WHITEPAPER
2024

4D Sensor technology and Deep Neural Network 2.0 in Oticon Intent™

Technical review and evaluation

ABSTRACT

Oticon Intent™ takes another big step in optimizing listening support for people with hearing loss – especially when listening in noise. This whitepaper goes through two major innovations in MoreSound Intelligence (MSI) 3.0: 4D Sensor technology and new Deep Neural Network (DNN) 2.0. The 4D Sensor technology estimates the listening intention of the user in a given situation and environment by combining signals from a new motion sensor and acoustic sensors. This combined information allows improved utilization of the help-in-noise system to provide the appropriate help based on the user's intentions in the specific listening situation. Additionally, the new DNN 2.0 has undergone improved training resulting in superior noise suppression compared to previous Oticon hearing aids. We also present technical evidence that shows that Oticon Intent outperforms Oticon Real. Whereas traditional technology can only provide a fixed level of support within a given sound environment, Oticon Intent provides listening support based on the user's listening intent, offering a 5-dB span of adaptation. Results also show that Oticon Intent provides 35% more access to speech cues than Oticon Real and is better at attenuating background noise and preserving speech details, resulting in a clearer sound scene for the user. Altogether, Oticon Intent shows superior performance in noisy environments. It provides the appropriate amount of support for the users when they need it the most.

EDITORS OF ISSUE

Mette Brændgaard*, Valentina Zapata-Rodríguez*, Ioana Stefancu, Raul Sanchez-Lopez, Sébastien Santurette
Centre for Applied Audiology Research, Oticon A/S

** Shared first authorship*

oticon
life-changing **technology**

Despite advancements in signal processing strategies in modern hearing aids, people with hearing loss still encounter challenges in understanding conversations in noisy environments. Traditional technology adjusts the level of support based on the acoustic complexity of the environment. However, the need for improved assistance in complex listening situations remains an area that requires further development (Picou, 2020).

Oticon Intent™, based on the new Sirius™ platform, introduces technology which for the first time brings the user's intention into play when processing sound. Integrated in MoreSound Intelligence™ 3.0 (MSI 3.0), this technology – 4D Sensor technology – combines four different types of sensor-input used by the hearing aid to understand and act on the user's listening needs. The four sensor-inputs are: head movement, body movement, conversation activity and acoustic sound scene analysis. Besides the new 4D Sensor technology, Oticon Intent utilizes a second generation Deep Neural Network (DNN 2.0) for even more clarity and contrast in difficult environments. This enables Oticon Intent to offer a targeted listening support by predicting the user's listening intent, which is a more comprehensive approach to the user's needs than only using the acoustic complexity of the environment and simple movement detectors to disable directionality. This paper will take you through the technical aspects and evaluation of MSI 3.0 including the new 4D Sensor technology and the DNN 2.0.

MoreSound Intelligence 3.0 processing

MSI 3.0 serves as the help-in-noise system in Oticon Intent. This section describes each part of the processing flow in MSI 3.0 (Figure 1) and consists of a brief introduction to the flow, followed by some details on each component.

Overview of the MSI 3.0 processing flow

The processing flow in MSI 3.0 begins with the Wind & Handling Stabilizer which receives the sound from the 2 microphones, cleans the signal, and removes uncomfortable and disturbing noises as needed. For example, the noise created by wind blowing over the microphone openings or handling noise from fingers or hair brushing against the microphone openings.

The next part of the flow is the 4D Sensor technology. This part processes information about sounds in the environment, the user's head and body movements, and their conversation activity. This new input, together with input from the level and SNR detectors, already known from existing versions of MSI, is analyzed, combined, and used to determine how the remaining parts of the help system should process the sound scene. The analysis is designed to determine the user's intention in the given situation and to ensure correct handling of the incoming sound based on their needs. Additionally, it is designed to engage more or less support depending on whether the environment is more challenging (Difficult path) or less challenging (Easy path) for a specific user. The amount of engagement is personalized and defined by the Oticon Genie 2 software. Within this personalized range, the output of the 4D Sensor technology will configure the help according to the complexity of the situation and the intention of the user.

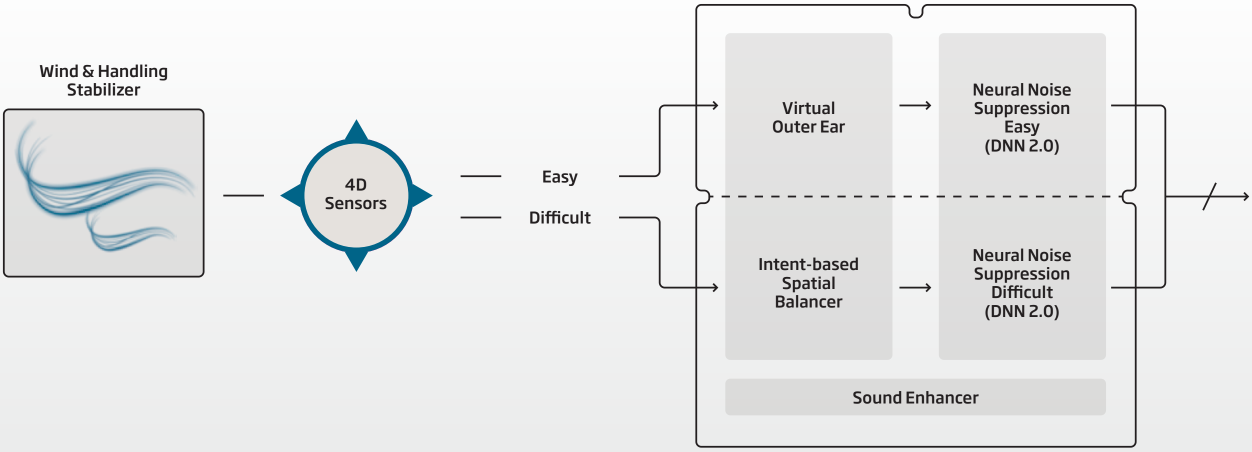

Figure 1. Overview of the processing flow in MoreSound Intelligence 3.0.

In the Easy path, Virtual Outer Ear recreates the spatial cues otherwise lost when the hearing aid is placed behind the ear. Next DNN 2.0 handles diffuse noise sources to increase clarity and user comfort.

