Resolving Precise Temporal Processing Properties of the Auditory System Using Continuous Stimuli

Edmund C. Lalor,1,2 Alan J. Power,2 Richard B. Reilly,1,2,3 and John J. Foxe4,5

1Trinity College Institute of Neuroscience, 2School of Engineering, and 3School of Medicine, Trinity College Dublin, Dublin, Ireland; 4Program in Cognitive Neuroscience, Departments of Psychology and Biology, City College of the City University of New York, New York; and 5The Cognitive Neurophysiology Laboratory, Nathan S. Kline Institute for Psychiatric Research, Program in Cognitive Neuroscience and Schizophrenia, Orangeburg, New York

Submitted 8 August 2008; accepted in final form 5 May 2009

Lalor EC, Power AJ, Reilly RB, Foxe JJ. Resolving precise temporal processing properties of the auditory system using continuous stimuli. J Neurophysiol 102: 349–359, 2009. First published May 13, 2009; doi:10.1152/jn.90896.2008. In natural environments complex and continuous auditory stimulation is virtually ubiquitous. The human auditory system has evolved to efficiently process an infinitude of everyday sounds, which range from short, simple bursts of noise to signals with a much higher order of information such as speech. Investigation of temporal processing in this system using the event-related potential (ERP) technique has led to great advances in our knowledge. However, this method is restricted by the need to present simple, discrete, repeated stimuli to obtain a useful response. Alternatively the continuous auditory steady-state response is used, although this method reduces the evoked response to its fundamental frequency component at the expense of useful information on the timing of response transmission through the auditory system. In this report, we describe a method for eliciting a novel ERP, which circumvents these limitations, known as the AESPA (auditory-evoked spread spectrum analysis). This method uses rapid amplitude modulation of audio carrier signals to estimate the impulse response of the auditory system. We show AESPA responses with high signal-to-noise ratios obtained using two types of carrier wave: a 1-kHz tone and broadband noise. To characterize these responses, they are compared with auditory-evoked potentials elicited using standard techniques. A number of similarities and differences between the responses are noted and these are discussed in light of the differing stimulation and analysis methods used. Data are presented that demonstrate the generalizability of the AESPA method and a number of applications are proposed.

INTRODUCTION

The auditory-evoked potential (AEP) is an electrical response recorded from the nervous system following presentation of a sound stimulus (Davis 1939). The AEP obtained from the scalp using electroencephalography (EEG) can be subdivided into three sequences of waves reflecting activity of cell populations at various levels along the sensory processing hierarchy. The first sequence is known as the brain stem response, because of its purported origins, and consists of seven positive peaks (labeled I–VII) occurring within the initial 8–12 ms. This is followed by the middle latency sequence, which is thought to be caused by activity in thalamic nuclei and neurons in primary auditory cortex and consists of three negative (N1, N2, and N3) and two positive (P1 and P2) peaks that occur from around 8–12 to 40–50 ms. The final sequence is known as the long-latency or cortical response and reflects activity in higher-order auditory and association cortex. This sequence is made up of two positive (P3 and P4) and two negative (N4 and N5) peaks proceeding from 50 to 500 ms (Celesia and Peachey 1999; Picton et al. 1974). This detailed componentry, which can be measured with great precision allied with the excellent temporal resolution of the event-related potential (ERP) technique, renders the AEP an extremely valuable and widely used tool in both research and clinical settings. For example, as well as being used to assess hearing function (for review see Celesia and Peachey 1999), specific components of the AEP have been used in the study of many neurological disorders including depression (Gallinat et al. 2000), Alzheimer’s disease (Cancelli et al. 2006), schizophrenia (Nagamoto et al. 1991), multiple sclerosis (Chiappa et al. 1980), and anxiety disorder (Drake Jr et al. 1991). In addition a number of cognitive processes such as selective attention (Bidet-Caulet et al. 2007; Picton and Hillyard 1974; Snyder et al. 2006) and auditory scene analysis (Bidet-Caulet et al. 2007; De Sanctis et al. 2008; De Sanctis et al. 2008; De Sanctis et al. 2008; Snyder et al. 2006; Sonnadara et al. 2006) have been studied using AEPs. Due to the low magnitude of the AEP relative to the ongoing EEG it can be estimated only by averaging responses over a large number of discrete, repeated trials using stimuli of generally short duration (e.g., 10–100 ms). This method elicits responses with high signal-to-noise ratios (SNRs) and good reproducibility across trials and subjects. However, given the broad range of complex auditory stimuli to which we are exposed in the natural environment (e.g., voices), the use of simplistic, discrete stimulation may not be optimal for a thorough analysis of the auditory system. Furthermore, the need to present discrete stimuli generally precludes the resolution of responses to more than one stimulus at a time. This limitation can severely hamper the design of environmentally valid electrophysiological experiments looking at cognitive processing of multiple audio streams, such as those investigating auditory attention and auditory scene analysis. This issue can be circumvented somewhat by using the auditory steady-state response (ASSR), which is a periodic frequency-following response typically elicited by an auditory stimulus that is amplitude modulated at a rate of about 40 Hz (e.g., Linden et al. 1987). However a major disadvantage of the ASSR method is the fact that it produces but a single measure of power at the stimulus frequency, with an accompanying loss of temporal
resolution in the form of the distinct components described earlier, which represent the major strength of the AEP technique.

