Proceedings of the 4th International
IEEE EMBS Conference on Neural Engineering
Antalya, Turkey, April 29 - May 2, 2009

SaA1.5

Extracting Separate Responses to Simultaneously Presented Continuous Auditory Stimuli: An Auditory Attention Study.

Alan J. Power, Edmund C. Lalor, Richard B. Reilly
School of Engineering,
Trinity College Dublin,
Dublin 2, Ireland.
powera3@tcd.ie

Richard B. Reilly
School of Medicine,
Trinity College Dublin,
Dublin 2, Ireland.

Abstract— **Auditory Evoked Potentials (AEPs) have been used extensively in the study of auditory attention. Some weaknesses of standard AEP paradigms include (i) stimulation is discrete in nature and thus not very ecological and (ii) the inability to extract separate responses to more than one simultaneous stimulus. Thus the ecological investigation of auditory attention is limited. Alternative to discrete stimulation the auditory steady-state response (ASSR) is employed but 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. A novel method, known as AESPA (Auditory Evoked Spread Spectrum Analysis), which overcomes these limitations has been previously established. Here we present data that show that separate AESPA responses to two simultaneously presented continuous stimuli can be elicited. We also show how this can be used to study auditory attention by investigating responses to attended and unattended stimuli.**

I. INTRODUCTION

Auditory Evoked Potentials (AEPs) have been used extensively in the study of auditory attention [1-8]. AEPs can be elicited by the repeated presentation of a click/pip stimulus [9] and extracted from the EEG using an epoch averaging procedure. By considering the brain in simplified form as a linear system, with isolated events as input and EEG as output, the average event-related potentials (ERPs) can be said to approximate the system's time-domain impulse response functions. This is the linear part of the event-related dynamics. However, in reality, events are not isolated. Rather, inputs occur in a rapid and continuous stream and their associated electrophysiological responses often overlap in time. Consequently, an ERP calculated as the average of many isolated inputs may obscure the brain's response dynamics.

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 utilizing 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 [8].However a major disadvantage of the ASSR method is the fact that it produces a single measure of power at the stimulus frequency, with the loss of temporal resolution in the form of the distinct components of the AEP.

The impulse response of the auditory system to a continuous stimulus, whose intensity is modulated by a signal with its power spread uniformly over the range 0-30Hz has recently been reported [10, 11]. The response generated using this spread spectrum approach is known as the AESPA (Auditory Evoked Spread Spectrum Analysis). Thus the continuous nature of the AESPA stimulus overcomes some of the ecological limitations of discrete stimuli found in the standard AEP method. Another advantage of the AESPA method is that the modulating waveform can be chosen to have any desired statistics required offering superior flexibility of stimuli compared to standard AEP methods.

In this study we use the AESPA method to extract temporally detailed responses to two simultaneously presented continuous stimuli. Many previous studies [1-7] present stimuli dichotically which limits the interpretation of the results in more natural settings where sounds do not reach each ear entirely separately. Thus with the intention of making the paradigm as ecological as possible, the stimuli here are presented binaurally as opposed to dichotically.

II. METHODS

A. Subjects:

Eight subjects aged 22-31 participated in the study. All had normal hearing.

B. Hardware

EEG data were recorded from 128 electrode positions, filtered over the range 0 – 134 Hz and digitized at the rate of
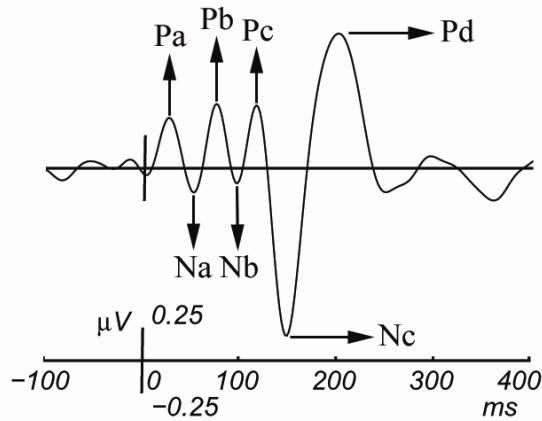
Figure 1. A AESPA response to a single stimulus obtained, at Fz , using a modulated broadband noise carrier with equal power in the range 0 - 22.05 kHz.

512 Hz using a BioSemi Active Two system (http://www.biosemi.com). Synchronization between the audio stimuli and the recorded EEG data was guaranteed by recording both stimuli and response simultaneously.

EEG data were digitally filtered with a high-pass filter, where the passband was above 2 Hz and with a -60 dB response at 1 Hz and a low-pass filter with a 0-35Hz passband and a -50 dB response at 45 Hz. The audio stimuli were generated with a SoundBlaster Extigy soundcard and presented to subjects using high fidelity Sennheiser HD650 headphones.

