

Phonak Field Study News.

Spheric Speech Clarity improves speech understanding and reduces both listening effort and fatigue

In this study, conducted at Hörzentrum Oldenburg in Germany, 22 hearing-aid users with moderate to severe hearing loss tested Spheric Speech Clarity (SSC) versus Speech in Loud Noise (SILN) including StereoZoom (SZ) via a variety of tests over a time-compressed auditory day in the lab. Results showed that in multi-talker listening situations in noise, SSC provided significantly better speech understanding as well as reduced listening effort and fatigue at the end of a time-compressed auditory day.

Latzel, M., Heeren, J., Lesimple, C., Gökalan, M., Puhlemann, L., Haf, C. October, 2025

Introduction

Speech understanding in noise is one of the most challenging situations for hearing-impaired people. Over the past 30 years, directional microphones (beamformers) and other sound cleaning technologies have been implemented in hearing aids to improve speech intelligibility or hearing comfort in noisy situations. The latest generation of Phonak hearing aids, Infinio, provides Spheric Speech Clarity (SSC), a

novel noise-reduction algorithm based on a deep-neural-network (DNN) approach. DNN-noise-reduction systems overcome the limitations of directional microphones by providing SNR improvements for speech from any direction (Haseman & Krylova, 2024). Hearing aids have been slow to adopt large-scale DNN systems due to a number of factors, not least of which are the limited capabilities of current hearing-aid processors. Early investigations indicated a

benefit of the denoising DNN in sound-quality ratings and speech intelligibility.

The premium representative of this platform is the Audéo™ I-Sphere. At the time of launch and at the time of release of this paper, it is the first and only hearing device to combine Phonak's new sound-processing chip ERA and the new DNN chip called DEEPSONIC. It is the new DNN chip that supports the new SSC feature, activated in speech-in-loud-noise environments. This SSC feature has been trained to amplify speech and to suppress noise, increasing the Signal-to-Noise Ratio (SNR), and thus improving speech understanding in noisy listening situations and satisfaction in complex listening situations (e.g. group conversations). SSC offers distinct advantages compared to a noise-reduction system based on binaural directional-microphone technology (StereoZoom), such as:

- SNR is improved not only for speech from front but also for speech from the side
- SNR is improved also if the noise and the target are spatially not separated
- SNR is improved instantly (i.e., no time delay for speech detection)

The motivation for developing this new technology was to enable the hearing aid user to understand speech better, even in complex listening situations when the target speaker is not always in front of the hearing aid user.

Some studies have shown this new algorithm to provide speech-understanding and listening-effort benefits for people with hearing impairment. These investigations took place in the lab where test conditions included speech from multiple directions (Wright et al., 2024) and speech with competing talkers in a complex communication situation (Tian et al., 2025), both in diffuse loud noise. Another study was conducted in real life, where hearing-impaired people were invited to test the new algorithm under real life conditions (Miller et al., 2024). All studies showed that SSC provides better speech intelligibility, less listening effort and was clearly preferred in comparison to the competitors in objective and subjective tests.

The latest study (Tian et al., 2025) demonstrated that SSC is able to improve speech intelligibility in a typical communication situation when one is listening to a communication partner in the presence of two interfering talkers. The current study was conducted to substantiate those outcomes in a confirmatory study at an external site. In addition, the study collected data on cognitive load, listening effort and fatigue when the SSC algorithm was used.

The objectives of this study were to evaluate whether, relative to SiLN using StereoZoom 2.0 (SZ), SSC would:

- Provide better spontaneous speech intelligibility for turn-taking talkers in complex noise
- Provide better spontaneous speech-reception thresholds in a complex listening situation with target talkers from different directions in the frontal hemisphere
- Reduce cognitive load in complex noise situations with competing talkers, leading to less listening fatigue at the end of a time-compressed auditory day
- Reduce listening effort in a complex noise situation with competing talkers

Methodology

Participants

22 experienced hearing aid users with moderate to severe, symmetrical hearing loss (figure 1) took part in the study (12 female, 9 male). Their ages ranged from 24–84 years with a mean of 70 years.

