Canadian Audiologist

The Official Publication of the Canadian Academy of Audiology

Improving Speech Understanding in Multiple-Speaker Noise

Published November 20th, 2017

Douglas L. Beck, AuD
Nicolas Le Goff, PhD

Reprinted with permission from
http://alliedweb.s3.amazonaws.com/hearingr/diged/201709/index.html

Introduction

The most common problem experienced by people with hearing loss and people wearing traditional hearing aids is not simply that sound isn't loud enough. The primary issue is understanding Speech-in-Noise (SIN). To be clear, 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 (i.e., 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), inner hair cell 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 (Killion, 2002, and Wilson, 2011).

To provide maximal understanding of speech in difficult listening situations, the goal of hearing aid amplification is two-fold; to make speech sounds audible and to increase the signal-to-noise ratio (Killion, 2002). 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 unintelligible, secondary sounds of other people speaking (i.e., "speech babble noise"), and other less meaningful sounds (background noise).

Signal-to-Noise Ratio (SNR):

Killion (2002) reported people with substantial difficulty understanding speech in noise may have significant "SNR loss." Killion stated the SNR loss is ***unrelated*** to, and cannot be predicted, by 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 (2011) evaluated more than 3400 veterans, and he, too, reported speech in quiet 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 that the Words in Noise (WIN) test should be used as the "stress test" for auditory function.

Beck & Flexer (2011) reported "Listening is Where Hearing Meets Brain." That is, the ability to

listen (i.e., to make sense of sound) is arguably a more important construct and reflects a more realistic representation of how people manage in the real world, than does a simple audiogram. They note many animals (including dogs and cats) “hear” better than humans. However, humans are the top of the food chain ***not*** because of their hearing, but due to their ability to listen (i.e., to apply meaning to sound).

As such, speech-in-noise (SIN) tests are more of a listening test than a hearing test. Specifically, the ability to make sense of speech in noise depends on a multitude of cognitive factors beyond loudness and audibility, and includes; speed of processing, working memory, attention and more.

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 an acoustic environment with multiple speakers.

TRADITIONAL STRATEGIES to MINIMIZE BACKGROUND NOISE:

As noted above, the primary problem for people with hearing loss is understanding speech in noise. As such, one major focus over the last 4 decades (or so) has been to reduce noise. Two processing strategies are often employed in modern hearing aid amplification systems to minimize background noise; digital noise reduction and directional microphones.

Digital Noise Reduction (DNR):

DNR systems can recognize and reduce the signature amplitude of steady state noise using various amplitude modulation (AM) detection systems. AM systems can detect differences in dynamic human speech, as opposed to steady state noise sources such as heating-ventilation-air-conditioning (HVAC) systems, electric motor hum, fan noise, 60 cycle noise etc (Mueller, Ricketts, Bentler, 2017, pg 171). 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., as the acoustic signature of people we desire to hear and the acoustic signature of other people in close proximity (i.e., speech babble noise) are essentially the same. Venema (2017, page 331) states the broad band spectrum of speech and the broad band spectrum of noise intersect, overlap and resultantly, are very much the same thing. Nonetheless, given multiple unintelligible speakers and physical distance of perhaps 7-10 meters from the hearing aid microphone, Beck & Venema (2017) report the secondary signal (i.e, speech babble noise) may present as (i.e., morph into) more of steady state “noise-like” signal and may be attenuated via DNR, providing additional comfort to the listener.

McCreery, Venediktov, Coleman & Leech (2012) reported their evidence-based systematic review to examine “the efficacy of digital noise reduction and directional microphones.” They searched 26 databases seeking contemporary publications (after 1980), resulting in 4 articles on DNR and 7 articles on directional microphone studies. McCreery and colleagues concluded DNR did not improve or degrade speech understanding.

Beck & Behrens (2016) reported DNR may offer substantial “cognitive” benefits such as; 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. Beck & Behrens suggest typical hearing aid fitting protocols should include activation of the DNR circuit.

