Phonak Field Study News.

Spheric Speech Clarity proven to outperform key competitors for clear speech in noise

Phonak introduces a paradigm shift to address the No. 1 client need – speech understanding in noise – with Phonak Audéo Sphere™ Infinio. The world's first hearing aid with a dedicated artificial intelligence (AI) chip that prioritizes speech over noise from any direction, in real-time. The Phonak proprietary chip architecture, DEEPSONIC™, has 53 times more processing power than current industry chip technology.* This paper outlines the results of technical measurements which prove that DEEPSONIC delivers a level of clean and clear speech far beyond current existing hearing technology.

Raufer, S., Kohlhauer, P., Uhlemayr, F., Kühnel, V., Preuss, M. & Hobi, S. August 2024.

Key highlights

- Instant all-around speech understanding in noise with Spheric Speech Clarity - with an unprecedented SNR improvement of up to 10 dB.
- Spheric Speech Clarity expands the access to speech from any direction.
- Spheric Speech Clarity and StereoZoom 2.0, respectively, outperform three competitor devices.

Considerations for practice

- Understanding speech in challenging complex environments remains a key communication need.
- Spheric Speech Clarity delivers significantly improved speech understanding from any direction and reduced listening effort. (Wright et al, 2024).
- Up to 3 times more likely to understand speech in challenging environments compared to two leading competitors (Wright et al., 2024).
- Both Spheric Speech Clarity and StereoZoom 2.0 are dedicated features known to improve the SNR and are instrumental in building success for clients.

* At time of launch

Introduction

Over the past 75 years, Phonak has been committed to address the No. 1 consumer need: speech understanding in noise. Speech understanding in noise is one of the most challenging listening situations for people with hearing loss and one of the most important predictors for hearing aid satisfaction (Abrams & Kihm, 2015). Hearing aid wearers require a better signal-to-noise ratio (SNR) compared to normal hearing listeners for the same speech intelligibility performance (Killion, 1997).

Directional microphone technology

An established method to improve speech understanding in noise is the use of directional microphone technology, also known as beamforming. Here, two microphones work together to increase the sensitivity to the front direction and reduce the sensitivity to all other directions, thus forming a virtual "beam" (Derleth et al., 2021).

The use of directional microphones results in SNR benefits to the front direction (Ricketts, 2005). The improved SNR translates to an improved speech intelligibility in noisy situations, where the target speech is located in the frontal hemisphere, and the noise is presented from the side and back (Mueller, 2000; Bentler, 2005). Phonak hearing aids have been using fixed directional microphone and UltraZoom technology in speech-in-noise programs for over two decades.

StereoZoom, introduced in 2012 with Phonak Quest products, is a binaural beamforming system which combines the signals from four microphones (two on the left and two on the right hearing aid) via a wireless link, resulting in a narrower beam pattern with an even greater SNR improvement to the front. A number of studies have demonstrated better speech intelligibility with StereoZoom compared to other directional microphone technologies in Phonak devices and competitor devices (Appleton & König, 2014, Latzel & Appleton-Huber, 2015).

StereoZoom 2.0 (SZ 2.0), introduced in 2022 with Phonak Lumity, was designed to balance spatial awareness at lower noise levels with speech focus at higher noise levels thanks to a noise level dependent transition from UltraZoom into StereoZoom 2.0. The strength of SZ 2.0 is adapted to the environment based on the noise level (increased focus with increased noise level). Compared to UltraZoom and StereoZoom, SZ 2.0 provides a better SNR in challenging listening environments. Phonak Audéo™ R Infinio hearing aids will continue to use StereoZoom 2.0 in speech-in-loud-noise programs.

DNN based speech separation from noise

Phonak Audéo Sphere™ Infinio is the world's first hearing aid with a dedicated artificial intelligence (AI) chip, DEEPSONIC™, which introduces a paradigm shift in hearing technology. The proprietary chip hosts a deep neural network (DNN) algorithm that significantly suppresses the background noise from any direction, leading to remarkable speech intelligibility improvements for individuals using hearing aids (Diehl et al., 2023). The technical capabilities to unleash the full potential of deep learning-based algorithms in hearing aid technology has arrived with DEEPSONIC.

