Speech-In-Noise Test Results for Oticon Opn

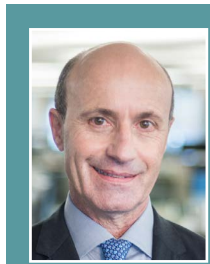
By DOUGLAS L. BECK, AuD, and NICOLAS LEGOFF, PhD

A discussion of signal-to-noise ratio problems and solutions, and a summary of findings of a study involving Oticon Opn hearing aids.

The most common problem experienced by people with hearing loss and people wearing traditional hearing aids is not that sound isn't loud enough. The primary issue is understanding speech-in-noise (SIN). Hearing is the ability to perceive sound, whereas listening is the ability to make sense of, or assign meaning to, sound.

As typical hearing loss (ie, presbycusis, noise-induced hearing loss) progresses, outer hair cell loss increases and higher frequencies become increasingly inaudible. As hearing loss progresses from mild (26 to 40 dB HL) to moderate (41 to 70 dB) and beyond, distortions increase, disrupting spectral, timing, and loudness perceptions. The amount of distortions vary, and listening results are not predictable based on an audiogram, nor are they predictable based on word recognition scores from speech in quiet.[1,2]

To provide maximal understanding of speech in difficult listening situations, the goal of hearing aid amplification is twofold: Make speech sounds audible and increase the signal-to-noise ratio (SNR).[1] That is, the goal is to make it easier for the brain to identify, locate, separate, recognize, and interpret speech sounds. "Speech" (in this article) is the spoken signal of primary interest and "noise" is the secondary sound of other people speaking (ie, "speech babble noise").




Douglas L. Beck, AuD, is the Executive Director of Academic Sciences at Oticon Inc, Somerset, NJ, and **Nicolas LeGoff, PhD,** is the Senior Researcher at Oticon A/S, in Kongebakken, Denmark.

The essence of the SIN problem is that the primary speech sounds and secondary sounds (ie, noise) are essentially the same thing! That is, both speech and speech babble noise originate with human voices with similar spectral and loudness attributes, rendering the "SIN" problem difficult to solve.

Signal-to-Noise Ratio (SNR) Loss

Killion[1] reported people with substantial difficulty understanding speech-in-noise may have significant "SNR loss." Of note, the SNR loss is unrelated to, and cannot be predicted from, the audiogram. Killion defines SNR loss as the increased SNR needed by an individual with difficulty understanding speech in noise, as compared to someone without difficulty understanding speech in noise. He reports people with relatively normal/typical SNR ability may require a 2 dB SNR to understand 50% of the sentences spoken in noise, whereas people with mild-moderate sensorineural hearing loss may require an 8 dB SNR to achieve the same 50% performance. Therefore, for the person who needs an 8 dB SNR, we subtract 2 dB (normal performance) from their 8 dB score, resulting in a 6 dB SNR loss.

Wilson[2] evaluated more than 3,400 veterans. He, too, reported speech in quiet testing does not predict speech-in-noise ability, as the two tests (speech in quiet, speech-in-noise) reflect different domains of auditory function. He suggested the Words-in-Noise (WIN) test should be used as the "stress test" for auditory function.

Beck and Flexer[3] reported, "Listening is where hearing meets the brain." They said the ability to listen (ie, to make sense of sound) is arguably a more important construct, and reflects a more realistic representation of how people manage in the noisy real world, than does pure-tone hearing thresholds reflected on an audiogram. Indeed, many animals (including dogs and cats) "hear" better than humans. However, humans are atop the food chain not because of their hearing, but due to their ability to *listen*—to apply meaning to sound. As such, a speech-in-noise test is more of a *listening test* than a hearing test. Specifically, listening in noise depends on a multitude of cognitive factors beyond loudness and audibility, and includes speed of processing, working memory, attention, and more.

McShefferty et al[4] report SNR plays a vital role for hearing-impaired and normal-hearing listeners. However, the smallest detectable SNR difference for a person with normal hearing is about 3 dB, and they report for more clinically-relevant tasks, a 6 to 8 dB SNR may be required.

