

Evaluation of the Efficacy of a Dual Variable Speed Compressor over a Single Fixed Speed Compressor

DOI: 10.3766/jaaa.17127

Francis Kuk*
Chris Slugocki*
Petri Korhonen*
Eric Soper†
Ole Hau‡

Abstract

Background: It has been suggested that hearing-impaired listeners with a good working memory (WM) should be fitted with a compression system using short time constants (i.e., fast-acting compression [FAC]), whereas those with a poorer WM should be fitted with a longer time constant (i.e., slow-acting compression [SAC]). However, commercial hearing aids (HAs) seldom use a fixed speed of compression.

Purpose: The performance of a variable speed compression (VSC) system relative to a fixed speed compressor (FAC and SAC) on measures of speech intelligibility, recall, and subjective report of listening effort and tolerable time was evaluated. The potential interaction with the listeners' WM capacity (WMC) was also examined.

Research Design: A double-blinded, repeated measures design.

Study Sample: Seventeen HA wearers (16 with greater than one year HA experience) with a bilaterally symmetrical, mild to moderately severe sensorineural hearing loss participated in the study.

Data Collection and Analysis: Participants wore the study HAs at three compression speeds (FAC, SAC, and VSC). Each listener was evaluated on the Office of Research in Clinical Amplification-nonsense syllable test (NST) at 50 dB SPL (signal-to-noise ratio [SNR] = +15 dB), 65 dB SPL (SNR = +5 dB), 80 dB SPL (SNR = 0 dB), and a split (80 dB SPL–50 dB SPL) condition. Listeners were also evaluated on a Repeat Recall Test (RRT), where they had to repeat six short sentences (both high- and low-context sentences) after each was presented. Listeners recalled target words in all six sentences after they were presented. They also rated their listening effort and the amount of time they would tolerate listening under the specific condition. RRT sentences were presented at 75 dB SPL in quiet, as well as SNR = 0, 5, 10, and 15 dB. A Reading Span Test (RST) was also administered to assess listeners' WMC. Analysis of variance using RST scores as a covariate was used to examine differences in listener performance among compressor speeds.

Results: Listener performance on the NST was similar among all three compression speeds at 50, 65, and 80 dB SPL. Performance with FAC was significantly better than SAC for the split condition; however, performance did not differ between FAC and VSC or between SAC and VSC. Performance on the NST was not affected by listeners' RST scores. On the RRT, there was no effect of compressor speed on measures of repeat, recall, listening effort, and tolerable time. However, VSC resulted in significantly lower (better) speech reception threshold at the 85% correct recognition criterion (SRT_{85}) than FAC and SAC. Listener RST scores significantly affected recall performance on the RRT but did not affect SRT_{85} , repeat, listening effort, or tolerable time.

Conclusion: These results suggest that the VSC, FAC, and SAC yield similar performance in most but not all test conditions. FAC outperforms SAC, where the stimulus levels change abruptly (i.e., split condition). The VSC yields a lower SRT_{85} than a fixed compression speed at a moderately high level with a favorable SNR. There is no interaction between compression speed and the participants' WMC.

*Widex Office of Research in Clinical Amplification (ORCA-USA), Lisle, IL; †University of Chicago Medicine, Audiology Program, Chicago, IL; ‡Widex A/S, Lyngø, Denmark

Corresponding author: Francis Kuk, Widex Office of Research in Clinical Amplification (ORCA-USA), Lisle, IL 60532; Email: Fkuk@widex.com

Key Words: compression speeds, cognition, fast-acting compression, slow-acting compression, variable speed compression

Abbreviations: ANOVA = analysis of variance; CR = compression ratio; CT = compression threshold; DR = dynamic range; FAC = fast-acting compression; FC = fast compression branch within VSC; HA = hearing aids; HC = high context; HL = hearing level; LC = low context; MoCA = Montreal Cognitive Assessment; NST = nonsense syllable test; ORCA = Office of Research in Clinical Amplification; RIC = receiver-in-the-canal; RRT = Repeat Recall Test; RST = Reading Span Test; SAC = slow-acting compression; SC = slow compression branch within VSC; SD = standard deviation; SNR = signal-to-noise ratio; SPL = sound pressure level; SRT = speech reception threshold; SRT₈₅ = speech reception threshold at 85% correct criterion; TFS = temporal fine structures; VSC = variable speed compression; WDRC = wide dynamic range compression; WM = working memory; WMC = working memory capacity

INTRODUCTION

The speed of gain change in a compression circuit (i.e., compression time constants) plays a significant role in the ultimate performance and acceptance of the system (Dillon 1996; Moore, 2008). In recent years, it has been suggested that people with a good working memory (WM) are better served with a fast-acting compression (FAC) circuit, whereas people with a poor WM may be better served with a slow-acting compression (SAC) circuit (Jenstad and Souza, 2005; Gatehouse et al, 2006; Jenstad and Souza, 2007; Lunner and Sundewall-Thorén, 2007; Rudner et al, 2011; Souza and Sirow, 2014; Ohlenforst et al, 2015). On the other hand, many commercial hearing aids (HA) use variable rates of gain change in response to changes in the instantaneous and/or overall input level. It is logical to question if an implementation that uses a combination of SAC and FAC would yield performances that may be similar to or better than the use of a fixed time constant (i.e., FAC or SAC alone). Furthermore, it would be meaningful to know if such an implementation would yield different degrees of benefit for people with different WM capacities (or WMC). In this study, we compared the performance of a dual variable speed compressor (VSC) to a single fixed speed compressor (FAC and SAC) implemented on a commercial HA on measures of speech intelligibility, recall, and measures of effort. We further examined if the WMC of the listeners, as evaluated on the Reading Span Test (RST, Van den Noort et al, 2008) interacted with the observations.

The timing characteristics of a compressor are defined by its attack and release times. By definition, attack time that is <20 ms is classified as fast, and anything longer than that is defined as slow. Release time that is 50 ms or shorter is classified as FAC, whereas a release time that is >500 ms may be classified as long or SAC (Dillon, 1996; Moore, 2008). Some release times can be as long as 20 sec (Kuk, 1998). It is generally acknowledged that attack time should be fast for the HAs to be responsive to sudden changes in the input sound pressure levels (SPLs) to protect

the listeners from loud sounds. The appropriate release time is more debatable. In general, a short release time or FAC is used to ensure consistent audibility of soft sounds in situations where the sound intensity fluctuates (Moore et al, 2004). The rapid changes in gain also reduce the dynamic variations in both the short-term and long-term amplitude envelope of the input signals. Despite the improvement in audibility, sound quality rating of FAC processed speech is often poorer than SAC processed speech (Neuman et al, 1995; 1998; Hansen 2002; Hau and Andersen, 2012). There may be two reasons for such an observation. First, the ambient noise between speech pauses become especially noticeable, giving rise to the “pumping” sensation and noisiness perception (Moore, 2008) over a SAC system. Second, the natural temporal envelope variations are reduced, leading to temporal smearing and more “muffled” sensation. This is especially so in multichannel compression systems or where a high compression ratio (CR) (>3:1) is used (Plomp, 1988; Boike and Souza, 2000; Souza et al, 2005).

An SAC preserves the short-term envelope of the waveform better through slower gain adjustment (Moore, 2008). Thus, less smearing of the temporal envelope (and thus a wider output dynamic range [DR]) is expected with a SAC than a FAC. Studies on the comparison between FAC and SAC usually reported that SAC processed speech was more natural and less noisy (Neuman et al, 1995; 1998; Hansen, 2002; Hau and Andersen, 2012). However, Gilbert et al (2008) showed that not all hearing-impaired listeners were able to discriminate between long and short release times once the overall level and noise floor differences were compensated. On the other hand, because of the slower gain change, SAC may be limited in ensuring the audibility of soft sounds, especially after a moderately loud sound. Many researchers have summarized the advantages and limitations of FAC and SAC compressors (Kuk, 1996; Souza, 2002; Gatehouse et al, 2006; Moore, 2008).

The limitation of each compression speed suggests that any commercial compression HAs, which until recently were mostly single compressor design using either FAC or SAC, must seek ways to minimize the

limitations of the fixed compression speed. For example, multichannel wide DR compression (WDRC) HAs using FAC typically are restricted to a small CR (<3:1) to minimize the smearing that could result from the exclusive use of a fast release (Moore, 2008). In addition, the use of a fewer number of compression channels, or linking nearby channels, could also minimize the amount of temporal distortion when using FAC. Some devices use fast response times in the low frequencies and slow response times in the high frequencies to minimize the spread of masking effect of the low frequencies on the high frequencies, and to avoid distortion of the temporal envelope in the high frequencies (Lunner et al, 1997).

Moore and Glasberg (1988) suggested a system referred to as a “dual front-end compressor” that used SAC at most times. The FAC was activated temporarily when a sudden increase occurred in the input levels. Such an implementation has reportedly yielded good performance for HA (Stone et al, 1999) and cochlear implant wearers (Boyle et al, 2009). Contrarian approaches such as using FAC as the primary compressor and SAC when needed were also evaluated (Moore et al, 2004). Thus, even when commercial compression HAs reported using fixed compression speeds, many were designed to minimize the limitations of a fixed compression speed. Effectively, few modern commercial HAs use a single compressor or a fixed compression speed.