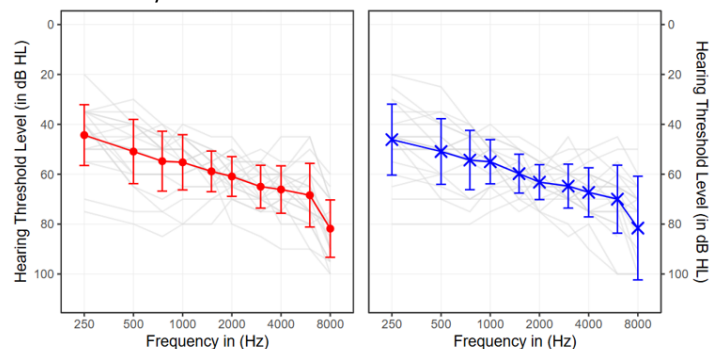


Figure 1. Pure tone hearing thresholds of study participants. Light gray lines are the individual thresholds. Red and blue lines are the mean hearing thresholds of the participants for the right and left ears respectively (+/- 1 standard deviation).

Devices

Participants were fit with Audéo I90-Sphere hearing aids equipped with SlimTips without vents. The fitting formula used was Adaptive Phonak Digital (APD) 3.0, with the gain level set to 100% and frequency lowering disabled. Hearing aids were fit with two programs:

1. Spheric Speech in Loud Noise (in the following called SSC). SSC set to 5, microphone set to fixed directional (12).
2. Speech in Loud Noise (in the following called SiLN). Dynamic Noise Cancellation and Noise Block set to 8. StereoZoom set to 24.

Procedures

The methodological approach made use of the time-compressed-auditory day (TCAD, Blümer et al., 2024), which is a sequence of lab tests that concentrates listening challenges that occur during a typical day into a 2.5 hour test session. It aims for a stimulation of listening fatigue with a high degree of ecological validity, while having controlled test conditions in a lab.

Measurement of speech intelligibility

Two tests of speech intelligibility were conducted. To account for complex situations with turn taking, the Concurrent Oldenburger Sentence Test (CCOLSA) (Heeren et al., 2022) was used. This test is based on the standard OLSA, but includes three different talkers speaking simultaneously. One female and two male speech corpora were presented against a diffuse dinner-party-noise recording, which was presented at 75 dB SPL via 16 loudspeakers using ambisonics. The speech was presented from the front and from minus/plus 67.5 ° (figure 2).

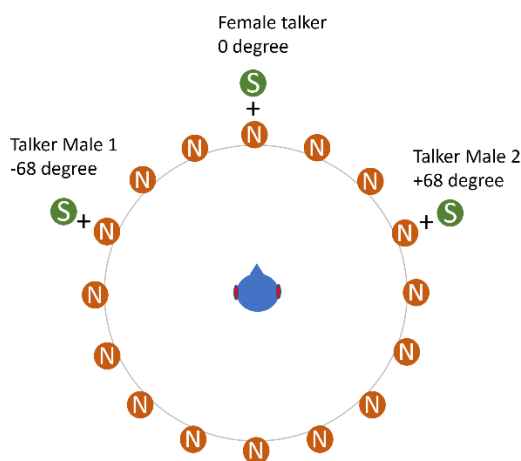


Figure 2. Speaker setup of the CCOLSA test with the participant in the center of a circle of 16 loudspeakers creating a diffuse noise environment. Speech was presented from speakers from the front and from minus/plus 67.5 °.

