Phonak Insight.

AutoSense OS™ 7.0 adapts 24% more precisely to listening situations.

With AutoSense OS 7.0, Phonak is introducing their retrained, AI-based environmental classifier, utilizing a new and larger set of sound samples to increase its accuracy in recognizing sound scenes.

Sanchez, C., Giurda, R., Hobi, S., Preuss, M., August 2025

Key highlights

- AutoSense OS is the foundation for steering the signal processing and applying the most appropriate setting for the client based on the acoustic environment.
- AutoSense OS 7.0: AI trained with 18 times more sound environments than the previous generation, boosts speech understanding in noise (Wright, 2025) for adults with mild to severe hearing loss.
- Optimal sound quality in every listening environment is always the goal of hearing aid manufacturers (Kochkin, 2010).

Considerations for practice

- Market research shows that 86% of clients value a hearing aid that automatically adjusts to the listening environment (Knorr, 2022). AutoSense OS meets clients' needs by ensuring that the hearing aids automatically and seamlessly adapt in real-time.
- A more accurate classification of scenes where both speech and noise are present allows AutoSense OS to activate the appropriate blend of programs to deliver the balance of comfort, environmental awareness, speech understanding and sound quality.
- Selective activation of sound cleaning features (e.g., NoiseBlock, Spheric Speech Clarity, StereoZoom 2.0) helps to ensure optimal outcomes. AutoSense OS 7.0 manages the automatic activation and deactivation of features, removing the need for manual user adjustments.

25 years of automated classification of sound environments

Phonak has a 25 year history of leveraging artificial intelligence (AI) in hearing instruments. Indeed, developing a sound scene classifier started with the idea of Bregman (Bregman, 1990) who introduced the concept of Auditory Scene Analysis in 1990, proposing that the brain organizes sounds into streams based on cues like timing, pitch, and spatial location (concept applied to hearing aids for instance by Kates, 1995).

Phonak introduced its first sound classification system with AutoSelect in Phonak Claro hearing aids. The addition of a more intelligent implementation, incorporating machine learning, was introduced with AutoPilot, in Phonak Savia and then subsequently replaced with SoundFlow in Phonak Exélia. 2014 saw the first version of AutoSense OS introduced to the market. AutoSense OS (ASOS) was the first automatic system in the industry that was capable of incorporating settings from multiple sound classes to optimally address a client's listening needs.

AutoSense OS taking a two-step approach to automatic functionality

Today's world is a busy and 'acoustically dynamic' place that makes it challenging to listen, understand, and actively engage. For people with hearing loss, this challenge is significant and may limit social engagement (National Academies of Sciences, Engineering, and Medicine, 2020). The use of hearing instruments may address these social limitations (Reed NS, Chen J, Huang AR, et al, 2025) and sound cleaning features in hearing instruments has been shown to improve communication.

The key is knowing when, and to which degree, to activate these features. Relying on clients to manually activate individual features or programs is not only cumbersome for the clients but also could result in suboptimal performance. ASOS is our proprietary algorithm that takes a two-step approach:

1. Accurate Auditory Scene Analysis and the classification of the sound environment. This first step is critical in determining how to steer the features.
2. Real-time system steering consisting of activation and/or blending of programs is then based on the classified environment. This ensures the appropriate activation of features for optimal sound quality, balancing comfort and speech understanding.

Training a machine learning based acoustic scene classification system

At the core of ASOS is the acoustic scene classifier—a type of machine learning system extensively trained to recognize different kinds of sounds to enable it to classify the sound environment or scene.

The training processes starts by inputting ASOS with thousands of labeled audio samples (e.g. traffic noise, female speech, classical music). As part of the training process, ASOS extracts over 30 unique characteristics from each audio sample (e.g. modulations, frequency pitch) – almost like an individual fingerprint. Once these characteristics are identified, each audio sample is organized on to a map called the „acoustic embedding space"(Fig. 1.). Although each audio sample has its unique spot, audio samples with similar characteristics are clearly clustered in a region, as seen by the colored regions in Figure 1. For example, there would be a cluster that represents all the variations of "speech in noise", however within this cluster there will be groups that have similar characteristics (e.g. female versus male voices, traffic noise versus restaurant noise).

The resulting map, after training, is used by ASOS when it is exposed to a new sound or sound environment, which it was not exposed to during training. ASOS is able to extracts the characteristics of the new acoustic scene, and uses the map to determine the cluster to which the sound type's centroid it is closest to. And thus classifying the sound and/or sound environment.