These observations have implications on the recent suggestion that compression speeds need to be optimized for people with different WMCs. By definition, WM is defined as the “*retention of information in conscious awareness...for its manipulation and use in guiding behavior*” (Postle, 2006). In other words, it describes an individual’s capacity to retain information in the short-term storage and to manipulate that remembered information for processing or problem solving. Thus, individuals with a good WMC have a larger capacity to store information for processing, whereas those with a poorer WMC have a smaller capacity for such functions. In relation to the use of amplification, Gatehouse et al (2003) reported that hearing-impaired listeners with a poor WMC performed better on a speech recognition test when stimuli were processed with a 2-channel compressor using a fixed long time constant than a fixed short time constant. The individuals with a good WMC did better when the test materials were processed with a shorter time constant than a longer time constant. Similar findings were replicated by other researchers from the same group (Gatehouse et al, 2006) and from other teams (e.g., Lunner and Sundewall-Thorén, 2007; Souza and Sirow, 2014; Ohlenforst et al, 2015). This has led to the suggestion that SAC may be used for people with a poor WMC, whereas FAC may be used for people with a good WMC. In other words, HA signal processing would need to be tailored to the

individual’s specific cognitive capacity. A logical question that ensues is the choice of optimal compression when modern commercial HAs are not restricted to a single compressor using a fixed speed of compression.

A practical solution is to design compression systems that are appropriate for people of all WMCs. A precursor is to understand how speed of compression affects the acoustic signals and how different parts of the acoustic signals are important to people with different cognitive abilities (or WM). As described by many researchers (e.g., Souza, 2002; Moore, 2008) and summarized previously, both FAC and SAC, at the same overall gain and constant input level, achieve similar long-term output. However, the short-term output and temporal envelope resulting from these two compression speeds differ. FAC achieves better audibility of soft sounds, especially after louder sounds or in widely fluctuating soundscapes, than SAC. This provides an opportunity for listeners to glimpse or hear in “dips” in backgrounds of fluctuating noise (Moore, 2008). This could be beneficial for the individuals who have the ability to glimpse. This ability to hear in “dips” in fluctuating noise background is dependent on the listener’s ability to use the temporal fine structures (TFS) of the sounds. People with a hearing loss had a reduced ability to use TFS (Lorenzi et al, 2006). Such reduced ability may be mediated from a poorer WM or from a direct result of hearing loss or both. The unfortunate side effect of FAC is that the temporal envelope of the input signal is distorted. A person with a poor WM, despite the availability of the audible cues from FAC, is affected negatively by the temporal distortion because of their reliance on the temporal envelope. However, people with good WM may be less affected by temporal distortion than those with a poorer WM (Arehart et al, 2013). Thus, people with good WM perform well with FAC, whereas people with poor WM may not because of their difference in reliance on the temporal envelope.

SAC has less of an effect on the temporal envelope of the input than FAC. Temporal envelope is an important cue in auditory scene analysis (Bregman, 1990). It also carries information on manner and voicing, as well as prosody and suprasegmentals of speech (Rosen et al, 1990). The ability to preserve the temporal envelope resulted in better sound quality judgment (Neuman et al, 1995; 1998; Hansen 2002; Souza, 2002) and aided in speech understanding (Van Tasell and Trine, 1996; Fogerty and Entwistle, 2015). The dependence on temporal envelope cue for speech understanding depends on the degree of hearing loss (Van Tasell and Trine, 1996), the availability of contextual cues from the input and other sources (Cox and Xu, 2010), and the cognitive capacity of the individual (Souza, 2000). Thus, it is a cue that is used by all, but its dependence varies among individuals. People who have a more severe hearing loss

and those with a poorer cognitive capacity are more dependent on this cue. An unfortunate side effect of preserving this cue is that audibility of soft sounds in soundscapes of fluctuating intensity levels (>30–40 dB) may not be ensured. However, the consequence is likely dependent on whether the DR of the HA output exceeds the residual DR of the listeners. If it does, this may limit audibility cues and affect speech understanding. In situations where the DR of the HA output is within the residual DR of the listener, audibility is ensured and one should not expect any differences in speech recognition regardless of SAC or FAC processing (Moore et al, 2004; Jenstad and Souza, 2005; Shi and Doherty, 2008). Rather, because of the preserved temporal envelope with SAC, one may expect better sound quality ratings than if the processing were FAC.

The aforementioned speculation suggests that although audibility is important; audibility achieved at the expense of temporal envelope distortion could have different consequences to people with good and poor WMCs. Rather than restricting the use of a fixed speed of compression (FAC versus SAC) for listeners with a specific WMC, a parsimonious solution is to preserve the temporal envelope as much as possible and provide the needed audibility cues when the situation demands (such as changes in input levels, modulations etc). The preservation of the temporal envelope will be beneficial to listeners with both good and poor WMCs. The extra audibility provided at critical input transitions may or may not benefit the listeners depending on their TFS resolution. However, having such cues, even if they are not used, should not be harmful. This view suggests that individuals with good and poor WMC can be best served by the same form of compression as long as compression achieves the same audibility with minimal temporal envelope distortion. Achieving such a goal could minimize the need for individualizing compression speeds based on cognitive abilities, which is not routinely measured during audiological assessment.

The suggestion by Moore and Glasberg (1988) of using SAC as the primary compression speed and changing the release time to fast when the situation demands fits into this consideration. The use of SAC above a conversational level as the primary speed is reasonable in that the DR of sounds occurring in any listening environments and at any one vocal effort is typically between 30 and 40 dB (Pearsons et al, 1977). This is well within the residual DR of most hearing-impaired persons with hearing losses <70 dB HL (Pascoe, 1988). Assuming that most of the input is at or above a conversational level, the processed output is likely audible even though slower gain change (or no gain change as in linear amplification) is used to preserve the short-term intensity relationship among sounds. When a large change in input level occurs, the release time may be changed adaptively to account for the level

changes. Such adaptive mechanism was used in earlier versions of Widex HAs (such as the DREAM), where the single compressor changes its release time according to the duration and intensity of the incoming signals. Unfortunately, in a single compressor that uses an adaptive speed, gain transition would invariably lag behind changes in input levels. This suggests that performance at the lower input levels (softer sounds) may still have some issues with audibility, whereas performance at the higher input level may still have issues with temporal smearing. In addition, it is difficult to adapt to a faster speed when a high CR is used.

One solution is to use two compressors together to achieve the stated benefits of FAC alone or SAC alone, while minimizing their limitations. In such a dual compressor system, one compressor uses a long time constant (slow compressor [SC]) and the other uses a short time constant (fast compressor [FC]). Both of these compressors run in parallel simultaneously in each of the multiple compression channels. The contribution from each compressor could vary based on specified rules. The output from each compressor is added to form the overall output of the HA. This could provide the flexibility and responsiveness over a single adaptive speed compression system. Pittman et al (2014) simulated the action of an adaptive 4-channel compression system by using two level estimators to track the changes in input levels and adjust the time constants accordingly. When no large changes in input level occurred, the longer release time was used. When a large increase in the input level occurred, a short time constant was used for instantaneous gain reduction. When a large decrease in the input level occurred, the slow time constant was again applied. The authors compared this adaptive compression system with FAC alone (attack—10 ms, release—40 ms) and SAC alone (attack—20 ms, release—800 ms) in children and adults with normal hearing and a hearing loss. The listeners identified target words and environmental sounds processed under the three compression speed conditions while engaged in a visual pattern task. For normal hearing listeners, SAC was significantly better than FAC and adaptive compression. For the hearing-impaired listeners, performance with the adaptive compression was better than SAC and FAC. The authors suggested the use of adaptive release time to benefit both children and adults with hearing losses. As cautioned by the authors, the study used simulated recordings and such results may not be consistent with current commercial multichannel compression devices.

Widex recently introduced a dual compressor design using variable speeds (VSC) in our Widex EVOKE™ HA (Widex A/S, Lyngby, Denmark). This compression system uses one compressor with a long time constant (SC) and the other with a short time constant FC. Both compressors operate at the same time and the final

output depends on the contributions from both compressors. At or above a steady conversational level, gain supplied by the SC forms the dominant contribution to the overall output. Contribution from the FC increases for softer sounds, highly modulated signals, and when large level changes in the input signals occur. This is intended to provide the needed audibility for softer speech sounds. This way, the SC provides gain while preserving the temporal envelope and the FC supplies the extra audibility cues when the need arises in permissible situations (i.e., highly modulated inputs).

As suggested previously, the advantage of a dual compressor over a single compressor with adaptive speed is that the dual compressor may be more responsive to the required gain changes than a single adaptive speed compressor. This was demonstrated in a previous study (Kuk and Hau, 2017), where the dual compressor used in the UNIQUE HA (an earlier version of the Widex EVOKE™ dual compressor) was compared with the single adaptive compressor used in the DREAM HA in a matched gain comparison. Nonsense syllables at a 50-dB SPL input level were presented following a carrier phrase presented at 80 dB SPL. The results showed a significant 5% point improvement in absolute consonant scores with the dual compressor over the single adaptive speed compressor. Additional details on the dual compressor can be found in Kuk and Hau (2017).

It is hypothesized that this dual VSC system in the Widex EVOKE™ could achieve the audibility benefits of a FAC system by providing the audibility cues with minimal alteration of the temporal envelope (or DR) of the input signals. It is conceivable that it will yield performance that is at least equal to, if not better than the use of a single FAC or SAC compressor in many situations. Thus, this study sought to compare the performance of the dual VSC to a single fixed speed compressor (using long and short time constants, SAC and FAC, respectively) implemented in the same HA. Specifically, we measured speech recognition ability using a nonsense syllable test at three steady and one fluctuating input levels. In addition, a sentence test that measured repeat and recall abilities in quiet and in noise at different signal-to-noise ratios (SNRs) was also conducted. Subjective report of listening effort and tolerable time were also measured during the sentence test to examine any potential perceptual differences resulting from the difference in signal processing. A secondary question was to determine whether the listeners' WMC, as reflected on a RST, would interact with signal processing to result in differences in performance.