The presentation level of the speech from the three different angles was equal to the participant's Speech Reception Threshold (SRT) plus 5dB. SRTs were individually determined for each participant using a standard OLSA procedure. In CCOLSA, sentences are presented from the three talkers in a random order, but with a defined overlap time between sentences of different talkers. This overlap time was set individually to a value where a participant reached 50% correct responses in the SSC condition (determined using the adaptive method described in section "Assessment of listening fatigue"). Participants were instructed to listen for sentences which started with the word 'Kerstin'. When they heard these target sentences, they needed to turn their attention towards that speaker and repeat the last words of all sentences from that talker until one of the other talkers

said the word 'Kerstin', at which point they would turn the attention to the new talker. Thus, subjects are performing a dual task including target word identification and speech recognition. There were 60 test sentences, with 20 items per speaker. The CCOLSA was measured twice for both test conditions: SSC and SiLN. The percentage of correctly repeated words resulted in a speech recognition score.

The spatial OLSA was included as a measure of spontaneous speech from random directions. It is a version of the OLSA (Wagener et al., 1999) where each of the five words of the matrix sentences are randomly assigned to the loudspeakers at 0, 45, 90, 270, and 315 degrees azimuth (figure 3).

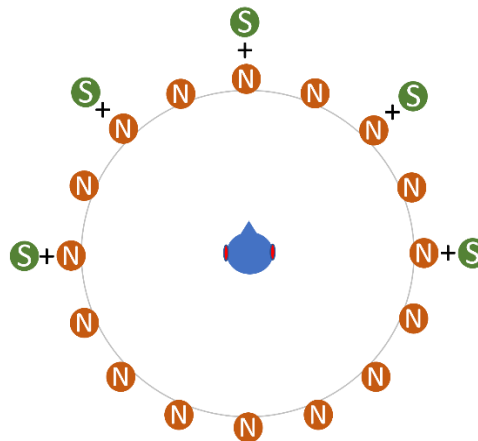


Figure 3. Setup of the spatial OLSA test with the participant in the center of a circle of 16 loudspeakers creating a diffuse noise environment. Speech was presented from speakers 0, 45, 90, 270, and 315°.

Results are recorded in SRTs. The SRT in noise is defined as the signal-to-noise ratio needed to understand 50% of the presented words. Each measurement consisted of 20 sentences, and measurements were made twice for each condition during the TCAD (SiLN and SSC).

Assessment of listening fatigue

To assess listening fatigue, another version of CCOLSA with adaptive overlap of the sentences was used (Heeren et al., 2023). The speaker setup was the same as that in figure 2 with participants performing the same task. Again, there were 60 test sentences, with 20 items per speaker. The test was measured twice with both tested conditions during the TCAD: SSC and SiLN. The difference to the classic CCOLSA (as described above) was that the overlap time between the sentences adapted throughout the test. The overlap adapted until the response accuracy for the participant was 50% correct. The resultant score was the overlap time. The more overlap of sentences, the more cognitively difficult the task. Therefore, higher overlap time meant a better cognitive performance. The adaptive CCOLSA was performed at the beginning and then towards the end of the TCAD when the participant was expected to be fatigued. This could give an

indication of whether either condition SiLN or SSC results in a different amount of listening fatigue.

Assessment of listening effort

The ACALES Test (Krüger et al., 2017) was used to assess subjective listening effort. Based on the OLSA speech corpus, participants rated the perceived listening effort on scale from 1 (no effort) to 13 (extremely strenuous) on a touchscreen. SNR was adaptively adjusted so that ratings for all categorical units of the scale (1 to 13) were collected. Speech was presented from the side against the diffuse dinner-party noise at 75 dB SPL (figure 4). The ACALES test was performed once for each condition: SiLN and SSC.