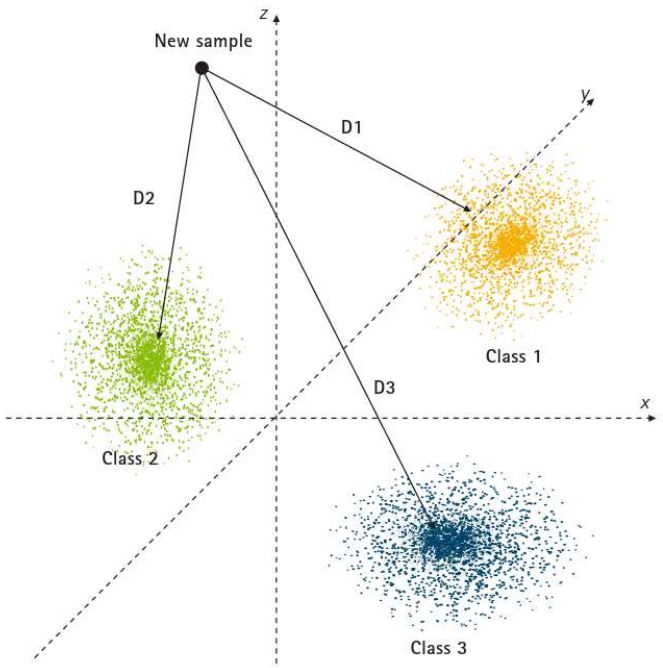

Figure 1: Example of projection space. The distance of the new sample from the centroid of each class is used to define the class probability.

Importance of large data sets for retraining

The accuracy of any machine learning models is correlated with the size of the data set it is trained on. The more examples a model sees—especially variations of the same condition—the better it becomes at making accurate predictions in the future.

ASOS 7.0 was retrained with new audio data, resulting in an update of the acoustic embedding space the locations and shapes of sound categories in this map— to reflect the new information. ASOS 7.0 was retrained with millions of sound recordings, thanks to building a large data set of audio files, resulting in 18 times more sound samples being used than in the original ASOS training. The goal of retraining ASOS 7.0 was to further increase the accuracy in identifying challenging environments that are broad in their acoustic spectral characteristics, specifically: speech in noise and music. The heterogeneous nature of speech in noise situations can make it challenging for a system to determine if the scene is just noise, and therefore activate a program for comfort, or a scene where speech is present which requires a program aimed at improving the Signal-to-Noise (SNR) ratio for speech understanding. Music genres vary considerably in their acoustic properties. Some music styles can show close feature similarity to noise, resulting in a program activation that could impact the perceived sound quality of the music. Exposing the ASOS classifier to 18 times more sound samples results in a more robust acoustic scene classification especially for these challenging heterogeneous environments.

In order to verify the result of our retraining we assessed ASOS 7.0 with:
1. technical verification: to verify that the retraining resulted in more accurate identification and classification of speech in noise and music scenes.
2. field trial verification: to verify that the more accurate classification leads to increased activation of the appropriate ASOS program in real life environments
3. clinical study: having confirmed the above a clinical study was conducted to measure the benefit that more accurate classification can offer

Technical verification of the retraining

To verify that the retraining increased classification accuracy technical assessments were conducted. Assessing an automatic system consists of subjecting the system to a varied and complex sound parkour. ASOS 7.0 underwent an extensive verification parkour, consisting of a variety of acoustic scenes that simulate real life environments such as variations of quiet situations, noise only, speech in noise and music. Each scene was presented for a minimum of 10 seconds. The accuracy in the classification of speech in noise scenes increased by 24%. When tested with the music parkour, consisting of 1000 different music pieces, we also observed a 24% increase in accuracy.

Field trial verification

To assess the classification difference between ASOS 6.0 and ASOS 7.0 in real life, a field trial verification was conducted where study participants were fitted asymmetrically: one hearing aid with old classification, one hearing aid with newly retrained classification. Thanks to data logging capabilities of the hearing aids, we could read out and compare the classification of the two versions of ASOS, knowing that the sound environment triggering the classification was the same for both devices. Eighteen subjects provided us with over 6,600 hours of logged wearing data (i.e. over a 2 week wearing period). We analyzed the classification, ASOS program activation, noise floor estimate and Signal-to-Noise ratio of the actual environment the subject was in. The newly retrained ASOS classifier showed an increase classification of speech in noise over the day which correlated to the measured noise floor and SNR of the actual environment. The field trial data supported our hypothesis that due to the retraining and the increased accuracy (as measured in the technical verification), we could expect a higher activation of Speech in Noise/Speech in Loud noise.