The ability to obtain electrophysiological responses to complex, continuously presented stimuli with full temporal resolution would thus be very useful. Such a method was recently described in the visual domain (Lalor et al. 2006), wherein the impulse response of the visual system is obtained using a stimulus whose luminance or contrast is smoothly modulated by a stochastic signal with its power spread over a range of frequencies. This impulse response is known as the VESPA (visual-evoked spread spectrum analysis). In the present study, an auditory equivalent of this spread spectrum approach, called the AESPA (auditory-evoked spread spectrum analysis), is presented. This is accomplished by smoothly, but stochastically modulating the amplitude of an auditory carrier stimulus and estimating the linear impulse response from the recorded EEG. We compare the profile of the AESPA with that of the AEP elicited using standard methods and discuss their similarities and the possible cellular underpinnings of their differences. We further demonstrate the flexibility of the AESPA method using a stimulus that consists of continuous bursts of sounds separated by intervals of silence. The results suggest the utility of the AESPA in a number of research and clinical experiments where the use of the standard AEP is not possible.

METHODS

Subjects

Participating in the study were 12 subjects (one female; aged 22–35 yr), all of whom had normal hearing. The experiment was undertaken in accordance with the Declaration of Helsinki. The Ethics Committee of St. Vincent’s University Hospital in Dublin approved the experimental procedures and each subject provided written informed consent.

Stimuli

Two types of carrier signal were used in this study:

1) A Gaussian broadband noise (BBN) waveform, with energy limited to a bandwidth of 0–22.05 kHz.

2) A 1-kHz pure tone (TONE).

The amplitude of these carrier stimuli was modulated using Gaussian noise signals with uniform power in the range 0–30 Hz, i.e., at a rate of 60 modulations/s. This rate was chosen based on the fact that EEG power >30 Hz is typically very low. Modulating signals with the desired statistical properties were precomputed and stored. Taking into account the logarithmic nature of auditory stimulus intensity perception, the values of these modulating signals ($x$) were then mapped to the amplitude of the audio stimulus $x'$, using the following exponential relationship

$$x' = 10^{x/10}$$

and normalized to between 0 and 1. It was expected that this would result in a more linear perception of audio intensity modulation. The modulating noise signal was then interpolated to give a smooth transition from one modulation amplitude to the next and stored. An example of the amplitude-modulated BBN and TONE signals can be seen in Fig. 1.

The audio stimuli were generated with a SoundBlaster Extigy soundcard and presented to subjects using high-fidelity Sennheiser HD650 headphones.

Experimental procedures

All subjects underwent ten presentations each of 120-s duration for both the amplitude-modulated BBN and amplitude-modulated TONE stimuli. Standard AEPs were also obtained from each subject during

![Fig. 1. A 0.5-s segment of (A) an amplitude-modulated broadband noise (BBN) stimulus and (B) an amplitude-modulated 1,000-Hz TONE stimulus. C: a 50-ms 1,000-Hz tone pip with 10-ms rise and fall times. D: a composite of the BBN stimulus and the standard auditory-evoked potential (AEP) stimulus (i.e., a modulated BBN stimulus incorporating forced onsets and offsets).](http://jn.physiology.org/)}
ten 120-s runs using repeated presentations of a discrete stimulus. These stimuli were presented at an average rate of 1 Hz with the interstimulus interval randomized between 750 and 1,250 ms and consisted of a 1-kHz tone, 50 ms in duration, which included 10-ms ramp-up and ramp-down.

To demonstrate the flexibility of the AESPA method, a third stimulus was also tested, which represented something of a hybrid between the fully discrete AEP stimulus and the fully continuous stimuli otherwise used for the AESPA. It consisted of an amplitude-modulated BBN carrier within each second of which a silent interval was randomly inserted. Each interval was also randomly chosen to be between 50 and 200 ms in duration. As with the AEP stimulus, these onsets and offsets consisted of 10-ms raised cosine ramps. Thus this stimulus contained bursts of continuous stimulation with multiple onsets and offsets and is referred to as a composite stimulus. Two subjects underwent 30 presentations each of 120-s duration for this stimulus. To directly compare the continuous AESPA data with those obtained using the composite stimulus, the same two subjects undertook a further 20 presentations each of 120-s duration using the original continuous BBN stimulus, bringing their total number of continuous BBN runs to 30.