The DEEPSONIC chip architecture is 53 times more powerful than current industry standards and dedicated to deliver unprecedented speech clarity in noise. DEEPSONIC enables Spheric Speech Clarity, the new DNN-based denoising feature from Phonak that revolutionizes signal processing in the most challenging acoustic environments to tackle the infamous cocktail party problem.

AutoSense OS 6.0, in Phonak Audéo Sphere Infinio hearing aids, is able to identify challenging acoustic environments to automatically transition into the Spheric Speech in Loud Noise program. Spheric Speech in Loud Noise is the unique combination of directional microphone technology and Spheric Speech Clarity.

This Field Study News describes a systematic technical study, examining how Phonak Audéo Sphere Infinio with Spheric Speech Clarity and Phonak Audéo R Infinio with StereoZoom 2.0 perform in challenging listening environments in comparison to key competitor products.

Results of technical measurements

Figure 1 presents the progression of Phonak technologies and shows the signal-to-noise ratio (SNR) benefit compared to an omnidirectional microphone mode. The higher the SNR, the better the hearing aid is at separating speech from noise, resulting in clearer speech for the client.

With a fixed directional microphone setting, 2.9 dB SNR benefit can be achieved in realistic listening situations for a speaker from the front. With the introduction of noise-cancelling algorithms, like NoiseBlock and Dynamic Noise Cancellation (DNC), the SNR benefit could be improved from 2.9 dB to 5.1 dB.

StereoZoom 2.0, in combination with NoiseBlock and DNC, offers an additional SNR improvement, totaling 6.4 dB.

This underlines that the well-known state-of-the-art directional microphone and noise reduction technologies from Phonak are instrumental in building success for clients, who will use Phonak Infinio hearing aids.

Spheric Speech Clarity goes beyond existing hearing technology. Spheric Speech Clarity, in combination with a fixed directional microphone, offers an unprecedented SNR improvement of 9 dB at default strength and 10.2 dB at maximum strength. The strength of Spheric Speech Clarity can be adjusted by the hearing care professional in Phonak Target and by the client using myPhonak.

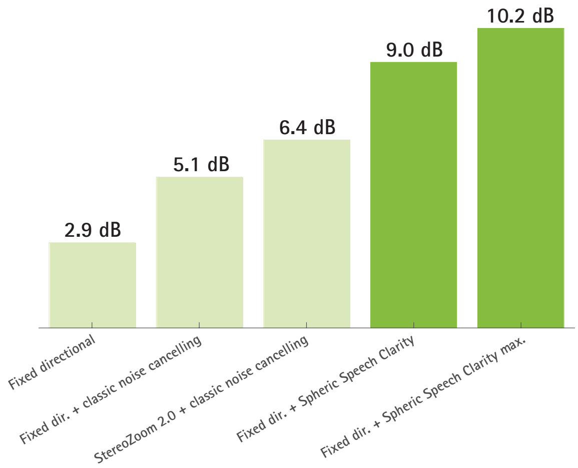


Figure 1. Progression of Phonak technologies: Speech intelligibility index (SII) weighted signal-to-noise ratio (SNR) improvements with respect to the omnidirectional microphone mode for Phonak technologies averaged across three realistic scenes, including a café scene, dinner party scene and food court scene. Results are presented for a fully occluded coupling and for the better ear. The speech signal was always presented from 0° azimuth.

Spheric Speech Clarity is not only Phonak's best performing technology to date, but it also outperforms three key competitors, allowing the client to focus on what's relevant while having a conversation in a challenging listening environment.

Industry leader in delivering clear speech

Figure 2 compares the SNR performance of the Spheric Speech in Loud Noise program in the Phonak Audéo I90-Sphere with the Speech in Loud Noise program (with StereoZoom 2.0) in the Phonak Audéo I90-R and to three competitor hearing aids (premium level), with and without AI technology (available on the market as of March 26, 2024. See methods for details of fitting). Here, the SNR improvement is shown relative to an unaided condition.