In the Difficult path, Intent-based Spatial Balancer handles distinct and spatially separated sound sources and delivers optimal clarity and contrast between speech and other sound sources. Then, processing by DNN 2.0 handles the remaining diffuse noise sources while further optimizing contrast and clarity. All of this happens while still providing access to sounds all around.

Sound Enhancer is the final part of the MSI 3.0 processing flow. It runs adaptively on the noise-suppressed signal and ensures the presence of subtle details in the frequencies important for speech appropriate for the user.

The following sections will go a bit more into details about each part of MSI 3.0.

Wind & Handling Stabilizer

Wind & Handling Stabilizer (WHS) detects and prevents wind and handling noise from entering the sound processing in the hearing aid.

Blowing wind creates turbulence which is detected as it moves across the hearing aid microphone. This turbulence causes unacceptable noise in the hearing aid, similar to blowing air across a handheld microphone.

WHS monitors for the presence of uncorrelated noise created by wind or handling in each hearing aid microphone and determines which microphone is receiving the most noise. WHS monitors for changes in the presence of wind and handling noise 500 times/second, as the detector is constantly active. The attenuation for wind and handling noise is dynamic, meaning it is only active when wind or handling noise is detected. For all situations, other than direct wind and handling, the system uses both microphones across the whole range of frequencies. However, when wind is detected shutting down one microphone is preferred. The microphone receiving less noise is prioritized and the microphone with more turbulence is deactivated momentarily. WHS shuts down one microphone only in the time and frequency intervals needed, while simultaneously ensuring that the dual-microphone input signal is retained as much of the time and in as many frequency channels as possible.

Traditional wind noise management systems are efficient at attenuating wind noise in the low frequencies up to around 1500 Hz which ensures user comfort. However, an additional advantage of WHS is the precise and efficient attenuation of wind noise in frequencies above 1500 Hz, which is significantly improved compared to previous wind management systems.

For more information on WHS see Oticon whitepaper Gade et al. (2023).

Sensor technology

Hearing aids have been able to detect information about environmental sound levels and signal-to-noise ratio (SNR) for many years. This information is used to determine how to process the sounds around the user. Thus, users with the same settings in Oticon Genie 2 were given the same processing and same help in the same environment no matter their listening intentions. This is changed with Oticon Intent with the added 4D Sensor technology – see detailed description in '4D Sensor technology' further below.

Virtual Outer Ear

Being able to localize sound sources in the spatial environment is an important ability which becomes more difficult when hearing loss is present (Akeroyd, 2014). We all have different ear sizes and pinna shapes, and sound will therefore be modified in different ways when it enters the ear canal, depending on the anatomy of the ear. For example, due to the shape of the outer ear, some people will have more or less frontal focus than others. When we place the hearing aid microphones behind the ear, the ability to utilize the natural spatial cues provided by the pinna is eliminated. This ability needs to be recreated by signal processing in the hearing aid.

Virtual Outer Ear (VOE) and Intent-based Spatial Balancer help recreate this spatial sensation in easy and difficult environments, respectively. VOE contains three different true-to-life pinna models that recreate spatial cues that help the user recreate spatial awareness in easy environments.

In recent studies where we characterized 130 sets of ears, we found that most people get a natural amplification of around 0.5-1 dB in the 2-5 kHz area from the outer ear. Based on these characteristics we created a pinna model that is as natural and accurate as possible. Our measurements show that the effect the pinna has on sound can vary between ears. This implies that sound perception is dependent on outer ear anatomy. To account for individual differences, the VOE has three different settings with slightly more or less frontal focus. These can be set in the Oticon Genie 2 fitting software according to user preference. The slightly more frontal focus is created by a slight frequency-specific reduction in level from the back, while the one with more awareness allows the user to perceive more sounds from the back.

Intent-based Spatial Balancer

Intent-based Spatial Balancer is a more powerful feature than VOE when it comes to difficult environments. Intent-based Spatial Balancer quickly balances distinct sound sources in the environment by using both an omnidirectional signal and a back-cardioid signal from the two microphones. The omnidirectional signal provides all the sounds in the sound scene including frontal sounds. The back-cardioid signal does the same but excludes frontal sounds. The two signals are constantly compared to define precision placement of noise sources. Intent-based Spatial Balancer uses a minimum-variance distortion less response (MVDR) beamformer to create the optimal balance for the given sound scene by creating bigger contrast between meaningful and less meaningful (often noise) sounds.

Intent-based Spatial Balancer increases the SNR by suppressing individual noise sources (by reducing the level in the direction with the noise source), placing them in the background and thereby creating a balanced sound scene.

For more information on VOE and Spatial Balancer see Brændgaard (2020).

Deep Neural Network 2.0

DNN 2.0 is trained to recognize what should be emphasized (sounds of interest with a lot of information) and what should be less noticeable (sounds of less interest with less information). The training has been improved to create better clarity and contrast between sound sources. The detailed description of the improved training is described in the section 'New generation DNN' further below.

Sound Enhancer

Normally the maximum effect of a noise suppression system is a compromise that works reasonably well for most users. The sound processing in the hearing aid needs to make sure that the user can handle environmental sounds while retaining the overall feel of the sound scene.

Sound Enhancer provides dynamic sound detail when noise suppression is active and allows output to be individualized via three distinct sound profile settings. The settings are designed to enhance noise suppression or clarity of speech or in the default, Balanced setting, a combination of these. In all settings Sound Enhancer provides added detail in the 1-4 kHz range which are the primary frequencies for speech sounds.

For more information on Sound Enhancer see Brændgaard (2020).

The new technology in Oticon Intent

4D Sensor technology

Oticon Intent introduces the 4D Sensor technology which includes new sensors to the help-in-noise system while retaining the traditional level and SNR detectors to determine the sound environment. This new 4D Sensor technology provides input from four dimensions: body movement, head movement, conversation activity, and acoustic environment. This information is used to determine the user's listening intention in a given moment. The interpretation of the user's intention is used by the remaining part of MSI 3.0 and ensures that appropriate help is provided in each unique listening situation according to the user intention. This section describes the different sensors.

Motion sensor

Studies have found that listeners tend to orient their bodies in a certain way in communication situations (Hadley et al., 2019, 2020; Hadley and Culling, 2022) and that our listening intentions are shown by head and body movements (Higgins et al., 2023). By adding motion sensor data to the environmental data also collected by the hearing aid, sound processing and applied help can happen on a much stronger foundation and better support how the brain makes sense of sound (Bianchi/Eskelund et al., 2024).

