Canadian Audiologist
The Official Publication of the Canadian Academy of Audiology

To the Brain and Back: Auditory Attention Decoding

Published April 8th, 2026

Brandon T. Paul, PhD

Attention is an essential cognitive ability that lets us filter through the rich amount of sensory information we encounter every day. An important way attention supports speech communication is through “cocktail party listening,” [1] the ability to focus on one person speaking while others talk in the background. In people with normal hearing, the brain and ear work together to achieve this in an almost seamless, automatic process. With hearing loss, cocktail party listening is more difficult because the acoustic features that would distinguish one talker are less easily perceived [2]. Modern hearing aid processors do well to suppress background noise, but hearing aid users still face attentional challenges, as do cochlear implant users. One reason for this is attention switching: a talker who was once part of the background noise becomes the desired signal when the focus of attention changes. There are no simple rules that signal processors use to handle this demand [3]. Can neuroscience research help?

Enter *neuro-steered* hearing aids or cochlear implants – hearing devices that adapt their speech signal processing based on brain activity [3]. Neuro-steered hearing aids are an active area of development in biomedical sciences and brain–computer interfaces. This article will describe how attentional modulation of the electroencephalogram (EEG) can control hearing aid processors. EEG records voltage fluctuations across the scalp that reflect ongoing sensory and cognitive brain activity. For decades, we’ve known that auditory evoked potentials recorded by EEG have larger amplitudes for attended stimuli than for unattended stimuli [4], suggesting that attention enhances neural representations. More recent methods in the last 15 years can show similar attentional enhancement for continuous speech in realistic listening situations, such as at the cocktail party [5]. Several continuous speech processing methods in EEG rely on the synchronization of electrical brain activity to the speech envelope, the slow changes in amplitude of a speech signal (See a previous “Brain and Back” for more information).

Auditory Attention Decoding

A*uditory attention decoding* (AAD) refers to algorithms that use physiological data, such as EEG, to estimate which talker a person is attending to. For example, consider a simple scenario in which a listener must attend to one talker and ignore the other (Figure 1). Signal analysis steps inside the processor can separate the talker mixture (e.g., using beamforming), and apply additional denoising steps. EEG electrodes on the device pick up brain activity, and the processor compares the EEG to each speech signal. Once the attended talker is identified, their speech will be enhanced, and other signals will be attenuated. In short, a person’s brain can guide the hearing aid based on their attentional state.

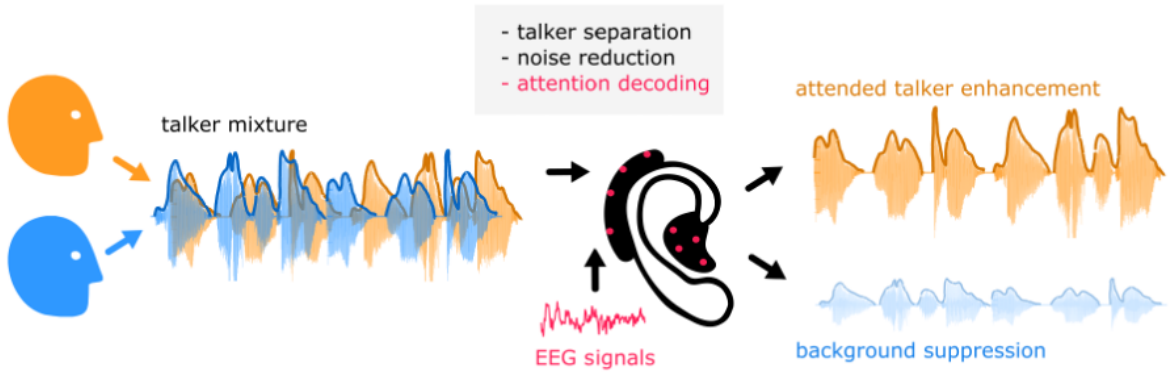


Figure 1. Simple AAD diagram. A mixture of talkers (orange and blue, left side) would be picked up by the hearing aid or CI microphones. Signal processing steps inside the processor (grey box) would separate the talkers and denoise them. EEG signals recorded from sensors on the processor or earpiece (magenta) would be compared with the speech signals, through either linear methods like envelope reconstruction, or through non-linear methods like deep neural nets. Once the attended talker is identified, the attended speech is then boosted by the hearing aid, while background talkers are suppressed. The speech envelope is shown as the boldface line that outlines the speech waveforms.

To build an AAD, the algorithm must first learn the idiosyncrasies of each person's brain activity and how they deploy attention. This model training step is crucial and is done in a controlled, supervised manner, such as in a basic two-talker setup described above, where a person is explicitly instructed to attend to one of multiple talkers [3]. The algorithm is given the ground truth (which talker is being attended to) so it can refine its parameters to best match the listener's brain signatures of attention. The goal of training is to "generalize" the algorithm so it accurately identifies the attended talker in new listening situations without needing to know the ground truth in advance. However, future AAD algorithms may bypass this step. Recent AAD algorithms can adapt to a listener "on the go" in an unsupervised manner without prior training [6].

There are a variety of algorithms available to perform AAD. Simpler, linear methods rely on synchronizing brain activity with speech envelopes. They are called "linear models" because they involve linear transformations of data and linear correlations between model outputs and real speech [5]. To explain, brain activity will partly align with the speech envelopes of multiple talkers, but as mentioned, the alignment to the attended talker is stronger. AAD algorithms create linear models that predict the speech envelope based on current brain activity. These predicted speech envelopes are known as "reconstructed" envelopes. The reconstructed envelopes are then correlated with each real speech envelope, and the attended talker is the one that produces the strongest correlation. In other words, the speech envelope with the best model reconstruction accuracy likely belongs to the attended talker. Once the AAD algorithm makes this determination, the processor amplifies the attended speech and suppresses the background noise.

