preserving temporal cues than fast-acting compression. The average improvement in SRTs in noise was 0.9 dB, which corresponds to approximately 10% improvement in speech intelligibility for the MHINT (Wong, Soli, et al., 2007). However, this improvement was small and may not be clinically noticeable by some listeners.

Sound Quality Comparisons

Paired comparisons were performed at +1 and +4 dB S/N. Based on the SRTs and informal communications, we assumed that most participants could understand the content of the continuous discourse at +4 dB S/N and thus were able to make reliable judgments of sound quality.

Slight preferences for dual compression at +4 dB S/N were found for listening effort, listening comfort, speech clarity, and overall sound quality. This may be attributed to the slow-acting compressor incorporated in the dual compression setting in this study. Previous studies reported that slow-acting compression was likely to be preferred over fast-acting compression for listening effort, listening comfort, and overall sound quality (Cox & Xu 2010; Hansen, 2002). The positive findings in this study add to the literature comparing dual and fast-acting compression and may help clinicians to decide whether to adopt this type of dual compression.

Loudness judgments were not correlated with sound-quality preferences, suggesting that loudness is a dimension different from the sound-quality attributes. This was also found in our previous study (Wong et al., 2018). Although the dual compression used here was matched to the fast compression for RMS level, signals processed with the dual compression was perceived to be slightly louder than those processed with fast compression at +1 dB and +4 dB S/N. This may be explained by the fact that peak levels were somewhat higher with dual compression than with fast-acting compression (Figure 3) and the loudness of dynamic signals is known to be strongly influenced by peak levels (Glasberg & Moore, 2002).

For the paired comparisons performed at +1 dB S/N, there was no preference for dual compression over fast-acting compression in listening effort, listening comfort, speech clarity, or overall sound quality. The lack of sound-quality preferences may have been due to participants not being able to discriminate the speech in noise well enough to make judgments. Smeds et al. (2015) showed that listening in noise mostly occurs in positive S/N settings; thus, one may question whether measuring sound quality at a low S/N is ecologically valid. The findings of this study suggest that measuring sound quality at a low S/N ratio may not provide additional information for the evaluation of compression in terms of sound-quality preferences. HA users can make judgments of sound quality when they are able to hear the speech in noise (e.g., at +4 dB S/N), but they may not be able to do so at a low S/N (e.g., +1 dB S/N or below). This finding is consistent with those of Preminger and Van Tasell (1995) and Souza et al. (2013), who suggested that low intelligibility would dominate other quality

attributes. Therefore, speech quality should only be evaluated when speech intelligibility is acceptable.

Compression is often implemented in combination with NR algorithms to improve the S/N by reducing HA gain for background noises while preserving gain for speech (Wong et al., 2018). The NR algorithm can be viewed as dynamic range expansion, often with comparable smoothing time constants to the compression, and hence, potentially counteracting the effects of compression. The combined effects, which may vary greatly across commercially available HAs (Brons et al., 2015), must be considered. Kortlang et al. (2018) evaluated three combined implementations of NR and compression. They showed that parallel operation of NR and compression denoted as (p) in their publication, reduced noise annoyance for listeners with hearing loss more effectively than a serial implementation. This parallel implementation is comparable to the one used in this study (Figure 1), where the NR was applied to the signal independent from the compression, and little interaction between the NR and the compression was expected. In addition, the NR had a short RT (<10 ms), which helped to decouple the NR and the compression.

The Effects of Hearing Thresholds, Education Level, Previous Experience With Dual Compression and Working Memory

Although previous studies have shown correlations between hearing thresholds and SRTs (Souza & Sirow, 2014), and between education level and SRTs in noise (Chen et al., 2020), such correlations were not found in this study. This may be partially attributed to the relative narrow range of hearing thresholds and education level, as most participants ($n = 27$) exhibited moderately severe to severe hearing loss (i.e., 56–82 dB), and 30 out of the 31 participants were graduates of junior middle school or below (i.e., ≤9 years of education).

