Auditory Evoked Potentials

J R Melcher, Harvard Medical School, Boston, MA, USA
© 2009 Elsevier Ltd. All rights reserved.

Introduction

When sounds reach our ears, they are transduced by peripheral hearing structures into electrical signals in the auditory nerve. The nerve transmits these signals to the central auditory system, where they are processed by neurons in the brain stem, thalamus, and cerebral cortex. The result of this processing is the perception of sound. Many thousands of neurons are engaged in the transformation from sound to perception. Neuronal activity produces in the surrounding brain tissue a voltage that can be recorded from electrodes on the surface of the head. A measurement of the voltage produced collectively by the active neurons is an auditory evoked potential (AEP).

AEPs Produced by Brief Stimuli

Many sounds produce a measurable AEP, but AEPs produced by brief sounds, such as acoustic pulses (called clicks), brief tones, and short bursts of broadband noise, are especially robust. Figure 1 shows a stylized AEP in response to a click. The response consists of a series of voltage fluctuations continuing for hundreds of milliseconds following the sound stimulus. In response to a single stimulus presentation, most, if not all, of these fluctuations are swamped by the voltage produced by ongoing activity in the brain (i.e., the electroencephalogram). An AEP is extracted from this ongoing signal by averaging the voltage following many presentations of the same stimulus.

The AEP illustrated in Figure 1 is based on recordings in people. However, AEPs have also been recorded in many other species, including mice, guinea pigs, rats, cats, and dolphins. The AEP waveform differs across species, reflecting, for instance, interspecies differences in head size, axonal path lengths, physical location and orientation of cells in the head, and the specific auditory cell types present in the auditory pathway. In all cases, however, there is a major commonality: the AEP comprises a time-varying voltage following the stimulus.

The AEP produced by a brief stimulus is generally divided into three parts:

1. The auditory brain stem response (ABR) consists of the shortest latency fluctuations occurring within approximately 10 ms of the stimulus (green in Figure 1). The voltage fluctuations of the ABR are each ~1 ms in width. As the name of this response indicates, these fluctuations are generated by auditory neurons in the brain stem.
2. The middle latency response (MLR) consists of a series of voltage fluctuations immediately following the ABR and preceding the long latency response (red in Figure 1). The voltage fluctuations comprising the MLR are much broader than are the fast fluctuations of the ABR. The MLR partly reflects cellular activity in the thalamus.
3. The late latency response (LLR) begins 50–100 ms after the stimulus (blue in Figure 1). It includes one of the most widely studied components of the AEP, the N1 or N100. The complete response consists of a complex of positive and negative fluctuations generated by the cortex.

Dependencies on Stimulus Parameters

The ABR, MLR, and LLR all depend on physical aspects of the sound stimulus. For instance, increases in the rate of stimulus presentation generally decrease response amplitude. Increases in stimulus intensity generally increase the amplitude of responses and decrease the latency (i.e., shorten the time between stimulus and response). These rate and intensity dependencies are partly attributable to the behavior of the auditory nerve, which generates the first wave of the ABR and provides input, directly or indirectly, to neurons contributing to the remainder of the AEP. For example, individual auditory nerve fibers are more likely to discharge in response to high-intensity stimuli and will do so with a shorter latency. Wave I of the ABR, which reflects the summed response from all auditory nerve fibers, is correspondingly greater in amplitude and shorter in latency, the typical intensity dependence for AEPs.

Dependencies on Arousal and Attention

The MLR and LLR differ from the ABR in their sensitivity to whether individuals exposed to the stimulus are asleep or awake, whether they are paying attention to the stimuli, and whether they are anesthetized. For instance, the LLR can change substantially across different stages of sleep. It also shows distinct changes in amplitude when individuals perform a task that requires concentrated listening (i.e., attention) to the stimuli used to evoke the LLR. The ABR is highly resistant to changes in state, persisting, for instance,
even when individuals are anesthetized deeply enough for surgery. The resilience of the ABR has made it an effective tool for monitoring the health of the auditory system during neurosurgery (e.g., to remove a tumor from the vestibulocochlear nerve) and for infant hearing tests (ideally performed during sleep).

