

Hearing, listening and deep neural networks in hearing aids

Abstract

Hearing aids have undergone vast changes in the last 30 years from basic analog sound processing techniques, to advanced digital technology, to Deep Neural Networks (DNNs) “on-the-chip” providing real-time sound processing. In addition to making sounds audible, advanced hearing aids with DNN on-the-chip are better able to provide clearer understanding of speech in noise, improve recall, maintain interaural loudness and timing differences, and improve the wearer’s ability to selectively attend to the speaker of choice in challenging listening situations. These improvements are delivered without acoustic feedback and with very high sound quality.

Volume 13 Issue 1 - 2021

Douglas L Beck

Vice President of Academic Sciences, Oticon Inc., Adjunct Professor of Communication Disorders and Sciences, State University of New York at Buffalo & Vice Chair of the Cognition in Hearing SIG, USA, Tel 732-673-4048, Email dbec@oticon.com

Correspondence: Douglas L Beck, Vice President of Academic Sciences, Adjunct Professor of Communication Disorders and Sciences, State University of New York at Buffalo & Vice Chair of the Cognition in Hearing SIG, USA, Tel 732-673-4048, Email dbec@oticon.com

Received: January 15, 2021 | **Published:** February 04, 2021

Introduction

The most common complaint received from patients with sensorineural hearing loss (SNHL) is *not that they cannot hear*, it is that they *cannot understand*. Unfortunately, understanding speech in a noisy background is even more challenging. To be clear, hearing is simply perceiving or detecting sound. The majority of people with SNHL hear lower speech frequencies (i.e., 100-350 Hz) at conversational levels, and so, their observation is that they can hear, which makes intuitive sense. However, most people with mild-to-moderate SNHL are not aware of the vast spectral components of speech, and of course, they are unaware of that which they cannot (and do not) hear. As such, many people with SNHL are unaware they have hearing loss in the mid-frequencies (i.e., 1500 to 3000 Hz) and/or high frequencies (i.e., 3000 to 6000 Hz) of speech sounds. Of note, the most important information for speech recognition (consonants, fricatives, sibilants, etc.) is conveyed in the mid and high frequencies.¹ Therefore, and understandably, many people with SNHL reach a similar conclusion. That is, from their viewpoint, the reason they cannot understand is because people don’t speak clearly, or people mumble, both of which may be reasonable conclusions for the person with a typical mild-to-moderate SNHL. However, neither conclusion is likely to be the primary impediment to understanding speech in noise.

Hearing and Listening and Sensorineural Hearing Loss

There is much more to *listening* than simply *hearing*, as listening is built upon hearing. Hearing is critically important and arguably among the two most important senses (vision and hearing). I am not dismissing the extraordinarily important contribution of hearing. I am saying that simply perceiving sound (i.e., simply hearing) is not enough. For language to develop and for humans to make sense of language involves incredibly human-only cognitive abilities, which are, to a large degree, reflected in our language acquisition and use. Clearly, many mammals (dogs, cat, horses, dolphins, whales, apes,

gorillas...) hear better than humans. What separates us from other mammals is not our hearing ability, it’s our listening ability. Listening is a brain function. Indeed, “Listening is where hearing meets brain.”² The entire brain is involved as we decode, recognize and attribute meaning to speech sounds. There is much more to listening than the common line-labeled definitions (i.e., the auditory nerve sends information to the temporal lobes where it is processed...). Listening involves the hippocampus, the amygdala, the frontal lobe, the occipital lobes (speech reading and facial recognition), the corpus callosum, the anterior commissure, the brain stem and more, all working in a uniquely human way, to attribute meaning to sound. The idea that sensory inputs are processed in specific unimodal cortices (i.e., line-labelled) is outdated³ and the importance of “multi-modality” as a fundamental factor in human brain organization and re-organization is rapidly emerging. Glick and Sharma⁴ demonstrated multiple visual evoked potential (VEP) changes (and more) associated with hearing loss due to cross-modal cortical neuro-plastic changes which were “striking” and remarkably, were reversed for some patients following six months of hearing aid use. Multiple reports⁵⁻⁸ have correlated hearing loss with dementia and/or cognitive decline, consistent with the multi-modal perspective of hearing loss as a potential pre-cursor or contributor to negative neurocognitive changes and outcomes. The correlation between cognition and listening is also noted at the other end of the age spectrum, as Seeto, Tomlin and Dillon⁹ report one cannot assume that pediatric tests of auditory processing abilities and tests of cognitive abilities measure separate abilities.