METHODS

Participants

A priori power analysis using G*Power (v3.1.9.2; Düsseldorf, Germany) based on the data from Lunner

and Sundewall-Thorén (2007), Souza and Sirow (2014) indicated that a sample of 16 subjects could detect a compressor difference of 1.5 dB (SD [standard deviation] = 2 dB) at 80% power for $\alpha = 0.05$. Thus, seventeen adults (12 females and five males) with bilaterally symmetrical (± 5 dB from 250 to 8000 Hz) sensorineural hearing loss were recruited. They ranged from 33 years to 83 years in age (mean = 65.5 years, SD = 14.6). All were native English speakers. Their four-frequency pure-tone averages were 50.2 dB HL (SD = 12.4) for the right ear and 49.0 dB HL (SD = 11.2) for the left ear (see Figure 1). Participants had worn bilateral HAs for an average of 15.2 years, although this ranged from no experience (one participant) to 40 years of experience. Five participants wore behind-the-ear, seven receiver-in-the-canal (RIC), two in-the-canal, and two completely-in-the-canal HAs as their own HAs. All were digital WDRC HAs, although the exact time constants used in each was difficult to ascertain. Before the study, each participant was assessed with the Montreal Cognitive Assessment (MoCA) test—a cognitive screening tool with high sensitivity and specificity for detecting mild cognitive impairment (Nasreddine et al, 2005). All but three participants scored below 26 on the MoCA. Of those three participants, two scored 25 and one scored 22. The data on these three participants were also included in the analysis. The participants were informed of the purpose of the study along with its benefits and risks before giving their written consent. All participants were compensated financially.

Hearing Aids

Each participant was fitted bilaterally with the experimental version of the Widex EVOKE™ Fusion

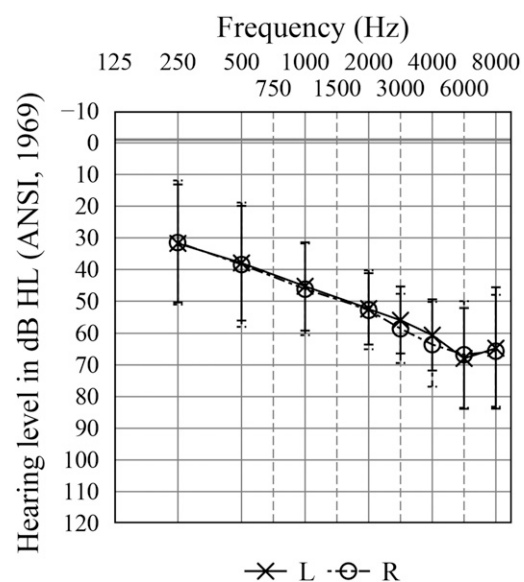


Figure 1. Average audiogram of participants at right (X) and left (O) ears. Error bars indicate one standard deviation.

440 RIC HAs. The HAs were programmed using the manufacturer's proprietary development tools that allowed the adjustment of HA parameters beyond what is available commercially. The study aid is a 15-channel WDRC HA with a compression threshold (CT) as low as 0 dB HL. The sampling rate of the analog-to-digital converter is 33.1 kHz and the input resolution is 18 bits. The input DR is linear between 5 and 113 dB SPL. This instrument includes a multichannel fully adaptive directional microphone, speech intelligibility index based noise reduction, wind noise management algorithm, active feedback cancellation, and a sound classification and learning system. Other than the feedback cancellation system, all the other features were deactivated during the study even though they may not affect the outcome of the present study given the stimulus and test conditions. An omnidirectional microphone was used to minimize any potential changes in microphone polar pattern from unintentional head movement.

Hearing Aid Conditions

The present study compared the efficacy of the VSC with a single fixed-speed FAC and a single fixed-speed SAC, all implemented on the same study HA and programmed using developmental software. The VSC uses two separate signal-processing blocks simultaneously to determine the overall gain and output. It combines a SC that adjusts the gain according to the overall input signal level and the hearing loss of the listener, and a FC that adjusts the gain based on fast changes and the modulation characteristics of the input signal. A greater modulation indicates greater amplitude variations of the signal, suggesting speech (instead of noise) or a speech-in-noise situation with a favorable SNR. Figure 2 shows the functional block diagram of the VSC.

The SC block within the VSC determines the gain (G_{SLOW}) based on the long-term average level of the input signal in each single 1/3-octave frequency channel. The goal of this processing block is to adjust the gain and the overall output level to compensate for the hearing loss so that the output signal is within the listeners'

DR. The attack time of the SC block is approximately 1.5 sec and the release time is approximately 17 sec. Because of the slow rate of gain adjustment, this processing block should preserve the dynamic variations of speech occurring at the syllable level. The exact CR varies with the individual's hearing loss but is, in general, between 2:1 and 3:1, with a higher CR in the higher frequencies. A CT between 27 and 38 dB SPL was used across frequencies for the average hearing loss.

The FC block within the VSC determines the gain (G_{FAST}) based on the differences between the short-term and the long-term average levels of the input signal. The short-term average level is determined in each of all 15 frequency channels by averaging the signals from three adjacent frequency channels (1-octave wide). Channel linking is used to minimize temporal and spectral smearing. When the short-term average level of the input signal is higher than the long-term average level of the input signal, G_{FAST} gain is reduced in a graded manner. Conversely, when the short-term average level of the input is lower than the long-term average level of the input signal, G_{FAST} increases its gain proportionally. In addition, the amount of gain from the FC varies with the SNR or modulation of the input signals. A greater modulation receives a higher G_{FAST} . This is carried out to provide extra audibility for soft sounds in quiet, while preserving the amplitude envelope of sounds in noisy situations. The calculated attack time of the FC block is approximately 12 ms and the release time is 130 ms. The CR varies between 1.1 and 1.4 and is similar across frequencies. The CT ranges between 39 and 45 dB SPL across frequencies. The overall instantaneous gain or G_{TOTAL} applied to each 1/3-octave channel is the arithmetic sum of the gain determined by the SC block (G_{SLOW}) and the FC block (G_{FAST}), that is, $G_{\text{TOTAL}} = G_{\text{SLOW}} + G_{\text{FAST}}$.

The FAC uses only one compressor and is set to an attack time of 5 ms and a release time of 50 ms to approximate the values used in previous studies. The SAC also uses only one compressor and is set with an attack time of 12 ms and a release time of 3.2 sec. In addition, the CR

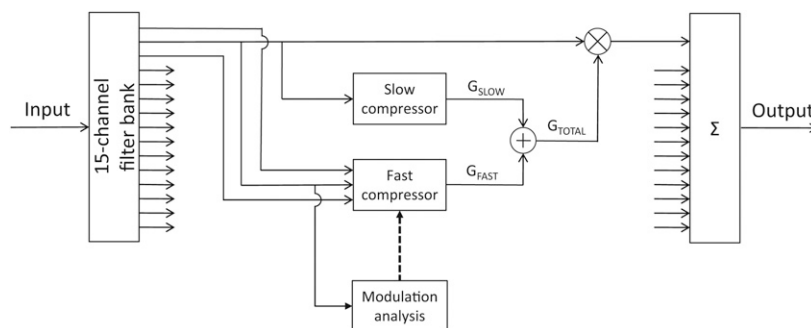


Figure 2. Functional block diagram of the VSC showing the signal flow for the second 1/3-octave channel of the input signal.

and CT of the SAC and the FAC also match the values used by the SC of the VSC. Hence, any difference between the VSC and the FAC/SAC in listener performance should only reflect the difference in time constants. A summary of the compression settings among these three programs (VSC, FAC, and SAC) is provided in Table 1.

Hearing Aid Fitting

HAs were coupled to the appropriate receiver (m- or p-receivers) and fitted using fully occluding “double-dome” instant-fit ear-tips to minimize the influence of direct sounds mixing with the processed sounds. The participants’ in situ thresholds (or sensograms) for both ears were measured with the HAs on the participants’ ears. The initial feedback path and gain limit was estimated using the feedback test. The target gain on the HAs was set based on the National Acoustics Laboratory-Nonlinear fitting target version 2 rationale. All fittings were verified via the fitting software using a simulated speech mapping measure, which has been demonstrated to provide accurate approximation (within 2 dB in 80% of cases) of real-ear output (Oeding and Valente, 2013). Figure 3 shows the average 2cc coupler output for the three compressor conditions measured for soft (50 dB SPL), medium (65 dB SPL), and loud (80 dB SPL) sounds using the Audioscan Verifit2 (version 4.6.1). The averaged hearing threshold of the participants was used as a reference. Two types of stimuli available on the Verifit2 were used—a pink noise and the speech passage (carrot passage). The output for the VSC, SAC, and FAC processing was similar for the steady, pink noise stimulus (left) at each input level. However, the output for the modulated speech stimulus was different among the three compressor speeds with FAC yielding a higher output than VSC and SAC. The SAC had the lowest output. The difference was especially marked at the 50 dB SPL and 65 dB SPL input levels between 1 and 4 kHz. A total of 3–4 dB difference was noted between FAC and SAC. Thus, compression speed could lead to a difference in overall output depending on the modulation of the stimulus. Because the intention of the study was to examine the natural effect of com-

pression speeds, such output difference between SAC and FAC/VSC was not compensated.

To verify that the three compressor settings resulted in a difference in output DRs, the output of the study HAs for the speech stimuli used in this study (nonsense syllable test [NST] and Repeat Recall Test [RRT] plus appropriate noise) was measured at all three compressor settings. For this analysis, the HAs were programmed using the average hearing thresholds of the participants (noted as dotted line in Figure 4). HAs were inserted into the Knowles Electronic Manikin for Acoustic Research (GRAS 45BB, IEC 60318-4:2010; Holte, Denmark) ear canal using fully occluding earmolds. The manikin was positioned at the same position as the participants would be seated during data collection. The Knowles Electronic Manikin for Acoustic Research output was recorded at 44.1 kHz sampling frequency and with a 32-bit depth. DR was determined by comparing the RMS of the output recordings in a 10-ms sliding window (rectangular) to a known reference for dB SPL conversion. The analysis was conducted at 1-octave wide bands centered at 0.5, 1, 2, and 4 kHz. DR was reported as the dB range between the 5th and 95th percentiles.