C. Generation of Audio Stimuli

As mentioned above a standard AESPA stimulus consists of a carrier stimulus modulated by a spread spectrum signal. In this case two root mean square (RMS) normalised band-pass noise carriers of bandwidth 1kHz, centred at 1kHz and 5kHz respectively, were employed. These centre frequencies were chosen on the basis that 1kHz and 5kHz tones are perceived with approximately the same loudness [12] and also they are far enough apart in frequency such that they are perceived separately. These will be referred to as the LOW and HIGH streams respectively. These carriers were then modulated by independent spread spectrum signals. The resulting stimuli were then added together to give a combined stimulus consisting of high and low frequency streams.

D. Spread Spectrum Signals

The spread spectrum modulating signals consisted of Gaussian white noise with energy limited to 0-30 Hz. 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:

$$x' = 10^{2x},$$

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. Using the Nyquist sampling theorem and given that EEG power above 30 Hz is very low, the modulation rate of each signal was set to be 60 Hz.

E. Experimental Procedure

Subjects were instructed to affix their gaze to a marker ('+') on a monitor in front of them in order to keep eye movements to a minimum. Subjects were also instructed to keep the number of eye-blinks and all other types of motor activity to a minimum during each session.

Each subject undertook ten 120 second trials where they were randomly asked to attend to the HIGH stream (attend-HIGH condition) or LOW stream (attend-LOW condition). The stimuli were presented binaurally. Stimuli were presented at an intensity level deemed comfortable by the subject before beginning the experiment.

In order to monitor each subject's progress targets and distracters were inserted randomly in each stream. These events consisted of a specific pattern of amplitude modulation imposed on the random process. Targets consisted of a modulation level of 0.75 for 25.5ms followed by 0.5 for 16ms followed 0.75 for 25.5ms, giving a total length of 67ms, whereas distracters consisted of a flat modulation of 0.5 for 67ms.

Subjects were instructed to respond by pressing a button only when a target in the attended stream was heard. EEG was recorded for later analysis where both the responses to the HIGH and LOW streams for each condition were extracted. No event could occur within one second of another, both within and between streams.

F. Signal Processing

We obtain the AESPA by performing a linear least squares fit of the response model

$$y(t) = w(\tau) * x(t) + noise$$

Where $y(t)$ is the measured EEG response $x(t)$ is the intensity waveform of the stimulus, the symbol * indicates convolution, $w(\tau)$ is the impulse-response function to the amplitude of the stimulus, and the noise is assumed to be Gaussian [10, 11].

G. Quantification of data

The responses to the HIGH and LOW streams in the attend-HIGH and attend-LOW conditions we plotted for comparison.

SNR values were calculated by defining noise as the mean squared values in the 100ms interval immediately preceding the stimulus, and the signal was defined as the mean of the squared values in the interval 0-300ms post-stimulus.

d' values, which give a measure of observer sensitivity uncontaminated by response bias, were used to assess if a subject was adequately attending to the required stream [13]. A d'>1 was considered adequate attention.