The hearing aid has a built-in accelerometer to detect head and body movement of the user accounting for two of the four dimensions. The accelerometer is a small and power-efficient motion sensor, which makes it an optimal choice for use in a hearing aid.

The accelerometer measures acceleration of the user's movements. The faster and more vigorous the movement, the bigger the impact on the accelerometer. The accelerometer measures movement along three different axes: X, Y, and Z (see Figure 2).

The accelerometer is calibrated based on the placement on the ear which makes it important that the hearing aid is placed correctly. The tracking of the user's movements done by the accelerometer is used in the interpretation of the user intent:

- Movement on the Z-axis indicates full body motion (like walking and running) which can indicate awareness of the surroundings is essential.
- Movement on the X- and Y-axes indicates nodding and head turns which can indicate participation in a conversation with multiple people.
- Limited movement on the X- and Y-axes means that the user is still, which can indicate that full attention is needed to participate in an intimate conversation.

The type of motion together with the Conversation activity and Acoustic environment (described below) are combined to interpret the user intention.

Conversation activity

Conversation activity is the third dimension in 4D Sensor technology. Conversation activity contributes to the determination of the user's intention by detecting if speech is present. If there is no detectable speech in the frontal half plane, there is no ongoing conversation.

Conversation activity is determined based on the detected modulation in the signal and the calculation of a signal-to-noise ratio (SNR) approximately within the SNR range where normal-hearing people can understand speech.

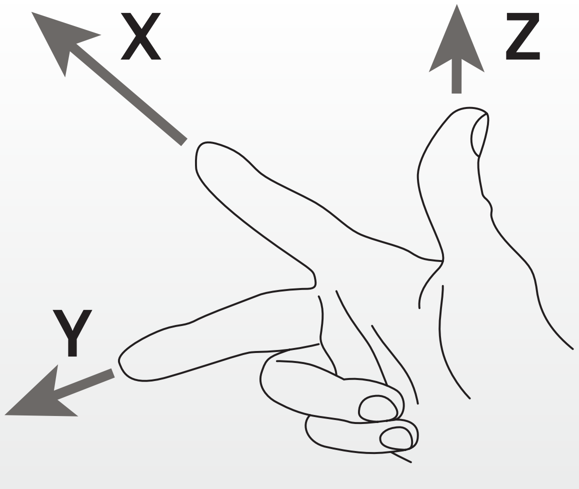

Figure 2. The three axes – X, Y, and Z – measured by the motion sensor.

Again, this analysis is used to interpret the user intention.

Acoustic environment

The fourth dimension in the 4D Sensor technology is Acoustic environment. This sensor is part of the system to ensure help is provided in the environments where it is needed. In easy environments, with low noise levels and good SNR, people normally communicate fine without any additional help from the hearing aid besides amplification.

This sensor detects sound levels in dB SPL. The determination of whether a sound environment is easy or difficult is individualized by the settings in Oticon Genie 2. These settings accommodate the user's unique sense of easy or difficult communication environments. The settings are personalized based on the personalization questions or the Audible Contrast Threshold (ACT™) value as an assessment of speech in noise abilities of the individual. Depending on the overall sound pressure level the sensor promotes the noise suppression system to provide additional noise suppression or it does not. The threshold for engaging additional noise suppression falls between about 45 and 60 dB SPL depending on the personal setting. The maximum effect is seen between about 62 and 72 dB SPL again depending on the personal setting.

4D Sensor technology outcome

The sensors constantly monitor movements, conversation activity, and sound in the environment to provide the correct input to the system. Each sensor brings the result of their own analysis. These are combined to form a single refined set of parameters for the remaining features in MSI 3.0 (Intent-based Spatial Balancer and DNN 2.0) to act upon. This sets the level of support required by the user intent and the sound environment (see Figure 3). The help system provides the appropriate amount of help as different support is needed from the hearing aid depending on what the user is doing in the specific situation. When the user is, for instance, navigating the room, more awareness of the surroundings is useful, when participating in a dynamic group conversation, it is important to jump in at the right time, or when engaging in an intimate conversation, full attention is focused on the conversation partner. Irrespective of the interpreted intention and the help provided, sounds from all directions are available at all times, but the balance between different sounds can vary.

The system performs an adaptation of support where it transitions slowly and smoothly within the range without any jumps between fixed modes to ensure clear and comfortable sounds for the hearing aid user. The range in which the help system can transition depends on the individual settings chosen in Oticon Genie 2 and on the sound environment.

Figure 4 shows the hearing aid SNR enhancement in dB measured with default settings. The range within which the adaptation of support provided by the 4D Sensor technology can transition is depicted as the light blue area. The range of output SNR enhancement on the vertical-axis (more contrast between speech and surrounding sounds vs. more awareness of surrounding sounds) varies as a function of input SNR on the

Figure 3. The process of the 4D Sensor technology: The 4D sensors gather input in the four dimensions, the input is evaluated to interpret the user intention and combined into a single set of parameters which define the appropriate amount of help for the remaining features in MSI 3.0 to act upon.

horizontal-axis (complexity of listening environment). The provided help depends on the user's intentions as interpreted by the analysis of head and body movements, conversation activity, and acoustic environment. The dark blue solid lines reflect the maximum and minimum adaptation of support. The blue dotted line is the curve corresponding to the same default settings in Oticon Genie 2 but with sensor technology Off, hence, disregarding user intent.

For more explanation on the measurements presented in Figure 4 see the section 'Larger adaptation of support to individual listening needs' further below.

Fitting with 4D Sensor technology

4D Sensor technology is available in the General and Speech in Noise programs when fitting in Oticon Genie 2. For 4D Sensor technology to be available, the adaptive functionality in MSI 3.0 must be activated. The adaptive functionality and the sensor technology are both on by default.

New generation DNN

The structure of a DNN is inspired by how our brain is organized, namely the neurons and their corresponding synapses. The neural network uses the iterative learning from a huge quantity of real-world data to learn about sound and how to process it. The iterative learning of the DNN replaces the old ways of processing sound which were based on a strict set of pre-established, man-made algorithms. Our DNN approach takes sound processing and noise handling out of the lab and into the real world.