Pittman, Stewart, Willman & Odgear (2017) report (consistent with previous reports) DNR provides little or no benefit with regard to audiologic measures (such as improved word recognition in noise). However, McCreery and Walker (2017) note DNR circuits are routinely

recommended for the purpose of improving listening comfort and re-stated that with regard to school age children with hearing loss, DNR neither improved nor degraded speech understanding.

Directional Microphones (DMs):

DMs are the only technology proven to improve the signal-to-noise ratio (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 (Beck & Venema, 2017). Directivity indexes (DIs) indicating 4-6 dB improvement are generally not “real world” measures. That is, DIs are most often measured on manikins, in an anechoic chamber, based on pure tones and, DIs quantify and compare sounds coming from the front, to 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 (see Chasin, 2013 and see Taylor & Mueller, 2017, and see Venema, 2017, pg 327).

In their extensive review of the published literature, McCreery, Venediktov, Coleman & Leech (2012) reported that in controlled optimal situations, directional microphones did improve speech recognition, yet they cautioned the effectiveness of directional microphones in real-world situation was not yet well documented and additional research is needed prior to making conclusive statements.

Brimijoin and colleagues (2014) stated directionality potentially makes it difficult to orient effectively in complex acoustic environments. Picou, Aspell & Ricketts (2014) reported directional processing (i.e., cue preserving bilateral beamformer) did reduce Interaural Loudness Differences (ILDs) and localization was disrupted in extreme situations without visual cues. Best and colleagues (2015) reported narrow directionality is only viable when the acoustic environment is highly predictable, which questions the effectiveness of such technology in real-life noisy environments that are, by nature, dynamic and unpredictable. Mejia, Dillon & van Hoesel (2015) report as beam width narrows, the possible SNR enhancement increases. However, as the beam width narrows, the opportunity increases that “listeners will misalign their heads, thus decreasing sensitivity to the target…” Geetha, Tanniru & Rajan (2017) 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:

Shinn-Cunningham & Best (2008) reported “the ability to selectively attend depends on the ability to analyze the acoustic scene and to form perceptual auditory objects properly.” If analyzed correctly, attending to a particular sound source (i.e., voice) while simultaneously suppressing background sounds may become easier as one successfully increases focus and attention.

The purpose of a new strategy for minimizing background noise should be to facilitate the ability to selectively attend to one primary sound source and to switch attention when desired. This is what Mutli Speaker Access Technology has been designed to do.

Multi Speaker Access Technology:

In 2016, Oticon introduced Multiple Speaker Access Technology (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, indeed, it maintains access to all distinct speakers. MSAT is built on three stages of sound processing including; 1- Analyze, 2- Balance and

3- Noise Removal (NR).

1. Analyze provides identify distinct speech sources and noise sources. An enhanced noise estimate is obtained by analyzing the acoustic environment with two views. One view is a 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 re-balanced 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 provides very fast NR removes noise between words and up to 9 dB of noise attenuation. Of note, if speech is detected in any band, Balance and NR systems are "frozen" so as to not isolate the front talker, but to preserve distinct talkers, regardless of their location around the end-user.

The Oticon OPN system uses MSAT and is, therefore, neither directional nor omnidirectional. Its most important improvement over current directional technologies is, beyond improved speed and accuracy, the improved capacity of the system to estimate and remove noise, while simultaneously identifying and preserving distinct speech (see above, the *Analyze* stage).

Maintaining Spatial Cues:

Of significant importance to understanding SIN in a scenario with multiple speakers, 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 (Beck & Sockalingham, 2010). The specific spatial cues required to improve the ability to understand SIN are interaural level differences (ILDs) and Interaural timing differences (ITDs), such that the left and right ears receive unique spatial cues.

Sockalingham & Holmberg (2010) 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 (2015) reported spatial hearing allows us to identify the origin/location of sound in space and to attend to a "primary sound source in difficult listening situations" due to binaural summation, while ignoring secondary sounds due to binaural squelch. He reports knowing "where to listen" allows the human brain to maximally listen in difficult listening situations, as the brain compares and contrasts unique sounds from the left and right ears—in real time, so the brain can better determine where to focus attention.