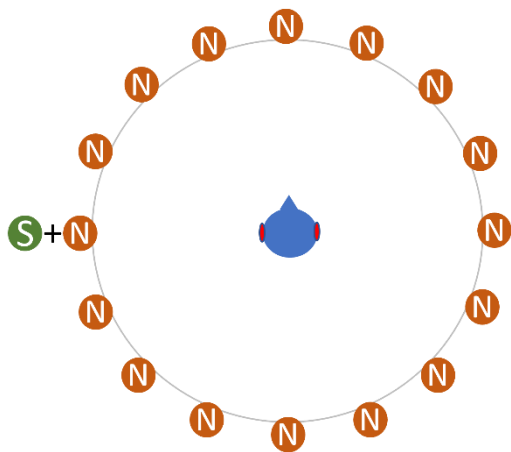


Figure 4. Setup of the ACALES test with the participant in the center of a circle of 16 loudspeakers creating a diffuse noise environment. Speech was presented from the 90° speaker on the better ear side (left or right).

Results

Speech intelligibility

CCOLSA scores represent the percentage of target words correctly repeated. The within-subject effect of condition (SSC versus SiLN) for test and re-test can be seen in figure 5. For both test and re-test, accuracy was better with SSC than with SiLN. SSC was found to provide 19.9% points better speech recognition for speech in noise ($p < 0.001$, $d = 1.6$) than SiLN. This difference corresponds to a relative performance change of 50 % with SSC taking the SiLN condition as reference.

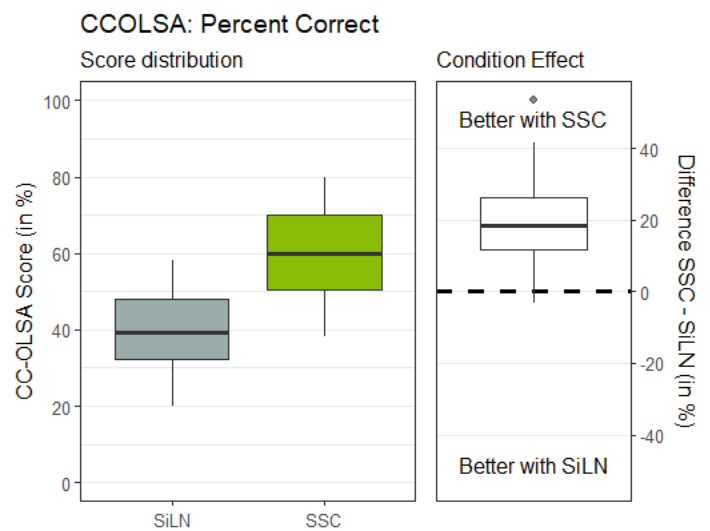


Figure 5. Data distribution and within subject effect of condition for CCOLSA scores

Spatial OLSA provides an SRT where lower values stand for better outcomes (same performance in more difficult conditions). Figure 6 shows the within-subject effect of condition for test and re-test of spatial OLSA. SSC was found to provide significantly lower SRTs (-3.2 dB) for speech in noise ($p < 0.001$, $d = 1.8$). This difference corresponds to a relative performance change of 44% with SSC based on the range of SRTs for both conditions.

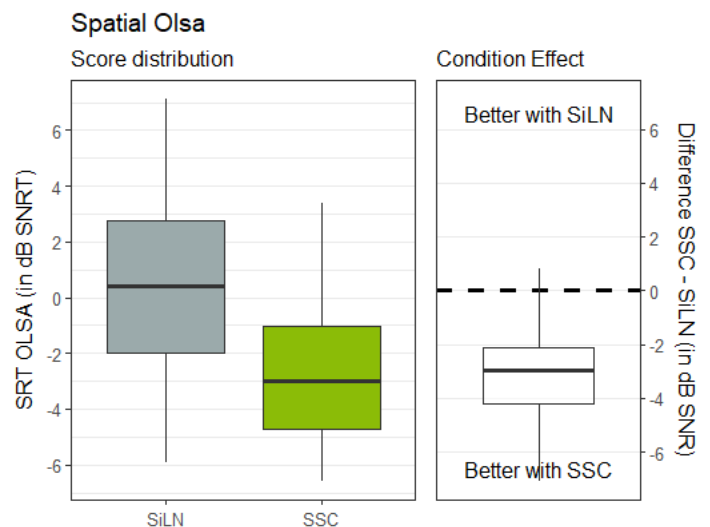


Figure 6. Data distribution and within subject effect of condition for spatial OLSA

