Phonak Field Study News.

Spheric Speech Clarity enhances interference control and working memory while listening in noise

This study, conducted in Shanghai, China, investigated the effect of a Deep Neural Network (DNN) signal processing noise reduction feature, Spheric Speech Clarity, on cognitive function. The feature, introduced with the Phonak Infinio platform, showed significant effects on interference control and working memory capacity while listening in noise.

Huang, HY., Huang, W., Tian, X, Guan, J., Kuhnel, V., & Appleton-Huber, J., April 2025

Key highlights

- Spheric speech clarity (SSC) was found to improve interference control in challenging auditory environments.
- A significant improvement was observed in working memory capacity in noise for the Auditory Backward Digit Span (BDS) performance with SSC compared to the Real Ear Sound (RES) condition.
- SSC was found to significantly improve speech intelligibility in multi-talker conversations in noise when compared to the deactivated condition.

Considerations for practice

- Although many studies have concluded that hearing aids might enhance cognitive performance in domains like executive function, little research has focused on how specific hearing aid features impact cognitive performance in noisy environments.
- Phonak introduced its advanced noise reduction feature, Spheric Speech Clarity, with the launch of the Infinio platform. This technology leverages a Deep Neural Network (DNN) that is specifically designed and trained to extract speech from noisy environments.
- As well as improving speech intelligibility, data from this study show that SSC improves cognitive performance in noisy environments, i.e. working memory capacity.

Introduction

Several studies have investigated the relationship between cognitive performance and hearing loss. For instance, Mamo and Helfer (2021) found that individuals with mild hearing loss and mild cognitive impairment exhibited poorer working memory performance and reduced speech intelligibility when exposed to speech-masking noise compared to modulated noise.

Building on the research linking cognitive performance and hearing loss, several studies have investigated the relationship between cognitive performance and hearing aid use. Doherty and Desjardins (2015) demonstrated that working memory performance improved significantly after six weeks of hearing aid use, as measured by two auditory-based tests: the Listening Span Test and N-back test. Sanders et al. (2021) reviewed the literature on the relationship between hearing aid interventions and cognitive performance. Their analysis found that hearing aids use was significantly associated with improvements in working memory and executive function. Yang et al. (2022) reviewed 15 studies and concluded that the use of hearing aids might enhance cognitive performance in domains like executive function, particularly in individuals without dementia. Chen et al. (2021) examined 49 hearing aid users and 46 non-users with mild to moderate hearing loss among Mandarin speakers. They found a strong correlation between speech-in-noise performance and cognitive function in both groups, with working memory and executive function being especially relevant. However, their study also noted that cognitive performance is influenced by additional factors, such as education level and age.

Comparatively, there is limited research focusing on how specific hearing aid features impact cognitive performance in noisy environments. Windle et al. (2023) investigated the association of various hearing aid functions with cognitive performance. They found that noise reduction improved working memory function in noise for individuals with hearing impairments. However, they also pointed out that excessive noise reduction may impair cognitive performance and speech perception, particularly for users with lower cognitive abilities, due to signal distortion. Therefore, Windle recommended setting noise reduction to moderate levels to achieve an optimal balance between performance and perception.

With the introduction of the Infinio platform Phonak introduced a novel DNN-based denoising feature, known as Spheric Speech Clarity (SSC). This technology leverages a proprietary large DNN that is specifically designed and trained to extract speech from noisy environments. As a result of the novelty of this technology, there are open research questions regarding the possible benefits of SSC across a number of domains in field of cognitive hearing science. Of specific interest are the effects on cognitive functions, such as selective attention, working memory and executive function. Investigating whether this DNN-based denoising method influences cognitive performance in noisy environments may yield valuable clinical insights.