Phonak Audéo I90-Sphere provides a staggering SNR improvement of 5.9 dB in realistic noise scenes, followed by Phonak Audéo I90-R with an SNR improvement of 4.8 dB. Spheric Speech Clarity and StereoZoom 2.0, respectively, outperform the three competitor devices, who show only SNR improvements between 2.2 and 3.3dB.

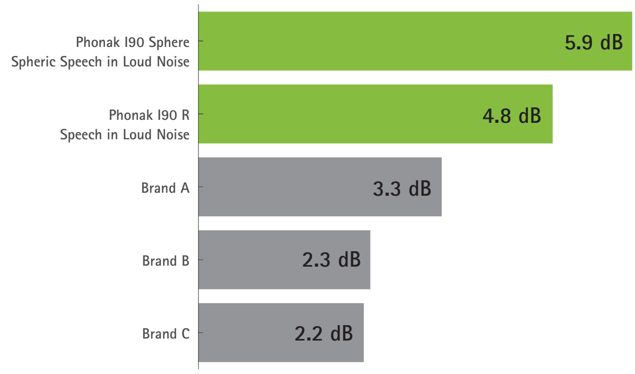


Figure 2. Competitor comparison: SII-weighted signal-to-noise ratio (SNR) improvements with respect to unaided, averaged across three realistic scenes. Results are presented for a fully occluded coupling and for the better ear. The speech signal was always presented from 0° azimuth. All devices were tested in manual programs at their respective default settings.

Spheric Speech Clarity expands the access to speech from any direction

Humans rely on their hearing to know where to look and how to position their body in the surrounding environment, which helps to form a mental representation of the auditory world (Derleth et al., 2021). This enables a focus on conversations, but also an awareness of indirect speech and sounds in the surrounding environment. While our eyes are focused to the front, our ears pick up important information from all directions.

Listeners with hearing loss may spend more effort on maintaining awareness of their surroundings than listeners with normal hearing. Similar to compensating for reduced speech intelligibility by "filling in the gaps," extra effort spent on auditory tasks such as environmental awareness may compromise the availability of cognitive resources for other purposes (Edwards, 2016). Spheric Speech Clarity was investigated in a clinical study conducted at Phonak Audiological Research Center (PARC) looking at speech understanding and listening effort. Results from this study shows that the SNR benefit measured in this technical study translates into a perceptual hearing benefit in subjects with moderate to moderately-severe hearing losses.

Study results from a complex speech intelligibility task, with stimuli drawn from the Coordinate Response Measure (CRM) Task corpus, indicated that participants were two times more likely to understand every word from any direction (angle tested 60, 120, 240 and 300 degrees) compared to feature off (Wright et al, 2024). When compared to two leading competitor devices, Spheric Speech Clarity was able to provide 2 to 3 times more chance of speech understanding (Wright et al, 2024).

Significant reduction of listening effort with Spheric Speech Clarity (compared to feature off) was also shown in this study using the Adaptive Categorical Listening Effort Scaling (ACALES) test (Wright et al, 2024).

Figure 3 presents SNR improvements for a speech signal from different incident angles, ranging from 0° to 180°. Here the SNR improvements are attributed to Spheric Speech Clarity alone, without the effect of directional microphone filtering. The graph shows that the performance of Spheric Speech Clarity does not depend on the incident angle of the speaker and allows access to speech from any direction. The SNR improvements range from 5.8 dB to 6.9 dB and are within 1 dB of the average all-around performance.