In this article, we'll address concepts and ideas associated with understanding speech-in-noise, and, importantly, we'll share results obtained while comparing SIN results with Oticon Opn and two major competitors, in a realistic acoustic environment.

Traditional Strategies to Minimize Background Noise

As noted above, the primary problem associated with hearing loss and hearing aid amplification in understanding speech, is noise. The major focus over the last four decades or so has been to reduce background noise. Two processing strategies have been employed in modern hearing aid amplification systems to minimize background noise: digital noise reduction and directional microphones.

Digital Noise Reduction (DNR). Venema[5] reports (p 335) the goal of DNR is noise reduction. DNR systems can recognize and reduce the signature amplitude of steady-state noise using various amplitude modulation (AM) detection systems. AM systems can identify differences in dynamic human speech as opposed to steady state noise sources, such as heating-ventilation/air-conditioning (HVAC) systems, electric motor and fan noise, 60 cycle noise, etc.[6] However, DNR systems are less able to attenuate secondary dynamic human voices in close proximity to the hearing aid, such as nearby loud voices in restaurants, cocktail parties, etc, because the acoustic signature of people we desire to hear and the acoustic signature of other people in close proximity (ie, speech babble noise) are essentially the same.

Venema[5] states (p 331) the broadband spectrum of speech and the broadband spectrum of noise intersect and overlap, and, consequently, are very much the same thing. Nonetheless, given multiple unintelligible speakers and a physical distance of perhaps 7-10 meters from the hearing aid microphone, Venema and Beck[7] report the secondary signal (ie, speech babble noise) may present as (or morph into) more of steady state "noise-like" signal and be attenuated via DNR—providing additional comfort to the listener.

McCreery et al[8] reported their evidence-based systematic review to examine "the efficacy of digital noise reduction and directional microphones." They searched 26 databases seeking contemporary publications (published after 1980), resulting in four articles on DNR and seven articles on directional microphone studies. Of tremendous importance, McCreery and colleagues[8] concluded DNR did not improve or degrade speech understanding.

Beck and Behrens[9] reported DNR may offer substantial "cognitive" benefits, including more rapid word learning rates for some children, less listening effort, improved recall of words, enhanced attention, and quicker word identification, as well as better neural coding of words. They suggested typical hearing aid fitting protocols should include activation of the DNR circuit.

Pittman et al[10] reported DNR provides little or no benefit with regard to improved word recognition in noise. Likewise, McCreery and Walker[11] note DNR circuits are routinely recommended for the purpose of improving listening comfort, but restated that, with regard to school-age children with hearing loss, DNR neither improved nor degraded speech understanding.

Directional Microphones (DMs). DMs (or "D-mics") are the only technology proven to improve SNR. However, the likely perceived benefit from DMs in the real world, due to the prevalence of open canal fittings, is often only 1-2 dB.[7] Directivity Indexes (DIs) indicating 4-6 dB improvement are generally not "real world" measures. That is, DIs are generally measured on manikins, in an anechoic chamber, based on pure tones, and DIs quantify and compare sounds coming from the front versus all other directions.

Nonetheless, although an SNR improvement of 1-2 dB may appear small, every 1 dB SNR improvement may provide a 10% word recognition score increase.[5,12,13]

In their extensive review of the published literature, McCreery et al[8] reported that, in controlled optimal situations, DMs *did* improve speech recognition; yet they cautioned the effectiveness of DMs in real-world situations was not yet well-documented and additional research is needed prior to making conclusive statements.

Brimijoin and colleagues[14] stated directionality potentially makes it difficult to orient effectively in complex acoustic environments. Picou et al[15] reported directional processing did reduce interaural loudness differences (ILDs), and localization was disrupted in extreme situations without visual cues. Research by Best and colleagues[16] found narrow directionality is only viable when the acoustic environment is highly predictable. Mejia et al[17] indicated, as beam-width narrows, the possible SNR enhancement increases; however, as the beam-width narrows, the possibility also increases that "listeners will misalign their heads, thus decreasing sensitivity to the target..." Geetha et al[18] reported "directionality in binaural hearing aids without wireless communication" may disrupt interaural timing and interaural loudness cues, leading to "poor performance in localization as well as in speech perception in noise..."