The objectives of this study were to investigate whether activating SSC (SSC-on), compared to having it deactivated (SSC-off) and to using Real Ear Sound (RES), would improve:

- behavioral performance in control of executive function and processing speed
- behavioral performance in working memory capacity in noise
- speech intelligibility in multi-talker environments with noise

Methodology

Participants

Eighteen native Mandarin-speaking participants (5 female) were recruited for the study, with ages ranging from 65-82 years, and screening with the Chinese version of the MoCA-Basic (Chen et al., 2016) revealed no significant cognitive impairment among them (passing score: 24/30). Participants in this study had mild to moderate sensorineural hearing loss (figure 1) and were experienced hearing aid wearers.

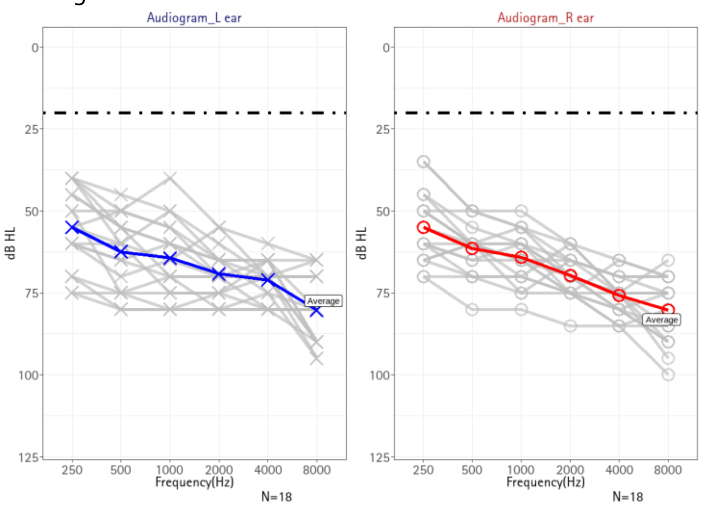

Figure 1. Pure Tone Audiograms of the test participants. Gray lines show individual audiograms. Blue and red lines show the mean audiogram of the participants for the left and right ears respectively.

Devices

Participants were fit with Phonak Audéo I90-Sphere hearing aids, equipped with individually customized Slim Tips. Vent sizes were determined based on Target's recommendations. The fitting formula used was Adaptive Phonak Digital (APD) Contrast 3.0, with the gain level set to 100% and frequency lowering disabled.

Three manual programs were created for this study:

1. Spheric Speech in Loud Noise: Fixed directional set to level 12 and Spheric Speech clarity set to level 5, with other features deactivated (SSC-on).
2. Speech in noise: Fixed directional (12), with other features deactivated (SSC-off).
3. Speech in noise: RES (4), with other features deactivated (RES).

Procedures

Auditory Stroop test

In the Auditory Stroop test (Kestens, 2021; Morgan, 1989), participants were required to discriminate between the semantic meaning and the acoustic properties of a word they heard. Once the stimulus word stopped playing, two buttons labelled 'High' and 'Low' immediately appeared for the participant to choose from (see figure 2). Three types of stimuli were used during the test:

- Congruent: semantic meaning and acoustic properties matched; for example, the word "high" was played at a high pitch.
- Incongruent: semantic meaning and acoustic properties differed; for example, the word "high" was played at a low pitch.
- Neutral: the word carried irrelevant semantic meaning; for example, the word "square" was played at either a high or a low pitch.

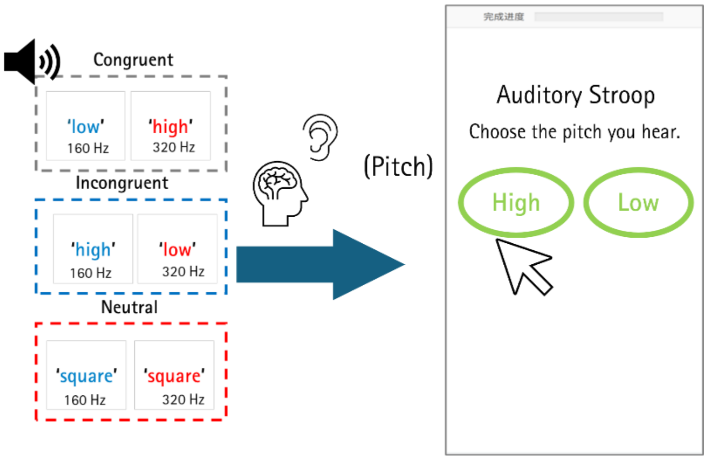

Figure 2. Visual representation of the Auditory Stroop test showing on the left what the participant heard and on the right what they could see on the screen where they could also select their answer.