Neural networks are based on deep learning algorithms. Deep learning algorithms take large amounts of data, referred to as training samples, and develop a system that can learn from them. The uniqueness of neural networks stems from their architectural similarity to the brain. Within neural networks, there is a basic unit called the neuron. A neuron's purpose, much like a relay neuron in the brain, is to receive information, store it, and finally pass it on to the next neuron. A group of neurons form a layer. Multiple specialized, interlinked layers form the neural network consisting of an input layer at the start, hidden layers in the middle, and output layer at the end.

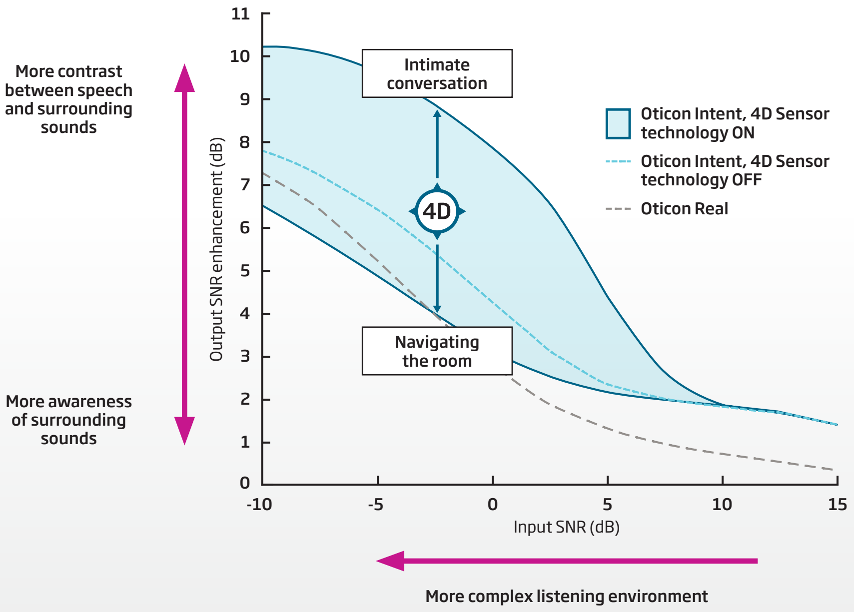

Figure 4. SNR enhancement measured in Oticon Intent and Oticon Real with default settings. The light blue area shows the adaptation of support as a flexible SNR enhancement with sensor technology On in Oticon Intent, and the blue dotted line shows the fixed SNR enhancement with sensor technology Off in Oticon Intent. The grey dashed line shows the fixed SNR enhancement in Oticon Real.

The output layer produces a result that is an acoustic signal that we can hear. This is the most basic class of neural networks. The input and output layers have 24 neurons corresponding to the 24 processing channels. Noise suppression is applied adaptively according to input from the different sensors mentioned previously.

DNN 2.0 is the new generation of deep neural network developed with completely new training rounds. To give insights into how the new training was performed this section will review the four steps in the training process. The steps are as shown in Figure 5: Input (A) – the DNN receives the input on the sound scene, Forward propagation (B) – the DNN processes the sound scene, Output (C) – the DNN produces the audible output, and Backward propagation (D) – feedback is provided for further improvement of the processing in the DNN.

Our goal was to train the DNN on sound scenes so that it could solve the task of balancing sound sources by preserving cues and attenuating noise. The large amount of data needed for this training was recorded in different sound scenes across a wide array of listening environments representing sound scenes that listeners would typically be exposed to in their everyday lives. We used a specialized spherical microphone, capable of capturing 360 degrees of sounds to provide the DNN with a spatially accurate and detailed sound scene and to train it on the full sound scene.
Compared to the training of the previous generation DNN steps A, C, and D in the training process were improved, which leads to improvement of step B.

During the input step (A), neurons receive information of one sound scene and store them. The sound samples used for training of DNN 2.0 are more diverse than the ones used for training of the previous generation DNN. DNN 2.0 training primarily used complex sound scenes to ensure a better response to the sound environments the DNN will have to process in the real world.

Next, forward propagation (B) takes the data input from each neuron and passes it on to the next layer. The volume of information passed on is dependent on the inter-neuron connection strength. DNN 2.0 has 24 channels in the input and output layers as does the rest of the processing scheme in MSI 3.0. What happens in this step is defined by the DNN, and it is working with no restrictions from the developers. So, by providing new training rounds with updates to the other steps in the training process this step automatically gets updated.

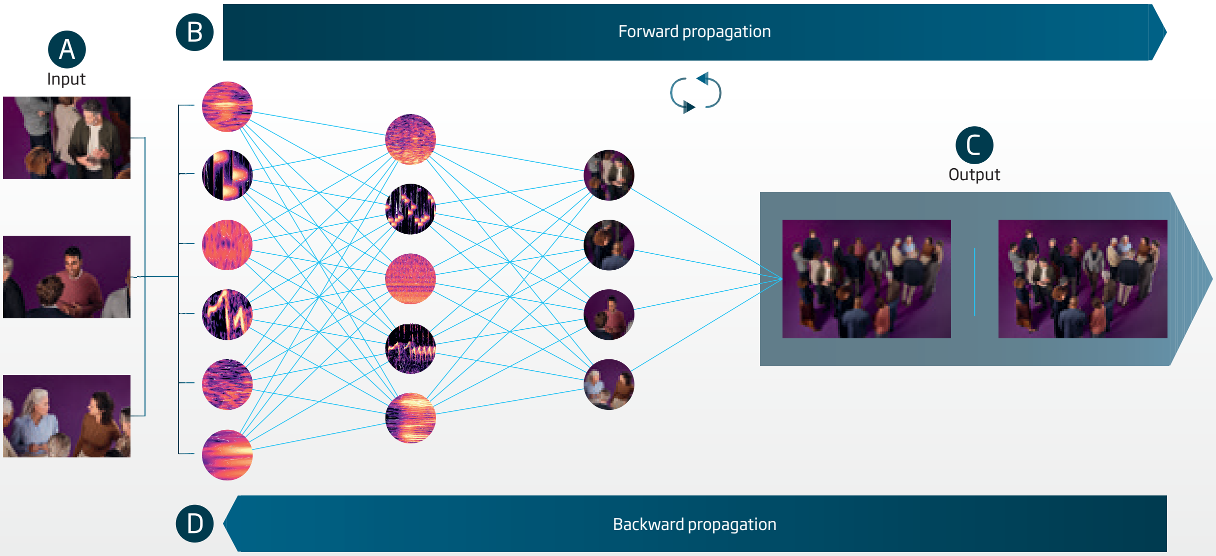

Figure 5. The flow in the training of a DNN. See text for explanation.