Geetha, Tanniru & Rajan (2017) state speech in noise is challenging for people with sensorineural hearing loss (SNHL). To better understand speech in noise, the responsible acoustic cues are Interaural Timing Differences (ITDs) and Interaural Loudness Differences (ILDs). They report "…preservation of binaural cues is…crucial for localization as well as speech understanding…" and they report directionality in binaural amplification without synchronized wireless communication can disrupt ITDs and ILDs, leading to "…poor performance in localization…(and) speech perception in noise…"

Oticon OPN uses Speech Guard™ LX to improve speech understanding in noise by preserving

clear sound quality and speech details such as spatial cues, in particular Interaural Loudness Differences (ILDs). Speech Guard LX uses adaptive compression and combines linear amplification and fast compression, to provide up to a 12dB dynamic range window to preserve amplitude cues in speech signals and between the left and right ear sounds (ILDs). 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 to identify where sound originates.

The Oticon OPN SNR Study

The purpose of this study was to compare the results obtained using Oticon Opn to two other major manufacturers, with regard to listener's ability to understand speech in noise in a lab-based, yet realistic, background noise situation with multiple speakers.

Pittman, Stewart, Willman and Odgear (2017) reported "Like the increasingly unique advances in hearing aid technology, equally unique approaches may be necessary to evaluate these new features…"

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:

Twenty-five native German speaking participants (i.e., listeners) with an average age of 73 years (standard deviation 6.2 years) with mild-moderate symmetric sensorineural hearing loss underwent listening tasks. During the listening tasks, each participant wore Oticon Opn 1 miniRITE with OSN set to the strongest noise reduction setting. The results obtained with Opn were compared to the results from two other major manufacturers' (referred to as Mfr A and Mfr B) solutions, using directionality and narrow directionality/beam forming. All hearing aids were fitted using the manufacturer's fitting software and Power Domes (or equivalent). 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 which reflects the SIN level at which the listener correctly identifies 50% of the sentence-based key words correctly.

For example, an SRT-50 of 5 dB indicates the listener correctly repeats 50% of the words when the signal-to-noise ratio (SNR) is 5 dB. Likewise, if the SRT is 12 dB, this would indicate the listener requires an SNR of 12 dB, to achieve 50% correct.

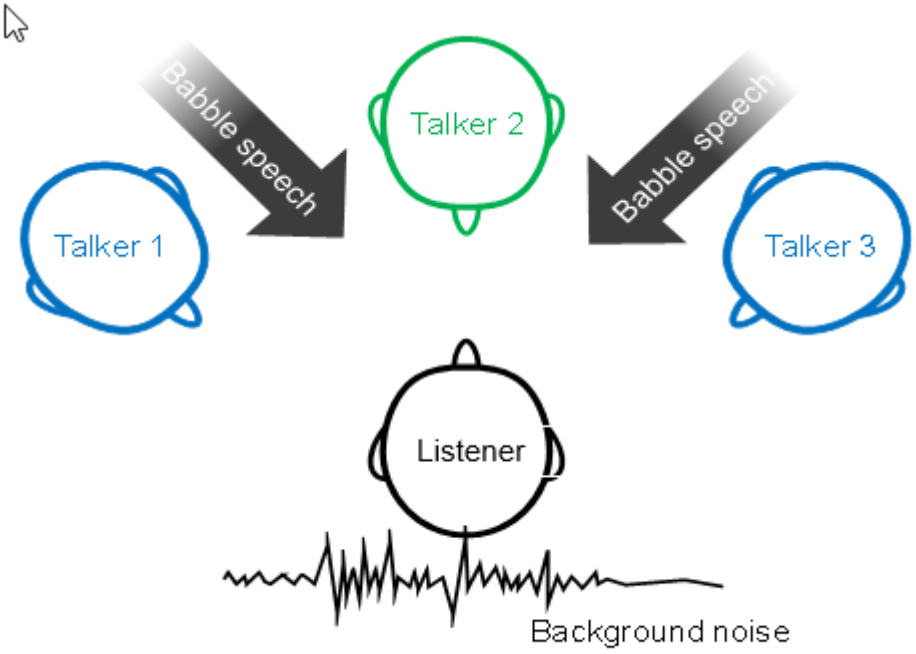


Figure One: Schematic representation of the acoustical conditions of the experiment.