For the incongruent condition, participants needed to suppress false semantic information and focus on the correct acoustic properties. This gives an indication of the ability of controlling interference and is a measure of cognitive inhibition and cognitive processing speed. Each test consisted of 36 trials, including 12 of each stimulus type (congruent/incongruent/neutral). The target stimuli were always played from 0°, while competing noise (cafeteria noise fixed at 65 dB SPL, corresponding a SNR of 5 dB) was played from 0°, 60°, 180° and 300° (figure 3). Participants completed two training lists with hearing aids set to RES, followed by three tests with different hearing aid settings, in a randomized order (SSC-on, SSC-off and RES). Only the response times from trials with correct responses were included in the analysis. Data from two participants, whose performance was around the chance level, were excluded.

The final analysis included 16 participants whose accuracies were above 80% across all conditions (chance level = 50%).

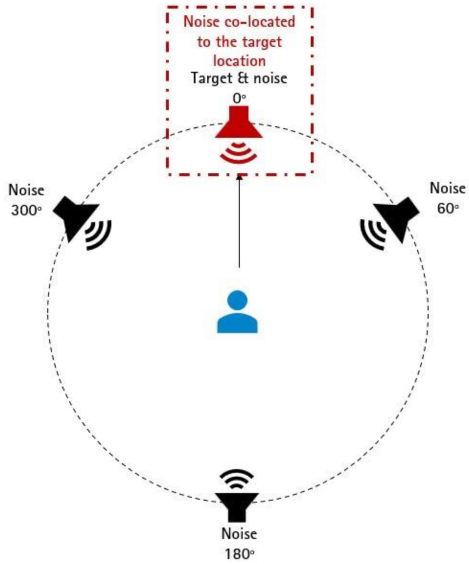

Figure 3. Speaker setup for the Cognitive processing speed test.

Working memory capacity in noise test

The Auditory Backward Digit Span task primarily assesses an individual's auditory working memory and working memory capacity. It also involves complex attention and executive function.

Participants sat in the center of a speaker setup identical to that of the cognitive processing speed test (figure 3). Both the cognitive processing speed test and Auditory Backward Digit Span tests follow the same procedure and use the same sound field settings.

To better evaluate the impact of SSC on working memory capacity, the number of digits retained by participants, rather than recall accuracy scores, was calculated. Each digit sequence was evaluated through two trials (same digit length but different digits). In each trial, participants listened to a series of digits, with each digit presented one at a time, at a rate of one per second. Once all the digits had been presented, participants were asked to enter the digits in reverse order using the keyboard (backward spans; figure 4). If participants correctly recall the sequence in at least one of the trials, the length of the digit sequence was increased by one digit. This pattern continued until participants failed to recall the digit sequence in both trials at a specific digit length (Hester et al. 2004, Tripathi et al. 2019). A longer Backward Digit Span (BDS) indicates increased memory capacity.

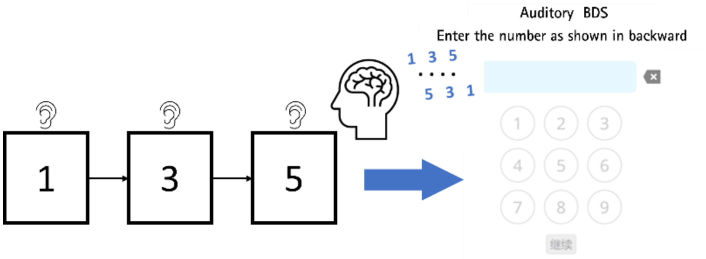

Figure 4. Illustration of test procedure for the Auditory Backward Digit Span test.