Subjects were instructed to keep their eyes open and to keep eye movements and blinks to a minimum for the duration of each run. Subjects were also instructed to limit all other types of motor activity during each run.

EEG acquisition

EEG data were recorded from 130 electrode positions, filtered over the range 0–134 Hz, and digitized at the rate of 512 Hz using a BioSemi Active Two system. Synchronization between the audio stimuli and the recorded EEG data was ensured by including the signal on the parallel port of the presentation computer, indicating the onset and offset of the stimuli, among the recorded signals. EEG data were digitally filtered with a high-pass filter, where the passband was >2 Hz and with a −60-dB response at 1 Hz and a low-pass filter with passband <35 Hz and a −50-dB response at 45 Hz. The data at each channel were rereferenced to the average of the two mastoids.

Signal processing

The AESPA method follows on from a broad variety of system identification methods that are used in auditory system analysis. Although the majority of these methods have been applied to neural spiking data (Eggermont 1993; Klein et al. 2000) some applications to EEG using M-sequenced presentation of binary stimuli have been reported (Eysholdt and Schreiner 1982; Shi and Hecox 1991). The AESPA can be considered as being most closely related to the Volterra–Wiener approach to system modeling (Marmarelis and Marmarelis 1978).

Specifically, the estimation of the AESPA is based on the assumption that the output EEG, \( y(t) \), consists of a convolution of the audio amplitude-modulated (AM) signal \( x(t) \), with an unknown impulse response \( w(\tau) \), plus noise (see Fig. 2), i.e.

\[
y(t) = w(\tau) \ast x(t) + \text{noise}
\]

Given the known audio AM signal and the measured EEG, we obtain \( w(\tau) \) (i.e., the AESPA), by performing linear least-squares estimation. This can be carried out analytically using the following equation

\[
w = (x, x^T)^{-1}(x, y)
\]

where \( x \) is a column vector of input stimulus intensity values in a certain window of time around \( t \), \( y \) is the EEG value at time \( t \), and \((\cdot)^{-1}\) denotes mean over \( t \). This involves inversion of the input signal’s autocorrelation matrix \( (x, x^T) \) (see Appendix A in Lalor et al. 2006 for details). By adding a regularization matrix with terms along or

![FIG. 2. Flow diagram of auditory-evoked spread spectrum analysis (AESPA) acquisition. A colored noise signal is used to modulate the amplitude of a carrier signal. The electroencephalogram (EEG) is modeled as a sum of overlapping scaled impulses in response to the amplitude modulations of the stimulus, plus noise. Three such scaled impulse responses are shown, corresponding to stimulus values A, B, and C.](http://jn.physiology.org/)

*J Neurophysiol* • VOL 102 • JULY 2009 • www.jn.org
adjacent to the diagonal of the input signal’s autocorrelation matrix before inversion one can increase the bias of the estimate, but reduce overall estimation error by greatly reducing the variance (see Appendix B in Lalor et al. 2006). As in that study, the regularization parameter $\lambda$, used in the current study, was empirically chosen to be $4.4 \times 10^{-3}$. This resulted in good reduction of off-sample error without affecting the height or the latency of the response peaks.

It should be noted that for the estimation step, the values of $x(t)$, were assumed to be constant across each 16.67-ms modulation period. Furthermore the initial modulation values, i.e., the linear values obtained prior to the exponential mapping, were used. This seemed reasonable under the assumption that the exponential mapping would actually result in a more linear intensity perception.

The AESPA was estimated using a sliding window from 200 ms prestimulus to 400 ms poststimulus that was advanced sample by sample. This window was chosen to present the AESPA using an interval similar to that typically used for plotting the AEP. However, the meaning of the interval is slightly different because the AESPA, unlike the AEP, does not correspond to a specific discrete event occurring at time 0. Instead, each time point on the time axis can be interpreted as being the relative time between the EEG and the input intensity signal. Therefore the AESPA at $-100$ ms, for example, indexes the relationship between the input intensity signal and the EEG $100$ ms earlier; obviously this should be zero. As another example, the AESPA at $+100$ ms indexes how the input intensity signal affects the EEG $100$ ms later. The steps involved in generating the stimuli and the estimation of the AESPA are illustrated in Fig. 2.

AESPAs were determined by averaging the EEG in response to a stimulus in epochs from 200 ms prestimulus to 400 ms poststimulus. SNRs for both methods were determined by considering data in the interval $0–300$ ms as signal and $-100$ to 0 ms as noise. Grand-average AEPs and AESPAs were baseline corrected by subtracting the mean of the prestimulus values ($-100$ to 0 ms) for each channel.