Neural Generators of AEPs

While the ABR, MLR, and LLR differ in many ways (amplitude, latency, duration of individual voltage fluctuations) and are generated by different neurons, their mechanism of generation is fundamentally the same, since all represent the net potential produced by large populations of neurons. Several factors determine whether any given neuronal population will contribute substantially to the overall AEP:

1. The total number of neurons in the population.
2. The amplitude of the potential produced by individual neurons at the AEP recording electrodes. This amplitude can be dependent on neuronal morphology. For example, when current flows across the membrane of a neuronal cell body and returns through the dendrites, a so-called closed field results if the dendrites uniformly surround the cell body (i.e., zero potential at the AEP electrodes), whereas an open field results if the dendrites are distributed non-uniformly (i.e., nonzero potential).
3. Similarity of morphology and orientation across neurons. For example, two oppositely oriented neurons next to one another will produce equal but opposite (i.e., canceling) potentials.
4. The degree to which activity is temporally synchronized across neurons. The potentials produced by the individual neurons of a population will only add effectively if the neuronal responses to the AEP stimulus are similar in latency (i.e., temporally aligned with one another).

A variety of approaches have been used to identify neural generators of the AEP. One widely used approach involves making focal brain lesions in animals and examining the effect on the AEP. Lesioning techniques used for this purpose include making surgical cuts, passing electric current, or injecting a neurotoxin. The effects of neurotoxic lesions, combined with modeling, have led to the conclusion that most of the click-evoked ABR in cats is generated by neurons in two of the many pathways of the auditory brain stem. These pathways originate with the bushy cells of the cochlear nucleus, the most peripheral nucleus of the central auditory system and recipient of direct inputs from the auditory nerve. They extend centrally to the superior olive, nuclei of the lateral lemniscus, and inferior colliculus. In cats, there are two types of bushy cells, spherical and globular, which mark the beginning of the two pathways involved in ABR generation. The pathways emanating from bushy cells...
have specialized features for maintaining the timing of neuronal activity (e.g., large-diameter axons and unusually large terminal endings called endbulbs), and hence maintain the high degree of temporal synchronization across neurons present in the auditory nerve. This temporal synchronization is one of several features that make bushy cells (spherical and globular), as well as the direct and indirect targets of these cells, robust contributors to the cat ABR. In contrast to cats, the globular bushy cell population in humans is proportionately quite small, and according to some is nonexistent. Therefore, the human ABR (after wave II) may largely reflect activity in just one pathway consisting of spherical bushy cells and their primary and secondary targets (e.g., medial superior olive, inferior colliculus).

The potential contributed to the AEP by brain stem neurons likely reflects discharge activity on the cell body and axon. (This is consistent with the brevity of the fluctuations comprising the ABR, ~1 ms.) In contrast, the traditional view of cortically generated potentials (e.g., P1, N1, P2, of the LLR) is that they arise from synaptic currents flowing between the cell body and dendrites of pyramidal cells (a synaptic origin is consistent with the longer duration of the fluctuations comprising the LLR). The cell bodies of pyramidal cells, lying in the deeper layers of cortical gray matter, give rise to an apical dendrite extending into the superficial layers and oriented perpendicular to the cortical surface. The morphology of pyramidal cells (resulting in an open, or dipole, field) and the consistent dendritic orientation across pyramidal cells mean that large groups of synchronously activated pyramidal cells are capable of generating a robust evoked potential.