Speech recognition test

The concurrent Chinese Coordinate Response Measure (CC-CRM) task was used to measure speech recognition performance in noise. The CC-CRM replicates the concurrent OLSA (CC-OLSA) paradigm described by Heeren et al. (2022). As shown in Figure 5, the target stimulus was sequentially presented from three directions: 0°, +60°,300°, using two male voices and one female voice. Cafeteria noise was presented from 12 different directions ranging from 0° and 360° (figure 5). The noise was fixed at 70 dB SPL, while the speech level was based on the individual's SRT50 measured under the RES condition (speech-in-noise processing enabled and other advanced features disabled), with an additional 3 dB added.

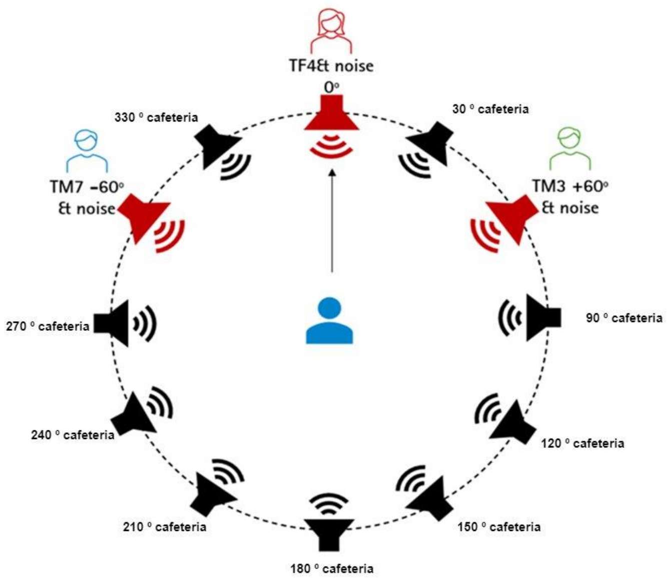

Figure 5. Speaker setup for the CC-CRM test

Each CC-CRM sentence consists of five words with the structure: Ready-name-go to-color-number. One of the names from the CC-CRM corpus, "Wang Li", was designated as the target name. During the CC-CRM task, participants were instructed to focus only on sentences containing this target name and ignore the distraction sentences with other names. A 0.6-second overlap interval was incorporated into the sequence of CC-CRM sentences, ensuring that the target word always aligned with the onset of the subsequent sentence spoken by the next speaker. This arrangement guaranteed that no more than two speakers were active at any given time (figure 6).

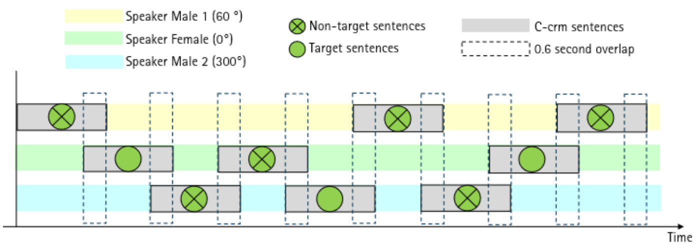

Figure 6. Sample sequences of C-CRM sentences; sentences are timed with a defined temporal overlap (0.6s) and are presented randomly from the three talkers; some sentences with name "Wang Li", which is defined as the call sign.

Participants were seated in front of a touchscreen, displaying 40 response buttons labeled according to the CRM response set, including combinations of four colors (blue, green, red, yellow) and 10 digits (1-10), such as "blue 1", "blue 2" etc. They were instructed to respond as quickly as possible by pressing the button that matched the color-number combination they heard in sentences featuring the target name. After two training lists, the speech reception threshold of 50% intelligibility (SRT50) in noise was obtained through the CRM adaptive procedure to define the optimum SNR for the CC-CRM tests. CC-CRM measurements were conducted at a fixed level 3 dB SNR higher than the participant's SRT50. A CC-CRM test was performed under three conditions: SSC-on, SSC-off and RES. The outcome is a speech intelligibility score expressed as percent correct for each condition.

