Phonak Field Study News.

AutoSense OS™ 7.0 improves speech understanding with highly rated sound quality for challenging listening environments

A clinical investigation conducted at the Phonak Audiology Research Center (PARC) found statistically significant improvements to speech understanding in real-world noise for next generation Infinio R and Sphere™ Infinio through improved accuracy of environmental classification.

Wright, A., Cox, M., Keller, M., Galster, J. August 2025.

Introduction

Phonak has a 25-year history of leveraging artificial intelligence (AI) in hearing instruments. The first commercial application was introduced in Phonak Savia, with the AutoSelect automatic system, and subsequently replaced with SoundFlow in Phonak Exélia. In 2014, the first version of AutoSense OS (ASOS) was introduced to the market, using real-time analysis of the sound environment to steer signal processing. A clinical investigation into the efficacy of the original ASOS found it provided a 1.3 dB signal-to-noise (SNR) benefit for speech understanding in noise compared to manual programs.[1]

Since its introduction, ASOS has undergone continuous improvements, further enhancing its capabilities by increasing sound classification accuracy and adding access to even more features to maximize user benefit and satisfaction—regardless of the acoustic environment. Today, ASOS includes six programs for acoustic input: Calm Situation, Speech in Noise, (Spheric) Speech in Loud Noise, Speech in Car, Comfort in Noise, and Music. Four of these six programs are defined by the presence of speech at various levels and types of background noise (including the absence of it), placing communication at the forefront of the design goals for ASOS.

ASOS 7.0 has been trained on 18 times more sound environments than the previous generation, making it the

most precise implementation of AutoSense to date—with a 24% increase in classification accuracy, according to technical measures.[2] To highlight the importance of classification accuracy, a clinical investigation evaluated the impact on speech understanding and sound quality if a speech in noise scene was mistakenly classified as just noise. This comparison reflects the two most probable choices for an automatic steering system in a noisy listening environment. A system that is well-trained with a high volume of diverse speech in noise recordings will identify the scene correctly, applying a sound class that reduces noise and improves speech understanding. Contrastingly, a less sophisticated system could have difficulty distinguishing the speech patterns in the noise, resulting in the application of a sound class designed for listening comfort, which reduces noise, but does not improve speech understanding.

Methodology

Participants

Seventeen experienced adult hearing aid users with moderate to moderately severe bilateral hearing loss aged 62 to 92 years ($m = 76 \pm 8$) were recruited for this study. Otoscopy was conducted prior to testing, as well as cerumen management on an as-needed basis.

Hearing instruments

A pair of Audéo Sphere™ Infinio 90 receiver-in-canal (RIC) hearing instruments (HIs) were programmed according to Target fitting software recommendations, including Phonak's proprietary fitting formula (APD 3.0) and the proposed dome type. Seven participants were fit bilaterally with power domes, six with vented, and four were fit with one power and one vented dome. Feedback testing and real-ear measurements, using a protocol that included real ear unaided response (REUR), real ear occluded response (REOR), real ear aided response (REAR) and maximum power output (MPO) measurements were conducted for all participants. No fine-tuning was performed.

Comparisons

Manual programs for Speech in Loud Noise (SiLN) and Spheric Speech in Loud Noise (SSiLN) were compared to a manual program of Comfort in Noise (CiN).

SiLN and SSiLN represent the sound classes selected by ASOS 7.0 for the tested sound scene within next generation Infinio 90R and Sphere™ Infinio hearing instruments. Previous iterations of ASOS were trained with smaller sets of real-world recordings and, consequently, may have over-classified speech in noise environments as CiN. Therefore, CiN represents the sound class chosen by a system that has not been trained with as many real-world recordings of speech in noise as ASOS 7.0.

These programs are differentiated by gain shaping and the application of advanced sound cleaning features such as directional microphones and noise reduction (NR). Table 1 summarizes the feature differences between the programs.

	CiN	SiLN	SSiLN
MicMode	Real-Ear Sound	StereoZoom 2.0	Fixed Directional
Dynamic Noise Cancellation	--	Yes	--
Spheric Speech Clarity	--	--	Yes
NoiseBlock	12	8	--

Table 1. Summary of the feature sets available in the three test programs.

CiN applies a global gain reduction across the frequency range for all input levels and uses a directional microphone configuration that simulates pinna effects (Real-Ear Sound). It also employs a moderate setting for NoiseBlock—an NR feature most effective against stationary noise sources.