The goal of this study was to measure the listening benefit provided to the participants in a real-life noisy acoustic situation, in which the location of the sound source (i.e., the person talking) could not be predicted. That is, three human talkers and one human listener were engaged in each segment of the study. There were a total of 25 separate listeners. Speech babble (ISTS) and background noise (speech-shaped) were delivered at 75 dB SPL. The German language Oldenburg sentence test (OLSA) 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, 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 plus and minus 30 degrees (relative to the listener) and simultaneously at 180 degrees, behind the listener, as illustrated in Figure One.

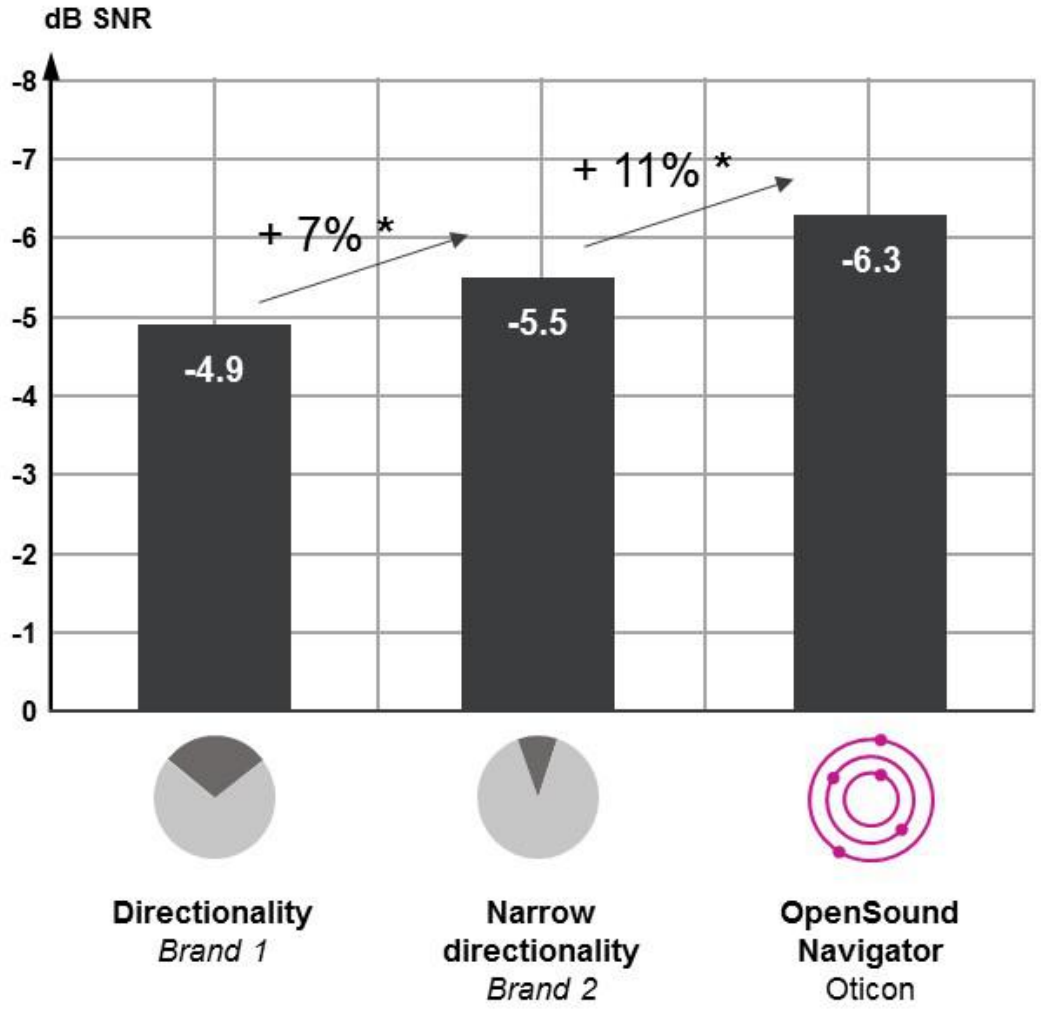
Each of the 25 listeners was seated centrally (see illustration #1) while three talkers were located at location #1 (in front of and 60 degrees to the left of the listener), location #2 (in front of), and location # 3 (in front of and 60 degrees to the right) of the participant. Target speech was randomly presented from one of the 3 locations. The listeners were free to turn their head if/as desired.