Results

Figure 7 shows the results of the Auditory Stroop task. Findings were as follows: A multifactorial analysis of variance (two-way ANOVA 2x2) was conducted to examine the effects of test condition (F(2,30)= 12.37, $p < .05$, $\eta^2 p$ = 0.452) and types of Stroop stimuli (F(2,30)=62.38, p < .05, $\eta^2\ p = 0.806$). The findings revealed a significant main effect for both test condition and Stroop stimulus types, with no observed interaction effect between the two variables. Following this, a post hoc analysis was performed using a pairwise t-test with Bonferroni correction to further investigate the effects of the three test conditions: SSC-on, SSC-off, and RES.

- The mean congruent response time did not show significant difference across hearing aid settings.
- The mean incongruent response time was significantly faster with SSC-on compared to SSC-off (t(15) = -3.546,

$p < .05$, Cohen's d = -0.816) and RES (t(15) = -5.095, $p < .05$, Cohen's d = -1.040).

- The mean neutral response time with SSC-on was significantly faster than RES (t (15) = -3.213, $p < .05$, Cohen's d = -0.953). However, there was no significant difference between SSC-on and SSC-off.

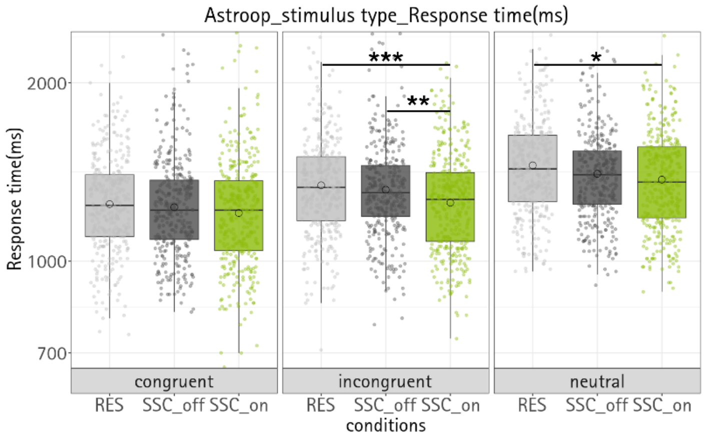

Figure 7. 1 Response times (ms) for the auditory cognitive processing speed test across congruent, incongruent, and neutral stimulus types under RES, SSC-off, and SSC-on conditions. significant differences between conditions (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

Backward Digit Span (BDS) performance can be seen in figure 8. A multifactorial analysis of variance (two-way ANOVA 2×2) was performed to investigate the effects of test condition (F(2,34) = 18.1, $p < .05$, η^2 $p = 0.516$) and test–retest (F(1,17) = 4.02, $p = 0.06$, η^2 $p = 0.192$). The results indicated a significant main effect for test condition; however, no significant effect was observed for test–retest, nor was there an interaction effect between the two variables. Subsequently, a post hoc analysis was conducted utilizing a pairwise t-test with Bonferroni correction to further explore the effects of the three test conditions: SSC-on, SSC-off, and RES.

- BDS in the RES condition was significantly shorter than in both the SSC-on (t (17) = 5.621, $p < 0.05$, Cohen's d = 1.151) and SSC-off (t (17) = 3.971, $p < 0.05$, Cohen's d = 0.581) conditions, with mean differences of 1.3 and 0.7 units, respectively. These findings suggest that SSC led to an improvement in span length, reflecting improved performance in working memory capacity in noise.
- There was no significant difference in BDS between the SSC-on and SSC-off conditions (t (17) = 2.651, p = 0.050), but the comparison showed a moderate effect size (Cohen's d = 0.569). This indicates that a trend that performance was slightly better in the SSC-on condition compared to SSC-off, even though the difference was not statistically significant.