Equipment and Setup

Testing was conducted in a double-wall sound-attenuating booth (Industrial Acoustics Company, Inc., Bronx, NY), with internal dimensions of 3 × 3 × 2 m (*W* × *L* × *H*). During sound field measurements, stimuli were presented using a single KRK-ST6 2-way passive loudspeaker placed 1 m in front of the participants. The loudspeaker had a flat frequency response with a deviation of $\leq \pm 2$ dB between 62 Hz and 20 kHz. The loudspeaker was placed on a stand 1 m above the floor. Target stimuli were generated using an Echo Audiofire12 12-channel digital audio interface (Echo, Santa Barbara, CA) and amplified using a Rotel RMB-1048 power amplifier (Rotel, North Reading, MA). All test software was implemented using VB.net programming language for Windows operating system (Microsoft, Redmond, WA). Monthly calibration of all test equipment was conducted in addition to daily calibration checks during the course of the study.

Table 1. Description of the (VSC; slow and fast branches), FAC, and SAC, including calculated attack times (AT) and release times (RT), CR, and CT (CT; dB SPL at the eardrum)

	AT (ms)	RT (ms)	CR _{soft}				CR _{loud}				CT (dB SPL at ED)			
			0.5 kHz	1 kHz	2 kHz	4 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz	0.5 kHz	1 kHz	2 kHz	4 kHz
VSC (slow)	1,500	17,000	1.9	2.3	2.7	2.5	1.6	1.9	2.5	3.8	38	27	29	28
VSC (fast)	12	130	1.1–1.3	1.1–1.4	1.1–1.4	1.1–1.4	1.1–1.3	1.1–1.4	1.1–1.4	1.1–1.4	45	39	44	39
FAC	5	50	1.9	2.3	2.7	2.5	1.6	1.9	2.5	3.8	38	27	29	28
SAC	12	3,200	1.9	2.3	2.7	2.5	1.6	1.9	2.5	3.8	38	27	29	28

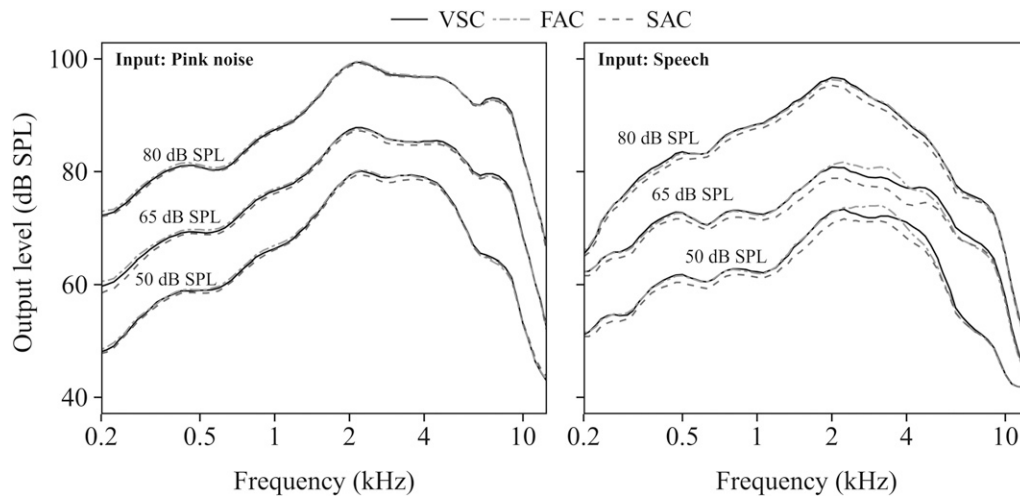


Figure 3. Coupler (2cc) output for soft (50 dB SPL), medium (65 dB SPL), and loud (80 dB SPL) input as measured for VSC, FAC, and SAC for a pink noise stimulus (left) and a speech stimulus (right).

Outcome Measures

A review of the literature suggests that many of the studies that reported an interaction effect between cognition and compression speeds used sentence-level materials presented in modulated noise (Gatehouse et al, 2003; Lunner and Sundewall-Thorén, 2007; Cox and Xu, 2010). In addition, speech materials with a low context (LC) may aid in differentiating between listeners with good and poor WM (Cox and Xu, 2010). Thus, the present study included nonsense syllables and high- and low-context sentences to optimize the evaluation of the effect of compression speeds and its interaction with the participants' WM.

Office of Research in Clinical Amplification-NST (ORCA-NST)

The Widex ORCA-NST was used as a measure of speech intelligibility in noise for speech materials lacking any contextual cues. The ORCA-NST is an open-set consonant-vowel-consonant-vowel-consonant test, containing 25 English consonants each appearing at least once in the initial, medial, and final word positions unless prohibited by phonotactic constraints. The present study used a 32-item version spoken by a female talker. A detailed description of the test can be found in Kuk et al (2010).

ORCA-NST was administered at three fixed input levels. In a first condition, speech was presented at 80 dB SPL and babble noise at 80 dB SPL (SNR = 0 dB). In a second condition, speech was presented at 65 dB SPL and babble noise at 60 dB SPL (SNR = 5 dB). In a third condition, speech was presented at 50 dB SPL and babble noise at 35 dB SPL (SNR = 15 dB). These stimulus conditions were chosen to

approximate the typical SNRs for speech at these presentation levels in real life (Smeds et al, 2015). A fourth stimulus condition, called the “split” condition, was included to provide widely fluctuating intensities to evaluate the relative efficacy of the VSC in matching the performance of the FAC/SAC in fluctuating environments. In this split condition, a carrier phrase (“please say the word”) was presented at 80 dB SPL in babble noise and followed by the target nonsense syllable presented at 50 dB SPL in quiet. This stimulus condition approximated real-life listening situations where a soft sound followed a loud sound (i.e., hearing soft speech after someone has been coughing for some time, or hearing soft speech in quiet interludes during a musical concert).

The babble noise used during the ORCA-NST testing was generated by mixing two streams of International Speech Test Signal where one of the signals was delayed by 10 sec (Holube and EHIMA-ISMADHA Working Group, 2006). Two-talker babble was chosen so that the noise stimulus would have a high degree of modulation. This was carried out to allow for potential “glimpsing” of speech in spectro-temporal regions where the speech signal was less affected by the background babble (Cooke, 2006). Various authors (e.g., Lunner and Sundewall-Thorén, 2007; Ohlenforst et al, 2015) also showed that the use of modulated noise allowed better differentiation between listeners with good and poor WM. Noise was presented from the same loudspeaker as the NST stimuli.

ORCA-RRT

We created a RRT (ORCA-RRT) to assess listeners' speech intelligibility at fixed SNRs while simultaneously measuring their auditory WM. A detailed description of the ORCA-RRT, as well as the rationale

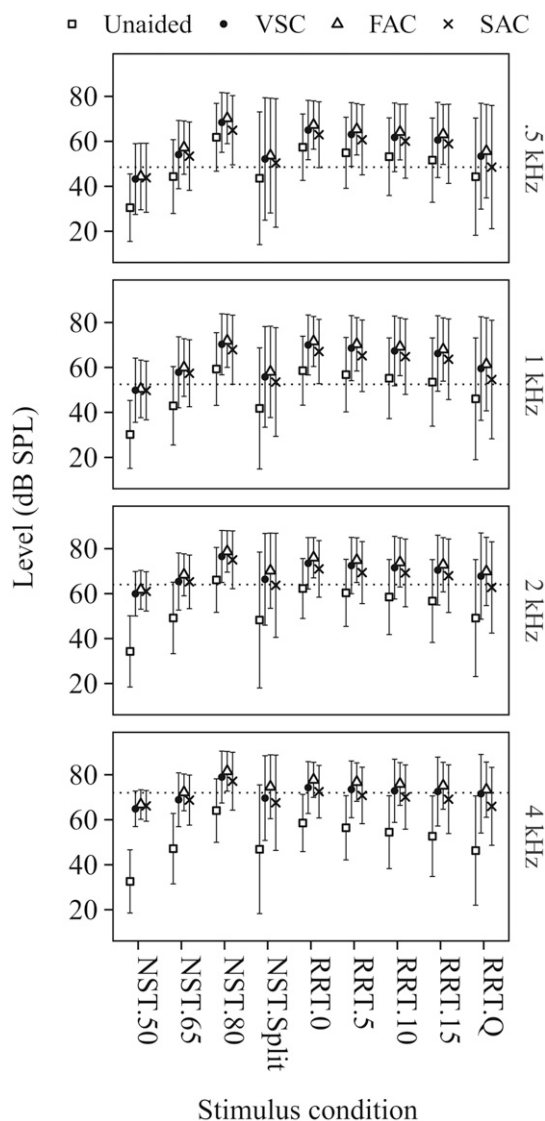


Figure 4. The DRs of the study stimuli measured in the unaided mode (\square), with the VSC (\bullet), FAC (\triangle), and SAC (\times) processing for the various stimulus test conditions. Symbols represent mean levels and error bars represent the range of levels between the 5th and 95th percentiles (i.e., DR). Analyses were carried out in 1-octave bands centered at 0.5, 1, 2, and 4 kHz (from top to bottom panels).

behind the development of the test, can be found in Slugocki et al (in submission). In short, the test involved two stages: a repeat stage and a recall stage. In the repeat stage, participants were asked to repeat prerecorded sentences that were spoken by a male speaker. Sentences were presented one at a time in quiet at 75 dB SPL and in noise at SNRs of 0, +5, +10, or +15 dB. The noise condition used a constant two-talker babble (male talkers) with the same long-term spectrum as the speech material presented from the same direction. Six sentences were presented per test block with an inter-sentence interval of 7 sec. The sentences were six to eight words in length and each contained either three or four target words that

related to a specific theme (such as food, sports, etc). Each block of six sentences contained 20 target words. The participants had no prior knowledge of the theme of the sentences or which words within the sentences were target words. They were instructed before the repeat stage that they had to recall as many words from the sentences after all six sentences were repeated. The test administrator recorded which target words were repeated correctly after each sentence. After all six sentences were repeated, participants were given one minute to freely recall (aloud) as many words from those six sentences in any order as they could remember. The test administrator recorded all target words that were recalled correctly. Only target words that were repeated correctly during the repeat stage were credited during the recall stage.