Listening fatigue

Cognitive load was assessed as the overlap time between talkers during the adaptive CCOLSA. A higher overlap time, indicating decreased cognitive load (same performance in a more difficult situation). The within-subject effect of condition (SSC versus SiLN) can be seen in figure 7. With SSC, the overlap of target speech and interfering talkers is significantly longer than when SiLN is activated (461.6 ms, $p < 0.001$, $d = 1.2$). This corresponds to 31% percent more

words per minute, indicating that cognitive load is lower when using SSC.

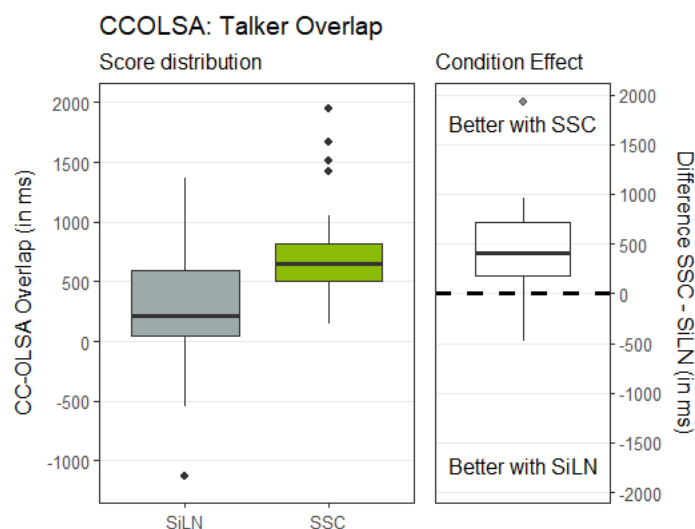


Figure 7. Data distribution and within subject effect of condition for adaptive CCOLSA

Figure 8 shows the effect of the repetition of adaptive CCOLSA further along the TCAD. When repeating the test, further along the TCAD, the overlap time shows a significant training effect (increase) for the SSC condition (200 ms, $p < 0.05$, $d = 0.3$), which is typical when measuring with OLSA speech (Schlueter et al, 2016). This increase in overlap time corresponds to 14% more words per minute. For the SiLN condition, the expected training effect was not observed, as listening fatigue counteracts the training.

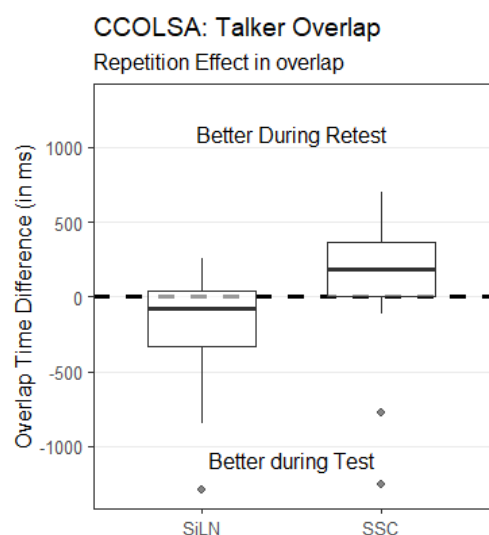


Figure 8. The repetition effect in overlap times of adaptive CCOLSA for the conditions SiLN and SSC.

Listening effort

ACALES is an adaptive test that assesses the SNR needed to achieve a predefined listening effort. Lower values stand for better outcomes, as the same subjective judgement is given

for a more difficult condition. The distribution of the ACALES scores is shown in figure 9.