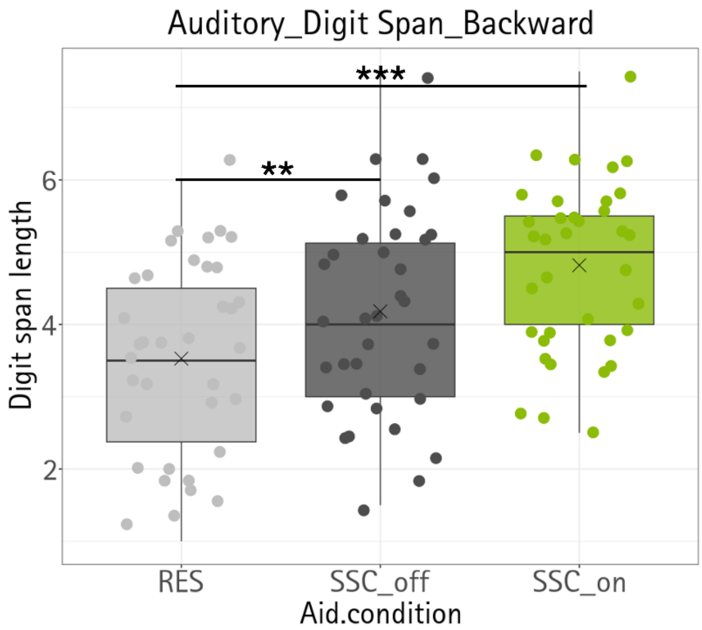

Figure 8. Comparison of Backward Digit Span (BDS) performance with RES, SSC-off, and SSC-on (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

Figure 9 shows the results of the speech intelligibility test. A multifactorial analysis of variance (two-way ANOVA 2×2) was conducted to examine the effects of test condition (F(2,30)= 12.37, $p < .05$, η^2 $p = 0.452$) and types of speaker angles (F(2,30)=62.38, $p < .05$, η^2 $p = 0.806$). The results indicated a significant main effect for both the test condition and the speaker angles, with no significant interaction effect detected between the two variables.

Significant differences in speech intelligibility were observed with SSC-on compared to SSC-off(60°: =t(16) = 2.767, p < .05, Cohen's d = .445; 300°=t(16) = 3.748, $p < .05$, Cohen's d = .616) and RES (60°: t(16) = 4.078, $p < .05$, Cohen's d = .707; 300°=t (16) = 7.589, $p < .05$, Cohen's d = 1.078) when speech came from angles of 60° and 300°. No significant difference was observed between the SSC-on and SSC-off conditions from an angle of 0°; however, in both conditions, speech intelligibility was significantly higher than the RES condition.

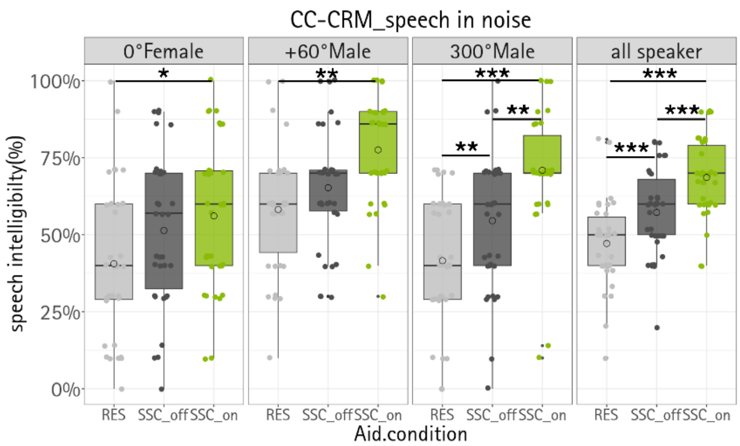

Figure 9. Speech intelligibility scores for the conditions RES, SSC-off, SSC-on when speech came from different angles (0, 60, 300 degrees and all angles (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

Conclusion

Phonak's introduction of Spheric Speech Clarity (SSC) brought novel DNN-based denoising to hearing instruments. As this is a new technology, there are open research questions regarding the effects of this denoising method on cognitive functions that include, auditory working memory, auditory selective attention, and executive function. The study reported here investigated the effects of several features within Phonak Infinio hearing aids and their influences on these domains of cognitive performance in noisy environments.