**Determination of AESPA components**

To ascertain the components of the AESPA response we applied the topographic pattern analysis methods described in Murray et al. (2008). This analysis method splits responses into distinct segments called microstates (Lehmann 1987), which represent periods of topographic stability, and was implemented in the present study using CarTool (http://brainmapping.unige.ch/cartool.htm). We used a $K$-means clustering algorithm that was constrained such that clusters that were $\geq 75\%$ correlated were merged and response segments less than or equal to three samples (5.9 ms) in duration were rejected. This is essentially a data-compression algorithm in which a number of template topographic maps that best account for the data are obtained. The distinct microstates are then identified by comparing the template maps to the responses by way of spatial correlation. This was carried out on a concatenated data set consisting of all three conditions (i.e., AEP, BBN AESPA, and TONE AESPA) as recommended in Murray et al. (2008). The choice of $75\%$ correlation for merging was based on the analysis of our AEP response into the standard $P_1$, $N_1$, and $P_2$ components (Picton et al. 1974). For consistency, we then used this same threshold for the analysis of AESPA responses. The choice of four samples as the minimum number needed to constitute a response was somewhat ad hoc and was chosen to avoid division of the AESPA into an unjustifiably large number of candidate components.

**RESULTS**

The group-average standard AEP and the group-average AESPAs obtained using the modulated BBN and TONE stimuli, along with the standard error (SE) across subjects, are shown in Fig. 3 for eight electrode locations distributed around the scalp, all referenced to the average of the mastoids. Clear responses can be seen for all three methods. The standard AEP exhibits clear $P_1$, $N_1$, and $P_2$ components at around 70, 120, and 200 ms, respectively, and is larger in amplitude at frontal sites compared with those in occipital areas. For the BBN AESPA at frontal locations a succession of small peaks and troughs can be seen in the range 0–100 ms along with prominent negative and positive components at 145 and 200 ms, respectively. In more posterior locations, the earlier components remain visible, but the large, later peaks are absent. Similarly for the TONE AESPA, there is some notable activity in the 0–100 ms range at all electrodes shown with large negative and positive later components being much more evident at frontal sites.

**AESPA components**

Figure 4 presents the group-average data in three ways. The bottom panels show the global field power (GFP; Lehmann and Skrandies 1980; Skrandies 1995) for each response, with the horizontal dashed line on this panel indicating twice the mean GFP in the interval $-100$ to 0 ms. The middle panel represents a butterfly plot showing all channels simultaneously. The microstates identified by the topographic analysis are indicated by vertical lines and their corresponding topographies are shown in the top panel. Not wanting to equate the componentry of the AESPA with that of the AEP, we tentatively designated seven BBN AESPA candidate components, based on the microstates identified by our analysis, as follows: $P_a$ (21–45 ms), $N_a$ (45–61 ms), $P_b$ (61–92 ms), $N_b$ (92–104 ms), $P_c$ (104–125 ms), $N_c$ (125–170 ms), and $P_d$ (170–238 ms). Although twice the mean GFP in the interval $-100$ to 0 ms was used as a component threshold, the response in the fifth interval ($P_c$) was also considered to be a candidate component, despite not exceeding this threshold, given that it was identified as a distinct microstate in the previous analysis. The same seven components are evident to a degree in the TONE AESPA, although the $N_p$, $P_c$, and $N_c$ components appear to be less easily distinguished and are, in fact, identified as a single microstate by the topographic analysis. The $P_c$ component is also present in the TONE response, but seems a little less well defined. Both of these results may be considered in light of the much lower SNR of the TONE AESPA response, which is discussed in the following text.

The butterfly and GFP plots in Fig. 4 and the topographic analysis strongly suggest the presence of these seven components, although we sought to submit this hypothesis to a further statistical test. Figure 5 consists of a pair of statistical cluster plots marking time points at which the BBN and TONE AESPAs differ significantly from zero for each channel. These plots were calculated using running $t$-tests across subjects at an alpha level of 0.05. In the case of the BBN AESPA, three highly distinct clusters are seen representing $P_a$, $N_c$, and $P_d$, whereas there appears to be some activity in central areas around 380 ms. The TONE AESPA also exhibits significant clusters for $P_p$, $N_c$, and $P_d$, although the $N_c$ activity is much more widely distributed in time and the $P_d$ activity is less evident over posterior and occipital scalp. The late activity around 380 ms, however, is much more evident. Although the four proposed components between $P_p$ and $N_c$ did not show up as significant for either response, the fact that the response has clearly begun by 20 ms—combined with evidence from the
butterfly and GFP plots of Fig. 4 and the microstates analysis—suggests that perhaps the t-tests were simply underpowered. It is also worth noting that the small local maximum evident at around 10 ms in the BBN GFP plot, which is suggestive of an even earlier component, does not show up as statistically significant. More data might help to elucidate whether this is in fact a real component.