New Strategies to Minimize Background Noise

Research by Shinn-Cunningham and Best[19] suggests the ability to selectively attend depends on the ability to analyze the acoustic scene and form perceptual auditory objects properly. If analyzed correctly, attending to a particular sound source (ie, voice) while simultaneously suppressing background sounds may become easier as one successfully increases focus and attention.

Therefore, the purpose of a new strategy for minimizing background noise should be to facilitate the ability to attend to one primary sound source and switch attention when desired—which is what Oticon's recently-released Multiple Speaker Access Technology (MSAT) has been designed to do.

Multiple Speaker Access Technology (MSAT). In 2016, Oticon introduced MSAT. The goals of MSAT are to selectively reduce disturbing noise while maintaining access to all distinct speech sounds and to support the ability of the user to select the voice they choose to attend to.

MSAT represents a new class of speech enhancement technology and is intended to replace current directional and noise reduction systems. MSAT does not isolate one talker; it maintains access to all distinct speakers. MSAT is built on three stages of sound processing:

1) **Analyze** provides two views of the acoustic environment. One view is from a 360 degree omni microphone; the other is a rear-facing cardioid microphone to identify which sounds originate from the sides and rear. The cardioid mic provides multiple noise estimates to provide a spatial weighting of noise.

2) **Balance** increases the SNR by constantly acquiring and mixing the two mics (similar to auditory brainstem response or radar) to obtain a rebalanced soundscape in which the loudest noise sources are attenuated. In general, the most important sounds are present in the omni view, while the most disturbing sounds are present in both omni and cardioid views. In essence, cardioid is subtracted from omni, to effectively create nulls in the direction of the noise sources, thus increasing the prominence of the primary speaker.

3) **Noise Removal** (NR) provides very fast removal of noise between words and up to 9 dB of noise attenuation. Importantly, if speech is detected in any band, Balance and NR systems are "frozen" so as to not isolate the front talker, but to preserve all talkers.

The Oticon Opn system uses MSAT and, therefore, it is neither directional nor omnidirectional. To be clear, the goal of DMs is to pick up more sound from the front of the listener, as compared to sounds from other angles, and D-mics do not help one tell the direction of sounds, nor do they increase the intensity of sounds coming from the front—or as Venema states (p 316), "they simply decrease the intensity of sounds coming from the sides and rear, or relative to sounds coming from the front..."[5] According to Geetha et al,[18] DMs are designed to provide attenuation of sounds emerging from the sides of the listener, and Venema[5] states omnidirectional microphones are "equally sensitive to sounds coming from all directions..." Thus, omnidirectional microphones theoretically have a DI of 0.

Therefore, MSAT is neither an omni or a directional system, but does represent a new technology.

Spatial Cues. Of significant importance to understanding SIN is the ability to know where to focus one's attention. Knowing "where to listen" is important with regard to increasing focus and attention. Spatial cues allow the listener to know where to focus attention and, consequently, what to ignore or dismiss.[20] The specific spatial cues required to improve the ability to understand SIN are interaural level