SiLN also utilizes NoiseBlock, but enhances it with Dynamic Noise Cancellation, which works in tandem with a directional microphone. In this case, the microphone mode is an adaptive binaural beamformer—StereoZoom 2.0—capable of improving the SNR by up to 6.4 dB.[3]

SSiLN pairs a fixed directional microphone with NR driven by a Deep Neural Network called Spheric Speech Clarity, which can improve SNR by up to 10.2 dB.[3]

All of these programs are designed to reduce noise and to provide listening comfort. However, only SiLN and SSiLN are specifically designed to simultaneously enhance speech understanding—SiLN through directionality and SSiLN through DNN-based noise reduction.

Speech understanding

Speech understanding in noise was measured using Higher Order Ambisonic noise presented at 70 dBA (summated) from all speakers in a 12-speaker array, with AzBio sentences[4] presented from 0° azimuth. The noise simulated a common and challenging listening environment—conversing on an airplane at cruising altitude. The recording was captured from a real airplane and included typical transient noises alongside the dominant engine noise.

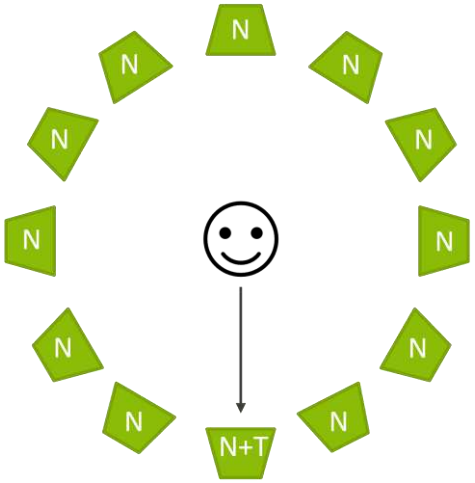

Figure 1. Schematic of test setup. A 12-speaker array was used in a sound booth. Ambisonic noise (N) played from all 12 speakers. Target speech (T) was collocated to the noise at 0° azimuth.

The sentences were presented at one of five fixed SNR levels, determined by individual performance on an adaptive consonant-nucleus-consonant (CNC) word task using the best aided condition—SSiLN in this case. The CNC task identified each participant's SRT-50, or the SNR threshold at which they achieved 50% correct. Each individual's SRT-50 was matched to the closest fixed SNR level from the five available options: –5, –3, 0, +3, and +5 dB. Fourteen out of seventeen participants were tested at –5 dB, with the remaining three tested at –3 dB.

After completing the adaptive CNC task and selecting the fixed SNR for testing, each participant completed one practice AzBio sentence list and two lists per test condition, for a total of seven lists per participant. The same seven lists were used for all participants, with conditions counterbalanced and lists counterbalanced across conditions to minimize order and list effects.

It was hypothesized that activating SiLN or SSiLN would lead to better speech understanding compared to when participants were using the CiN program. This hypothesis was tested using a linear mixed-effects model. Test conditions were controlled via connection to the myPhonak app. The participant and the investigator scoring the test were blinded to the condition.

In addition to measuring speech understanding with a behavioral test, participants were asked to provide a subjective rating of speech understanding after each test condition. Ratings were given on a 5-point scale, with 1 indicating that the speech was *not at all intelligible* and 5 indicating that the speech was *as intelligible as can be*.

Sound quality

Sound quality ratings were collected using an A/B comparison tool. Participants used a tablet to toggle between the hearing aid programs in each paired comparison (CiN vs. SiLN and CiN vs. SSiLN). They then rated the programs according to the following 5-point scale:

- A is much better than B
- A is slightly better than B
- A is the same as B
- B is slightly better than A
- B is much better than A

Participants provided these ratings for both comparison pairs across seven sound quality dimensions: speech clarity, speech naturalness, speech volume, noise intrusion, listening effort, overall sound quality, and preference. Conditions were counterbalanced within the paired comparisons, and each dimension was measured twice.

The same noise scene at the same presentation level was used for both the sound quality ratings and the speech understanding task. The SNR was also the same between the tasks. The target speech for the sound quality ratings was a natural speech passage delivered by a female talker using realistic vocal effort. The passage was under two minutes in length and looped if the participant needed more time.

Results

Speech understanding

SiLN outperformed CiN on the AzBio task, significantly increasing correct word identification by an average of 8 percentile points ($p < .001$).