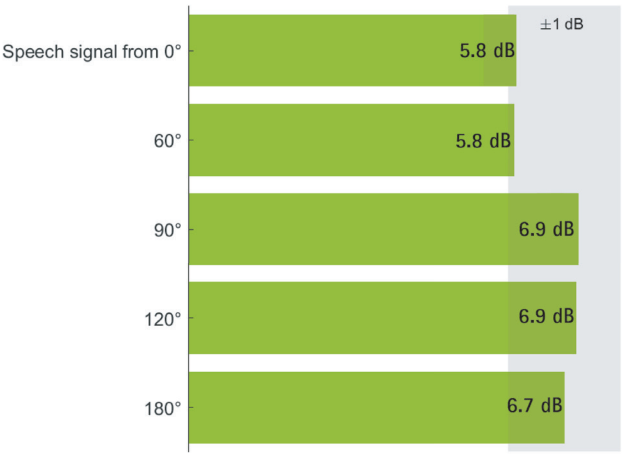


Figure 3. Spheric Speech Clarity benefit from different speech incident angles: SII-weighted signal-to-noise ratio (SNR) improvements across three realistic scenes for a speech signal presented from 0°, 60°, 90°, 120° and 180° azimuth. Results are presented for a fully occluded coupling and for the better ear. For 0° and 60°, the reference condition for Spheric Speech Clarity deactivated versus activated at maximum strength was a fixed directional microphone mode. For 90° to 180°, the reference condition for Spheric Speech Clarity deactivated versus activated at maximum strength was an omnidirectional microphone mode. The gray box shows the +-1 dB range from the average performance across all directions.

Conclusions

With Phonak Audéo Sphere Infinio, Phonak introduces a deep-learning based algorithm to address the No. 1 client need: speech understanding in noise. Phonak Audéo Sphere Infinio is the world's first hearing aid with a dedicated AI chip that instantly separates speech from noise. With 53 times more processing power compared to current industry chip technology at time of launch, the Phonak proprietary chip architecture, DEEPSONIC™ enables Spheric Speech Clarity. This groundbreaking technology delivers a level of clean and clear speech far beyond any existing hearing technology. As a result of the clean and clear speech, clinical objective testing indicates not only improved speech understanding and reduced listening effort versus feature off, but also better performance compared to two leading competitor devices (Wright et al, 2024). Spheric Speech Clarity and StereoZoom 2.0, respectively, are dedicated features known to improve the SNR and outperform three key competitors. Phonak Audéo Sphere Infinio and Phonak Audéo R Infinio are the latest premium hearing aids from Phonak that will build success for your client.

Methods

Measurement Setup

All measurements took place in an acoustically treated room with a reverberation (RT60) time of 0.15 seconds. 12 Genelec 8020D loudspeakers (Genelec, Finland), were arranged in a circle with a 1.4 m radius, leading to a 30° distance between two loudspeakers. A KEMAR mannequin (GRAS Sound & Vibration, Denmark) with anthropometric ears was placed at the center of the loudspeakers and oriented towards 0°. The center of the ear canal of the KEMAR mannequin was at the same height as the loudspeakers. An RME M-16 DA converter (Audio AG, Germany) was used to send the audio to the active loudspeakers; an RME Fireface 802 USB soundcard (Audio AG, Germany) was used to record the audio from the KEMAR ear-canal microphones.

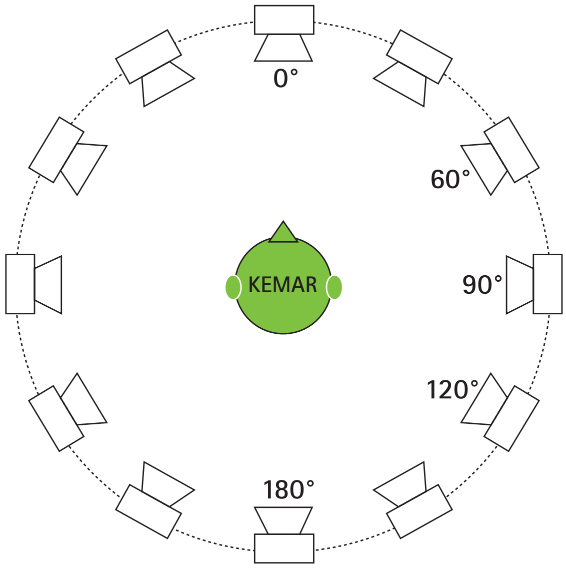


Figure 4. Top-down view of the setup: 12 Genelec 8020D loudspeakers were arranged in a circle with a 1.4m radius. A KEMAR mannequin with anthropometric ears was placed at the center of the loudspeakers and oriented towards 0°. The center of the ear canal of the KEMAR mannequin was at the same height as the loudspeakers.