Previous studies suggested that individuals with better WM had better in speech perception in noise with fast than with slow compression, while those with low WM performed better with slow compression than with fast compression (Cox & Xu, 2010; Ohlenforst et al., 2016; Souza & Sirow, 2014). This study failed to find a significant relationship between compression type and WM. There may be two reasons for this. First, previous studies tended to categorize participants as having either high or low WM on the basis of the median for the group (Cox & Xu, 2010; Ohlenforst et al., 2016; Souza & Sirow, 2014). This categorization was quite arbitrary and might have led to bias, especially when the sample was small. The WM scores were regarded as a continuous variable in this study as we were interested in comparing the effects of compression type on speech

perception and sound-quality preference when controlling the effects of WM. When WM is treated as a continuous variable, a relationship between WM and compression type may not be found (e.g., Kuk et al., 2019). In addition, the relationship between WM and speech perception was smaller for participants who has used HAs for a longer period (Ng & Rönnberg, 2020; Rählmann et al., 2017). According to the ELU Model (Rönnberg et al., 2019), WM comes into play when a mismatch exists between the perceptual input (e.g., phonology, prosody, syntax, and semantics) and the representation stored in LTM. Although signal processing in HAs could cause such a mismatch, after consistent exposure to distorted information via HAs, newly established and recalibrated internal representations could gradually supplement the existing LTM representations, weakening the WM–SRT relationship (Ng & Rönnberg, 2020). In this study, participants had used HAs for an average of 10.5 years and thus were experienced HA users. Therefore, the nonsignificant WM–SRT relationship is not surprising. This is consistent with the findings of Rählmann et al. (2017). Using a master HA, they found that the relationship between WM and SRT significantly weakened for listeners with more than 7 years of HA experience compared to those with less than 3.5 years of HA experience and became nonsignificant for listeners with the most experience, although the HA signal processing by Rählmann et al. (2017) was not the same as in the participants' own HAs.

The long-term use of HA in this study may also explain the nonsignificant relationship between previous experience with dual compression and SRTs, since familiarity with HA signal processing may have alleviated a possible mismatch between the speech input and LTM representations, even when the HA signal processing was not the same as for their own HAs (Rählmann et al., 2017).

Limitations

We did not compare fast-acting and dual compression with slow-acting compression. This makes it difficult to unambiguously allocate the improvements in SRTs and preference for dual compression to the slow-acting compression component or the combination of fast- and slow-acting compression.

It is worth noting that the NR function was turned on when the SRTs were measured in noise for better ecological validity. However, we expected little interaction between the NR and the compression, as discussed earlier. The findings of this study may not apply to HAs employing different NR functions or a different interaction between NR and compression.

The speech-shaped noise used for the SRT measurement might not have reflected the potential advantages

of the fast compression for backgrounds with amplitude fluctuations. Fast compression is better than slow compression at restoring the audibility of weak sounds rapidly following intense sounds, providing the potential for listening in the dips (Moore, 2008). In backgrounds such as multitalker babble, fast-acting compression allow better "glimpses" of the target sound (Moore et al., 1999). In addition, fast-acting compression improves the ability to detect a weak consonant following a relatively intense vowel (Moore, 2008). Chen et al. (2013) reported a 3:1 intelligibility advantage of vowel-only sentences over consonant-only sentences in Mandarin as compared to an intelligibility advantage of 2:1 in English, suggesting that consonants in Mandarin do not contribute as much to sentence intelligibility as consonants in English (Chen et al., 2017). Therefore, future studies are warranted to examine whether the advantages of slow or dual compression over fast compression are language dependent.

Conclusions

Dual compression, as implemented in this study, yielded slightly better SRTs for speech in speech-spectrum-shaped noise than fast-acting compression. In addition, participants slightly preferred dual compression over fast compression for sound quality at +4 dB S/N. Experience with dual compression, age, gender, and degree of hearing loss did not affect the results, suggesting that clinicians need not be concerned about these factors when switching from fast to dual compression. Whether these results predict performance at other S/Ns using different types of noise or in real-life situations and whether they could be applied to other types of fast and dual compression systems should be assessed in future studies.

Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Wong and Chen declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The other authors are employees of Sonova AG, which manufactured hearing aids used in the study.

ORCID iD

Yuan Chen ⓘD https://orcid.org/0000-0001-6706-0962

References

American National Standards Institute. (2017). *Methods for Calculation of the Speech Intelligibility Index*. ANSI/ASA S3.5-1997 (R2017).

American National Standards Institute. (2014). *Specification of Hearing Aid Characteristics*. ANSI/ASA S3.22-2014.

Bor, S., Souza, P., & Wright, R. (2008). Multichannel compression: Effects of reduced spectral contrast on vowel identification. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1315–1327. https://doi.org/10.1044/1092-4388(2008/07-0009)

Brons, I., Houben, R., & Dreschler, W. A. (2015). Acoustical and perceptual comparison of noise reduction and compression in hearing aids. *Journal of Speech, Language, and Hearing Research*, *58*(4), 1363–1376. https://doi.org/10.1044/2015_JSLHR-H-14-0347

Burkhard, M. D., & Sachs, R. M. (1975). Anthropometric manikin for acoustic research. *The Journal of the Acoustical Society of America*, *58*(1), 214–222. https://doi.org/10.1121/1.380648

Chen, Y., Wong, L. L., Qian, J., Kuehnel, V., Christina Voss, S., & Chen, F. (2020). The role of lexical tone information in the recognition of Mandarin sentences in listeners with hearing aids. *Ear and Hearing*, *41*(3), 532–538. https://doi.org/10.1097/AUD.0000000000000774

Chen, F., Wong, L. L., & Wong, E. Y. (2013). Assessing the perceptual contributions of vowels and consonants to Mandarin sentence intelligibility. *The Journal of the Acoustical Society of America*, *134*(2), EL178–EL184. https://doi.org/10.1121/1.4967453

Chen, Y., Wong, L. L., Zhu, S., & Xi, X. (2017). Vocabulary development in mandarin-speaking children with cochlear implants and its relationship with speech perception abilities. *Research in Developmental Disabilities*, *60*, 243–255. https://doi.org/10.1016/j.ridd.2016.10.010

Cox, R. M., & Xu, J. (2010). Short and long compression release times: Speech understanding, real-world preferences, and association with cognitive ability. *Journal of the American Academy of Audiology*, *21*(2), 121–138. https://doi.org/10.3766/jaaa.21.2.6

Davies-Venn, E., Souza, P., Brennan, M., & Stecker, G. (2009). Effects of audibility and multichannel wide dynamic range compression on consonant recognition for listeners with severe hearing loss. *Ear and Hearing*, *30*(5), 494–504. https://doi.org/10.1097/AUD.0b013e3181aec5bc

European Hearing Instrument Manufacturers Association. (2016). *Description and terms of use of the IFFM and IFnoise signals*. http://www.ehima.com/wp-content/uploads/2016/06/IFFM_and_IFnoise.zip

Füllgrabe, C., & Rosen, S. (2016). On the (un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology*, 7, 1268. https://doi.org/10.3389/fpsyg.2016.01268

Gatehouse, S., Naylor, G., & Elberling, C. (2006). Linear and nonlinear hearing aid fittings—2. patterns of candidature. *International Journal of Audiology*, 45(3), 153–171. https://doi.org/10.1080/14992020500429484

Glasberg, B. R., & Moore, B. C. J. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5), 331–342. https://doi.org/10.1121/1.2999342

Hagerman, B., & Olofsson, A. (2004). A method to measure the effects of noise reduction algorithms using simultaneous speech and noise. *Acta Acustica*, 90, 356–361. https://doi.org/10.1177/2331216518803198

Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear and Hearing*, 23(4), 369–380. https://doi.org/10.1097/00003446-200208000-00012

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an international speech test signal (ISTS). *International Journal of Audiology*, 49(12), 891–903. https://doi.org/10.3109/14992027.2010.506889