Comparison of AEP and AESPA

Clearly the AEP exhibits a very different morphology with only three components being evident in Fig. 4A. The prominence of the N1 and P2 components of the AEP and the Nc and Pd components of the AESPAs suggest that they may be related, although the timing is notably different. This is backed up somewhat by the topographical maps in the top panels of Fig. 4, which show that the AEP’s N1 exhibits as a strong frontal negativity. The Nc of the BBN AESPA in the sixth interval appears to be similarly distributed, as does the negativity of the TONE AESPA in its third interval. Furthermore, these three components were all represented by the same template map in the topographic analysis and show high spatial correlation (AEP/BBN: $r = 0.96$; AEP/TONE: $r = 0.98$; BBN/TONE: $r = 0.94$; all with $P < 0.01$). All methods also display a strong late positivity in frontal areas—i.e., in the third interval for the AEP, the fourth and fifth intervals for the TONE AESPA, and in the seventh interval for the BBN AESPA. Again, all were significantly spatially correlated (AEP/BBN: $r = 0.93$; AEP/TONE: $r = 0.55$; BBN/TONE: $r = 0.77$; all with $P < 0.01$). This suggests that these components may derive from similar generators.

Further comparison between responses can be seen in Fig. 6, which illustrates the values of the correlation coefficients at each electrode for each pair of methods. This analysis was aimed at providing information on how the latencies of the different responses differed at every electrode location. Somewhat surprisingly, the TONE AESPA and the AEP are seen to be the most strongly correlated pair, with $r > 0.6$ ($P < 0.01$) for all frontal and almost all central electrode locations. Both
AESPA responses are seen to be widely correlated with each other, although not as strongly as the TONE AESPA and AEP. The BBN AESPA and AEP exhibit significant, if not very strong, correlations in central and frontal locations only.

As is evident from the SE traces in Fig. 3, there is a discrepancy in the SNR between the AEP and the AESPA responses. SNRs of the GFP of the group-average responses

AESPA responses are seen to be widely correlated with each other, although not as strongly as the TONE AESPA and AEP. The BBN AESPA and AEP exhibit significant, if not very strong, correlations in central and frontal locations only.
for the AEP, BBN AESPA, and TONE AESPA were 25.18, 11.42, and 11.93 dB, respectively.

AESPA as a generalization of the AEP

Although the AESPA responses and the AEP exhibit clear differences, there are clearly some similarities. This is not surprising given that, as discussed in the study describing the visual analog to the AESPA (the VESPA; Lalor et al. 2006), the standard AEP can be considered a special case of the broader AESPA method. That is, although the AEP method is restricted to the use of discrete stimuli that are isolated in time, the AESPA stimuli can include such discrete events along with periodic bursts, temporal rate changes, and fully continuous stimulation.

To demonstrate this, we analyzed the continuous (AESPA), composite, and discrete (AEP) stimuli using both the AESPA analysis method and simple signal averaging for the two subjects who undertook the composite stimulus presentations. Figure 7 plots the corresponding responses at frontocentral electrode location Fz, referenced to the average mastoids. In the continuous stimulus case, signal averaging was carried out time-locked to positive intensity changes that exceeded 0.5, whereas in the composite case, we time-locked to all onsets. Although responses are evident when using signal averaging on the continuous and composite data, they exhibit much lower SNRs (3.92 dB for the continuous and 4.98 dB for the composite on the GFP) and the componentry is much less sharply defined than that in the case where the same data are analyzed using the AESPA method (SNRs: 10.95 dB for the continuous and 13.16 dB for the composite on the GFP). Furthermore, the responses produced by the AESPA method for both the continuous and composite responses are highly correlated ($r = 0.62, P < 10^{-16}$ at Fz), whereas those produced by the averaging method are significantly negatively correlated ($r = -0.33, P < 10^{-4}$ at Fz). This serves to illustrate that, for such nontrivial stimuli, the effect of the noise on the averaging procedure has been sufficient to render the AEP an inconsistent measure of neural activity, whereas the AESPA method continues to deliver a stable response. Finally, because they are mathematically equivalent, using a train of impulses as input to the AESPA analysis unsurprisingly results in a response to the discrete stimulus that is identical to that obtained using the signal-averaging method, reinforcing the notion that the AESPA method can be considered a generalization of the traditional AEP technique.

Source analysis of the AESPA

To further investigate the origins of the AESPA componentry, we performed source analysis using the BESA software suite (http://www.besa.de). Because of the greater SNR of the BBN AESPA, we began our source analysis with the group-average BBN response. We used the following strategy to arrive at a solution.

1) We first fit a window from 138 to 152 around the $N_c$ peak with a pair of symmetrically constrained dipoles. The fit is exceptional, explaining 98.9% of the variance in this window.