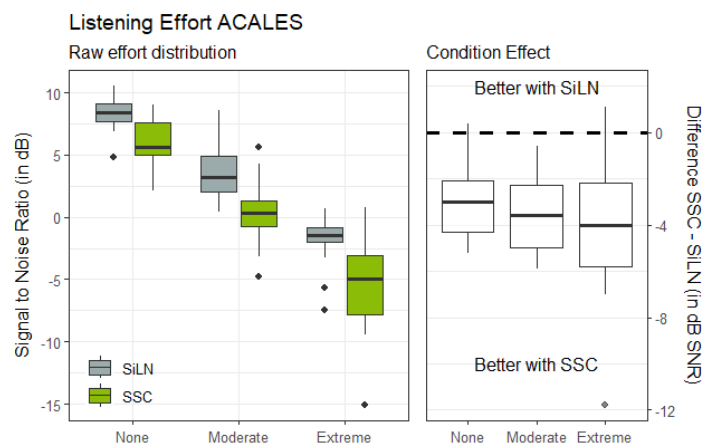


Figure 9. Distribution of individual SNRs (in dB) associated to a given listening effort degree (x-axis) from the ACALES test.

Results show that listening effort increases with decrease in SNR. With SSC, SNRs were 3.4 dB lower than for SiLN in order to achieve the same listening effort rating. This indicates that in a challenging listening situation, use of SSC results in significantly less listening effort in comparison to use of SiLN ($p < 0.001$, $d = 1.6$). This difference corresponds to a relative performance change of 37 % with SSC based on the range of SNR for both conditions.

Conclusion

Spheric Speech Clarity was introduced with the Phonak Infinio platform in August 2024. This study investigated benefits of SSC in complex listening conditions for speech recognition, listening effort, cognitive load, and listening fatigue. For a simulated three-talker group conversation in noise, it was found that SSC improved speech recognition over SiLN. In addition, participants were able to process more words per minute in this listening situation when using SSC than with SiLN. Towards the end of a time-compressed auditory day, participants were able to significantly improve their performance with SSC, whereas the performance stayed constant using SiLN. This suggests that SSC is able to reduce the effect of listening fatigue compared to SiLN. Lastly, use of SSC was also found to result in less listening effort than use of SiLN. These results indicate that when using Sphere hearing aids with SSC in complex noisy environments, users can expect to understand more speech, require less effort to listen and be less fatigued at the end of the day.

References

Blümer, M., Heeren, J., Mirkovic, B., Latzel, M., Gordon, C., Crowhen, D., Meis, M., Wagener, K., & Schulte, M. (2024). The impact of hearing aids on listening effort and listening-related fatigue-investigations in a virtual realistic listening environment. *Trends in Hearing*, 28, doi: 10.1177/23312165241265199

Hasemann, H., & Krylova, A. (2024). Artificial intelligence in hearing aid technology. Phonak Insight, available at <https://www.phonak.com/en-int/professionals/audiology-hub/evidence-library>

Heeren, J., Hohmann, V., Schulte, M., & Wagener, K.C. (2023). Adaptive CCOLSA: Cognitive overload thresholds in a multi-talker speech test. In Forum Acusticum.

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, doi: 10.1177/23312165221108257

Krueger, M., Schulte, M., Brand, T., & Holube, I. (2017). Development of an adaptive scaling method for subjective listening effort. *The Journal of the Acoustical Society of America*, 141(6), 4680-4693.

Miller, A., Wright, A., Zhu, X., Kuehnel, V., Latzel, M., & Seitz-Paquette, K., (2024). Phonak Audéo Sphere™ Infinio is preferred by clients during real-world use. Phonak Field Study News available at <https://www.phonak.com/en-int/professionals/audiology-hub/evidence-library>

Schlueter, A., Lemke, U., Kollmeier, B., & Holube, I. (2016). Normal and time-compressed speech: How does learning affect speech recognition thresholds in noise? *Trends in Hearing*, 20, 2331216516669889

Tian, X., Guan, J., Latzel, M., & Kuehnel, V. (2025). Spheric Speech Clarity significantly improves speech understanding in a multi-talker scenario and reduces spatial listening effort. Phonak Field Study News, available at <https://www.phonak.com/en-int/professionals/audiology-hub/evidence-library>

Wagener, K., Brand, T., & Kollmeier, B. (1999). Entwicklung und evaluation eines satztests für die deutsche sprache I-III: design, optimierung und evaluation des oldenburger satztests. *Zeitschrift für Audiologie*, 38(1-3), 4-15.

