Canadian Audiologist

The Official Publication of the Canadian Academy of Audiology

Speech in Brain in Noise: Neural Plasticity in the Auditory Brainstem

Published May 14th, 2015

Gabe Nespoli

When sound enters the ear, the cochlea and other inner ear structures convert the auditory signal into a neural signal, and so begins the process of perceiving the sound. The path from ear to brain is not a particularly long one, but the signal does make a few stops before arriving in auditory cortex (Figure 1). Brainstem structures such as the cochlear nucleus, the olivary complex and the inferior colliculus all see the signal pass through before it gets to the cortex.

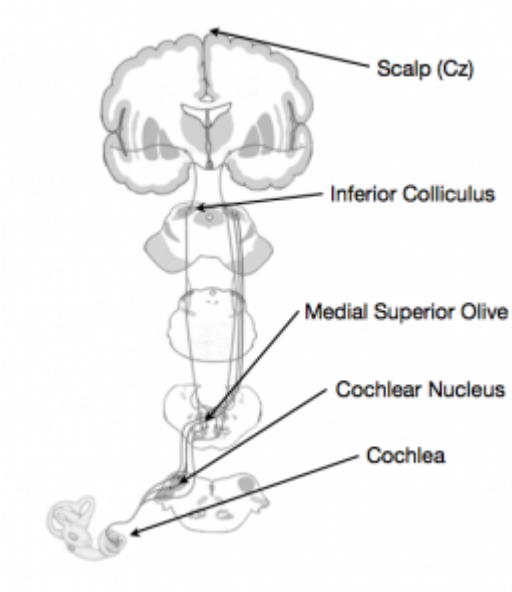


Figure 1. The auditory pathway.

At each of these stops, the similarity between the timing of the neural firing and the oscillation of the original signal wave is quite remarkable. Smith and colleagues measured electrophysiological activity in these neural structures (the auditory brainstem response, or ABR) in response to pure tones, by implanting electrodes in the brains of cats.[1] They also measured activity on the surface of the scalp. They noticed was that the response was oscillatory, with increases and decreases in activity repeating in a regular fashion in much the same way as a sound waveform. In fact the oscillations were occurring at the same frequency as the pure tone that was being heard (Figure 2). As a result they called this phenomenon the frequency following response (FFR), for its ability to mimic the periodic nature of a sound signal. In fact, if you play back the brain response as sound, it sounds like the original signal!