![Fig. 6](http://www.jn.org/)

**Fig. 6.** Correlation coefficients for the standard AEP and the BBN AESPA (A), the standard AEP and the TONE AESPA (B), and the BBN AESPA and TONE AESPA (C). The black, dark gray, and light gray circles indicate a significant correlation with $r > 0.6$, $0.3 < r < 0.6$ and $0 < r < 0.3$, respectively. The white circles indicate electrodes at which correlations were insignificant (i.e., $P > 0.01$). Circles with ‘‘−’’ in the center denote negative correlations with strengths indicated using the same shading scheme as above.

![Fig. 7](http://www.jn.org/)

**Fig. 7.** Responses to the 3 different stimuli used analyzed by both the AESPA method and the time-locked averaging method used in standard AEP extraction. These responses are the grand average of 2 subjects who each underwent 30 trials of each of the continuous and composite stimulus type and 10 trials of the discrete stimulus. Note the plots for the discrete stimulus are on a different scale from that of the other plots. There seems to be some activity evident for the continuous and composite stimuli analyzed using the AEP technique, but the AESPA method results in a much higher signal-to-noise ratio, highlighting its ability to handle a more versatile array of stimuli than the standard AEP method.
Looking at the model residual showed no remaining signal around the Nc period. These two dipoles, which are shown in red in Fig. 8, were localized very close to Heschl’s gyrus (X, Y, Z = −39, −15, 16). In fact, fixing these dipoles in the centroid of the auditory core (about −46, −24, 12), about 1 cm from where they landed when allowed to freely fit, and allowed their orientation to vary, the model still explained 98.6% of the variance in this window. As such, it is with good confidence that we say that the Nc is generated in and around the auditory core, consistent with the localization of the AEP’s N1 component in many previous studies (see Leavitt et al. 2007).

2) The residual waveforms indicated that a number of obvious components preceding the Nc timeframe remain unexplained by primary auditory generators. For instance, the Pa component at 30 ms is still fully evident and mapping the residual shows a bilateral posterior map. Fitting a pair of symmetrically constrained dipoles to this period (23–41 ms) resulted in localization to the inferior occipitotemporal surface. This seemed unlikely to be correct. Looking at the butterfly plot in Fig. 4, we noted a rather unipolar distribution for this peak, which fits the pattern expected of a deep source. This was confirmed by a reasonably good fit (explained variance 89.6%) using the addition of a single dipole in the brain stem (8, −39, −26). Although identification of brain stem sources using EEG is unlikely to be completely reliable, the fact that this component is significantly positive at almost every scalp electrode when referenced to the average of the mastoids and the short latency of the component strongly suggest a subcortical source. This source may actually be the thalamus, given previous work linking AEP components at this latency with thalamic sources (Picton et al. 1974).

3) The distribution of the residual over the Pa component indicated that bilateral cortical sources were likely involved. Fitting bilateral symmetric dipoles from 110 to 130 ms resulted in a pair of locations in the temporal lobe in the region of Brodmann Areas 21 and 22, between the superior and middle temporal gyri (−48, 1, −31), with a different orientation to the primary Nc generators. Explained variance in this timeframe was 94.6% and there was no obvious signal left in this timeframe in the residual waveforms.

4) Working backward in time, another relatively distinct component was evident with a peak at about 70 ms. The topography of the residual during this timeframe was multifocal and complex and was fit well by a pair of dipoles in the middle frontal gyrus (−37, −2, 40).

Between 21 and 238 ms, this simple model, shown in Fig. 9, accounted for 97.3% of the variance. Allowing this seven dipole fit to go free (with the exception of the primary auditory generators) across the 21- to 238-ms range showed that the solution was very stable, with no improvement in explained variance and very little movement in dipole location. Using this solution to fit the average BBN AESPAs on a subject-by-subject basis gave us an average explained variance of 84.1% with SD of 8.4%.

Given the difficulty in arriving at a stable solution for the TONE response as a novel tool for carrying out research on the nonlinear processing effects (Lalor et al. 2008). The notion that activity of a limited subset of cells is reflected in the AEP, with 98.7% of the variance of the group average AEP over the interval 21–238 ms.

FIG. 8. Source dipoles for the group average BBN AESPA response obtained using BESA software. Between 22 and 237 ms, this simple model accounted for 97.3% of the variance. On a subject-by-subject basis the average explained variance was 84.1% with SD of 8.4%. The TONE response was also very well accounted for by the same solution, with 94.2% of the variance explained. The average explained variance across subjects was 74.4% and SD of 7.3%. Interestingly, the same solution accounted for 98.7% of the variance of the group average AEP over the interval 21–238 ms.

DISCUSSION

We have described a method that allows for the estimation of a novel auditory potential, known as the AESPA, using continuous, amplitude-modulated stimuli. These responses display temporally detailed componentry with high SNR and can be elicited using stimuli with differing characteristics. This method may facilitate research that cannot be conducted using the traditional ERP technique of discrete stimulation and averaging.