Background noise and target speech

Three realistic sound scenes from the ARTE database were used as background noise (Weisser et al. 2019). The scenes were presented at the recorded SPLs. The café scene was used at an SPL of 71.7 dB SPL, the dinner party scene at 72.8 dB SPL, and the food court scene at 78.2 dB SPL.

The background noise was mixed with the ISTS speech signal (Holube et al. 2010) to achieve free-field SNRs of 0 dB SNR for the café scene, -3 dB SNR for the dinner party scene, and -6 dB SNR for the food court scene.

The ARTE scenes are Higher-order Ambisonics (HoA) recordings. The ARTE recordings were decoded and played back in the horizontal plane through 12 loudspeakers at a distance of 1.4m. To overcome drawbacks like spatial aliasing of standard HoA decoding, the Coding and Multi-Parameterization of Ambisonic Sound Scenes (COMPASS) method was used for decoding of the continuous background signals (Politis et al. 2018).

The ISTS signal was convolved with the impulse response of the respective scene. The room impulse responses (RIR) were decoded using the Higher-order Spatial Impulse Response Rendering (HO-SIRR) method by McCormack et al. (2020). The reverberation times (RT60) were 1.2 seconds for the café scene, 0.4 seconds for the dinner party scene, and 0.9 seconds for the food court scene. As described above, the room had an RT60 of 0.15 seconds, which was not accounted for in the decoding.

Stimulus presentation

Figure 5 shows an overview of the stimulus presentation. Before each measurements, 40 seconds of the ISTS speech signal and respective background noise were played to allow the hearing instruments to settle. The evaluation period was 10 seconds for each Hagerman phase-inversion block, i.e. the first 10 seconds of the ISTS speech signal and respective background noise were repeated with the respective Hagerman phase, leading to a total acquisition time of 70 seconds for one condition.

For data in Figure 1-3, the ISTS speech signal was presented from 0° azimuth. For Figure 3, the measurements were repeated with the speech signal presented from 0°, 60°, 90°, 120°, and 180° azimuth.

Figure 5. Stimulus presentation: Before each measurements, 40 seconds of the ISTS speech signal (S) and respective background noise (N) were played to allow the hearing instruments to settle. The evaluation period was 10 seconds for each evaluation Block, see methods.

Hagerman and Olofsson Phase-Inversion Method

Hagerman and Olofsson (2004) describe a method to extract the signal and noise components of a simultaneous presented audio signal. For the Hagerman and Olofsson test, multiple recordings take place, whereas the phase of one signal or noise is inverted between the measurements. This phase inversion method allows to separate the signal (S') and the noise (N') at the output of the hearing aid and the calculation of the signal-to-noise ratio (SNR').

Three evaluation blocks were played back:
Block A = S+N
Block B = S-N
Block C =-S-N

The recoded signals at the KEMAR mannequin after hearing aid processing are:
Block A' = S'+N'+ε
Block B' = S'-N'+ε
Block C' =-S'-N'+ε,
Where the apostrophe " ' " denotes the recorded audio.

The recorded speech signal S', noise signal N', and error ε are extracted as follows:
S' = (A'+B')/2
N' = (A'-B')/2
Error ε = (A'+C')/2

The signal-to-noise ratio (SNR) after the hearing aid processing and error terms are calculated as follows:
SNR' = 20*log10(rmsSII(S')/rmsSII(N'))
Signal error εS = 20*log10(rmsSII(ε)./rmsSII(S')
Noise error εN = 20*log10(rmsSII(ε)./rmsSII(N')

The Hagerman error was below -15 dB (SII-weighted) for all measurement conditions, including the competitor devices.