There are also non-linear methods that rely on deep neural networks, which are more computationally advanced algorithms that can learn highly complex patterns in brain data [7]. While linear methods rely on decoding the envelope, there is far more information available in an EEG recording that deep neural networks can leverage. Some deep neural network methods learn the spatial direction of a person's attention directly from brain activity, without calculating envelope correspondence. This can be advantageous when the processor struggles to cleanly extract individual envelopes. The downside of non-linear AAD models is that they are harder to train and often face reliability issues.

Implementation

How would a hearing aid or cochlear implant processor incorporate EEG? While most people think of EEG as a montage of sensors and wires spread across the scalp, it is possible to “miniaturize” electrodes and line them through parts of an earpiece or over-ear processor. Figure 2A and B show prototype earpieces containing small EEG sensors, which are easy to imagine as part of a hearing aid. For cochlear implants, EEG sensors could be embedded in the processor and earhook, as demonstrated in the “cEEGrid” (C-shaped EEG layout) in Figure 2D. Compared to an EEG montage across the scalp, an electrode array in and around the ear covers considerably less surface area. While this reduces data quality [8], it is possible to strategically place a small set of sensors to maintain good AAD performance [9]. For example, a recent study showed that a completely wireless EEG setup that combines electrode information from both ears produces reliable attention-decoding performance [10].



Figure 2. In- and around-ear EEG. Panel A and B show dry contact electrodes (black dots) throughout an ear piece. Panel C shows the sensor sizes in comparison to sensor mounts found on a typical EEG cap. D shows both a transparent around-ear montage (known as a cEEGrid) as well as the ear-EEG mould. Original image from Kappel, S. L., Makeig, S., & Kidmose, P. (2019). Ear-EEG forward models: improved head-models for ear-EEG. Frontiers in neuroscience, 13, 943. Reproduced here under CC BY International 4.0 license.

Challenges and Future Directions

Linear AAD algorithms must have access to the acoustical speech envelopes for talker separation and envelope reconstruction, but this is highly dependent on the microphones' orientation relative to the talkers, which makes envelope reconstruction difficult in some circumstances [3]. Nonlinear

deep neural network methods may be preferred in the long term to sidestep envelope reconstruction, although talker separation is still required to determine which speech signal to boost. Another challenge is that the algorithms are slow to keep up with rapid attention switching. AAD cannot determine attention instantaneously, and algorithms may need several seconds of data before attention is decoded. Non-linear methods may again have an advantage in this regard, needing only ~2 seconds of data to achieve ~80% accuracy [7]. Finally, performance in everyday circumstances is still unclear. Although AAD algorithms have shown robustness despite realistic barriers such as background noise and reverberation [11], AAD training and testing typically occur in idealized, well-controlled laboratory environments. Fortunately, in the last year, AAD has shown strong performance in natural conversational settings involving two or three people [12], a good first step. Nonetheless, a complete real-world characterization of a neuro-steered hearing device is needed before it's ready for the cocktail party.

References

1. McDermott, J. H. (2009). The cocktail party problem. Current Biology, 19(22), R1024-R1027.
2. Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. Trends in amplification, 12(4), 283-299.
3. Geirnaert, S., Vandecappelle, S., Alickovic, E., De Cheveigne, A., Lalor, E., Meyer, B. T., ... & Bertrand, A. (2021). Electroencephalography-based auditory attention decoding: Toward neurosteered hearing devices. IEEE Signal Processing Magazine, 38(4), 89-102.
4. Näätänen, R., & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: a critical review and reinterpretation. Biological psychology, 8(2), 81-136.
5. O'sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., ... & Lalor, E. C. (2015). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. Cerebral cortex, 25(7), 1697-1706.
6. Geirnaert, S., Francart, T., & Bertrand, A. (2021). Unsupervised self-adaptive auditory attention decoding. IEEE journal of biomedical and health informatics, 25(10), 3955-3966.
7. Vandecappelle, S., Deckers, L., Das, N., Ansari, A. H., Bertrand, A., & Francart, T. (2021). EEG-based detection of the locus of auditory attention with convolutional neural networks. Elife, 10, e56481.
8. Mirkovic, B., Bleichner, M. G., De Vos, M., & Debener, S. (2016). Target speaker detection with concealed EEG around the ear. Frontiers in neuroscience, 10, 349.
9. A. M. Narayanan and A. Bertrand, "Analysis of miniaturization effects and channel selection strategies for EEG sensor networks with application to auditory attention detection," IEEE Trans. Biomed. Eng., vol. 67, no. 1, pp. 234–244, 2020.
10. Geirnaert, S., Ding, R., & Bertrand, A. (2026). Selective Auditory Attention Decoding with a Two-Node Wireless EEG Sensor Network. bioRxiv, 2026-02.
11. Das, N., Bertrand, A., & Francart, T. (2018). EEG-based auditory attention detection: boundary conditions for background noise and speaker positions. Journal of neural engineering, 15(6),

066017.

12. Van de Ryck, I., Heintz, N., Rotaru, I., Geirnaert, S., Bertrand, A., & Francart, T. (2026). EEG-based Decoding of Auditory Attention to Conversations with Turn-taking Speakers. Hearing Research, 109539.