After each repeat and recall stage, participants were asked to rate how effortful they found the listening situation to be on a ten-point rating scale. Three anchors were provided to guide judgment: “1” = not effortful, “5” = moderately effortful, and “10” = very effortful. An off-scale response (“this task was too difficult”) was also accepted if the participants gave up on the task because it was too demanding. Such responses were coded as “10” for data reduction and analysis purposes. This occurred in 17 instances of the possible 510 trials (or 3.3%, at SNR = 0 dB). After each trial, participants were asked to provide an estimate of how long (in minutes) they would be willing to spend concentrating on the talker under the specific test condition.

Two versions of the RRT were used in this experiment. The high context (HC) version comprised sentences that were meaningful and related to each another by a shared topic (i.e., “sports”). The LC version comprised sentences, wherein the target words from the HC set were scrambled between sentences while keeping the same syntactic structure. Thus, the LC passages had the same syntactic structure and used the same words as the complementary HC set; but the sentences were not meaningful nor did the sentences relate to each other. The participants were always tested with the LC passage first before they were tested with the HC version. A different set of six sentences was used at each SNR condition and sets were not reused between HA compressor conditions. The different SNRs were tested in a random order.

Cognitive Measures—RST

The RST by Van den Noort et al (2008) was used to assess listeners' WMCs. The RST is a measure of verbal WM that involves reading aloud and retaining content from a series of visually presented sentences. One hundred sentences were divided into five sets of 20 sentences with each set containing two, three, four, five, and six sentences. A computer program presented the

sentences automatically in random order at a rate of one sentence per seven seconds. Participants were tasked with reading aloud each sentence within a set for its meaning as it was presented. After each set, they were asked to recall the last word of each sentence. Performance in the RST was quantified as the total number of words that were recalled correctly across all 100 sentences.

It should be noted that this version of RST was used instead of the original Daneman and Carpenter (1980) version that was used in several other WM studies. This was because the original RST was criticized for lacking control of the content materials when translated into different languages (Van den Noort et al, 2008). This revised version had sentence lengths ranging from 12 to 17 words with 20–22 syllables and 53–77 letters per sentence. The number of syllables and the frequency of occurrence of the final words, as well as the concreteness of the final words were controlled in each sentence series. The revised version was tested for plausibility and abstract concreteness in four languages, including English, Dutch, German, and Norwegian. Its authors reported high correlation between the RST (in Dutch) and performance on the relative clause task and the sentence reading task which supported the use of this test to measure WM. High internal consistency and test-retest reliability was also reported. Although scores on this RST may not be directly comparable with that of the Daneman and Carpenter's (1980), it was chosen because of its sound construction and standardization among languages.

Procedure

Listeners' performance at the three compressor settings (VSC, FAC, and SAC) was measured in a double-blinded, counterbalanced manner. Each compressor setting was tested on a different session. The participants spent three to four 2-h sessions to complete the required HA fitting and data collection.

Testing began with the administration of the ORCA-NST at the four stimulus conditions (50, 65, 80 dB SPL, and Split). The order of presentation of the ORCA-NST stimulus was randomized for each participant. The ORCA-RRT was then administered in quiet and at the four SNRs. The order of ORCA-RRT conditions was also randomized for each participant with the constraint that quiet was never used as the first condition. Both versions (LC and HC) of the ORCA-RRT were administered, with the LC version always followed by the HC version. Participants' performance on the RST was measured at the final visit.

RESULTS

All statistical analyses were performed in R Statistical Software (Core Team, 2016). As required for

parametric analysis, data on the ORCA-NST and ORCA-RRT were transformed into rationalized arcsine units according to the method described in Studebaker (1985). Unless otherwise specified, all figures of participant performance represent the average of the original nontransformed scores and error bars represent 95% confidence intervals of the mean adjusted for within-subjects measures using the method from Morey (2008). For posthoc analyses involving multiple paired samples *t*-tests, familywise error rates were controlled using the Holm–Bonferroni method (Holm, 1979).

DRs of Processed Stimuli

Figure 4 shows the mean and the DRs of the output for each ORCA-NST (NST.50, 65, 80, and split) and RRT (RRT.0, 5, 10, 15, and quiet) stimulus condition for VSC, FAC, and SAC at each one-octave band centered at 0.5, 1, 2, and 4 kHz. The DR of the unaided stimuli was also included for reference. Several observations are apparent. First, the DR was the widest for the unaided stimuli followed by the SAC, VSC, and finally FAC. This confirms that the different compressor settings resulted in a difference in output DRs. On the other hand, the mean output (where the symbols were) was similar between the VSC and FAC. The output for the SAC was slightly lower in all stimulus conditions. Finally, the mean output for the FAC and VSC was consistently above the average thresholds (shown by the dotted horizontal lines) of the participants in all but the NST50 and NST65 (at 2 and 4 kHz) conditions. In addition, the mean output at all compression speeds for the NST split and RRT high frequencies (2000 and 4000 Hz) was below the mean thresholds of the participants. This reaffirmed the observations in the previous figure that there were audibility differences in addition to temporal envelope differences among the three compressor settings under some test conditions.

Cognitive Profile

One participant was unable to complete the RST because of a scheduling conflict and her data were excluded from all related analyses. Of the 16 participants, the RST scores ranged from 47 to 85 with a mean of 64. Listeners' MoCA scores ranged from 22 to 30 with a mean of 26.3. The two measures were only mildly correlated ($r = 0.43$). Although the sample of participants covered a range of cognitive abilities, it should be noted that the mean RST score (64) was higher than the mean score of 40 reported by other investigators using other versions of the RST (Arehart et al, 2013; Ohlenforst et al, 2015). Although differences in RST versions could partially account for the difference in absolute scores, we cannot rule out the possibility that the cognitive backgrounds of this subject sample may be different from previous studies. Thus, the RST

scores were included as a covariate in subsequent analysis of variance (ANOVA).

ORCA-NST

Listener performance on the ORCA-NST was assessed at the phoneme level. Performance for VSC, FAC, and SAC settings was quantified separately for each ORCA-NST condition (i.e., 50, 65, 80, and Split) by dividing the number of correctly repeated phonemes by the total number of phonemes presented in the test condition. The average performance on the ORCA-NST is shown in Figure 5 for all listeners. A mixed design ANOVA was used to assess the within-subject factors of compressor speed (VSC/FAC/SAC) and test condition (50/65/80/Split) on listeners' phoneme scores. Performance on the RST was included in the model as a covariate. Mauchly's test indicated that the assumption of sphericity was violated for test condition and the interaction of test condition and RST ($W = 0.037, p < 0.001, \epsilon = 0.43$) and for the interaction of test condition and compressor speed ($W = 0.004, p < 0.001, \epsilon = 0.35$). Accordingly, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity.

The ANOVA revealed a significant main effect of test condition [$F_{(1.31,18.33)} = 44.01, p_{GG} < 0.001, \eta^2 = 0.36$] and a significant interaction between test condition and compressor [$F_{(2.10,29.45)} = 4.23, p_{GG} < 0.05, \eta^2 = 0.03$]. Posthoc paired samples' t -tests ($\alpha = 0.05$; Holm–Bonferroni corrected) revealed that listener performance in the split condition was significantly poorer with SAC than with FAC [$t_{(15)} = 2.40, p < 0.05$]. No other significant difference was observed between compressor speeds for any ORCA-NST conditions. Listener RST performance did not significantly interact with compressor

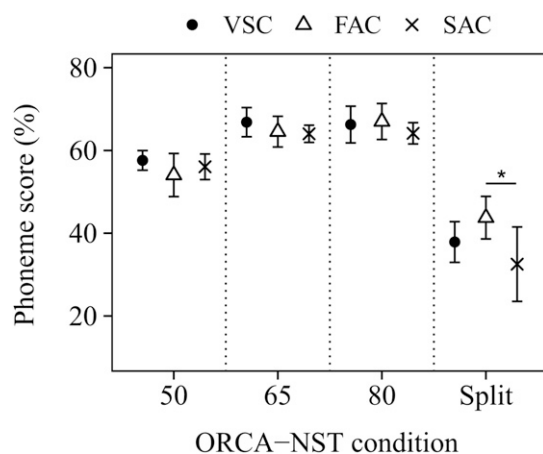


Figure 5. Average phoneme identification scores (percent correct; %) measured in four ORCA-NST conditions while listeners wore HAs programmed with VSC (●), FAC (△), or SAC (×). Values denoting ORCA-NST conditions reflect presentation levels of target speech. Error bars represent 95% confidence intervals of the mean (within-subjects adjusted).

type across any ORCA-NST conditions [$F_{(2,28)} = 0.14, p = 0.87$].

In summary, compressor speed did not result in any difference on the ORCA-NST scores when the stimuli were presented at a fixed level of 50, 65, and 80 dB SPL. Although FAC was significantly better than SAC in the split condition, both FAC and SAC were not significantly different from VSC. The listeners' WMCs did not interact with compressor speed.

ORCA-RRT

Repeat and recall performance on the ORCA-RRT was quantified by dividing the number of correctly repeated/recalled target words by the number of target words presented in each list (i.e., 20 words). Figure 6 shows the performance-intensity functions for repeat, recall, listening effort, and tolerable time measures as a function of SNR for high and LC passages. Four mixed design ANOVAs were used to assess the within-subject factors of compressor speed (VSC/FAC/SAC), test condition (0/5/10/15 dB SNR or Quiet), and passage context (High/Low) separately for each RRT measure: repeat, recall, listening effort, and tolerable time. Performance on the RST was included in these models as a covariate. Where Mauchly's test indicated that the assumption of sphericity has been violated, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity.