Activation of SSiLN also significantly increased correct word identification by an average of 22 percentile points ($p < .001$).

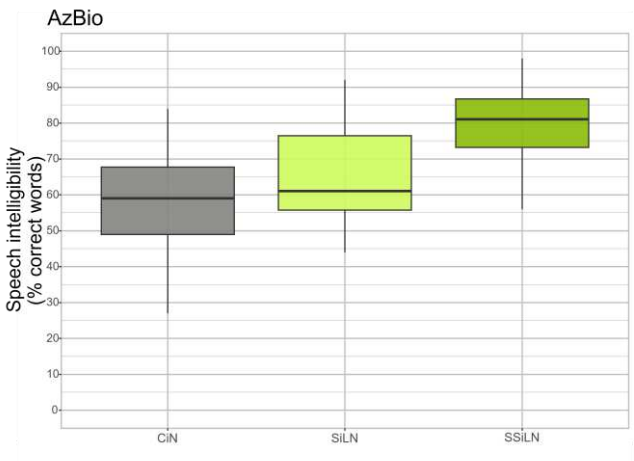

Figure 2. Speech understanding scores for all three conditions with 17 participants. CiN had an average performance of 58% (±15). SiLN had an average performance of 66% (±14). SSiLN has an average performance of 80% (±11).

For the ratings of subjective speech understanding, 88% of participants rated SiLN as intelligible as, or more intelligible than, CiN. This increased to 94% for SSiLN.

Sound quality

The majority of participants rated SiLN the same or more favorably than CiN across all sound quality dimensions. Speech volume and noise intrusion were rated the most favorably, indicating a perceived SNR improvement that is supported by the behavioral results.

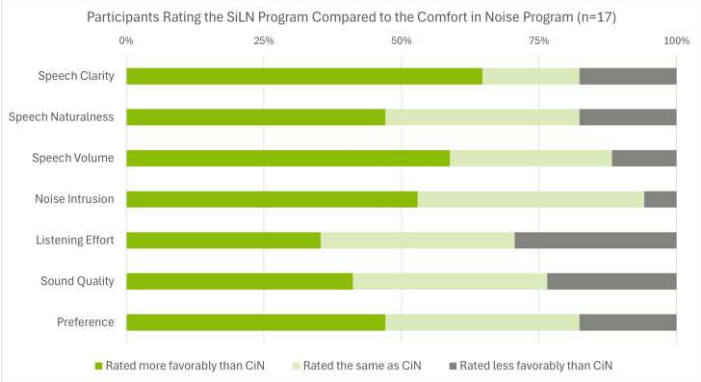

Figure 3. Sound quality ratings for the SiLN program compared to the CiN program.

SSiLN was also rated highly, with five out of seven dimensions unanimously rated the same or more favorably than CiN: speech volume, noise intrusion, listening effort, overall sound quality, and preference. Even more impressively, SSiLN was rated more favorably than CiN by 94% of participants for listening effort and sound quality. As expected with these ratings, preference was strong for SSiLN with 100% of participants preferring the SSiLN program over the CiN program.

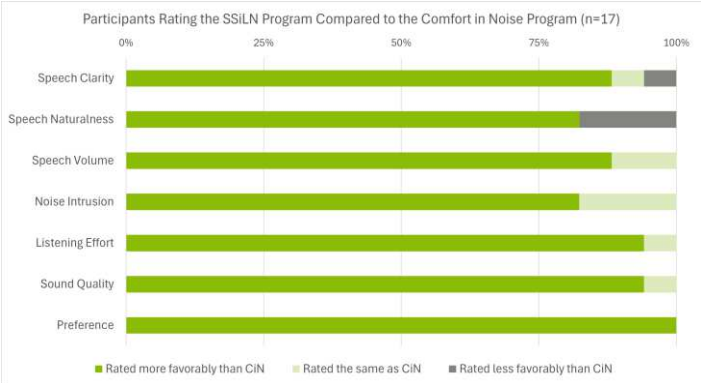

Figure 4. Sound quality ratings for the SSiLN program compared to the CiN program.

Conclusion

SiLN performed as expected, delivering sound quality that equaled or exceeded that of CiN, while providing additional benefits for speech understanding as a result of StereoZoom 2.0 and Dynamic Noise Cancellation. Since StereoZoom improves the SNR through directionality, it is expected that the speech understanding advantage of SiLN over CiN would be even greater when speech is spatially separated from the noise.