After the forward propagation step is complete, DNN 2.0 makes an output prediction (C) of the sounds it determines should be enhanced or suppressed in the sound scene. This output is analyzed and compared to a desired target: a real sound scene. The analysis is made in 256 channels (compared to 24 channels for the previous generation DNN) to ensure that every mistake made in the processing by DNN 2.0 during training is captured and corrected. The comparison to the target is performed using improved definitions of training parameters which emphasize preservation of speech, attenuation of noise, minimization of loss of energy from speech-like sounds, stable gain, keeping all sounds clear and undistorted, and that DNN 2.0 only activates when necessary, as indicated by the HCP during fitting (or by the Oticon Genie 2 default settings).

As the final step in the process, we teach DNN 2.0 to learn from its mistakes and adapt. This action drives the process of backward propagation (D), in which DNN 2.0 tweaks the individual connections between neurons to better suppress the correct sounds. The process is iterated for all sound scenes until a plateau is reached. The whole training process is then repeated with new sound scenes. This process teaches DNN 2.0 to identify the features of each sound and to better distinguish between them. Over time, DNN 2.0's ability to emphasize and suppress meaningful and non-meaningful sounds improves.

For a more detailed description of the construction and training of a DNN see Brændgaard (2020) and Andersen et al. (2021).

DNN 2.0 is improved compared to the previous generation DNN. It produces a clearer output while preserving more of the original cues. Due to the increased 256-channel analysis in the training process, DNN 2.0 can also provide more attenuation (up to 12 dB in difficult environments) without introducing distortion to the sound. For more information on the benefits of DNN 2.0 see the section 'Testing the upgraded DNN' further below.

DNN 2.0 is prescribed in Oticon Genie 2 under the Neural Noise Suppression section. The default setting is set to 2 dB for Easy environments and 10 dB for Difficult environments. The settings can be personalized together with other MSI 3.0 settings based on the personalization questions or the Audible Contrast Threshold (ACT) value (Santurette & Laugesen, 2023) or based on client input and needs. The possible settings are 0, 2, 4, and 6 dB for Easy environments and 6, 8, 10, and 12 dB for Difficult environments.

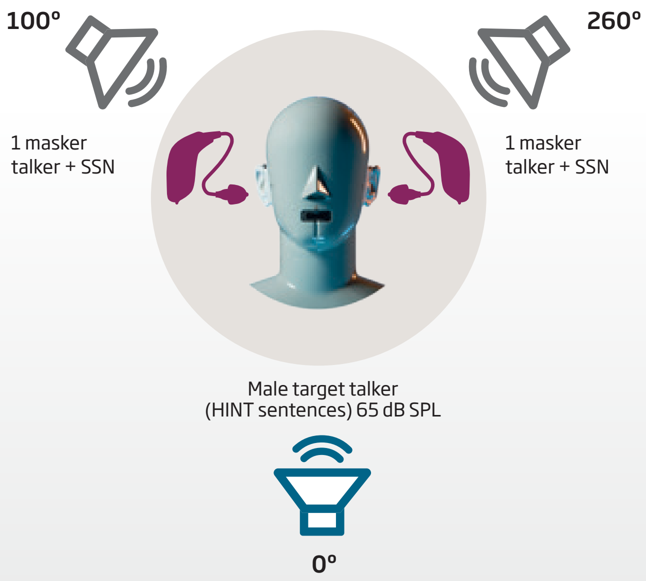

Figure 6. Measurement set-up with a target talker from the front and interfering talkers mixed with speech-shaped noise (SSN) from the sides.

Technical evaluation of Oticon Intent

To assess the performance of Oticon Intent, we conducted a systematic technical evaluation and compared it to our premium legacy device, Oticon Real. The primary goal was to evaluate the hearing aid's capability to provide listening support for speech in complex listening situations. For this purpose, we set up a controlled test procedure in an acoustically treated sound studio in which we simulated an intimate conversation. This involved positioning a head-and-torso simulator (HATS) at the center of the room, 1.6 meters away from a loudspeaker (0°) reproducing a target speech signal. Additionally, two loudspeakers were placed at 100° and 260°, as illustrated in Figure 6, and served as maskers. The masker loudspeakers presented interfering talkers mixed with stationary speech-shaped noise (SSN). The target was consistently played at 65 dB SPL, while the sound of the maskers varied from 50 to 75 dB SPL . This enabled us to explore a range of listening scenarios from very simple to very complex.

We recorded speech-in-noise signals using the HATS wearing either Oticon Intent or Oticon Real. We used closed MicroMolds fitted to the HATS' ear canal to minimize interference from sounds not processed by the hearing aid. Both hearing aids were adjusted to compensate for a moderate hearing loss based on the N3 standard audiogram (Bisgaard et al., 2010), for an adult with long-term experience set in Genie 2. All features were maintained at their default, prescription-based configurations, except for feedback management and binaural link which were deactivated. This was needed due to the highly specialized recording methods used and discussed below. Additionally, linear amplification was used to ensure that the results reflect the effect of MSI only. We examined the HATS recordings to investigate the hearing aid performance on the following:

Help in noise: We used output SNR measurements to quantify the contrast created between speech and noise to evaluate the hearing aid's effectiveness in enhancing speech clarity while reducing disruptive noise.

Access to speech cues: We used an objective speech intelligibility metric to quantify access to speech cues and spectrogram analysis to examine the hearing aid's ability to preserve speech details in the presence of noise.

Larger adaptation of support to individual listening needs

Oticon Intent with 4D Sensor technology can now address the user's listening needs across a broader range of listening situations by providing a more pronounced contrast between the desired signals and background noise. We evaluated this enhancement using output SNR measurements. The output SNR was calculated using the phase inversion method by Hagerman & Olofsson (2004). The SNR was then weighted across all frequency bands using Speech Intelligibility Index weights (ANSI S3.5, 1997) corresponding to each band's center frequency before computing the output SNR.