As expected, RRT outcome measures were significantly affected by the SNR of the test condition [repeat: $F_{(4,56)} = 545.71, p < 0.001, \eta^2 = 0.80$; recall: $F_{(4,56)} = 132.45, p < 0.001, \eta^2 = 0.48$; listening effort: $F_{(1.47,20.55)} = 105.12, p < 0.001, \eta^2 = 0.47$; tolerable time: $F_{(2.29,30.03)} = 126.32, p < 0.001, \eta^2 = 0.54$] and by passage context [repeat: $F_{(1,14)} = 116.45, p < 0.001, \eta^2 = 0.13$; recall: $F_{(1,14)} = 72.29, p < 0.001, \eta^2 = 0.17$; listening effort: $F_{(1,14)} = 128.78, p < 0.001, \eta^2 = 0.10$; tolerable time: $F_{(1,14)} = 83.49, p < 0.001, \eta^2 = 0.12$]. The ANOVAs also revealed a significant interaction of SNR and passage context on recall [$F_{(4,56)} = 6.78, p < 0.001, \eta^2 = 0.05$], listening effort [$F_{(4,56)} = 10.05, p < 0.001, \eta^2 = 0.03$], and tolerable time [$F_{(4,56)} = 6.74, p < 0.001, \eta^2 = 0.03$]. This interaction likely reflected poorer performance-intensity functions for low compared with HC passages. Taken together, these results suggest that listener performance on the ORCA-RRT was sensitive to the demands of the test conditions and the limited contextual cues. Furthermore, listeners' RST performance was found to be a significant covariate of recall performance [$F_{(1,14)} = 9.10, p < 0.01, \eta^2 = 0.13$]. However, the effect was qualified by an interaction with passage context [$F_{(1,14)} = 5.82, p < 0.05, \eta^2 = 0.02$], wherein the recall and RST performance correlated more strongly for high than for LC passages (Figure 7). These results

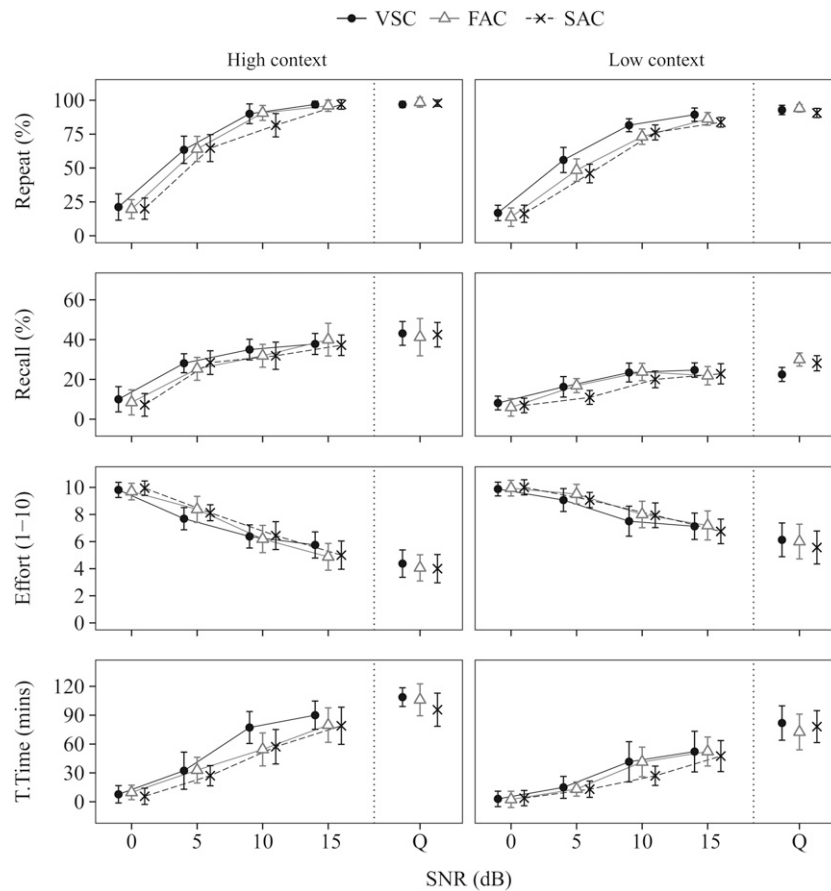


Figure 6. Comparison of VSC (●), FAC (△), and SAC (×) on mean listener performance (repeat, recall, listening effort, and tolerable time) measured with the ORCA-RRT. Stimuli were presented at SNR = 0, +5, +10, and +15 dB and in Quiet (Q). Results are shown for all listeners for HC (left) and low (right) context passages. Error bars represent 95% confidence intervals of the mean (within-subjects adjusted).

suggest that RRT recall and RST both capture meaningful differences in listeners' verbal WMCs. On the other hand, compressor speed did not significantly affect repeat or recall performance or self-reports of listening effort or tolerable time.

ORCA-RRT-Derived Speech Reception Threshold (SRT)

The SRT derived from the performance-intensity functions of the RRT repeat measure may offer another glimpse at performance difference among compressor speeds. Although the traditional criterion for SRT is set at 50% correct, recent work on realistic SNRs suggests that listeners likely operate at a higher criterion level under realistic communication situations. Specifically, the recent work by Smeds et al (2015) and Wu et al (2018) showed that hearing-impaired listeners chose to communicate in environments where the SNR typically ranged between 10 and 15 dB. This corresponds to a repeat rate of 80–95% when examining Figure 6. Recently, Lunner and Sundewall-Thorén (2007) examined differences in compressor speeds by

cognitive background and reported the greatest difference between groups at a SNR corresponding to an 80% correct criterion. Lunner et al (2016) examined the efficacy of a noise reduction algorithm on the Sentence Final Word Identification and Recall Test in a New Language (Ng et al, 2015) at a SNR corresponding to 95% correct. Thus, we decided to use a criteria level of 85% (i.e., SRT_{85}) as a realistic condition to examine potential effect of compressor speed. A higher criterion (such as 90% or 95%) was not possible because of the ceiling performance of some listeners.

Logistic functions were fit to individual listener repeat performance at SNR = 0, +5, +10, and +15 dB using the maximum likelihood method. To limit the biasing of parameter estimates by lapses in listener performance, lambda was permitted to vary as proposed by Wichmann and Hill (2001). Performance data were fit separately for high and LC passages while listeners were VSC, FAC, or SAC compressor settings. Psychometric fitting and threshold estimation were accomplished with the use of the quickpsy package in R (Linares and López-Moliner, 2016). Speech reception thresholds were estimated at 85% correct (SRT_{85}).

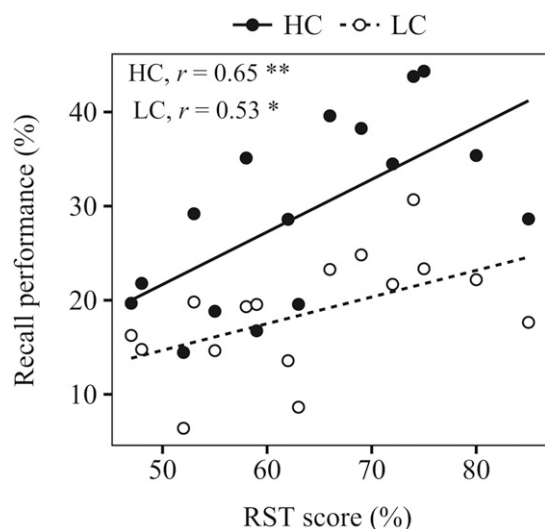


Figure 7. Scatter plot showing the correlation between Recall performance (HC and LC) and RST scores.

Mixed design ANOVAs were used to assess the within-subject factors of compressor speed (VSC/FAC/SAC) and passage context (High/Low) on estimated SRT_{85} . Listener RST performance was included as a covariate in these models. The results of the analysis revealed that passage context significantly affected estimated SRT_{85} [$F_{(1,14)} = 67.50, p < 0.001, \eta^2 = 0.34$]. Posthoc tests confirmed that SRTs were significantly lower for high than for LC passages. Furthermore, compressor speed significantly affected SRT_{85} [$F_{(2,28)} = 3.75, p = 0.03, \eta^2 = 0.09$]. As shown in Figure 8, SRT_{85} was significantly lower for VSC compared with FAC [$t_{(15)} = -2.46, p < 0.05$] and SAC [$t_{(15)} = -2.44, p < 0.05$]. There was no difference between FAC and SAC [$t_{(15)} = -0.70, p = 0.50$]. Again, WMC did not interact with any of the observed effects.

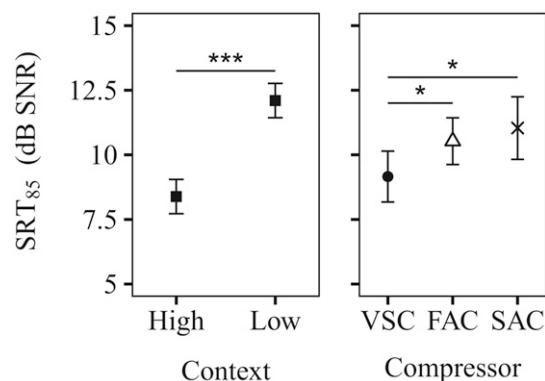


Figure 8. Effect of context (left) and compressor speed (right) on average speech reception thresholds at 85% correct (SRT_{85} ; dB SNR) across all listeners. Estimates of SRT_{85} were derived separately for VSC (●), FAC (△), and SAC (×). Context effect shows the average of all three compressor speeds. Compressor effect shows the average of HC and LC passages. Error bars represent 95% confidence intervals. Asterisks over bars denote significant contrasts.

In summary, the SRT_{85} derived from the RRT repeat scores averaged across all listeners and passages was 1.4 and 2 dB lower for VSC than for FAC and SAC, respectively. Recall performance on the RRT was not affected by compressor speed, although recall scores were highly correlated with RST scores. Neither listening effort nor tolerable time was significantly affected by compressor speed or WMC.