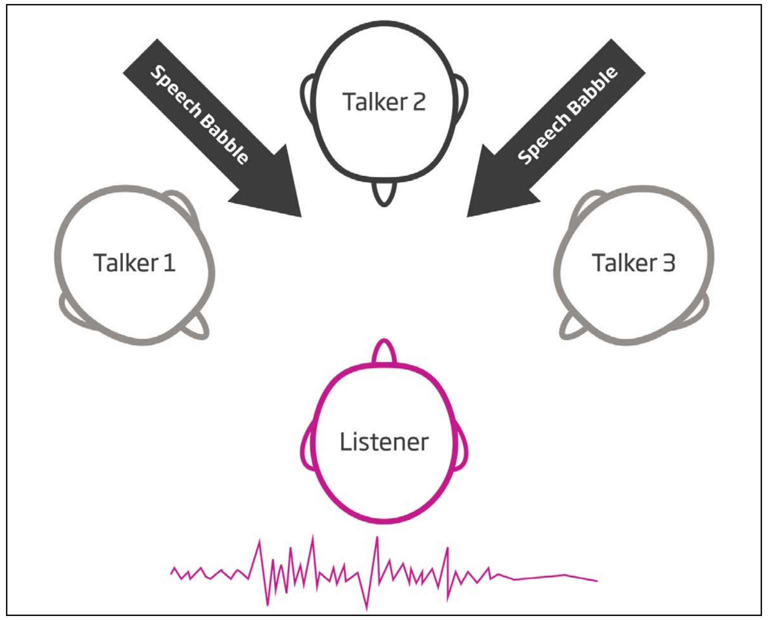


Figure 1. Schematic representation of the acoustical conditions of the experiment.

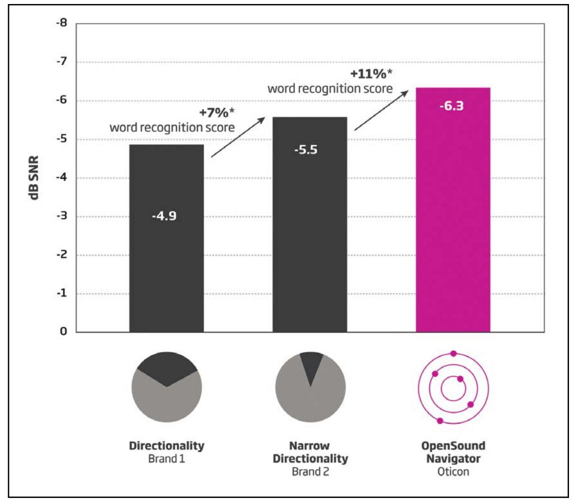


Figure 2. Overall results. Average improvement with respect to SNR and anticipated improvement in word recognition. Bar heights represent the average SRT-50 from 25 listeners. Each improvement bar is statistically different from the others at p<0.05. The one-way arrow (+7%) from -4.9 to -5.5 represents the improvement in word recognition scores from directionality to narrow directionality, and the one-way arrow (+11%) from -5.5 to -6.3 represents the anticipated improvement in word recognition scores from narrow directionality to OpenSound Navigator.

differences (ILDs) and interaural time differences (ITDs), such that the left and right ears receive unique spatial cues.

Sockalingham and Holmberg[21] presented laboratory and real-world results demonstrating "strong user preference and statistically significant improved ratings of listening effort and statistically significant improvements in real-world performance resulting from Spatial Noise Management."

Beck[22] reported spatial hearing allows us to identify the origin/location of sound in space and attend to a "primary sound source in difficult listening situations" while ignoring secondary sounds. Knowing "where to listen" allows the brain to maximally listen in difficult/noisy listening situations, as the brain is better able to compare and contrast unique sounds from the left and right ears—in real time—to better determine where to focus attention.

Geetha et al[18] stated speech in noise is challenging for people with sensorineural hearing loss (SNHL). To better understand speech in noise, the responsible acoustic cues are ITDs and ILDs. They report "...preservation of binaural cues is...crucial for localization as well as speech understanding..." and directionality in binaural amplification without synchronized wireless communication can disrupt ITDs and ILDs, leading to "...poor performance in localization...(and) speech perception in noise..."[18]

Oticon Opn uses Speech Guard™ LX to improve speech understanding in noise by preserving clear sound quality and speech details, such as spatial cues via ILDs and ITDs. It uses adaptive compression and combines linear amplification with fast compression to provide up to a 12 dB dynamic range window, preserving natural amplitude cues in speech signals. Likewise, Spatial Sound™ LX helps the user locate, follow, and shift focus to the primary speaker via advanced technologies to provide a more precise spatial awareness for identify where sounds originate.

The Oticon OPN SNR Study

In Pittman et al's research,[10] they stated, "Like the increasingly unique advances in hearing aid technology, equally unique approaches may be necessary to evaluate these new features..." The purpose of this study was to compare the results obtained using Oticon Opn to two other major manufacturers with regard to listeners' ability to understand speech in noise in a lab-based, yet realistic, background noise situation.