The output SNR was measured for Oticon Intent with 4D Sensor technology both On and Off, as well as for the previous generation premium device, Oticon Real. Figure 4 illustrates the output SNR enhancement for various input SNR levels, representing different complexities in the listening environment. The results demonstrate that Oticon Intent provides superior support to users, with an improvement of up to 5 dB compared to Oticon Real when having an intimate conversation in a noisy listening environment. When the 4D Sensor technology is deactivated, an improvement of up to 1.5 dB (comparing the dotted blue curve and dashed grey curve in Figure 4) is mainly attributed to the updated DNN 2.0 on the new Sirius platform. The results presented in Figure 4 also highlight that Oticon Intent with 4D Sensor technology enabled provides a larger adaptation of support of up to 5 dB based on the user's listening intent. A higher output SNR indicates improved speech clarity, while a lower output SNR allows for greater access to surrounding sounds. Therefore, Oticon Intent can consistently provide users with the appropriate level of support and access to surrounding sounds within the same listening environment, whether engaged in an intimate conversation or navigating the room.

Let's now examine in detail the results obtained at a 0 dB input SNR in Figure 4. In this challenging listening situation, where the intensity of the target speech signal is equal to that of the background noise, Oticon Intent outperformed Oticon Real. This is illustrated in Figure 7, where it is shown that the output SNR for Oticon Intent is 5 dB larger when 4D Sensor technology is On and 1.5 dB when it is Off, compared to Oticon Real. These results highlight that Oticon Intent excels in making the target signal stand out from the background noise, representing an enhancement in speech clarity that was found to significantly improve the user's ability to understand speech (Bianchi/Eskelund et al., 2024).

More access to speech cues for better speech understanding

We used the Speech Intelligibility Index (SII) (ANSI S3.5, 1997) to evaluate the impact of the output SNR enhancements on speech intelligibility. The SII is a measure of predicted speech intelligibility and estimates the degree to which speech can be comprehended by taking into consideration the audibility of the speech signal. The SII is derived by weighting factors like the clarity of the speech signal, the presence of background noise, and the hearing ability of the listener. A higher SII value indicates an increased access to speech cues, indicating a greater chance of intelligibility.

Figure 8 shows the SII in % calculated at 0 dB input SNR for Oticon Real and Oticon Intent with 4D Sensor technology On and Off. The results show SII values for Oticon Intent of 58% when 4D Sensor technology is On and 48% when it is Off. As a comparison, the SII calculated for a person with normal hearing thresholds without hearing aids would only reach a value of 40% in the same challenging situation . Compared to Oticon Real, the SII values obtained for Oticon Intent represent a substantial 35% relative increase in SII with 4D Sensor technology On and a 12% SII increase with 4D Sensor technology Off. Overall, these results demonstrate that Oticon Intent offers superior access to speech cues compared to Oticon Real, enhancing the user's ability to comprehend spoken words and communicate more effectively. This leads to an improved overall listening experience, even in challenging environments.

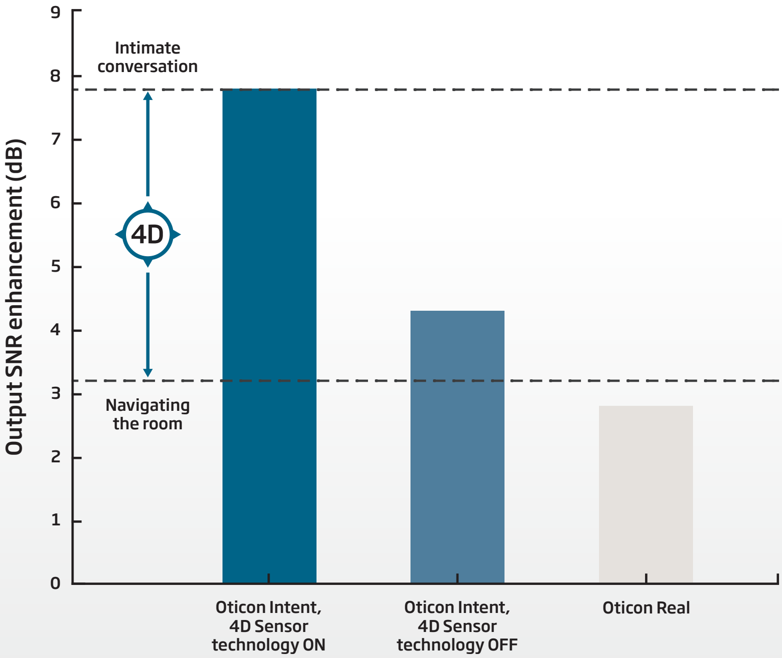

Figure 7. Output SNR enhancement measured at 0 dB input SNR for Oticon Intent with 4D Sensor technology On and Off, and for Oticon Real.

Better speech clarity and enhancement of speech details

To ensure optimal support for speech understanding, it is crucial that hearing aids accurately capture and preserve all details of the speech signal. This can be studied using a spectrogram, which provides a detailed time-frequency analysis of hearing aid recordings. Represented commonly as a heat map, the spectrogram illustrates the distribution of sound power (in dB) per frequency (on the vertical axis) across time (on the horizontal axis). In the spectrogram, the sound power magnitude is represented by brightness variations. Dark regions indicate low sound power (quiet areas) and bright regions indicate high sound power (loud areas).

We performed spectrogram analysis for a very demanding scenario: an intimate conversation in the presence of noise and competing talkers at a 0 dB SNR. Here, we wanted to compare the ability to preserve speech details in Oticon Intent and Oticon Real. Figure 9 displays the spectrogram of a sentence in noise recorded at the output of Oticon Real (Panel A) and Oticon Intent with 4D Sensor technology On (Panel B), as depicted in the experimental setup shown in Figure 7. Both hearing aids were fitted using default prescription settings.

When comparing the spectrograms, it can be observed that Oticon Intent provides a greater reduction in noise, as more dark areas are visible in the spectrogram between speech elements across the frequency range.

Additionally, the speech signal can be identified more clearly in the spectrogram obtained for Oticon Intent. Note for example the finer, more precise representation of the harmonics of the vowel "a" and a better preservation of its frequency content for Intent in Panel B. Hence, Oticon Intent better preserves the speech details while reducing more background noise than Oticon Real, thus providing users with improved speech clarity.