DISCUSSION

The present study examined the effect of realistic compression speeds on intelligibility, recall, and subjective ratings of different speech materials. In the past, most studies comparing compression speeds limited the comparison to fixed compression speeds (i.e., FAC versus SAC) in simulated or simple, two to six channel devices. Although these studies laid the foundation for our understanding of the effect of compression speeds on speech understanding, they may not reflect the current reality of commercial devices many of which use different means of achieving gain changes during compression. In this study, it was shown that FAC resulted in higher phoneme scores than SAC on the ORCA-NST only in the split condition. For the RRT, all compressor speeds were similar to one another on measures of recall, listening effort, and tolerable time. However, VSC resulted in significantly lower (better) speech reception threshold at 85% criterion (SRT_{85}) than fixed, FAC and SAC. The measures of WM, as determined using the RST, did not interact with compression speed during any comparison. These observations suggest that the VSC implemented on the study HA may improve speech intelligibility in some situations over the use of fixed FAC/SAC, possibly by ensuring audibility and temporal envelope cues for all listeners (and not only listeners with good or poor WMC).

The VSC in the present study uses the summed output of both compressors to reach the final output in each channel. It is designed such that the contribution from the SC forms the primary output under a steady and moderate-to-high input level. The contribution from the FC increases when a large fluctuation in input level occurs, at a low input level, and when the signal is highly modulated (such as speech alone or a favorable SNR). This suggests that when the ORCA-NST stimuli were presented at a low input level (of 50 dB SPL at SNR = 15 dB) and in the split condition (80 dB to 50 dB SPL), the FC was likely contributing more because of the fluctuating input levels and of the high modulation of the input. That is, the VSC circuit acted like a FAC circuit under those conditions. Indeed, the similar DRs and the similar phoneme scores between the VSC and the FAC at those two stimulus conditions confirmed that the VSC ensured the audibility of soft

speech like an FAC. The lower output of the SAC compared with the FAC and VSC for modulated signals may account for the poorer score of the SAC in the split condition.

On the other hand, for the 65-dB SPL (SNR = 5 dB) and the 80-dB SPL (SNR = 0 dB) conditions, it is likely that the primary contribution was from the SC of the VSC. The poorer SNRs resulted in a less modulated input. The lack of a difference in performance among VSC, FAC, and SAC at these levels is not surprising when the mean output level of the three sets of processed stimuli were above the thresholds of the participants for most frequencies despite their different DRs (see Figure 4). Because all three sets of stimuli were equally audible, the lack of a speech intelligibility difference would suggest that intelligibility measure may not be the most sensitive measure to differentiate among compression speeds. These results are also in line with the results of several studies that reported similar speech intelligibility scores between SAC and FAC using nonsense syllables (e.g., Moore et al, 2004; Jenstad and Souza, 2005) and sentence materials (e.g. Gilbert et al, 2008; Shi and Doherty, 2008). However, the results of this study are in contrast to other studies that showed a significant interaction between compression speed (SAC versus FAC) and WMC (e.g., Gatehouse et al, 2003; Lunner, 2003; Lunner and Sundewall-Thorén, 2007; Cox and Xu, 2010; Ohlenforst et al, 2015). This is despite the fact that many design elements (such as use of modulated noise, different contextual cues etc) of this study were inspired by the outcomes of previous studies.

The observation of similar identification scores on the ORCA-NST among VSC, SAC, and FAC underlines the fact that compression speeds may not affect intelligibility of nonsense syllables as long as the amplified sounds are sufficiently above the listeners' thresholds. Even the SAC used in this study was sufficient in providing adequate audibility despite the level variations in the nonsense syllables. In other words, it is not necessary to have FAC at all times to ensure consistent audibility. However, a rapid change in overall level where the mean output level was below threshold, like that seen in the split condition, could affect audibility and thus intelligibility. In this situation, the use of the VSC could provide the needed audibility while preserving the temporal envelope.

It is likely that the SC was the greater contributor during the ORCA-RRT because of the lower modulation for speech presented in noise. This is demonstrated in Figure 4, which shows the DR of the VSC stimuli was wider than that of the FAC stimuli but narrower than the SAC stimuli at all SNRs. The FC likely increased its contribution at the SNR = +10 and +15 because the more positive SNR may be recognized as greater modulation. This suggests that the audibility of any softer

sounds presented during the passage was enhanced over a fixed SAC condition, especially at a more favorable SNR. Any perceptual differences between VSC and FAC/SAC would likely be based on a combination of enhanced audibility and enhanced temporal envelope cues. The results of the statistical analyses revealed that any such enhancement did not significantly improve the recall of the sentence materials or subjective judgments of the listening conditions.

However, the analysis of the RRT showed a significant difference in SRT_{85} between compressor speeds and passage contexts despite no interaction with WMC (as measured with the RST). Even though a cognitive effect was not seen, the observation may support Lunner and Sundewall-Thorén's (2007) suggestion that a SNR corresponding to SRT_{80} may be a good test condition to examine differences among compressor speeds. When averaged across all listeners and contexts, the SRT_{85} was 9.1 dB for VSC, 10.5 dB for FAC, and 11.1 dB for SAC. Because the SRT_{85} was measured around SNRs of 10 and 15 dB, this suggests speech intelligibility at a moderately loud level and favorable SNRs would benefit the most from VSC processing instead of the fixed FAC/SAC processing (which do not differ from each other). It is unclear why this SRT criterion is sensitive to compressor speed. Clearly, at such a favorable SNR and speech level that most, if not all, of the speech sounds should be audible. Perhaps more subtle acoustic cues, including temporal ones, other cognitive factors, and/or attention may be involved. Nonetheless, this observation is noteworthy because this SNR is similar to the realistic SNRs that hearing-impaired persons chose to encounter in real life (Smeds et al, 2015; Wu et al, 2018).

It is reasonable to expect that communication between a talker and a listener occurring at SNRs that correspond to an SRT_{85} to be successfully. Conversely, if one understands only 50% of what is said, it is likely that any communication will be unsuccessful and terminate quickly. Thus, determining a listener's SRT_{85} may be more predictive of real-life communication difficulties. Furthermore, finding technologies that minimize the SNR requirement will be helpful. Although it is customary to consider the use of noise reduction and directional microphones to improve such SNR, the results of this study would add that the use of VSC may provide additional improvement (1–2 dB) over the use of a fixed speed compression, in at least some situations. Albeit small, it is similar in magnitude to that provided by a directional microphone (about 2 dB) fitted with an open earmold (Valente and Mispagel, 2008; Magnusson et al, 2013). In such fittings, the use of a directional microphone in behind-the-ear and RIC HAs is deemed indispensable.

Despite the observation that VSC implemented in the study HA yielded a better SRT_{85} than FAC and

SAC, one cannot conclude that all VSC or adaptive compression circuits would yield superior performance. This is because adaptive compression can be realized with just one compressor (such as the Widex DREAM), or with two compressors each using a different fixed or variable speed. Some dual compressors place more weight on the output from the SC (e.g., Pittman et al, 2014), whereas some place more weight on the FC (e.g., Moore et al, 2004). Some adaptive compression circuits change the CR instead of the release time (e.g., Lai et al, 2013) while others bypass a discussion of any time constants (Blamey, 2005). The potential variation among commercial devices underscore the importance of transparency on the part of the manufacturers so that clinicians can make informed choices on the right commercial product(s). Regardless of the approach, products that preserve as much of the temporal envelope while ensuring sufficient audibility and loudness across frequencies will likely be the most successful. The results from this study support the VSC implemented on the study HA as a viable option for a wide range of people with different cognitive abilities and hearing requirements.

Future investigation into the possible benefits of VSC would benefit from a larger sample of listeners that cover a wider range of cognitive abilities. A greater number of listeners might provide the power required to investigate interesting trends regarding the effect of compression speed on the RRT measures of listening effort and tolerable time. Moreover, compression speed might be found to interact with WMC when the latter measures include listeners who perform poorly on WM tests. By the same token, the use of additional WM tests may allow a more precise characterization of the listener's WMC. In addition, the present study was conducted with the HAs in situ; all the adaptive features within the HAs (such as noise reduction, directional microphones, sound classifier, and learning) were deactivated during the study when they were coupled with occluded earmolds. When the HAs are worn in real life with individualized earmolds and a full range of features, the temporal envelope and audibility of the processed signals would be altered. How such factors affect the results reported herein remains a topic for further investigation.

Nonetheless, the results of this study prompt the following considerations. First, the recommendation that people with good WMC should be fitted with FAC and those with poor WMC should be fitted with SAC may need reexamination. No difference between FAC and SAC processing was noted in our study over a range of test conditions even when considering WMC as a covariate. Furthermore, advances in compression technology have allowed compression systems to use two compressors to achieve the benefits of different fixed compression speeds while minimizing the limitations

associated with either fixed speed. With an advanced compression system, all wearers are provided with all the cues such that the wearer may not be as dependent on the test conditions as with the use of a single fixed speed system that may compromise the availability of specific cues (i.e., temporal or audibility) in certain situations. Moreover, an adaptive system should be at least as good as a fixed FAC/SAC in most situations, but better than the fixed system in some situations. If this is the case, it may not be necessary to measure the WMC of listeners for the purpose of fitting compression HAs. This is not to say that knowing the cognitive ability of the patient is not important; such knowledge would allow us to explain the reasons for patients' difficulty with speech in noise, to set realistic expectations, and to design the optimal rehabilitation approach (e.g., Shen et al, 2016). Because the implementation of the dual compressor is fundamental to the wearers' potential success, its operation should be transparent so clinicians can make informed choices for their patients.

CONCLUSIONS

The present study showed that FAC is similar to SAC on most test measures other than the split level condition used during the NST. The VSC algorithm implemented on the study HA was similar to a fixed FAC and SAC in most situations, but superior to such fixed systems at a moderately high input level with a favorable SNR. The WMC of the listeners in this study did not affect the outcome of the comparison.