503

2. (a) **HIGH responses**

— attend HIGH
--- attend LOW

Fpz

Cz

Pa Pb   Pd

Na
Nb   Pc   Nc

μV 0.25

−100   0   100   200   300   400
        −0.25   ms

2. (b) **LOW responses**

— attend LOW
--- attend HIGH
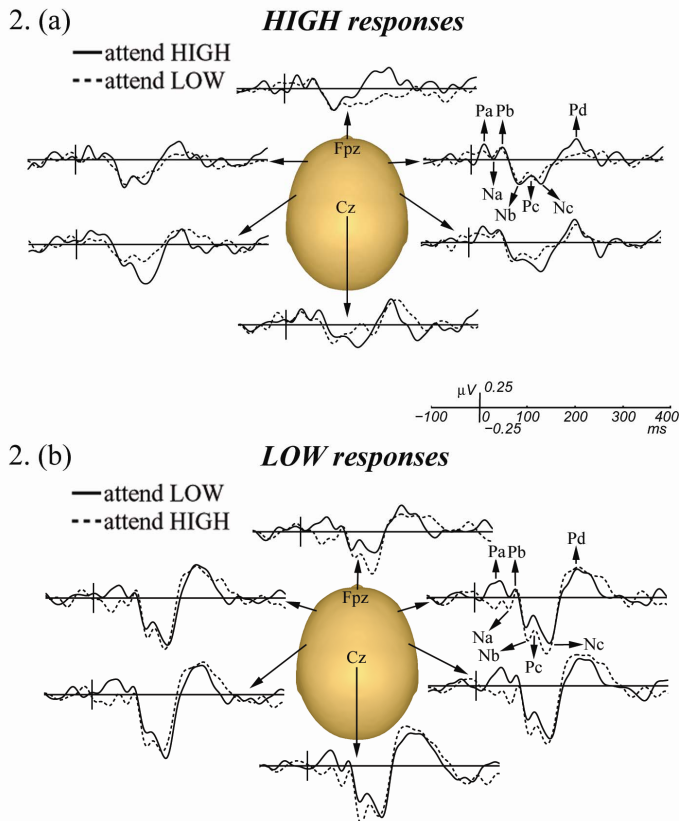

Fpz

Cz

Pa Pb   Pd

Na
Nb   Pc   Nc

Figure 2. The AESPA responses to (a) the HIGH stream for the attend-HIGH (solid lines) and the attend-LOW (dashed lines) conditions and (b) the LOW stream for the attend-LOW (solid lines) and the attend-HIGH (dashed lines) conditions at five representative electrode locations.

III. RESULTS

Sensitivity analysis using d' resulted in the inclusion of 60 of the 80 trials (i.e. 75%) for the attend-HIGH condition and 57 of the 80 trials (i.e. 71.25%) for the attend-LOW condition.

Comparing the responses in fig. 2 to the AESPA response to a single broadband carrier stimulus (fig. 1) we see that the signal power in the attentional responses is not as high as for the single stimulus case. From fig. 2 we can see that the Pc component never actually goes above the baseline for any of the responses and that the HIGH stream generally elicits a somewhat smaller response than the LOW stream.

As expected for the responses to the HIGH stream the SNR for the attend-HIGH condition was significantly greater than that of the attend-LOW condition ($t = 14.9$; $p<0.001$). Also for the responses to the LOW stream the SNR for the attend-LOW condition were significantly greater than for the attend-HIGH condition ($t = 6.72$; $p < 0.001$).

A. HIGH responses

Fig 2(a) shows the responses to the HIGH stream for the attend HIGH (solid lines) and the attend LOW (dashed lines) conditions at five representative electrode locations. From the plots we can see an early effect frontally and temporally on the

Pa component with the attend-LOW condition showing little or no activity.

A further effect is seen on the Nc component. Here we see three distinct effects depending on electrode location:

(i) the attend-HIGH condition is more negative than the attend-LOW condition in temporal and central areas
(ii) Both conditions have a similar response in lateral frontal areas
(iii) the onset of the larger Pd component in the attend-HIGH condition than the attend-LOW.

For the Pd component alluded to above we see that responses are generally similar centrally and temporally but moving frontally greater activity for the attend-HIGH condition gradually emerges. In fact at Fpz only slight activity for the attend-LOW condition is seen.

B. LOW responses

From Fig 2 (b) we see, as with the responses to the HIGH stream, a general early effect on the Pa component. In this case, however, instead of seeing little to no activity in the unattended stream (in this case the attend-HIGH condition) we see a definite negativity. An effect is also seen in the Na component which was not the case for the HIGH responses. Here we see greater negativity in the unattended response than the attended response.

Again contrary to the results for the HIGH responses we see an effect on the Nb component: a greater negativity in the unattended response than the attended at all electrode locations, most notably at Cz.

For the Nc component we again distinct effects depending on electrode location:

(i) A similar response for both conditions in temporal and central areas
(ii) A larger negativity for unattended condition in frontal areas (i.e. Fpz).

Fig. 2(b) shows an indication of a late affect at Fpz for the LOW responses but this is not as pronounced as was found for the HIGH responses.

IV. DISCUSSION

Firstly Fig. 2 shows that it is possible to extract temporally detailed responses to two simultaneously presented stimuli using the AESPA method.

Also given that the SNRs for the attended streams are significantly greater than those for unattended streams suggest that the AESPA is indeed be modulated by auditory attention.

The fact that the current responses are smaller in amplitude than the AEPSA obtained when only one stimulus is presented is to be expected as the standard AESPA was elicited using a BBN carrier wave with equal energy from 0 to 22.05 kHz while the current stimuli employ bandpass noise carriers which activate smaller regions of cortex. The fact that the

504

HIGH stream elicits smaller responses than the LOW stream may be due in part to the logarithmic nature of the tonotopic organization of auditory cortex [14] i.e. the higher the frequency the smaller the area of cortex devoted to it. Also smaller numbers of subjects took part in this experiment and not all trials from each subject were included due to the d' criterion. Thus the responses may be somewhat underpowered which may also account for the diminished nature of the Pc component.

A. HIGH responses

The enhancement of the early Pa component for the attended compared to the unattended stream is consistent with previous work [1-3] and seems to support the gating theory of auditory attention [1]. Our results emphasise the importance of inhibition of responses to unattended streams. Further work involving the use of a neutral attend-to-neither condition would be useful in clarifying the true nature of this early effect including the identification of whether facilitation of responses to attended streams also takes place. Furthermore it is generally seen that in standard AEP attention paradigms the latency of this early attention effect decreases with decreasing inter-stimulus interval (ISI) [4]. If we interpret the continuous stimuli as having an ISI of zero this early effect is to be expected.