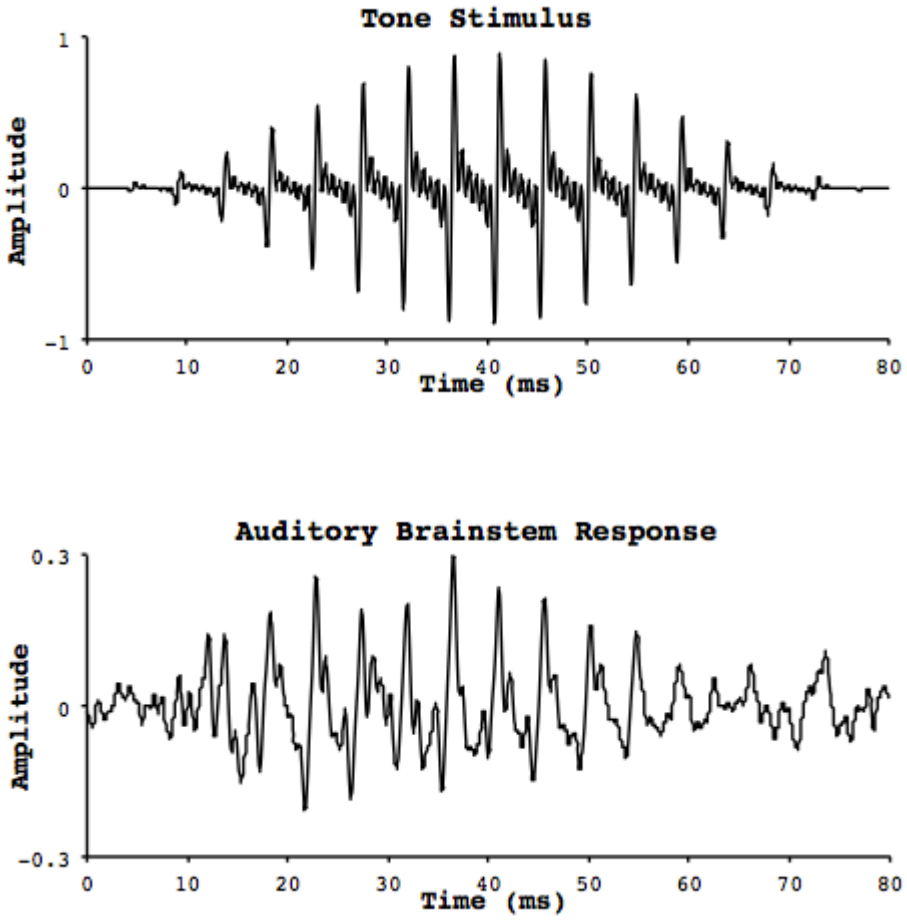


Figure 2. FFR stimulation and response.

The second thing they noticed was that as their recording site moved further along the processing stream – further away from the cochlea and closer to the cortex – the onset latency of the response increased. This is perhaps not surprising since neural transmission is not instantaneous, and so the further the signal travels the more time it will take. The latency was on the order of 6–10 milliseconds (ms) by the time the signal reached the inferior colliculus. Interestingly, the latency of the response on the surface of the scalp was approximately the same (6–10 ms). How could this be?

Whereas intracellular recordings are a direct measure of neural activity, scalp recordings are an indirect measure and thus represent a grand average of all electrical activity that reaches the electrode. Proximity to the electrode and strength of activity from a given neural source are factors that will affect a source's influence on the electrode. The activity measured from a scalp electrode represents the sum of activity from many different sources in the brain, and when many responses to the same stimulus are averaged together the result represents elements that were consistent in the brain's response across all presentations of that stimulus. In the case of Smith et al.'s cats,[1] they reasoned that this was the signal originating in the inferior colliculus, since the two responses had a similar latency. What a discovery! There appear to be structures in the brainstem that almost exactly mimic the oscillatory nature of auditory tones, and given enough trials this activity can be measured non-invasively from the surface of the scalp.[2]

These first studies only considered pure tones, but the ABR maintains a high degree of similarity to complex auditory stimuli as well, and preserves more than just the fundamental frequency. To be sure, we would expect as much from the ABR if it is to supply the cortex with all the information it

needs to perceive a signal. Features like the pitch contour, loudness, and spectrum are all preserved in the ABR signal. Notably though, the FFR has an upper limit of about 1500 Hz and its amplitude decreases with increasing frequency.

Like most other brain responses, the ABR is plastic and can undergo experience-based changes. Both musicians and tone-language speakers have received some attention in the literature in this regard, since being able to track pitch is tied to the understanding of musical or linguistic material. For example, FFRs elicited by Mandarin tones from native Mandarin speakers exhibited stronger representation of pitch and smoother pitch tracking than those of English speakers.[3] Similarly, FFRs recorded from musicians showed larger spectral peaks at certain harmonics than nonmusicians when listening to consonant and dissonant intervals.[4] Perhaps most interesting is a study by Bidelman et al, who compared FFRs and pitch discrimination of English speaking musicians and nonmusicians, and Mandarin speaking nonmusicians to musical stimuli (arpeggiated triads).[5] Both major and minor triads were played one note at a time starting with the lowest note (i.e., root-third-fifth). The only difference between a major and a minor triad is that the middle note (the third) is one semitone lower in pitch than the major triad (e.g., C-E-G vs. C-Eb-G). They also used a "detuned up" version (the third had a higher f0 than in the major triad) and a "detuned down" version (the third had a lower f0 than in the minor triad). While ABRs were being recorded, participants were presented with two arpeggios and asked to respond whether they were the same or different (possible combinations were major/minor, major/detuned up, and minor/detuned down). Spectral peaks in the FFR at the f0 of the middle tone were found to be significantly larger in musicians and Mandarin-speakers compared to nonmusicians. This was to be expected since both Mandarin-speakers and musicians discriminate pitch on a daily basis, leading to enhanced neural pitch tracking in the brainstem. In the discrimination task, though the Mandarin-speakers were able to reliably discriminate between major and minor triads, they performed just as poorly as nonmusicians for the detuned conditions. This suggests that while enhanced brainstem responses may relay a higher-fidelity signal, cortical areas responsible for making the discrimination decisions must also have adequate experience with the specific task at hand. In this case, embedding the pitch discrimination into a melodic sequence of pitches gave the musicians an advantage. We might expect a different pattern of results if the stimuli were Mandarin tones embedded in a sentence.