International Electrotechnical Commission. (2012). *Electroacoustics-hearing aids–part 15: Methods for characterizing signal processing in hearing aids with a speech-like signal. IEC 60118-15*. https://webstore.iec.ch/publication/793

Jenstad, L. M., & Souza, P. E. (2005). Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *Journal of Speech, Language, and Hearing Research*, 48(3), 651–667. https://doi.org/10.1044/1092-4388(2005/045)

Kortlang, S., Chen, Z., Gerkmann, T., Kollmeier, B., Hohmann, V., & Ewert, S. D. (2018). Evaluation of combined dynamic compression and single channel noise reduction for hearing aid applications. *International Journal of Audiology*, 57(sup3), S43–S54. https://doi.org/10.1080/14992027.2017.1300695

Kuk, F. (2016). Selecting the right compression. *Audiology Online*. www.audiologyonline.com

Kuk, F., Slugocki, C., Korhonen, P., Seper, E., & Hau, O. (2019). Evaluation of the efficacy of a dual variable speed compressor over a single fixed speed compressor. *Journal of the American Academy of Audiology*, 30(7), 590–606. https://doi.org/10.3766/jaaa.17127

Latzel, M. (2013). *Compendium Adaptive Phonak Digital (APD), Audiological background*. Phonak AG, Switzerland. https://www.phonakpro.com/content/dam/phonakpro/gc_hq/en/resources/evidence/white_paper/documents/technical_paper/Adaptive_Phonak_Digital_Compendium.pdf

Moore, B. C. J. (2008). The choice of compression speed in hearing aids: Theoretical and practical considerations and the role of individual differences. *Trends in Amplification*, 12(2), 103–112. https://doi.org/10.1177/1084713808317819

Moore, B. C. J. & Glasberg, B. R. (1988). A comparison of four methods of implementing automatic gain control (AGC) in hearing aids. *British Journal of Audiology*, 22(2), 93–104.

Moore, B. C. J., Glasberg, B. R. & Stone, M. A. (1991). Optimization of a slow-acting automatic gain control system for use in hearing aids. *British Journal of Audiology*, 25(3), 171–182. https://doi.org/10.3109/03005369109079851

Moore, B. C. J., Peters, R. W., & Stone, M. A. (1999). Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. *The Journal of the Acoustical Society of America*, 105(1), 400–411. https://doi.org/10.1121/1.424571

Moore, B. C. J., & Sęk, A. (2016). Preferred compression speed for speech and music and its relationship to sensitivity to temporal fine structure. *Trends in Hearing*, 20, 233121651664048. https://doi.org/10.1177/2331216516640486

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1995). Effect of release time in compression hearing aids: Paired-comparison judgments of quality. *Journal of the Acoustical Society of America*, 98(6), 3182–3187. https://doi.org/10.1121/1.413807

Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1998). The effect of compression ratio and release time on the categorical rating of sound quality. *The Journal of the Acoustical Society of America*, 103(5), 2273–2281. https://doi.org/10.1121/1.422745

Ng, E. H. N., & Rönnberg, J. (2020). Hearing aid experience and background noise affect the robust relationship between working memory and speech recognition in noise. *International Journal of Audiology*, 59(3), 208–218. https://doi.org/10.1080/14992027.2019.1677951

Novick, M. L., Bentler, R. A., Dittberner, A., & Flamme, G. A. (2001). Effects of release time and directionality on unilateral and bilateral hearing aid fittings in complex sound fields. *Journal of the American Academy of Audiology*, 12(10), 534.