Results:

While listening to conversational speech, the average scores obtained from all three talkers across participants using DIRECTIONALITY (via manufacturer 1) demonstrated an average SRT-50 of -4.9 dB. Participants with NARROW DIRECTIONALITY/Beam Forming (via manufacturer 2) demonstrated an average SRT-50 of -5.5 dB and participants wearing Oticon OPN had an average SRT-50 of -6.3 dB.

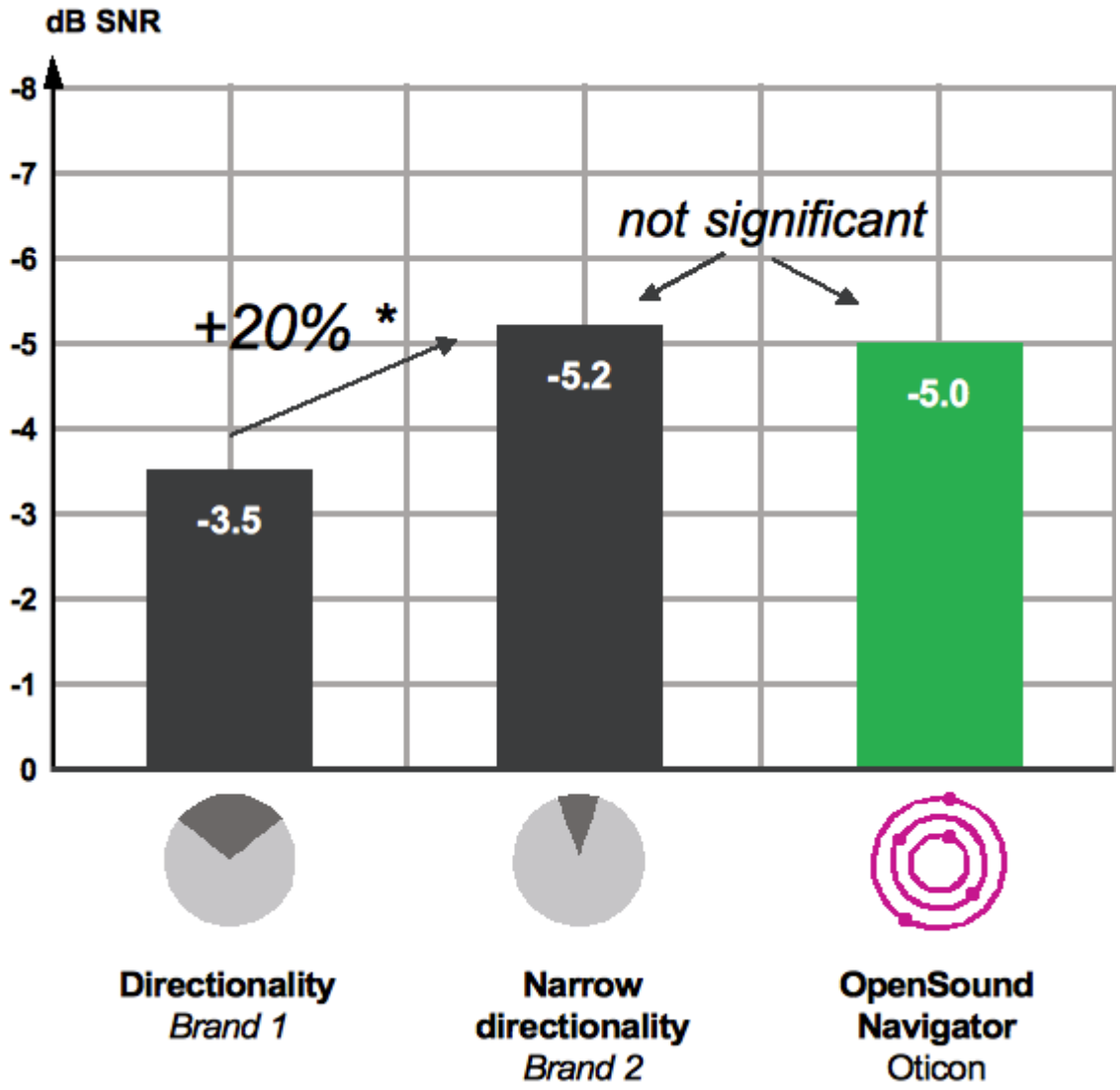
Of note, it is generally accepted that for each decibel of improvement in SNR, the listener likely gains some 10% with regard to word recognition scores.