Accordingly, SSiLN also outperformed CiN—and to an even greater degree than SiLN. By design, SSiLN provides speech understanding benefits regardless of spatial configuration. The extensive training of its Deep Neural Network also enables it to preserve more of the target speech during sound cleaning, as evidenced by the nearly perfect sound quality ratings for SSiLN.

These findings support the notion that not all noisy environments should be treated the same. When speech is present, a global reduction in sound without the activation of advanced sound cleaning features can lead to missed communication opportunities and suboptimal performance. Therefore, it is important that AI-based algorithms like AutoSense OS continue to be improved, ensuring that end-users consistently experience the full benefits of their hearing instruments.

References

1. Übelacker, E., Tchorz, J., Latzel, M. (2015). AutoSense OS: Benefit of the next generation of technology automation. Phonak Field Study News retrieved from https://www.phonak.com/evidence
2. Sanchez, C., & Giurda, R. (2025). AutoSense OS™ 7.0 adapts 24% more precisely to listening situations so that wearers can enjoy exceptional sound quality in every moment. Phonak Insight, available at https://www.phonak.com/en-int/professionals/audiology-hub/evidence-library
3. Raufer, S., Kohlhauer, P., Jehle, F., Kühnel, V., Preuss, M., Hobi, S. (2024). Spheric Speech Clarity proven to outperform three key competitors for clear speech in noise. Phonak Field Study News retrieved from https://www.phonak.com/evidence
4. Spahr, A. J., Dorman, M. F., Litvak, L. M., Van Wie, S., Gifford, R. H., Loizou, P. C., Loiselle, L. M., Oakes, T., & Cook, S. (2012). Development and validation of the AzBio sentence lists. *Ear and hearing*, *33*(1), 112–117. https://doi.org/10.1097/AUD.0b013e31822c2549

Authors and investigators

Internal investigator/Author
Ashley Wright, Au.D., Senior Research Audiologist Ashley

is a Senior Research Audiologist at the Phonak Audiology Research Center (PARC) in Aurora, IL. She completed her Au.D. at Rush University in Chicago and joined PARC in 2018. Her primary responsibilities include managing internal clinical studies with adults.

Internal investigator/Author
Marina Cox, Au.D., Research Audiologist

Marina is a Research Audiologist at the Phonak Audiology Research Center (PARC) in Aurora, IL. She completed her Au.D. at Northwestern University in Chicago and joined PARC in 2024.

Internal investigator/Author
Matthias Keller, Ph.D., Scientist, Audiology Research

Matthias Keller earned his PhD in Psychology from University of Zurich, Switzerland, researching age-related differences in neural processing of spoken language. After working for Phonak/Sonova in Switzerland since 2019 he transitioned to the US and became part of the PARC team in 2023.

Principal investigator
Jason Galster, Ph.D., VP Clinical Research Strategy

Jason works with a global network of clinical research teams on the development and scientific study of tomorrow's hearing technology. He is an internationally recognized author, lecturer, and adjunct professor in audiology whose research interests include statistical methods for characterizing individual variability and the measurement of hearing outcomes during daily life.

Phonak Field Study News.

One-page summary

AutoSense OS™ 7.0 improves speech understanding with highly rated sound quality for challenging listening environments

A clinical investigation conducted at the Phonak Audiology Research Center (PARC) found statistically significant improvements to speech understanding in real-world noise for next generation Infinio R and Sphere™ Infinio through improved accuracy of environmental classification.

Wright, A., Cox, M., Keller, M., Galster, J. July 2025.

Key highlights

- AutoSense OS 7.0 adapts 24% more precisely to listening situations so that wearers can enjoy exceptional sound quality in every moment.
- 100% of Sphere™ Infinio wearers prefer AutoSense OS 7.0 over the previous generation.
- The correct classification of the listening situation improved speech understanding by 8% for Infinio R and 22% for Sphere™ Infinio.

Considerations for practice

- Based on the findings of the present study, it is expected that clients who are fit with next generation Infinio R and Sphere™ Infinio hearing instruments will experience improved speech understanding in noise without compromising sound quality.
- HCPs can feel confident that their clients will be satisfied with AutoSense OS 7.0 as the startup program.
- Phonak believes in a continuous improvement model, meaning AI features like AutoSense OS will always be on the cutting edge.