The auditory Stroop test was selected as measure of cognitive performance that is reported in response time. Group data revealed that the mean incongruent response time was significantly faster in the SSC-on program as compared to the SSC-off program indicating that SSC improves auditory processing speed and control of interference in incongruent trials. This emphasizes that SSC may be beneficial for enhancing auditory selective attention in challenging listening environments.

The recall of digits (i.e. numbers) is a common measure of working or short-term memory. In this study, a task of backward digit span was selected to assess the effect of the three hearing instrument programs. With the RES condition as a control, significant improvements were seen for the two treatment conditions. Specifically, working memory was significantly improved for the Speech in Noise program (SSC-off) and the Speech in Loud Noise Program (SSC-on). These findings suggest that DNN denoising in SSC-on significantly improved working memory, allowing for more words to be recognized and remembered.

Study participants also completed a test of speech intelligibility that required auditory segregation across concurrent speech streams. Group data show that speech intelligibility improved significantly across test conditions, when compared to the RES condition. In other words, speech intelligibility improved significantly for the Speech in Noise program. For the Speech in Loud Noise program (SSC-on) speech intelligibility improved significantly from all directions, the average increase in speech intelligibility was 23 percentile points.

The results of this study show that programs for Speech in Noise and Speech in Loud Noise, in Phonak Infinio hearing instruments, provide significant benefits that relate to cognitive performance and speech intelligibility. Participants in this study experienced an improved ability to focus attention on important sounds in complex listening environments. The same participants demonstrated significantly improved memory for words in a challenging background of noise. Finally, speech intelligibility improved for both treatment conditions, even when speech was overlapping and presented from different locations around listener.

In sum, these findings show significant benefits of the Speech in Noise and Speech in Loud Noise programs. The denoising provided by SSC improved participants auditory attention by allowing them to select an auditory stream, understand that information, and store it in short term memory. These abilities are essential to speech understanding in complex environments. Finally, speech intelligibility was significantly improved for sounds from all directions, showing that listeners could selectively attend to a sound source from any direction.

References

Chen, K., Xu, Y., Chu, A., Ding, D., Liang, X., Nasreddine, Z. S., Dong, Q., Hong, Z., Zhao, Q., & Guo, Q. (2016). Validation of the Chinese version of Montreal Cognitive Assessment Basic for Screening Mild Cognitive Impairment. Journal of the American Geriatrics Society, 64(12) e285-e290.

Chen, Y., Wong, L. L. N., Chan, S.S., & Yu, J. (2021). Speech perception in noise is associated with different cognitive abilities in Chinese-speaking older adults with and without hearing aids. Frontiers in Psychology, 12. 640300.

Doherty, K. A., & Desjardins, J., L. (2015). The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults. Frontiers in Psychology, 6, 721.

Heeren, J., Nuesse, T., Latzel, M., Holube, I., Hohmann, V., Wagener, K., C., & Schulte, M. (2022). The concurrent OLSA test: A method for speech recognition in multi-talker situations at fixed SNR. Trends in Hearing, 26, 1–12.

Hester, R. L., Kinsella, G. J., Ong, B. (2004). Effect of age on forward and backward span tasks. Journal of the International Neuropsychological Society, 10(4), 475-81.

Kestens, K., Degeest, S., Miatton, M., & Keppler, H. (2021). An auditory Stroop test to implement in cognitive hearing sciences: development and normative data. International Journal of Psychological Research 14(2), 37-51.