ILLUSTRATION ONE: OVERALL RESULTS



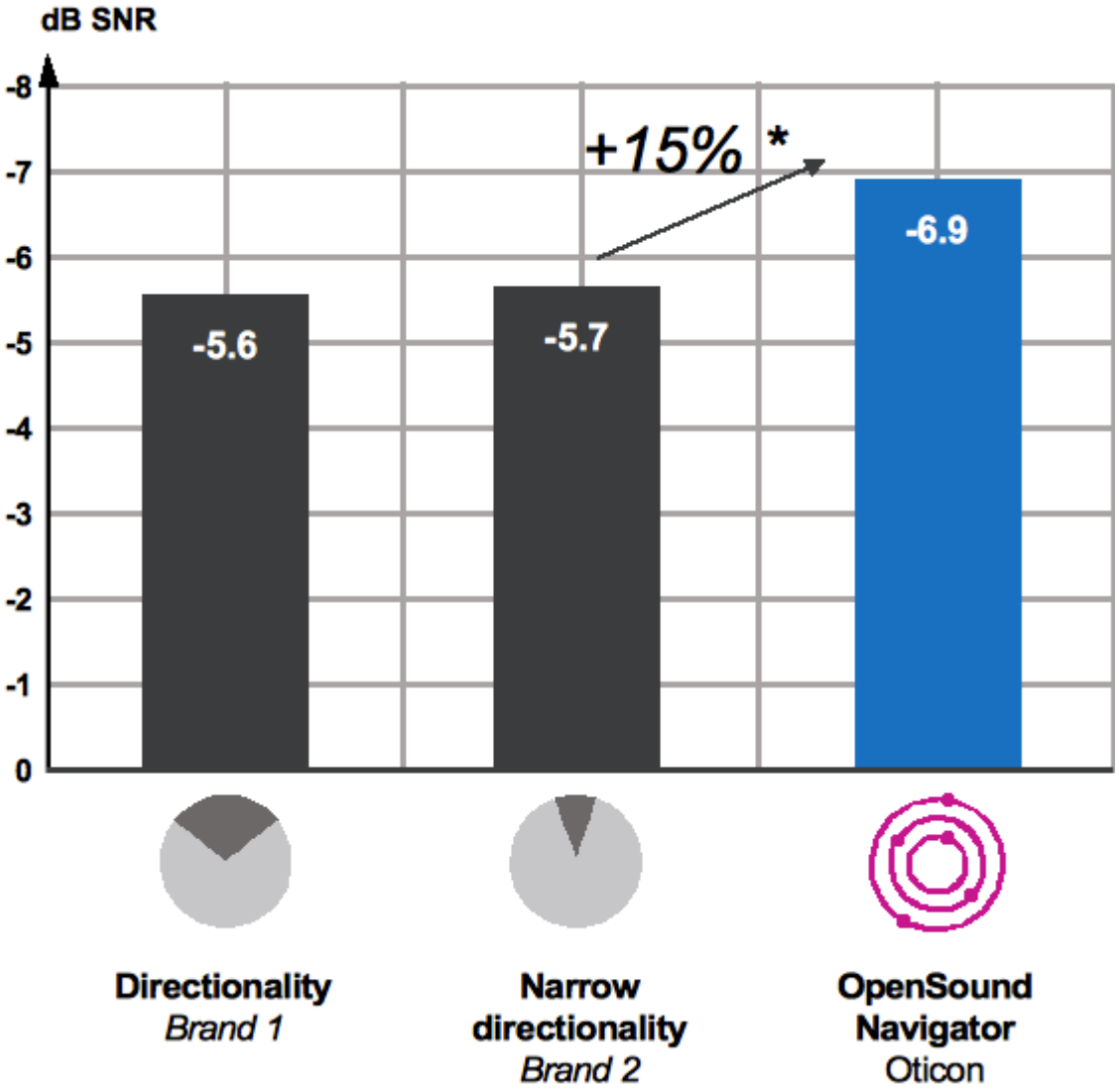
Overall improvement in SRT-50. As bar height increases, the SRT-50 decreases, demonstrating a reduction in the SNR required to correctly identify 50% of the key words correctly. Each improvement bar is statistically different from the others at $p<0.05$. The one-way arrow (+7%) from -4.9 to -5.5 represents 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.