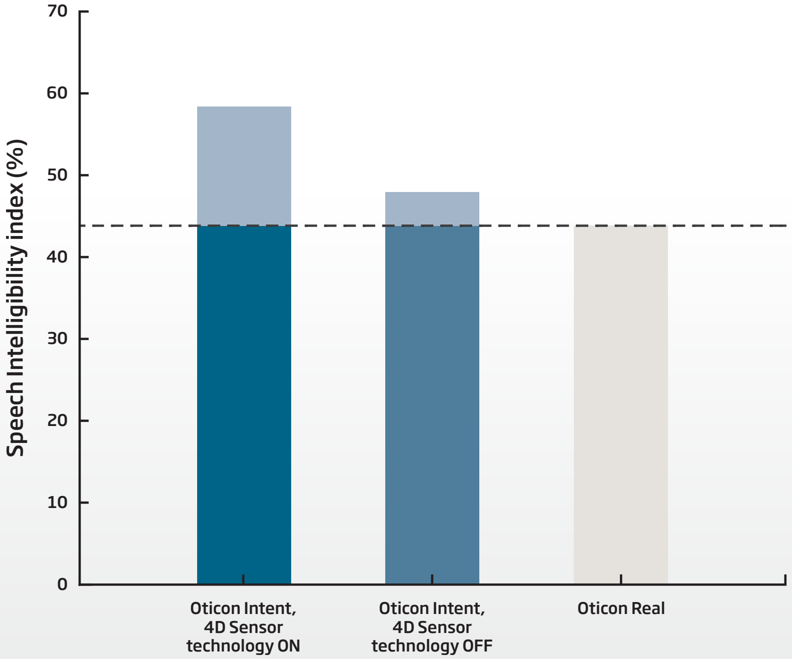

Figure 8. Speech intelligibility index measured at 0 dB input SNR for Oticon Intent with 4D Sensor technology On and Off, and for Oticon Real.

Testing the upgraded DNN

Finally, we put DNN 2.0, a standout innovation of Oticon Intent, to the test. To conduct a fair evaluation between DNN 2.0 and the previous generation DNN 1.0, we isolated their effects by simulating their sound processing using a 0 dB input SNR. Figure 10 presents spectrograms illustrating clean speech (Panel A) and speech in noise (Panel B) processed by the HA with gain only (MSI Off). Dark regions indicate low sound power (quiet areas), and bright regions indicate high sound power (loud areas), like Figure 9. The lower part of Figure 10 shows how the two versions of the DNN process different areas of the noisy speech signal when Neural Noise Suppression is turned on. The areas that are attenuated (red areas), preserved (white areas), and enhanced (blue areas) are illustrated for DNN 1.0 in Oticon Real (Panel C) and for DNN 2.0 in Oticon Intent (Panel D).

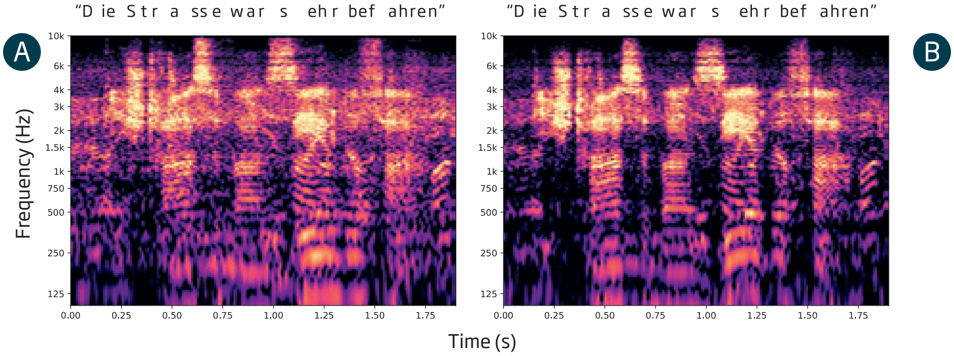

Figure 9. Spectrogram of the German sentence “Die Strasse war sehr befahren” (“The road was very busy”) recorded at 0 dB input SNR for Oticon Real (Panel A) and Oticon Intent with 4D Sensor technology (Panel B). Dark regions indicate low sound power (quiet areas) and bright regions indicate high sound power (loud areas)

The comparison between Oticon Real and Oticon Intent reveals differences in noise suppression capabilities. The spectrogram for DNN 1.0 in Oticon Real (Panel C) shows that noise suppression is more limited beyond 7.5-8 kHz than in Oticon Intent (Panel D). The contrast created between speech and noise is also larger with DNN 2.0 in Oticon Intent than with DNN 1.0 in Oticon Real: On the same scale, darker blue and darker red areas are visible in Panel D compared to Panel C. Moreover, some areas with speech information are not fully enhanced with DNN 1.0, while DNN 2.0 proves to be more effective in preserving speech details across the frequency range. As a result, the Intent spectrogram in Panel D mirrors the clean speech spectrogram of Panel A more closely. For example, see how the vowel “a” has a more consistent enhancement across frequency and the consonant “s” is enhanced towards higher frequencies in Panel D than in Panel C. Overall, DNN 2.0 thus provides better contrast between speech and noise as well as a more precise enhancement of speech details.