The effects seen on the Nc component in temporal and central regions and the Pd component in frontal areas is also consistent with effects found in previous work [1, 3]. This modulation of sensory components again supports the gating theory.

These results agree with [7] who also used a continuous stimulation paradigm in which transient evoked, sustained evoked and steady-state responses are obtained and demonstrate the reduction the neural representation of irrelevant information. They also report the enhancement of the neural representation of relevant sounds. Again further work with a neutral condition should show if this effect is present here.

B. LOW responses

Again the early effect is present and is consistent with previous work as outlined above. The effect in this case seems to be more consistent across all electrodes and the negative deflection in the unattended response further emphasises the importance of the inhibition of responses to irrelevant information.

The larger Nb and Nc components for the unattended stream are most probably a carrier over from the preceding negative "Pa". Otherwise it may be the separate modulation of the exogenous components. If so this result is at odds with the corresponding result for the HIGH responses (i.e larger Nc for the attended stream). The reason for this is unclear. One possible reason for this anomaly is that the responses are somewhat underpowered and thus extending the study to include further subjects may bring clarification.

The late effect for the LOW responses is less pronounced than that for the HIGH responses but it is evident from the Fpz trace in fig.2 (b).

V.   CONCLUSION

The AESPA method represents a novel tool for the study of auditory attention. The continuous nature of the stimuli and the ability to extract separate responses to concurrently presented stimuli represents a much more ecological stimulation paradigm and thus has a huge advantage over standard protocols. Our results support the gain/filter theory of auditory attention. Further work involving a neutral condition and different stimuli incorporating different modulating statistics will shed further light in the effects found here.

REFERENCES

[1]   S. A. Hillyard, R. A. Hink, V.L. Schwent, T.W. Picton, Electrical Signs of Selective Attention in the Human Brain, *Science*, vol.128, No. 4108, pp. 177-180, 1973.

[2]   M.G. Woldorff, C.C. Gallen, S.A Hampson, S.A. Hillyard, C. Pantev, D. Sobel, F.E. Bloom, Modulation of Early Sensory Processing in Human Auditory Cortex During Auditory Selective Attention, *Proc Natl Acad Sci USA*, vol. 90, pp. 8722-8726, 1993.

[3]   M. Woldorff, J.C. Hansen, S.A. Hillyard, Evidence for Effects of Selective Attention in the Mid-Latency Range of the Human Auditory Evoked Potential. In: Current Trends in Event-Related Potential research (EEG suppl. 40). Eds: J. Johnson, J.W. Rohrburgh, R Parasuraman, Elsevier Science publishers B.V., 146-154 (1987).

[4]   W. Teder, K. Alho, K. Reinikainen, R. Naatanen, Interstimulus Interval and the Selective-Attention Effect on Auditory ERPs.: "N1 enhancement" versus processing negativity. *Psychophysiol,* 30, 71-81 (1993).

[5]   R. Naatanen, A.W.K. Gaillard, S. Mantysalo, Early Selective-Attention effect on Evoked Potential Reinterpreted, *Acta Psychologica*, Vol.42, 4, pp. 313-329, 1978.

[6]   P.T. Michie H.M. Bearpark, J.M. Crawford, L.C.T. Glue, The Nature of Selective Attention Effects on Auditory Event-Related Potentials, *Biol Psychol*, 30 219-250, 1990.

[7]   A. Bidet-Caulet, C. Fischer, J. Besle, P. Aguera, M. Giard, O. Bertrand, Effects of Selective Attention on the Electrophysiological Representation of Concurrent Sounds in the Human Auditory Cortex., *J Neurosci*, 23(35); 9252-9261, 2007.

[8]   Linden RD, Picton TW, Hamel G, Campbell KB. Human auditory steady-state evoked potentials during selective attention. *Electroencephalogr Clin Neurophysiol* 66:145-59, 1987.

[9]   Davis PA. Effects of Acoustic Stimulation on the Waking Human Brain. *J Neurophysiol* 2:494-499, 1939.

[10] E.C. Lalor, A.J. Power, J.J. Foxe, R.B. Reilly, Resolving the Precise Temporal Processing Properties of the Auditory System Using Continuous Stimuli, *Under review (2nd)*

[11] A.J. Power, E.C. Lalor, R.B. Reilly, Eliciting Audio Evoked Potentials Using Continuous Stimuli. *Proc. of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* , Lyon, August 2007.

[12] ISO 226:2003.

[13] G. Levine, S. Parkinson, *Experimental Methods in Psychology*, Lawrence Erlbaum Associates, 1994.

[14] G.L. Romani, S.J. Williamson, L. Kaufman, Tonotopic organization of the human auditory cortex, Science, Vol 216, Issue 4552, 1339-1340, 1982.