SII-weighting

The signal-to-noise ratio (SNR) benefit was frequency weighted based on the band importance function according to the speech intelligibility index (SII, ANSI S3.5/1997, Table 3).

Hearing aid fittings

For the hearing aid fitting, a standard N3 audiogram was used (Bisgaard et al. 2010) as well as a fully occluded coupling. The real ear occluded gain (REOG) was measured using the international female noise (IF noise, Holube et al. 2010).
For each device the REOG was >15 dB at frequencies between 100 Hz to 20 kHz. Additionally, the insertion between 100 Hz to 10 kHz was measured to make sure that the occlusion effect and insertion gain were symmetric between the left and right ears.

For each device, the proprietary fitting formula was used and the highest experience level was chosen. The feedback management system in each device was deactivated, so that the Hagerman and Olofsson (2004) method could be performed. Impulse noise reduction and wind noise reduction was also deactivated for each device. For Figure 2, the program options were left at the default settings. The maximum power output (MPO) was set to maximum for all devices. A manual program was applied for all devices, including competitors.

References

Abrams, H. B. & Kihm, J. (2015). An introduction to MarkeTrak IX: A New Baseline for the Hearing Aid Market. Hearing Review, 22(6).

ANSI S3.5 1997. Methods for calculation of the Speech Intelligibility Index.

Appleton, J. & König, G. (2014). Improvements in speech intelligibility and subjective benefit with binaural beamformer technology. Hearing Review, 21(11), 40-42.

Bentler, R. A. (2005). Effectiveness of Directional Microphones and Noise Reduction Schemes in Hearing Aids: A Systematic Review of the Evidence. J Am Acad Audiol, 16(07), 473–484.

Bisgaard, N., Vlaming, M. S. M. G., & Dahlquist, M. (2010). Standard Audiograms for the IEC 60118-15 Measurement Procedure. Trends in Hearing, 14(2), 113-120.

Derleth, P., Georganti, E., Latzel, M., Courtois, G., Hofbauer, M., Raether, J., & Kuehnel, V. (2021). Binaural Signal Processing in Hearing Aids. Seminars in Hearing, 42, 206 - 223.

Diehl, P. U., Zilly, H., Sattler, F., Singer, Y., Kepp, K., Berry, M., Hasemann, H., Zippel, M., Kaya, M., Meyer-Rachner, P., Pudszuhn, A., Hofmann, V. M., Vormann, M., & Sprengel, E. (2023). Deep learning-based denoising streamed from mobile phones improves speech-in-noise understanding for hearing aid users. Frontiers in Medical Engineering, 1:1281904. doi: 10.3389/fmede.2023.1281904

Diehl, P.U., Singer, Y., Zilly, H., Schönfeld, U., Meyer-Rachner, P., Berry, M., Sprekeler, H., Sprengel, E., Pudszuhn, A., & Hofmann, V. M. (2023). Restoring speech intelligibility for hearing aid users with deep learning. Sci Rep 13, 2719. https://doi.org/10.1038/s41598-023-29871-8

Edwards, B. (2016). A Model of Auditory-Cognitive Processing and Relevance to Clinical Applicability. Ear and Hearing, 37(1), 85-91.

Hagerman, B., & Olofsson, A. (2004). A Method to Measure the Effect of Noise Reduction Algorithms Using Simultaneous Speech and Noise. Acta Acustica united with Acustica, 90, 356-361.

Holube, I., Fredelake, S., Vlaming, M., Kollmeier, B. (2010). Development and analysis of an international speech test signal (ISTS). International Journal of Audiology, 49(12), 891–903.

Kates, J. M., Arehart, K. H. (2014). The Hearing-Aid Speech Quality Index (HASQI) Version 2. Journal of the Audio Engineering Society, 62, 3, 99-117.