Speech sounds must be audible before they can be processed. However, they must be audible at a signal-to-noise ratio (SNR) which is acceptable and beneficial to the patient. Recently, Golub et al.¹⁰ reported that even among people within the “normal hearing” category (0-25 dB HL), those with 0 dB thresholds performed better than those with 25 dB thresholds regarding measurable aspects of cognition. That is, people with better thresholds perceived sound at a greater SNR than people with worse thresholds. Gaete and colleagues¹¹ similarly reported reduced audibility negatively impacts scores on the Mini Mental State Examination, even in cognitively intact people.

Unfortunately, there is no “one size fits all” with regard to hearing loss, brains and amplification. Each person (even with similar or identical SNHL) varies based on the specific factors associated with their particular SNHL, their cognitive ability, their genetics, their personal listening preferences, the acoustic environment, the type and model of hearing aid, the fitting protocol and more. Further, the all-encompassing term “sensorineural hearing loss” includes and identifies vastly different etiologies, as well as anatomic and physiologic variations from age related hearing loss (ARHL), noise induced hearing loss (NIHL), auditory neuropathy spectrum disorders (ANS), cochlear synaptopathy (CS), hidden hearing loss (HHL), to neuro-cognitive disorders, etc.

Selective Attention

Shinn-Cunningham & Best¹² reported hearing loss degrades the auditory signal (i.e., neural code) thereby making it more difficult for the person with hearing loss to selectively attend to the auditory signal. New research¹³⁻¹⁶ reveals that for the human brain to recognize, untangle, and comprehend speech sounds (i.e., to listen) after hearing sound in the acoustic environment, the brain must orient and *then* focus on the sound of maximal interest. The ability to attend to the sound of maximal interest, particularly in noise, appears to be based on multiple key factors. Among them, the quantity and quality of the neural code (bio-electric neural activity) sent from the cochlea to the brain, and “Selective Attention.” Selective attention is the central nervous system process through which people can focus on the sounds of maximal interest in the “foreground,” while dismissing sounds which are essentially “background.” Typically, the listener’s primary foreground interest includes speech sounds, whereas the sounds of heating and ventilation systems, computer fans, fluorescent lights and similar noises are most often relegated to the background.

Selective attention in hearing is somewhat analogous to the ability to change the image on the fovea of our retina by simply looking elsewhere with regard to vision. That is, vision is volitionally and centrally controlled by the person’s desire as to where to focus their visual attention, pupillary gymnastics, the 3rd, 4th and 6th cranial nerves and more. None of which we think about while scanning the page, the room, or the horizon. We simply refer to this as where we focus our (visual) attention. Selective Attention via audition is surprisingly analogous to visual focus.¹⁷ That is, the ability to switch our focus of auditory attention to a specific person in a noisy background (i.e., cocktail party) given relatively normal hearing and listening ability is an example of selective attention. To maximally selectively attend to the voice or sound we choose, the brain benefits from an excellent, highly representative neural code. As expected, a complete sound scene which is dynamically and faithfully represented by an excellent neural code is easier to decode, requires less effort and energy, and allows for more highly advantageous listening. Further, if the information contained within the neural code does not contain a complete acoustic sound scene, it is unlikely the brain can decode acoustic information which was not provided.