Musicians also see benefits in language – at least in their native language – notably performing better on speech-in-noise (SIN) tasks. Music requires one to continually parse certain signals apart from others, such as a violin melody amongst a string orchestra, which is more-or-less the same task as SIN. Parbery-Clark et al. investigated brainstem responses of musicians and nonmusicians in quiet and in noise, comparing the neural measures with behavioural scores on the Hearing in Noise Test (HINT).[6] For ABR measurements, participants listened to a synthesized speech sound (/da/), which can be characterized by (1) an onset, (2) a transition period between the consonant /d/ and the vowel /a/, and (3) the steady-state vowel portion /a/. Importantly, the frequency content present in the signal changes rapidly during the transition period (~50 ms), and is held constant during the steady-state portion (~120 ms). Parbery-Clark et al. found that the da-in-noise was much better represented in the musicians than the nonmusicians, with the response waveform correlating better with the stimulus waveform. There was increased response latency during the onset and transition periods in noise compared to quiet for both groups, but this latency increase was smaller in musicians.[6] During the steady-state portion, the f0 was equally well-represented in all participants, but the upper harmonics (H2-H10) were better represented in musicians. Most interestingly though, stimulus–response correlations and transition latencies were correlated with

better HINT scores, suggesting that the precision of neural timing has behavioural consequences for SIN.

Hearing declines with age, and so does the ABR, but musicians tend not to see the same extent of decline. In fact, age-related decline in the ABR has been linked with SIN performance. Parbery-Clark and colleagues found that the ABRs of younger and older nonmusicians also differed during the transition portion of the signal (but not the steady-state portion); the latency of the response was increased in the older adults.[7] This same pattern was not seen for the younger and older musicians, with both groups showing comparable latency for both transition and steady-state portions of the ABR.

So the experience-driven SIN benefits seen in older musicians are also seen in the ABR, suggesting that these neural structures early on in the processing stream might play an active role in shaping the perception of sound.

Taken together, these data seem to shed light on the brainstem's role in perceiving sound. ABRs in noise seem to affect the high-frequency portions of the signal the most, such as formant transitions and upper harmonics.[6] The low-pass nature of age-related hearing loss seems to be reflected in the ABR. High-frequency components of speech also tend to be most affected during SIN[6]; perhaps the neural deficit in resolving high-frequency content underlies the SIN deficit that comes with age. Playing music is apt for improving SIN performance, and this is reflected in the experience-dependent nature of the brain including the ABR.[5] Musicianship also appears able to stave off SIN deficits in old age, as well as preserve ABRs,[6] suggesting that these two things might be related in some way.

Synchronization in the auditory brainstem is the brain's way of representing the auditory world, and the precision of these structures has profound effects on the perception of sound. The ABR is a valuable method for evaluating the ability to represent sound, and offers insight into the noisy world of the human brain.

References

1. Smith JC, Marsh JT, and Brown WS. Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. Electroencephalograph Clin Neurophysiol 1975;39:465–72.
2. Aiken SJ and Picton TW. Human cortical responses to the speech envelope. Ear Hear 2008;29(2):139–57.
3. Krishnan A, Xu Y, Gandour J, and Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. Cogn Brain Res 2005;25(1):161–68. doi:10.1016/j.cogbrainres.2005.05.004
4. Lee KM, Skoe E, Kraus N, and Ashley R. Selective subcortical enhancement of musical intervals in musicians. J Neurosci 2009;29(18):5832–40. doi:10.1523/JNEUROSCI.6133-08.2009.
5. Bidelman GM, Gandour JT, and Krishnan, A. Brain Cognition 2011;77(1):1–10. doi:10.1016/j.bandc.2011.07.006.
6. Parbery-Clark A, Skoe E, and Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci 2009;29(45):14100–107. doi:10.1523/JNEUROSCI.3256-09.2009
7. Parbery-Clark A, Anderson S, Hittner E, and Kraus, N. Musical experience offsets age-related delays in neural timing. Neurobiol Aging 2012;33(7):1483.e1–1483.e4. doi:10.1016/j.neurobiolaging.2011.12.015.