Killion, M. C. (1997). The SIN report: Circuits haven't solved the hearing-in-noise problem. Hearing Journal, 50(10), 28-32.

Latzel, M., & Appleton-Huber, J. (2015). StereoZoom – Adaptive behaviour improves speech intelligibility, sound quality and suppression of noise. Field Study News available at https://www.phonak.com/en-int/professionals/audiology-hub/evidence-library

McCormack, L., Pulkki, V., Politis, A., Scheuregger, O. & Marschall, M., (2020). Higher-Order Spatial Impulse Response Rendering: Investigating the Perceived Effects of Spherical Order, Dedicated Diffuse Rendering, and Frequency Resolution. Journal of the Audio Engineering Society, 68(5), 338-354.

Mueller, G. H., Ricketts, T. A. (2000). Directional-microphone hearing aids: an update. Hearing Journal, 53(5), 10–19.

Politis, A., Tervo S., & Pulkki, V. (2018). COMPASS: Coding and Multidirectional Parameterization of Ambisonic Sound Scenes. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Ricketts, T. A. (2005). Directional hearing aids: then and now. J Rehabil Res Dev 42(4, Suppl 2), 133–144.

Weisser, A., Buchholz, J. M., Oreinos, C., Badajoz-Davila, J., Galloway, J., Beechey, T., & Keidser, G. (2019). The Ambisonic Recordings of Typical Environments (ARTE) database. Acta Acustica united with Acustica, 105, 4, 695-713.

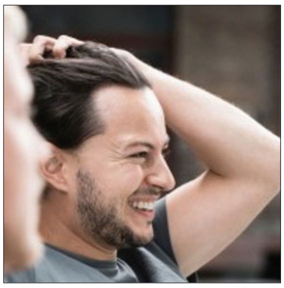
Wright, A., Kuehnel, V., Keller, M., Seitz-Paquette, K., Latzel, M. (2024). "Spheric Speech Clarity applies DNN signal processing to significantly improve speech understanding from any direction and reduce the listening effort ." Phonak Field Study News retrieved from https://www.phonak.com/evidence.

Authors and investigators

Stefan Raufer, Audiological Researcher

Stefan joined Sonova in 2021 and is developing audiological concepts for new hearing aid features. Together with other teams, Stefan brings latest technology to our products, making sure that new technologies translate to user benefits. Stefan holds a PhD in speech and hearing sciences from Harvard University.

Frédérick Uhlemayr, Audio & Development Engineer



Frédérick joined Sonova in 2021 and is currently working on perceptual models and tools for technical measurements. Frédérick graduated as Diplom-Tonmeister from the University of Music and Performing Arts in Vienna. Frédérick has a background in professional audio reproduction as well as in automotive audio.

Michael Preuss, Audiology Manager at Phonak HQ

Michael joined Phonak HQ as an Audiology Manager in 2020. Benefiting from his lecturing background at the Academy of Hearing Acoustics in Lübeck and living with hearing loss, Michael delivers comprehensive expert training sessions and provides audiological input during product development. Michael studied Hearing Acoustics at the University of Applied Sciences in Lübeck, Germany.

Shin-Shin Hobi, Senior Product Manager for Audiological Performance at Phonak HQ

Shin-Shin joined Phonak HQ in 2006 and has worked on various projects as an Audiology manager. In her current role as Senior Product manager for Audiological Performance, she ensures end user and hearing care professional needs are taken into account during product development of audiological features. Originally from Australia, Shin-Shin earned her Audiology qualifications from University of Melbourne. She gained wide clinical experience in private practice, in Perth, before making the move to Switzerland.

Volker Kuehnel, Principle Expert Hearing Performance

Volker Kuehnel, 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. Dr. Dr. 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.

Philipp Kohlhauer, Audiological Researcher

Philipp joined Phonak HQ as an Audiological Researcher in 2021. With his background as a hearing aid specialist, he supports the development of new features to ensure their quality and optimum performance. Philipp did his Master of Science in Auditory Technology at the University of Lübeck.