Admittedly, despite the fact that no lab-based protocol perfectly replicates the real world, this study endeavored to realistically simulate what a listener experiences as three people speak sequentially from three locations, without prior knowledge as to which person would speak next.

Methods. A total of 25 German native-speaking participants (ie, listeners) with an average age of 73 years (SD: 6.2 years) with mild-to-moderate symmetric SNHL underwent listening tasks. Each participant wore the Oticon Opn 1 miniRITE with Open Sound Navigator set to the strongest noise reduction setting. Power domes were worn with each of the three hearing aids.

The results obtained with Opn were compared to the results from two other major manufacturers' (Brand 1 and 2) solutions, using directionality and narrow directionality/beamforming (respectively). All hearing aids were fitted using the manufacturer's fitting software and earmold recommendations, based on the hearing loss and other gathered data. The level of amplification for each hearing aid was provided according to NAL-NL2 rationale.

The primary measure reported here was the Speech Reception Threshold-50 (SRT-50). The SRT- 50 is a measure that reflects the SIN level at which the listener correctly identifies 50% of the sentence-based keywords correctly. For example, an SRT-50 of 5 dB indicates the listener correctly repeats 50% of the words when the SNR is 5 dB. Likewise, if the SRT is 12 dB, this indicates the listener requires an SNR of 12 dB to achieve 50% correct.

The goal of this study was to measure the listening benefit provided to the listeners in a real-life noisy acoustic situation in which the location of the sound source (ie, the person talking) could not be predicted. That is, three human talkers and one human listener were engaged in each segment of the study (Figure 1). There were 25 separate listeners.

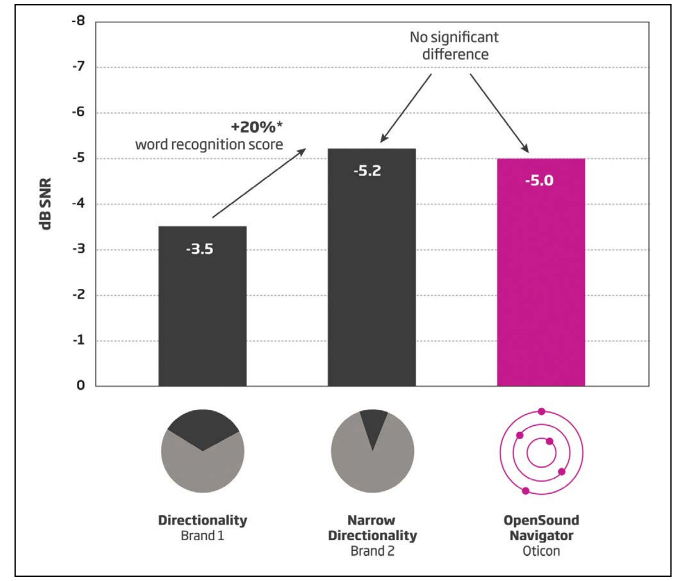


Figure 3. Average speech understanding of center speaker. The average SRT-50 for the hearing aid with directionality was significantly lower than the average SRT obtained with the two other hearing aids. The average SRT-50 obtained with the narrow directionality and OpenSound Navigator were not significantly different from each other, although narrow directionality and Open Sound Navigator were each statistically and significantly better than directionality. The right-facing arrow indicates an approximate 20% word recognition improvement using either narrow directionality or Open Sound Navigator, as compared to directionality.