Wright, A., Kuehnel, V., Keller, M., Seitz-Paquette, K., & Latzel, M. (2024). Spheric Speech Clarity applies DNN signal processing to significantly improve speech understanding from any direction and reduce the listening effort. Phonak Field Study News available at <https://www.phonak.com/en-int/professionals/audiology-hub/evidence-library>

Authors and investigators

External investigators



Jan Heeren studied Physics at the University of Oldenburg, Germany, and graduated in the Medical Physics group in 2014. From 2012, he worked on several projects in the field of hearing aid evaluation and virtual acoustics at the university and at

Hörzentrum Oldenburg gGmbH, where he is employed as a project manager since 2016. Apart from his scientific activities, he has conducted more than 500 events as a freelancing audio engineer since 2008.



Müge Gökalan has been working as a medical-technical assistant at the Hörzentrum Oldenburg since 2000, focusing on audiological hearing system evaluation, special audiological diagnostics, cross-project organization and subject

acquisition.



Lisa Puhlemann has been working as a hearing-care professional since 2012. In 2022, she started as an investigator for audiological studies at the Hörzentrum Oldenburg.



Chiara Haf started as a hearing-care professional in 2016 and holds a master craftsman's certificate since 2023. In 2024, she joined the audiological-studies team at the Hörzentrum Oldenburg as an investigator.

Internal investigators



Dr. Matthias Latzel studied electrical engineering in Bochum and Vienna in 1995. After completing his Ph.D. in 2001, he carried out his Postdoc from 2002 to 2004 in the Department of Audiology at Giessen University. He was the head of the Audiology department at Phonak Germany from 2011. From 2012 to 2022 he has been working as the Clinical Research Manager for Phonak AG, Switzerland. Since 2022 he is senior expert clinical studies for Sonova AG.



Christophe Lesimple studied music in Stuttgart, audiology in Lyon, and statistics in Paris and Bern. He is working as a research audiologist and contributes to various aspects of development including concepts, supporting clinical trials, and analyzing data. Besides his activities with Sonova, he teaches audio analytics for machine learning at the University of Applied Science in Bern, hearing aid verification at the Akademie Hören Schweiz, and volunteers for a hearing-impaired association.

Phonak Field Study News.

One-page summary

Spheric Speech Clarity improves speech understanding and reduces both listening effort and fatigue

In multi-talker listening situations in noise, Spheric Speech Clarity (SCC) provided significantly better speech understanding as well as reduced listening effort and fatigue at the end of a time-compressed auditory day, in comparison to Speech in Loud Noise (SiLN).

Latzel, M., Heeren, J., Lesimple, C., Gökalan, M., Puhlemann, L., Haf, C. October, 2025

Key highlights

- In a diffuse noise environment with multi-talkers, SCC was found to provide 19.9% better speech recognition for speech in noise than SiLN with SZ. This difference corresponds to a relative performance change of 50% with SCC taking the SiLN condition as reference.
- When talkers partially overlapped, participants showed a higher cognitive performance with SCC than with Speech in Loud Noise (SiLN). This improvement was even more noticeable at the end of a time compressed auditory day, indicating that SCC prevents listening fatigue.
- Subjective listening effort was found to be lower with SCC than with SiLN with SZ.

Considerations for practice

- SCC reacts without noticeable delay and thus enables clients to understand spontaneous speech from lateral directions.
- In group conversations, neither turn taking nor interruptions will be missed.