Ideally, it seems an enhanced neural code should:

1. Provide information specific to each ear including interaural loudness differences (ILDs, the difference between the left and right ears across the speech spectrum can be 20-22 dB at 5000, 6000 and 7000 Hz) and interaural timing differences (ITDs, aka phase indicates which ear heard the sound first).
2. Provide a very high-quality sound.
3. Provide substantial noise reduction and a prioritized balance of background sounds, while allowing volitional access to the foreground and the background.
4. Contain no acoustic feedback (often described as a whistling sound, a common problem with traditional hearing aids).
5. Allow the brain to exert less listening effort.
6. Facilitate a maximal SNR.

If these challenges are accomplished prior to creating and sending the neural code to the brain, the brain’s task of decoding the neural code is easier, more efficient, and thus requires less energy and less effort. Miller et al.¹⁸ report that hearing aid amplification has been shown to alter the neural code secondary to hearing aid use, based on cortical responses to aided versus unaided fricative stimuli.

Traditional hearing aids and speech in noise

To be clear, traditional hearing aid technology has been beneficial for millions of people across the globe, especially in quiet environments, when background noise is not an issue. However, the primary complaint of people with SNHL, as well as those with traditional hearing aid technologies, is that they can hear, but they cannot understand, particularly in noise. Unfortunately, traditional hearing aids tend to inadvertently limit and reduce some acoustic information. For example, traditional hearing aids use multiple forms of compression to maintain the overall sound level between comfortable and uncomfortably loud. Although at first glance this seems like a beneficial idea, it does mean that when a (typical) 2:1 compression ratio is chosen, the dynamic range of speech (the sound pressure level [SPL] difference between the quietest and the loudest speech sounds) changes from the very dynamic and typical 30 dB range, to a 15 dB range, thus attenuating loudness cues and attenuating the neural code.

A 3:1 compression ratio delivers only a 10 dB dynamic range for the listener, further reducing the overall quality and quantity of the neural code. Many commercially available hearing aids use fixed or adaptive directionality (or beam-forming) to focus amplification on the talker most likely to be the primary person the listener would like to attend to, based on amplitude modulation (i.e., loudness variation) and other acoustic factors. However, in cocktail party and restaurant-like acoustic scenarios and other noisy scenarios in which an excellent neural code is needed the most, the loudest sounds may not be the sounds one wishes to attend to. The most significant amplitude modulation may originate from the in-house background music, a live band, a patron(s) at a neighboring table, a waiter, or elsewhere. That is, traditional hearing aids may maximally amplify the wrong person (not the one the listener chooses to selectively attend to) or perhaps the wrong people (for example a nearby table with loud voices) thereby resulting in a louder confusing sound scene, but not a clearer neural code.

Unfortunately, many traditional hearing aid designs assume the most important sounds are often in front of the wearer, non-speech sounds are unwanted, a reduced dynamic range is desirable, a smaller window to the surrounding acoustic sound scene (from directional microphones or beam-forming) is desired or required, and acoustic feedback needs to be reduced by lowering gain.¹⁹ However, these assumptions tend to attenuate the neural code rather than enriching it, and provide fewer natural, acoustic cues to the brain.

Deep neural networks

Deep Neural Networks (DNNs) have significantly improved technical abilities in recent years. DNNs are the most sophisticated processors within the world of artificial intelligence (AI). What makes DNNs so impressive is what they do and how they do it. DNNs try to find patterns among vast amounts of information, and identify and decode those patterns, similar to how a human brain would. DNNs process huge data sets data without specific permutation programming. That is, DNNs operate without a specific event-driven program. DNNs are the processors of self-driving cars and DNNs are the intelligence of voice and face recognition (used by Google, Amazon, Facebook and more). DNNs determine the most and least likely path of hurricanes and other weather events. When Amazon suggests you might buy X, Y or Z, based on your previous purchases, or when Netflix suggests a movie, DNNs drive those suggestions. DNNs derive a solution without a specific written protocol for each and every decision point and they self-check to make sure their solution (the output) is the most representative solution, based on the incoming data.