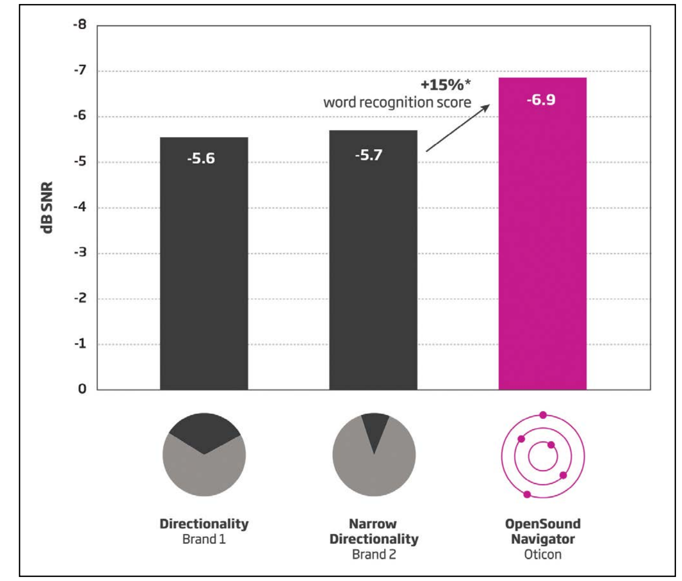


Figure 4. Left and right speaker results. The average SRT-50s obtained from speakers located at ±60° using directionality and narrow directionality were not significantly different from each other. The average SRT-50 obtained with Open Sound Navigator was statistically and significantly better than those obtained with other technologies. The right-facing arrow indicates 15% likely improvement in word recognition with Open Sound Navigator as compared to directionality or narrow directionality.

Speech babble (ISTS) and background noise (speech-shaped) were delivered at 75 dB SPL. The German-language Oldenburg sentence test (OLSA)[23] was delivered to each participant, while wearing each of the three hearing aids.

The OLSA Matrix Tests are commercially available and are accessible via software-based audiometers. We conducted these tests with an adaptive procedure targeting the 50% threshold of speech intelligibility in noise (the speech reception threshold, or SRT). As noted above, speech noise was held constant at 75 dB SPL while the OLSA speech stimuli loudness varied to determine the 50% SRT using a standard adaptive protocol. Each of the 25 listeners was seated centrally and was permitted to turn his/her head as desired to maximize their auditory and visual cues. The background speech noise was a mix of speech babble delivered continuously to the sound-field speakers at ±30° (relative to the listener) and simultaneously at 180° behind the listener, as illustrated in Figure 1.

Each of the 25 listeners was seated centrally while three talkers were located in front of, as well as ±60° (left and right) of the listener. Target speech was randomly presented from one of the three talker locations. Listeners were free to turn their heads as desired.

Results. While listening to conversational speech, the average scores obtained from all three talkers by the 25 listeners using *directionality* (Brand 1) demonstrated an average SRT of -4.9 dB. Listeners using narrow *directionality/beamforming* (Brand 2) demonstrated an average SRT of -5.5 dB. Listeners wearing Oticon OPN had an average SRT of -6.3 dB (see Figures 2-4).

Of note, it is generally accepted that for, each decibel of SNR improvement, the listener likely gains some 10% with regard to word recognition scores.

Discussion

Realistic listening situations are difficult to replicate in a lab-based setting. Nonetheless, the lab-based setup described here is relatively new and is believed to better replicate real-life listening situations than is the typical "speech in front" and "noise in back" scenario.

Obviously, important speech sounds occur all around the listener while ambulating through work, recreational, and social situations. The goal of the amplification system, particularly in difficult listening situations, is to make speech sounds more audible while increasing the SNR to make it easier for the brain to identify, locate, separate, recognize, and interpret speech sounds.

Oticon's Open Sound Navigator (OSN) with Multi Speaker Access Technology (MSAT) has been shown to allow (on average) improved SNR-50s and improved word recognition scores in noisy situations, such as when speech and noise surrounds the listener. OSN with MSAT allows essentially the same word recognition scores as the best narrow directional protocols when speech originates in front of the listener, while dramatically improving access to speakers on the sides and, therefore, delivering an improved sound experience. OSN with MSAT effectively demonstrates improved word recognition scores in noise when speech originates *around the listener*.

The results of this study demonstrate that Oticon's Open Sound Navigator provides overall improved overall word recognition in noise when compared to directional and narrow directionality/beamforming systems. We hypothesize these results are due to Multiple Speaker Access Technology, as well as the maintenance of spatial cues and many other advanced features available in Oticon Opn.

CORRESPONDENCE can be addressed to *HR* or Dr Beck at: DBEC@oticon.com

REFERENCES can be found in the online version of this article at www.hearingreview.com.