The model for pyramidal cells just described is generally assumed by one of the most extensively used approaches for identifying cortical generators of AEPs. The approach involves (1) measuring the AEP at many finely spaced locations over the head (e.g., 100 or more), (2) modeling the underlying cortical activity using the dipole model suggested by the morphology and dendritic orientation of pyramidal cells, (3) assuming an electric model for the head, and (4) determining the strength, location, and orientation of dipole(s) that yield a best fit between the measured AEP and AEP calculated based on the head and dipole models. The location of AEP generators estimated by this approach can be further refined using information from other imaging modalities, including structural and functional magnetic resonance imaging (MRI). The results are vivid spatial mappings of cortical activity unfolding over time.

Magnetic Analog of the AEP

When neurons responding to sound produce an AEP, they also produce a magnetic field recordable outside the head. While small (about 100 000 times less than the magnetic fields in our everyday surroundings), these auditory evoked magnetic fields (AEMFs) can be measured routinely when specialized, low-noise magnetic detectors are used. Both AEPs and AEMFs are produced directly by neuronally generated currents in the brain. This distinguishes AEP and AEMF recordings from other modalities for measuring brain function, such as functional MRI (fMRI) and positron emission tomography (PET), which detect hemodynamic changes accompanying changes in neuronal activity, rather than neuronal activity itself.

While both AEPs and AEMFs are a direct reflection of neuronal activity, the two types of recordings differ and are, in many respects, complementary. For instance, AEMFs, but not AEPs, are preferentially sensitive to cortical activity in pyramidal cells having a particular dendritic orientation relative to the surface of the head (dendrites parallel to the head surface, creating what is called a tangential dipole source). Much of the gray matter on the human superior temporal lobe contains pyramidal cells with this orientation, including the crown of Heschl’s gyrus, where primary auditory cortex is typically located, and planum temporale, which contains nonprimary areas. Because the orientation of auditory cortex is well suited for magnetic recordings, AEMFs have been used extensively to investigate human auditory cortical function.

Other, Specialized AEPs

In order to obtain a specific indicator of activity in limited neuronal populations, responses are sometimes derived from multiple AEP measurements. The binaural difference potential (BDP) illustrates the selectivity provided by a derived response; it specifically reflects activity involved in comparing sounds between the two ears. A BDP is derived by summing the ABRs produced by left and right monaural stimulation and subtracting the ABR produced by stimulating both ears simultaneously. For click stimuli, the summed monaural and binaural ABRs show no difference during waves I–III, but diverge during waves IV/V. This difference is the BDP. A BDP would not be detected if sounds at the two ears were processed independently; the BDP reflects an interaction between the two ears. The coincidence of the BDP with the ABR indicates that it is generated by brain stem neurons.
Animal lesion data and the prominence of a particular binaural nucleus in the brain stem of humans, the medial superior olive (MSO), have led to suggestions that MSO neurons contribute substantially to the human BDP.

While AEPs can be elicited by discrete stimuli presented in silence, they can also be elicited by brief changes in ongoing sound. One example is the lateralization shift response (LSR), a cortically generated potential produced by changing the interaural time difference of binaurally presented continuous noise. A paradigm for eliciting this response involves initially presenting identical noise to the two ears (the noise is heard diffusely, but centered in the head), then, over a brief interval, temporally shifting the noise at one ear relative to other (by ~1 ms), and then quickly shifting it back (the diffuse, centered noise is still heard without interruption, but a brief sound lateralized to the ear with the leading noise is also heard). The result is an evoked response 100–300 ms after the shift, the LSR. Since the noise presented to either ear alone continues without disruption, it does not produce an AEP. The LSR only occurs when the noise is presented to both ears. Thus, like the BDP, the LSR specifically reflects binaural processing within the auditory central nervous system, although at a cortical rather than brain stem stage.