Although the goal of this study is to present the AESPA response as a novel tool for carrying out research on the auditory system and not to equate it with the traditional AEP, comparison between the two methods can help in characterizing this new response. In fact, although the AESPA responses show significant correlation with the standard AEP, they are undoubtedly distinct, particularly in posterior regions. One obvious reason for this is that the AESPA signal processing assumes a linear relationship between the input stimulus intensity and the output EEG, whereas no such assumption is made in calculating the AEP. As such, the AESPA likely reflects activity of a limited subset of auditory cells for which this assumption holds true. That said, it should be noted that the analysis outlined in this study can easily be extended to include higher-order processing effects (Lalor et al. 2008). The notion that activity of a limited subset of cells is reflected in the
AESP also speaks to the issue of the differing morphologies of the TONE and BBN AESPA responses. Specifically, given the tonotopic nature of auditory cortex (Howard et al. 1996; Morel et al. 1993; Pantev et al. 1988), it is likely that the response to the amplitude-modulated TONE stimulus reflects activity of an even more specific subset of cells than that of the BBN stimulus. Tonotopy of auditory cortex may also explain why the TONE AESPA is so highly correlated with the standard AEP (Fig. 4), given that both were elicited by 1-kHz tones in the current study.

The dissimilar nature of the stimulus methods used to elicit the AEP and AESPA is another likely reason for the differing responses. It has been reported that only a limited number of auditory cortical neurons display stimulus-synchronized responses (Lu et al. 2001). As a result cortical representations of rapidly occurring stimuli are likely to be quite distinct from those of isolated, discrete stimuli. Specifically, it is thought that two processes in auditory cortex account for the representation of sound modulation. Slower modulation frequencies (≤30 Hz) are thought to be represented by neuronal discharges that are temporally synchronized to the stimulus, whereas higher modulation frequencies are represented by nonsynchronized rate-based discharges (Lu et al. 2001). Further evidence of this dichotomy was shown in a study using amplitude-modulated sound in squirrel monkeys (Bieser and Müller-Preuss 1996). There it was reported that 78.1% of all acoustically driven neurons encoded the envelope of the AM sound with most neurons using a combination of two distinct modes. One mode involved spikes that followed the AM envelopes in a phase-locked manner, whereas the other involved significant changes in spike rate with the changing stimulation. Based on this mixed encoding of AM modulation in nonhuman primates, it is unclear what percentage of human auditory cortical cells would respond to an AM signal in such a way as to be reflected in the AESPA. Bieser and Müller-Preuss (1996) also note that the 21.9% of neurons not encoding AM displayed simple on, on–off, or off responses at the beginning or the end of a stimulus sound. The activity of cells such as these would be reflected in a standard AEP and not in an AESPA obtained to an AM signal in such a way as to be reflected in the AESPA. Bieser and Müller-Preuss (1996) also note that the 21.9% of neurons not encoding AM displayed simple on, on–off, or off responses at the beginning or the end of a stimulus sound. The activity of cells such as these would be reflected in a standard AEP and not in an AESPA obtained to a continuous stimulus. The fact that the SNR values for the AESPA responses obtained to the continuous and composite stimuli in this study are so similar and that the two responses are so highly correlated, however, suggests that further work is required to investigate this line of reasoning.

The AEP and the AESPA responses are also notably different in terms of their amplitude. There are a number of possible reasons for this. First the aforementioned likelihood that the AESPA reflects activity from a much more specific subset of cells is probably a factor. Second, as can be seen in Fig. 1, the intensity of the AESPA stimuli in this study rarely reached that used in generating the AEP which, given the dependence of response amplitude on stimulus intensity (Beagley and Knight 1967), likely plays a role. Third, and perhaps most important, acquiring the AEP by signal averaging involves the assumption that a stimulus of nonzero duration (in our case 50 ms) is essentially an ideal unit impulse. Thus the only stimulus attribute that is taken into consideration when calculating an AEP is the stimulus onset time, whereas other features such as duration of ramp-up/down and offset time are typically ignored. The AESPA needs to make no such assumption because the flexibility of the method allows the experimenter to take into account the fact that the auditory stimulus has its energy spread in time when calculating the response. Thus because the AESPA represents a function that produces an estimate of the EEG output based on a broader window of input stimulus values, it is not surprising that it differs in amplitude from that of the AEP, which represents a response to an assumed ideal impulse. Figure 9 plots the results of the AESPA analysis on the discrete data in the case where we assume the stimuli to be ideal impulses and the case where we take into account the full 50-ms profile of the stimulus. As expected, taking account of the fact that the stimulus power is spread across 50 ms results in a smaller response function with a shorter latency. Interestingly, some subcomponents of the P2 wave also become more evident when using the full stimulus profile, suggesting that the assumption of a 50-ms stimulus as an impulse might result in some blurring of neighboring componentry.