In the biology-based world, there are many examples of DNNs. There are vast multitudes of remarkable things which happen without written step-by-step instructions. For example, birds learn to fly, fish learn to swim, and of course, babies learn to walk and talk, generally after some 12-18 months of tremendous sensory input. After trillions of sensory-based bio-electric signals stimulate the brain, the brain organizes the information and initiates motoric and logical and philosophical solutions, all without written step-by-step instructions. DNNs in biology appear to be driven, and are perhaps embedded within, “programmed” genetic codes, instincts and biological markers. DNNs in technology attempt to mimic biologic DNN using human-made digital tools to better process data and arrive at the most likely beneficial solution to the problem under consideration.

Lesica²⁰ reported hearing aids should ideally restore the neural activity patterns sent to the brain to decrease the effort and energy exerted by the brain while processing a degraded neural code. Ronnberg et al.²¹ noted that when the brain uses less energy to process sound, it has more energy left to store what was received in memory. As above, to comprehend speech sounds (i.e., to listen) after hearing occurs, the brain must orient and then focus on the sound of maximal interest. Selective Attention facilitates the brain’s ability to separate sound into foreground (maximal attention) and background (awareness of, but only slight attention to). An ideal DNN processor should facilitate a complete and balanced sound scene in which the most important sounds appear in the foreground, while attenuating (yet still availing) background sounds.¹⁹

DNNs in hearing aids

In 2021, Oticon Inc. released the world’s first commercially available hearing aid with “on-the-chip” DNN technology. The Oticon More™ DNN has been trained on 12 million sound samples to facilitate improved speech in noise ability, to help improve recall/memory, to deliver a very high sound quality, and to improve selective attention. It is founded on the idea that improving the neural code enhances the brain’s ability to make sense of sound. The DNN within Oticon More represents a revolutionary way of processing sounds for use in personalized hearing aids. Research based on 64 channel EEG studies¹³ demonstrates that Oticon More improves selective attention. Specifically, 31 experienced hearing aid wearers (mean age 65 yrs) with mild-to-moderate SNHL were instructed to attend to two

different talkers, each at 73 dB SPL in a highly challenging acoustic environment with four-talker babble originating from four separate locations, each at 70 dB SPL. With regard to the amplitude of the EEG recordings, the DNN improved the brain’s ability to “track all the objects in the full acoustic sound scene” by 60% (DNN enabled versus disabled). Further, (as compared to our best previous hearing aid technology, Oticon Opn S) the DNN allowed 30% better access to the full acoustic sound scene. The DNN hearing aid wearers demonstrated overall improvements in the ability to understand speech in noise with less effort, and they showed an improved ability to recall spoken words.²²

Conclusion

In 2021, there are well-known and clearly defined differences between hearing and listening. However, professionals and patients often convolute the two concepts, which is disadvantageous for all. As professionals, it seems appropriate to expand and clarify our thoughts and explanations from simply making things louder/audible (i.e., hearing) to using commercially available life changing technology to enable a better understanding of speech in noise (i.e., listening) while enhancing selective attention and reducing listening effort secondary to an improved neural code. Basic analog sound processing techniques and protocols from the last century were an excellent and timely starting point to address hearing loss through enhanced audibility. Yet, in 2021, we can provide vastly more sophisticated hearing amplification which provides an enriched neural code to better support listening (i.e., the ability to comprehend sounds). Specifically, Deep Neural Networks (DNNs) “on-the-chip” have been shown to provide real-time sound processing, a clearer understanding of speech in noise, improved recall, can maintain interaural loudness differences and can improve the wearer’s ability to selectively attend to the speaker of choice in challenging listening situations. These improvements are delivered without acoustic feedback and with very high sound quality. DNNs represent life changing technology in hearing aids. Early studies have demonstrated that wearers of DNN-based hearing aids were provided more access to the complete sound scene, allowing the wearer to better focus on the sounds they choose in the foreground, while not losing access to meaningful background sounds.¹⁹

Acknowledgment

None.

Conflict of interest

None.

Funding

None.