Clinical benefits of AutoSense OS

Ultimately the goal is to be able to show tangible benefit from a retrained system. A clinical study (Wright, 2025) was conduced to show that the technical and field trial results translated into a perceptual benefit in people with a hearing loss in the dimensions of sound quality and speech understanding. A study, conducted at Phonak Audiology Research Center (PARC), assessed the speech understanding and preference rating from 17 adult subjects with moderate to severe hearing loss. The study aimed to show the importance of correct classification, by comparing the clients' performance in a comfort in noise program against their performance in a Speech in Loud Noise program and Spheric Speech in Loud Noise program, respectively. Results showed not only a statistically significant improvement in speech understanding but also an overall preference across various sound quality dimensions (e.g. speech volume, noise intrusion).

This study demonstrated that an increased accuracy in classification, results in optimal activation of features for the classified sound environment, which in turn can lead to a perceptual preference in sound quality and measurable improvements in speech understanding for individuals with a mild to moderate hearing loss.

Conclusion

For over 25 years, Phonak has led the way in automatic sound environment classification, using machine learning models back to the release of the Phonak Claro, beginning with AutoSelect and culminating in the current ASOS 7.0 system. What began as a vision to relieve hearing aid users of the burden of manual adjustments has evolved into an advanced machine learning-based classifier that adapts seamlessly to the complexity of modern and realistic acoustic environments and hearing experiences. With its latest retraining — targeting critical environments like Music and Speech in Noise — ASOS has achieved measurable gains in classification accuracy and user benefit, including a 24% increase in both music and speech in noise environments detection leading to measurable clinical benefit of improved speech understanding and perceptual preference of sound quality (Wright, 2025). These improvements are not just incremental—they reflect Phonak's deep commitment to innovation, usability, and user-centered design. By leveraging its machine learning model and expanding its dataset, the system is now more capable than ever of delivering a truly automatic, immersive listening experience. AutoSense OS 7.0 is the cornerstone of a personalized hearing care for the next generation of hearing aids orchestrating the best fitting strategies for the clients.

References

Bregman, A.S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press, Cambridge, MA, USA.

Büchler, M, Allegro S, Launer S, Dillier N. (2005) Sound Classification in Hearing Aids Inspired by Auditory Scene Analysis. EURASIP Journal on Advances in Signal Processing.

Kates, J.M. (1995). Classification of background noises for hearing aid applications. The Journal of the Acoustical Society of America, 97 (1): 461 - 70

Knorr, H. (2022) Market Research ID # 4535 Please contact marketinsight@phonak.com if you are interested in further information.

Kochkin, S. (2010) MarkeTrak VIII: Consumer satisfaction with hearing aids is slowly increasing, Hearing Journal, 63(1), 11 – 19.

National Academies of Sciences, Engineering, and Medicine. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. National Academies Press; 2020.

Reed NS, Chen J, Huang AR, et al. Hearing Intervention, Social Isolation, and Loneliness: A Secondary Analysis of the ACHIEVE Randomized Clinical Trial. JAMA Intern Med. Published online May 12, 2025. doi:10.1001/jamainternmed.2025.1140

Sonova proprietary research (2025). Project ID #4736, n=85. Please contact marketinsight@phonak.com if you are interested in further information.

Wright, A. et al. (2025) AutoSense OS 7.0 improves speech understanding with highly rated sound quality for challenging listening environments. Phonak Field Study News. Retrieved from https://www.phonak.com/evidence.

Author

Clément Sanchez, Senior Audiology Manager

Clément is Senior Audiology Manager in the Phonak RIC team at Phonak HQ. He joined Sonova in 2024 and his collection of background includes his clinical audiology degree from the ILMH of Brussels (2001), an MBA in technology management from the Technical University of Denmark (DTU) and a Certification in Radical Innovation from the MIT of Cambridge. He has been working in both ENT and hearing aids dispensing practices, as university lecturer in France, and since 2005 he has been working in the hearing industry in training and education, product development and customer support.

Co-Authors

Ruksana Giurda, Multi-modal Classification Expert

Ruksana is expert in Multimodal Classification in Sonova. She joined Sonova in 2020 after completing her PhD on source localization and classification with hearing devices at ETH (Zurich). She holds a master degree in Acoustic Engineering from the Technical University of Denmark (DTU).

Shin-Shin Hobi, Senior Product Manager for Audiological Performance

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.

Michael Preuss, Senior Audiology Manager

Michael is Senior Audiology Manager and joined Phonak HQ in 2020. Drawing on his experience as a lecturer at the Academy of Hearing Acoustics in Lübeck and his personal experience with hearing loss, Michael provides audiological input during product development and conducts in-depth expert training sessions. He holds a B.Sc. in Hearing Acoustics from the University of Applied Sciences in Lübeck, Germany.