ILLUSTRATION TWO: CENTER SPEAKER RESULTS



AVERAGE SRT-50 of CENTER SPEAKER. The average SRT-50 for the hearing aid with directionality was significantly lower than the average SRT-50 obtained with the other hearing aids. The average SRT-50 obtained with narrow directionality and OpenSound Navigator were *not* significantly different from each other. The right-facing arrow indicates an approximate 20% word recognition improvement using narrow directionality or Open Sound Navigator, as compared to directionality.

ILLUSTRATION THREE: LEFT & RIGHT SPEAKER RESULTS



The average SRT-50s obtained from hearing aids with directionality and narrow directionality technologies were not significantly different from each other. The average SRT-50 obtained with Open Sound Navigator was statistically significantly better than other technologies. The right-facing arrow indicates 15% improvement in word recognition with Open Sound Navigator, as compared to either directionality of narrow directionality

---

Results & Discussion:

Realistic listening situations are difficult to replicate in a lab-based setting. Nonetheless, the lab-based set-up described here is relatively new, and is believed to better replicate real-life listening situations, than is the typical “speech in front” with “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 audible while increasing the signal-to-noise ratio (SNR), to make it easier for the brain to identify, locate, separate, recognize and interpret speech sounds. There are, however, different strategies to improve the SNRs, as represented here by 3 hearing aids from 3 different manufacturers using 3 different approaches.

Oticon’s Open Sound Navigator (OSN) with Multi Speaker Access Technology (MSAT) has been shown to allow (on average) significantly improved word recognition. OSN with MSAT allows essentially the same word recognition scores as the best narrow directional protocols when speech originates directly in front of the listener, while dramatically improving access to speakers on the

side, and therefore delivering an improved overall 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 improved 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 other advanced features available in Oticon Opn, as demonstrated above, in lab-based simulated difficult listening situations such as three people speaking from different, random locations in the presence of noise sources, also arriving from three separate locations.

SPECIAL THANKS to Karl Strom, Editor in Chief, Hearing Review, for allowing us to minimally revise and essentially re-publish "Speech in Noise Test Results for Oticon Opn" http://www.hearingreview.com/2017/08/speech-noise-test-results-oticon-opn/

We are grateful to Karl Strom and Hearing Review for allowing us to provide this version for the membership at CAA.

REFERENCES & Recommended Readings:

Beck DL. (2015): BrainHearing: Maximizing hearing and listening. Hearing Review. 2015;21(3):20.

Beck, DL. & Behrens, T (2016): The Surprising Success of Digital Noise Reduction. Hearing Review. May.

Beck DL, Flexer C. Listening is where hearing meets brain…in children and adults. Hearing Review. 2011;18(2):30-35.

Beck, DL. & Venema, TH. (2017): Hearing Review Interview by Douglas L. Beck, with Ted Venema. INSERT LINK HERE.

Beck DL, Sockalingam R. (2010): Facilitating spatial hearing through advanced hearing aid technology. Hearing Review. 17(4):44-47.