The effect of taking into account the detailed nature of the stimulus profile is also evident in Fig. 7, where the AESPA method clearly outperforms the signal-averaging technique for both the continuous and composite stimuli. This demonstrates the flexibility of the AESPA method in terms of producing high-quality responses to stimuli with diverse acoustic properties. Given the purported variation in the response of auditory cells to stimulus modulation with different temporal statistics and the prevalence of complex stimulus methods in the auditory processing literature as well as in the natural world, this flexibility suggests the AESPA as a useful addition to the palette of research tools currently used. Future work will aim to examine the characteristics of AESPA responses at more points along the “continuous–discrete” stimulus spectrum using a variety of carrier signals.

The topographic analysis carried out in this study provides another means for comparing the AESPA responses and the AEP. Again, it is not the aim of the current study to equate the

![](https://example.com/figure9.png)

**FIG. 9.** AESPA responses due to the discrete stimulus analyzed with an impulse and a 50-ms window (dashed plot). The flexibility of the AESPA method allows the experimenter to take into account the fact that the auditory stimulus is not an ideal impulse but has its energy spread in time when calculating the response. As expected, stimulus power spread across 50 ms results in a smaller response function with a shorter latency.
AESPA with the AEP, but to use the standard AEP to assist in characterizing the AESPA response. The high spatial correlations of the large late negativity and the large late positivity between each of the three methods and the fact that the same template map was assigned to the negativity for each of the three methods lends significant weight to the notion that they may be generated in similar cortical areas. The source solution generated for the BBN AESPA in this study provides additional support for some commonality in the cortical generators. The fact that the same simple solution accounted for 97.3, 94.2, and 98.7% of the variance for the BBN AESPA, TONE AESPA, and AEP, respectively, strongly suggests that similar neural pathways are involved in generating the responses—a result that does not come as a surprise.

Unlike their visual analogs, the VESPA and visually evoked potential (VEP; Lalor et al. 2006), the AESPAs shown in this study display lower SNR than that of the traditional AEP given a comparable amount of testing time. This is particularly the case for the TONE AESPA. The likely limited number of cells for which the AESPA’s assumption of linearity hold true and perhaps a less than optimal choice of carrier signals are possible reasons for this. It should also be noted that no artifact rejection was performed on the AESPA data in this study, which would assuredly influence the SNR. The reason for this is that the continuous nature of the stimulus makes it difficult to know which data to reject when an eyeblink or movement artifact is identified. Because such artifacts are independent of the stimulus signal against which we are regressing the EEG, they do not have any systematic effect on the AESPA response and would influence the response only through a slight reduction in SNR. This can be somewhat addressed using independent component analysis as was previously carried out on VESA data (Lalor et al. 2007) or by using shorter trials and averaging responses across them.

A number of research applications can be proposed for which the AESPA may be better suited than the traditional AEP. As mentioned earlier, the temporally constrained and simplistic nature of the stimuli required by the AEP method surely makes the method suboptimal for a system that has evolved to analyze an infinitely broad range of sounds. Much of today’s research on auditory processing focuses on using complex, species-specific stimulation (see Suga 1995) for which the traditional ERP method is not entirely suited. The flexibility afforded by the complex AESPA stimuli may make it a very useful tool for analysis of the temporal processing of speech or speechlike sounds in humans.

Another potential advantage of the AESPA method is in simultaneously obtaining responses to multiple stimuli. Lalor et al. (2006) demonstrated the ability to obtain VESA responses to two concurrently presented stimuli in different parts of the visual field. If we assume an analogy between tonotopy in the auditory domain and retinotopy in the visual domain, it is highly likely that AM of multiple carrier waves that do not overlap in frequency using independent modulating signals would enable the estimation of separate AESPA responses to each. This would render the method useful in various applications such as research on auditory scene analysis and on the effects of attention on the temporal processing of audition. The VESA method has already been used to demonstrate the timing of endogenous visual spatial attention, something that is very difficult to do using suddenly onsetting VEP stimuli (Lalor et al. 2007). Similarly, questions that are difficult to address using isolated suddenly onsetting AEP stimuli may be investigated using the AESPA.

In summary, evoked responses with detailed temporal profiles, known as AESPAs, have been obtained using complex, continuous auditory stimuli. To begin characterizing these responses, comparisons have been made with the standard AEP and similarities and differences between the two responses have been discussed. The demonstrated flexibility of the AESPA method gives it a number of advantages over the traditional AEP technique and suggests its applicability to fundamental research on auditory processing as well as studies of attention, auditory scene analysis, and perhaps in a number of neurological disorders where auditory processing is affected.

REFERENCES