REFERENCES

- Arehart K, Souza P, Baca R, Kates J. (2013) Working memory, age, and hearing loss: susceptibility to hearing aid distortion. *Ear Hear* 34:251–260.
- Blamey P. (2005) Adaptive dynamic range optimization (ADRO): a digital amplification strategy for hearing aids and cochlear implants. *Trends Amplif* 9:77–98.
- Boike K, Souza P. (2000) Effect of compression ratio on speech recognition and speech-quality ratings with wide dynamic range compression amplification. *J Speech Lang Hear Res* 43(2):456–468.
- Boyle P, Buchner A, Stone M, Lenarz T, Moore B. (2009) Comparison of dual time constant and fast acting automatic gain control (AGC) systems in cochlear implants. *Int J Audiol* 48:211–221.
- Bregman A. (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.
- Cooke M. (2006) A glimpsing model of speech perception in noise. *J Acoust Soc Am* 119(3):1562–1573.
- R Core Team. (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Cox R, Xu J. (2010) Short and long compression release times: speech understanding, real-world preferences, and association with cognitive ability. *J Am Acad Audiol* 21(2):121–138.

- Daneman M, Carpenter P. (1980) Individual differences in working memory and reading. *J Verbal Learn Verbal Behav* 19: 450–466.
- Dillon H. (1996) Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times? *Ear Hear* 17(4):287–307.
- Fogerty D, Entwistle J. (2015) Level considerations for chimeric processing: temporal envelope and fine structure contributions to speech intelligibility. *J Acoust Soc Am* 138(5):459–464.
- Gatehouse S, Naylor G, Elberling C. (2003) Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol* 42(1, Suppl):S77–S85.
- Gatehouse S, Naylor G, Elberling C. (2006) Linear and nonlinear hearing aid fittings-2. Patterns of candidature. *Int J Audiol* 45(3): 153–171.
- Gilbert G, Akeroyd M, Gatehouse S. (2008) Discrimination of release time constants in hearing aid compressors. *Int J Audiol* 47(4):189–198.
- Hansen M. (2002) Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear Hear* 23(4):369–380.
- Hau O, Andersen H. (2012) Hearing aid compression: effects of speed, ratio and channel bandwidth on perceived sound quality. www.audiologyonline.com. Accessed April 2, 2012.
- Holm S. (1979) A simple sequentially rejective multiple test procedure. *Scand J Stats* 6(2):65–70.
- Holube I, EHIMA-ISMADHA Working Group. (2006) Short Description of the International Speech Test Signal (ISTS). Oldenburg, Germany: Center of Competence HörTech and Institute of Hearing Technology and Audiology.
- Jenstad L, Souza P. (2005) Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *J Speech Lang Hear Res* 48(3):651–667.
- Jenstad L, Souza P. (2007) Temporal envelope changes of compression and speech rate: combined effects on recognition for older adults. *J Speech Lang Hear Res* 50(5):1123–1138.
- Kuk F. (1996) Theoretical and practical considerations in compression hearing aids. *Trends Amplif* 1(1):5–39.
- Kuk F. (1998) Rationale and requirements for a slow acting compression hearing aid. *Hear J* 51(6):45–53.
- Kuk F, Hau O. (2017) Compression speed and cognition: a variable speed compressor for all. *Hear Rev* 24(3):40–48.
- Kuk F, Lau C, Korhonen P, Crose B, Peeters H, Keenan D. (2010) Development of the ORCA nonsense syllable test. *Ear Hear* 31(6): 779–795.
- Lai Y, Li P, Tsai K, Chu W, Young S. (2013) Measuring the long term SNRs of static and adaptive compression amplification techniques for speech in noise. *J Am Acad Audiol* 24(8): 671–683.
- Linares D, López-Moliner J. (2016) Quickpsy: an R package to fit psychometric functions for multiple groups. *R J* 8(1):122–131.
- Lorenzi C, Gilbert G, Carn C, Garnier S, Moore B. (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci U.S.A.* 103: 18866–18869.
- Lunner T, Hellgren J, Arlinger S, Elberling C. (1997) A digital filter-bank hearing aid: three digital signal processing algorithms—user preference and performance. *Ear Hear* 18:373–387.
- Lunner T. (2003) Cognitive function in relation to hearing aid use. *Int J Audiol* 42(1, Suppl):S49–S58.
- Lunner T, Sundewall-Thorén E. (2007) Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J Am Acad Audiol* 18(7):604–617.
- Lunner T, Rudner M, Rosenbom T, Agren J, Ng E. (2016) Using speech recall in hearing aid fitting and outcome evaluation under ecological test conditions. *Ear Hear* 37(1, Suppl): 145S–154S.
- Magnusson L, Claesson A, Persson M, Tengstrand T. (2013) Speech recognition in noise using bilateral open-fit hearing aids—the limited benefit of directional microphones and noise reduction. *Int J Audiol* 52:29–36.
- Moore B. (2008) The choice of compression speed in hearing aids: theoretical and practical considerations and the role of individual differences. *Trends Amplif* 12(2):103–112.
- Moore B, Glasberg B. (1988) A comparison of four methods of implementing automatic gain control (AGC) in hearing aids. *Br J Audiol* 22:93–104.
- Moore B, Stainsby T, Alcantara J, Kuhnel V. (2004) The effect on speech intelligibility of varying compression time constants in a digital hearing aid. *Int J Audiol* 43:399–409.
- Morey RD. (2008) Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial Quant Methods Psychol* 4(2):61–64.
- Nasreddine Z, Phillips N, Bédirian V, Charbonneau S, Whitehead V, Collin I, Chertkow H. (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53(4):695–699.
- Neuman A, Bakke M, Mackersie C, Hellman S, Levitt H. (1995) Effect of release time in compression hearing aids. *J Acoust Soc Am* 98(6):3182–3187.
- Neuman A, Bakke M, Mackersie C, Hellman S, Levitt H. (1998) The effect of compression ratio and release time on the categorical rating of sound quality. *J Acoust Soc Am* 103(5):2273–2281.
- Ng E, Rudner M, Lunner T, Ronnberg J. (2015) Noise reduction improves memory for target language speech in competing native but not foreign language speech. *Ear Hear* 36:82–91.
- Oeding K, Valente M. (2013) Differences in sensation level between the Widex SoundTracker and two real-ear analyzers. *J Am Acad Audiol* 24(8):660–670.
- Ohlenforst B, Souza P, MacDonald E. (2015) Exploring the relationship between working memory, compressor speed, and background noise characteristics. *Ear Hear* 37(2):137–143.
- Pascoe D. (1988) Clinical measurements of the auditory dynamic range and their relation to formulas for hearing aid gain. Hearing Aid Fitting Theoretical and Practical Views. 13th Danavox Symposium 129–151. Copenhagen, August 1988.
- Pearsons K, Bennett R, Fidell S. (1977) Speech levels in various noise environments. EPA report#600/1-77-025. Washington DC.

- Pittman A, Pederson A, Rash M. (2014) Effects of fast, slow, and adaptive amplitude compression on children's and adults' perception of meaningful acoustic information. *J Am Acad Audiol* 25:834–847.
- Plomp R. (1988) The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *J Acoust Soc Am* 83(6):2322–2327.
- Postle B. (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139(1):23–38.
- Rosen S, Faulkner A, Smith D. (1990) The psychoacoustics of profound hearing impairment. *Acta Otolaryngol Suppl* 469:16–22.
- Rudner M, Rönnberg J, Lunner T. (2011) Working memory supports listening in noise for persons with hearing impairment. *J Am Acad Audiol* 22:156–167.
- Shen J, Anderson M, Arehart K, Souza P. (2016) Using cognitive screening tests in audiology. *Am J Audiol* 25:319–331.
- Shi L, Doherty K. (2008) Subjective and objective effects of fast and slow compression on the perception of reverberant speech in listeners with hearing loss. *J Speech Lang Hear Res* 51(5):1328–1340.
- Slugocki C, Kuk F, Korhonen P. (2018) Development of an integrated repeat and recall test (RRT). (submitted).
- Smeds K, Wolters F, Rung M. (2015) Estimation of signal-to-noise ratios in realistic sound scenarios. *J Am Acad Audiol* 26(2):183–196.
- Souza P. (2000) Older listeners' use of temporal cues altered by compression amplification. *J Speech Lang Hear Res* 43(3):661–674.
- Souza P. (2002) Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends Amplif* 6(4):131–165.
- Souza P, Jenstad L, Folino R. (2005) Using multichannel wide-dynamic range compression in severely hearing-impaired listeners: effects on speech recognition and quality. *Ear Hear* 26(2):120–131.
- Souza P, Sirow L. (2014) Relating working memory to compression parameters in clinically fit hearing aids. *Am J Audiol* 23(4):394–401.
- Stone M, Moore B, Alcantara J, Glasberg B. (1999) Comparison of different forms of compression using wearable digital hearing aids. *J Acoust Soc Am* 106:3603–3619.
- Studebaker GA. (1985) A rationalized arcsine transform. *J Speech Hear Res* 28:455–462.
- Valente M, Mispagel K. (2008) Unaided and aided performance with a directional open-fit hearing aid. *Int J Audiol* 47(6):329–336.
- Van den Noort M, Bosch P, Haverkort M, Hugdahl K. (2008) A standard computerized version of the reading span test in different languages. *Eur J Psychol Assess* 24(1):35–42.
- Van Tasell D, Trine T. (1996) Effects of single-band syllabic amplitude compression on temporal speech information in nonsense syllables and in sentences. *J Speech Hear Res* 39(5):912–922.
- Wichmann F, Hill J. (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys* 63(8):1293–1313.
- Wu Y, Stangl E, Chipara O, Hasan S, Welhaven A, Oleson J. (2018) Characteristics of real world signal-to-noise ratios and speech listening situations of older adults with mild to moderate hearing loss. *Ear Hear* 39(2):293–304.