Ohlenforst, B., Souza, P., & MacDonald, E. (2016). Exploring the relationship between working memory, compressor speed, and background noise characteristics. *Ear and Hearing*, 37(2), 137–143. https://doi.org/10.1097/AUD.0000000000000240

Phonak. (2000). *Claro TM Digital perception processing*. Whitepaper published By Phonak AG, Switzerland. https://www.researchgate.net/publication/348729682_Claro_TM_Digital_Perception_Processing_TM

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *The Journal of the Acoustical Society of America*, 83(6), 2322–2327. https://doi.org/10.1121/1.396363

Preminger, J., & Van Tasell, D. (1995). Quantifying the relation between speech quality and speech intelligibility. *Journal of Speech and Hearing Research*, 38(3), 714. https://doi.org/10.1044/jshr.3803.714

Rählmann, S., Meis, M., Schulte, M., Kießling, J., Walger, M., & Meister, H. (2017). Assessment of hearing aid algorithms using a master hearing aid: The influence of hearing aid experience on the relationship between speech recognition and cognitive capacity. *International Journal of Audiology*, *57*(sup3), S105–S111. https://doi.org/10.1080/14992027.2017.1319079

Reinhart, P. N., & Souza, P. E. (2016). Intelligibility and clarity of reverberant speech: Effects of wide dynamic range compression release time and working memory. *Journal of Speech, Language, and Hearing Research*, *59*(6), 1543–1554. https://doi.org/10.1044/2016_JSLHR-H-15-0371

Rönnberg, J., Holmer, E., & Rudner, M. (2019). Cognitive hearing science and ease of language understanding. *International Journal of Audiology*, *58*(5), 247–261. https://doi.org/10.1080/14992027.2018.1551631

Salorio-Corbetto, M., Baer, T., Stone, M. A., & Moore, B. C. J. (2020). Effect of the number of amplitude-compression channels and compression speed on speech recognition by listeners with mild to moderate sensorineural hearing loss. *The Journal of the Acoustical Society of America*, *147*(3), 1344–1358. https://doi.org/10.1121/10.0000804

Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology*, *26*(2), 183–196. https://doi.org/10.3766/jaaa.26.2.7

Souza, P., Arehart, K., Kates, J., Croghan, N., & Gehani, N. (2013). Exploring the limits of frequency lowering. *Journal of Speech, Language and Hearing Research*, *56*(5), 1349–1363. https://doi.org/10.1044/1092-4388(2013/12-0151)

Souza, P., & Sirow, L. (2014). Relating working memory to compression parameters in clinically fit hearing aids. *American Journal of Audiology*, *23*(4), 394–401. https://doi.org/10.1044/2014_AJA-14-0006

Stone, M. A., & Moore, B. C. J. (2004). Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task. *Journal of the Acoustical Society of America*, *116*(4), 2311–2323. https://doi.org/10.1121/1.1784447

Stone, M. A., & Moore, B. C. J. (2008). Effects of spectro-temporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task. *The Journal of the Acoustical Society of America*, *123*(2), 1063–1076. https://doi.org/10.1121/1.2821969

Wong, L. L. N., Chen, Y., Wang, Q., & Kuehnel, V. (2018). Efficacy of a hearing aid noise reduction function. *Trends in Hearing*, *22*, 2331216518782839. https://doi.org/10.1177/2331216518782839

Wong, L. L. N., Ho, A. H. S., Chua, E. W. W., & Soli, S. D. (2007). Development of the Cantonese speech intelligibility index. *Journal of the Acoustical Society of America*, *121*(4), 2350–2361. https://doi.org/10.1121/1.2431338

Wong, L. L. N., Soli, S., Liu, S., Han, N., & Huang, M. (2007). Development of the mandarin hearing in noise test (MHINT). *Ear and Hearing*, *28*(2 Suppl), 70S–74S. https://doi.org/10.1097/AUD.0b013e31803154d0

Yu, J., Li, J., & Huang, X. (2012). The Beijing version of the Montreal cognitive assessment as a brief screening tool for mild cognitive impairment: A community-based study. *BMC Psychiatry*, *12*(1), 156. https://doi.org/10.1186/1471-244X-12-156

Zhong, N., Jiang, H., Wu, J., Chen, H., Lin, S., Zhao, Y., Du, J., Ma, X., Chen, C., Gao, C., & Hashimoto, K. (2013). Reliability and validity of the CogState battery Chinese language version in schizophrenia. *PLoS One*, *8*(9), e74258. https://doi.org/10.1371/journal.pone.0074258