Best, V., Mejia, J., Freeston, K., van Hoesel, RJ., Dillon, , H. (2015): "An Evaluation of the Performance of Two Binaural Beamformers in Complex and Dynamic Multitalker Environments", International Journal of Audiology 54(10):727-35.

Brimijoin, WO., Whitmer, WM., McShefferty, D., Akeroyd, MA. (2014): The Effect of Hearing Aid Microphone Mode on Performance in an Auditory Orienting Task. Ear and Hearing Sep-Oct;35(5):e204-12.

Chasin, M. (2013): Back to Basics: Slope of PI Function Is Not 10%-per-dB in Noise for All Noises and for All Patients. October, Hearing ReviewGeetha, C., Tanniru, K. & Rajan, RR. (2017): Efficacy of Directional Microphones in Hearing Aids Equipped with Wireless Synchronization Technology. The Journal of International Advanced Otology. 13(1):113-7.

Killion, MC. (2002): New Thinking on Hearing in Noise – A Generalized Articulation Index. In Seminars in Hearing. Volume 23, Number 1, Pages 57-75

Littmann, V. & Hoydal, EH. (2017): Comparison Study of Speech Recognition Using Binarual Beamforming Narrow Directionality. Tech Topic I Hearing Review. May.

Mejia, J., Carter, L., Dillon, H. and Littmann, V. (2017): Listening Effort, Speech Intelligibility, and Narrow Directionality. Tech Topic in Hearing Review, Jan.

McCreery, RW., Venediktov, RA., Coleman, JJ. & Leech, HM. (2012): An Evidence-Based Systematic Review of Directional Microphones and Digital Noise Reduction Hearing Aids in School-Age Children With Hearing Loss. Am J Audiol. 2012 Dec; 21(2): 295–312.

McCreery, RW. & Walker, EA. (2017): Hearing Aid Candidacy and Feature Selection for Children. In “Pediatric Amplification – Enhancing Auditory Access.” Published by Plural Publishing, San Diego, Ca.

Mejia, J., Dillon, H., van Hoesel, R. et al. (2015): Loss of speech perception in noise – causes and compensation, Page 209, Proceedings of ISAAR 2015: Individual Hearing Loss – Characterization, Modelling, Compensation Strategies. 5th symposium on Auditory and Audiological Research. August. The Danavox Jubilee Foundation.Mueller, HG., Ricketts, TA. & Bentler, R. (2017): Speech Mapping and Probe Microphone Measurements. Plural Publishing, San Diego, Ca.

Picou, EM., Aspell, E. & Rickettes, TA (2014): Potential benefits and limitations of three types of directional processing in hearing aids. Ear Hear. 2014 May-Jun;35(3):339-52.Pittman, AL., Stewart, EC., Willman, AP. & Odgear, IS. (2017): Word Recognition and Learning – Effects of Hearing Loss and Amplification Feature. Trends in Hearing. Vol 21, pages 1-13.

Shinn-Cunningham, B. & Best, V. (2008): Selective Attention in Normal and Impaired Hearing. Trends in Amplification. Vol 12, No 4. Pages 283-299.

Smeds, K., Wolters, F. & Rung, M. (2015): Estimation of Signal-to-Noise Ratios in Realistic Sound Scenarios. J Am Acad Audiol. Feb; 26(2):183-96. doi: 10.3766/jaaa.26.2.7.

Sockalingam R, Holmberg M. (2010): Evidence of the effectiveness of a spatial noise management system. Hearing Review 17(9):44-47.

Taylor, B. & Mueller, HG. (2017): Fitting and Dispensing Hearing Aids. 2nd edition. 2017. Plural Publishing. San Diego.Venema, TH. (2017): Compression for Clinicians. A Compass for Hearing Aid Fittings. Third Edition. Plural Publishing, San Diego, Ca.

Wilson, RH (2011): Clinical experience with the words-in-noise test on 3430 veterans: comparisons with pure-tone thresholds and word recognition in quiet. J Am Acad Audiol. Jul-Aug;22(7):405-23.