Knight, S., & Heinrich, A. (2017). Different measures of auditory and visual stroop interference and their relationship to speech intelligibility in noise. Frontiers in Psychology, 8(MAR), 1–21.

Mamo, S. K. & Helfer, K., S. (2021). Speech understanding in modulated noise and speech maskers as a function of cognitive status in older adults. American Journal of Audiology 30(3), 642-654.

Morgan, A. L. & Brandt, J. F. (1989). An auditory Stroop effect for pitch, loudness, and time. Brain and Language, 36(4), 592-603.

Sanders, M. E., Kant, E., Smit, A. L., & Stegeman, I. (2021). The effect of hearing aids on cognitive function: A systematic review. PLoS One, 16(12), e0261207.

Tripathi, R., Kumar, K. J., Bharath, S., Palaniappan, M., Rawat, V. S., & Varghese, M. (2019). Indian older adults and the digit span A preliminary report. Dementia & Neuropsychologia 13(1), 111-115.

Windle, R., Dillon, H., & Heinrich, H. (2023). A review of auditory processing and cognitive change during normal ageing, and the implications for setting hearing aids for older adults. Frontiers in Neurology, 14, 1122420.

Yang, Z., Ni, J., Teng, Y., Su, M., Wei, M., Li, T., Fan, D., Lu, T., Xie, H., Zhang, W., Shi, J., & Tian, J. (2022). Effect of hearing aids on cognitive functions in middle-aged and older adults with hearing loss: A systematic review and meta-analysis. Frontiers in Aging Neuroscience, 14, 1017882.

Authors and investigators

Hsuan Yun Huang, **Senior Clinical Audiology Engineer**

Huang Hsuan Yun completed her Master's in Audiology from the Department of Speech-Language Pathology and Audiology at the National Taipei University of Nursing and Health Sciences in 2019. Since then, she has been working as a Clinical Audiology Engineer at the Sonova Audiology Research Center in Shanghai.

Wanting Huang, PhD, Senior Hearing Researcher

Wanting Huang has been conducting research on speech and auditory processing from a cognitive neuroscience perspective. She earned her Ph.D. in Speech and Hearing Sciences from the University of Hong Kong in 2021. From 2021 to 2023, she engaged in postdoctoral research at the Southern University of Science and Technology. Since 2024, she has been working as a Senior Hearing Researcher at Sonova Audiology Research Center in Shanghai

Xin Tian , Clinical Audiology Manager

Xin Tian joined the Sonova Audiology Research Center in Shanghai as a Clinical Audiology Engineer in 2015. With a background in audiology, he has worked on various research and development projects related to hearing aid technologies, collaborating closely with both HQ and local teams. His expertise spans product testing, clinical evaluation, and optimizing user experience.

Jingjing Guan, PhD., Senior Director ARC Shanghai

Jingjing holds a Ph.D. degree from University of Texas at Austin and worked in Texas Tech University Health Sciences Center as an assistant professor and a licensed audiologist. Jingjing started her journey in Sonova Audiology Research Center in Shanghai in 2018. Her main research interests include psychoacoustics and hearing aid outcome measures.

Volker Kühnel, PhD., Principal Expert Hearing Performance

Volker Kühnel, PhD, holds a doctorate in physics and completed his studies in 1995. From 1995 to 1997 he worked in Oldenburg as a research assistant in the Medical Physics group of Prof. B. Kollmeier. Since 1998, he has been working at Phonak/Sonova in product development on the audiological design at the interface between hearing aid algorithms and fitting software. His work focuses on the audiological quality of hearing systems to achieve maximum customer benefit.

Jennifer Appleton-Huber, Scientific Audiology and Education Manager

Jennifer Appleton-Huber received her M.Sc. in Audiology from the University of Manchester in 2004. Until 2013, she worked as an Audiological Scientist mainly in the UK and Switzerland, where she worked with adults and pediatrics, in the areas of hearing aids and cochlear implants. She has worked at Phonak Headquarters in Switzerland since 2013.