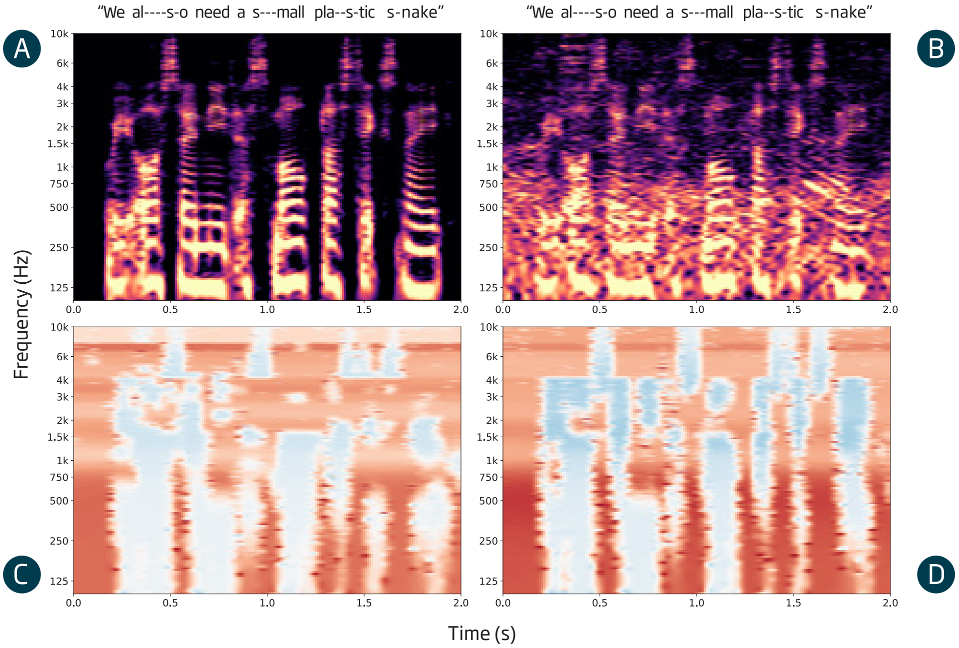

Figure 10. Spectrogram of the English sentence "We also need a small plastic snake" simulated at 0 dB input SNR. Panel A shows the spectrogram of the clean speech signal. Panel B shows the spectrogram of the same speech signal in babble noise processed by the hearing aid without noise reduction. Panels C and D show the contrast between speech and noise created by DNN 1.0 (C) and DNN 2.0 (D) for the noisy signal shown in Panel B. In Panels A and B, dark regions indicate low sound power (quiet areas), and bright regions indicate high sound power (loud areas). Panels C and D show the time frequency areas that are attenuated (red areas), preserved (white areas), and enhanced (blue areas) by DNN 1.0 and DNN 2.0, respectively.

Conclusion

This whitepaper presented the two major technological advances of MSI 3.0 in Oticon Intent, 4D Sensor technology and DNN 2.0. 4D Sensor technology provides the user with better personalized support based on their listening needs in any given listening situation. DNN 2.0 provides better contrast between the target signal and background noise in the sound scene and better enhancement of speech details.

The results of a technical evaluation clearly showed that Oticon Intent outperforms Oticon Real:

- Oticon Intent with 4D Sensor technology provides an adaptation of support of up to 5 dB within the same sound environment based on the user's listening intent.
- Oticon Intent provides up to 1.5 dB more output SNR than Oticon Real with the inclusion of the Sirius platform and the DNN 2.0, and a 5-dB increase in output SNR with the addition of 4D Sensor technology. Hence, Oticon Intent provides users with better speech clarity and more contrast between speech and background noise.
- Oticon Intent gives 35% more access to speech cues than Oticon Real, providing users with clearer sound scenes where the speech details are better preserved and more background noise is attenuated.

Now with added personalized support in every sound scene, these technical benefits empower users to more easily attend to and focus on the task of interest, as proven by the numerous BrainHearing™ benefits observed in clinical studies with hearing aid users (Bianchi/Eskelund et al., 2024; Andersen et al., 2021; Alickovic et al., 2021).

References

1. Akeroyd, M. A. (2014). An Overview of the Major Phenomena of the Localization of Sound Sources by Normal-Hearing, Hearing-Impaired, and Aided Listeners. Trends in Hearing Vol. 18, pp. 1-7.
2. Alickovic, E., Ng, E. H. N., Fiedler, L., Santurette, S., Innes-Brown, H., & Graversen, C. (2021). Effects of hearing aid noise reduction on early and late cortical representations of competing talkers in noise. Frontiers in neuroscience, 15, 636060.
3. Andersen, A. H., Santurette, S., Pedersen, M. S., Alickovic, E., Fiedler, L., Jensen, J., & Behrens, T. (2021). Creating clarity in noisy environments by using deep learning in hearing aids. Seminars in Hearing 42(3), 260-281.
4. ANSI S3.5. (1997). American national standards methods for the calculation of the articulation index. American National Standards Institute.
5. Bianchi, F./Eskelund, K., Zapata-Rodriguez, V., Sanchez Lopez, R., & Gade, P. (2024). Oticon Intent™ - Clinical evidence. BrainHearing™ benefits of the 4D Sensor Technology. Oticon whitepaper.
6. Bisgaard, N., Vlaming, M. S., & Dahiquist, M. (2010). Standard audiograms for the IEC 60118-15 measurement procedure. Trends in Amplification, 14(2), 113-120. doi:10.1177/1084713810379609
7. Brændgaard, M. (2020). An introduction to MoreSound Intelligence™. Oticon tech paper.
8. Gade, P.A., Brændgaard, M., Flocken, H., Preszcator, D., & Santurette, S. (2023). Wind & Handling Stabilizer – Evidence and user benefits. Improved wind and handling noise removal for better clarity. Oticon whitepaper.
9. Hadley, L. V., Brimijoin, W. O., & Whitmer, W. M. (2019). Speech, movement, and gaze behaviours during dyadic conversation in noise. Scientific reports, 9(1), 1-8.
10. Hadley, L. V., Whitmer, W. M., Brimijoin, W. O., & Naylor, G. (2020). Conversation in small groups: Speaking and listening strategies depend on the complexities of the environment and group. Psychonomic Bulletin & Review, 28(2), 632-640.
11. Hadley, L. V., & Culling, J. F. (2022). Timing of head turns to upcoming talkers in triadic conversation: Evidence for prediction of turn ends and interruptions. Frontiers in Psychology, 13.
12. Hagerman, B., & Olofsson, Å. (2004). A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. Acta Acustica United with Acustica, 90(2), 356-361.
13. Higgins, N. C., Pupo, D. A., Ozmeral, E. J., & Eddins, D. A. (2023). Head movement and its relation to hearing. Frontiers in Psychology, 14.
14. Picou, E. M. (2020). MarkeTrak 10 (MT10) survey results demonstrate high satisfaction with and benefits from hearing aids. In Seminars in hearing (Vol. 41, No. 01, pp. 021-036). Thieme Medical Publishers.
15. Santurette, S., & Laugesen, S. (2023). Audible Contrast Threshold (ACT™). A language-independent diagnostic test to quantify real-life speech-in-noise ability and personalize help-in-noise settings in hearing aids. Oticon whitepaper.

24-569952 15500-0299/02.24

oticon.com

Oticon is part of the Demant Group.

oticon
life-changing **technology**