References

1. DePaolis RA, Janota CP, Frank T. Frequency Importance for Words, Sentences and Continuous Discourse. *Journal of Speech, Language and Hearing Research (JSLHR)*. 39(4): 714–723.
2. Beck DL, Flexer C. Listening is Where Hearing Meets Brain in Children and Adults. *Hearing Science*. 2011;18(2): 30–35.
3. Pereira-Jorge MR, Andrade KC, Palhano-Fontes FX, et al. Anatomical and Functional MRI Changes after One Year of Auditory Rehabilitation with Hearing Aids. *Neural Plasticity*. 2018:13.

4. Glick HA, Sharma A. Cortical Neuroplasticity and Cognitive Function in Early-Stage, Mild-Moderate Hearing Loss: Evidence of Neurocognitive Benefit from Hearing Aid Use. *Front Neurosci.* 2020;14:93.
5. Beck DL, Clark JL. Audition Matters More as Cognition Declines and Cognition Matters More as Audition Declines. *Audiology Today. American Academy of Audiology.* 2009.
6. Beck DL, Harvey M. Issues in Cognition, Audiology and Amplification. *Hearing Review.* 2021.
7. Amieva H, Ouvrard C, Giulioli C, et al. Self-reported hearing loss, hearing aids, and cognitive decline in elderly adults: A 25-year study. *J Am Geriatr Soc.* 2015;63(10):2099–104.
8. Amieva H, Ouvrard C. Does Treating Hearing Loss in Older Adults Improve Cognitive Outcomes? A Review. *J Clin Med.* 2020;9(3):805.
9. Seeto M, Tomlin D, Dillon H. The Relations Between Auditory Processing Scores and Cognitive, Listening and Reading Abilities. *Ear Hear.* 2021.
10. Golub JS, Brickman AM, Ciarleglio AJ, Schupf N, Luchsinger JA. Association of subclinical hearing loss with cognitive performance. *JAMA Otolaryngol Head Neck Surg.* 2020;146(1):57–67.
11. Gaeta L, Azzarello J, Baldwin J, Ciro CA, Hudson MA, Johnson CE, John AB. Effect of reduced audibility on Mini-Mental State Examination scores. *J Am Acad Audiol.* 2019;30(10):845–855.
12. Shinn-Cunningham BG, And Best V. Selective Attention in Normal and Impaired Hearing. *Trends Amplif.* 2008;12(4):283–299.
13. Alikovic E, Ng EHN, Fiedler L, et al. Effects of hearing aid noise reduction on early and late cortical representations of competing talkers in noise. *Currently in revision for Frontiers in Neuroscience.* 2021.
14. O’Sullivan J, Herrero J, Smith E, et al. Hierarchical Encoding of Attended Auditory Objects in Multi-Talker Speech Perception. *Neuron.* 2019;104(6):1195–1209.
15. Hausfeld L, Riecke L, Valente G, et al. Cortical Tracking of Multiple Streams outside the Focus of Attention in Naturalistic Auditory scenes. *NeuroImage.* 2018;181:617–626.
16. Puvvada, KC, Simon JZ. Cortical Representations of Speech in a Multi-talker Auditory Scene. *J Neurosci.* 2017;37(38):9189–9196.
17. Beck DL. Cognition, Audition and Amplification. Keynote presentation to the Missouri Hearing Society. Saint Louis, Mo. 2020.
18. Miller SE, Zhang Y. Neural Coding of Syllable-Final Fricatives with and without Hearing Aid Amplification. *J Am Acad Audiol.* 2021;31(8): 566-577.
19. Santurette S, Behrens T. The Audiology of More. Whitepaper. 2020.
20. Lesica NA. Why Do Hearing Aids Fail to Restore Normal Auditory Perception? *Trends Neurosci.* 2018;41(4):174–185.
21. Ronnberg J, Lunner T, Zekveld A, et. Al. The Ease of Language Understanding (ELU) Model: Theoretical Empirical and Clinical Advances. *Front Syst Neurosci.* 2013;7:31.
22. Santurette S, Ng EHN, Jensen JJ, et al. Oticon More Clinical Evidence. *Whitepaper.* 2020.