In addition to depending on the physical features of the stimuli used to evoke them, evoked potentials can also depend on the context in which the stimuli are presented. An example of a context-dependent AEP is the mismatch negativity (MMN). An MMN is elicited in a paradigm that involves presenting a standard stimulus (e.g., a tone burst) at regular intervals but, every so often, replacing the standard with a deviant sound (e.g., a tone burst of slightly different frequency; also called a rare stimulus). Responses to the deviant sound presented in the context of this paradigm (called an oddball paradigm) are different from the responses to the same stimulus presented over and over again on its own. The difference is the MMN (difference between the solid trace, which is the stimulus presented on its own, and the dashed trace, which is the stimulus presented as deviant, in Figure 1). The MMN, generated largely by the auditory cortex, can be thought of as a form of sensory memory, 'sensory' because it originates within cortex that specializes in encoding one of our senses, and 'memory' because a history of the stimuli appears to be retained by cortex; each incoming sound is compared to its predecessors and an MMN is produced when they differ. An MMN can be recorded even when individuals are distracted from the stimuli (e.g., by watching a movie), indicating that the MMN reflects, at least in part, neuronal processing prior to conscious perception.

**AEP Correlates of Auditory Perception**

Since AEPs can be recorded in awake people while they listen to sounds, they offer a way to understand the neural processing that determines how sounds are perceived. Two examples of this important role for AEPs follow.

The first example concerns the binaural difference potential and its correlation with certain perceived aspects of binaurally presented clicks. When clicks of equal intensity are presented simultaneously to the two ears, they are perceived as a single, fused object that is centered with respect to the head. As the clicks presented to one ear are made greater in intensity or shorter in latency, compared to the clicks presented to the other ear, the perceived object moves to the side of the more intense or leading clicks. As the interaural click delay is increased beyond about 1 ms, the object stops moving, and eventually splits in two such that two separate clicks are perceived. Studies have compared these perceptual properties with measurements of the BDP in human listeners. They showed that (1) the BDP was only detectable when the clicks at the two ears were fused into a single object and (2) the latency of the BDP covaried with the perceived location of the object. These correlations between percept and BDP indicate that the neuronal activity underlying the BDP contains a representation of both the fusion and the location of binaurally presented sounds. This implies that binaural fusion and perceived sound location may be at least partially encoded in the brain stem, where the BDP is generated, and, interestingly, well before cortical processing stages, where the activity dictating our perceptions is usually assumed to reside.

The second example concerns the N1 of the LLR and its correlation with the perceived rate of successive sounds and the segregation of sounds into distinct perceptual streams. A nice illustration of stream segregation, and the changes in perceived rate that sometimes accompany it, is provided by a sequence of tone bursts alternating in frequency (ABAB...), where A and B denote different frequencies. When the frequency difference between tones is small, a single, high-rate sequence is heard. However, when the frequency difference is large, the tones perceptually segregate into two lower rate sequences, comprising the high- and the low-frequency tones, respectively. For this and similar tone sequences, N1 evoked by the tones changes with increasing frequency separation in a manner predictable from the perceived rate and typical AEP rate dependencies described earlier (see the
section titled ‘Dependencies on stimulus parameters’). In particular, N1 is greater for large frequency separations (perceived rate is low), and less for smaller ones (perceived rate is high). Especially interesting are the changes in N1 that occur when there is a change in percept without a corresponding change in physical stimulus. Such changes can be seen using tone sequences with an intermediate frequency separation, chosen such that the percept spontaneously fluctuates between ‘one high-rate stream’ and ‘two low-rate streams.’ In experiments selectively averaging N1 during each of these two perceptual conditions, N1 (or more specifically, its AEMF analog) was found to be greater during the perception of the low-rate streams. In other words, N1 was again greater during the perception of low-, as compared to high-rate sounds, but this time the difference in N1 occurred in the absence of physical alterations to the stimulus. The implication is that some of the neural activity underlying N1 specifically encodes the percept elicited by sound, rather than physical attributes of sound.

These examples are two of many in which auditory evoked responses have been used to identify, and study, the neural activity leading to human sound perception.

See also: Auditory Cortex Structure and Circuitry; Auditory System: Central Pathways; Evoked Potentials: Recording Methods; Sound Localization: Neural Mechanisms.

Further Reading


