ORIGINAL ARTICLE

Electrophysiological Characterisation of Envelope-Following Responses

Eleina Mijares Nodarse,* María Cecilia Pérez Abalo, Alejandro Torres Fortuny, Mayrim Vega Hernández

Departamento de Audición y Lenguaje, Centro de Neurociencias de Cuba, La Habana, Cuba

Received 22 April 2011; accepted 8 June 2011

KEYWORDS
Temporal auditory processing; Human envelope-following responses; Language; Human auditory temporal resolution; Subject state; Auditory system

Abstract
Introduction: The auditory ability to discriminate rapid changes in the envelope of language sounds is essential for speech comprehension. This ability is deteriorated in some neurological diseases such as multiple sclerosis, auditory neuropathy, sensorineural hearing loss, presbycusis, and primary developmental language disorder. Envelope-following responses (EFRs) in humans are useful in objective measurement of temporal processing in the auditory nervous system.

Objectives: To evaluate EFRs in healthy younger subjects and to investigate the effects of subject states on the EFRs recorded.

Methods: Eleven young subjects were included; 6 of them were awake and 5 were asleep. EFRs were evoked by white noise carrier stimuli with a sweep of modulation frequencies from 20 to 200 Hz presented at 50 dB HL.

Results: The EFRs we recorded were similar in all subjects. There were two principal components. During both subject sleep and wakefulness, the first component (located between 30 and 50 Hz) was significantly larger than the second component (located between 80 and 110 Hz). There was also a significant effect of sleep on the EFR amplitude for the modulation frequencies between 88 and 110 Hz, 155 and 165 Hz, and 190 and 200 Hz. However, there were no significant effects of sleep on the principal EFR components.

Conclusions: These results corroborate the usefulness of the EFR technique for objective measurement of human auditory temporal processing.

© 2011 Elsevier España, S.L. All rights reserved.

* Corresponding author.
E-mail addresses: eleina.mijares@cneuro.edu.cu, eleinamijares@yahoo.com (E. Mijares Nodarse).
Caracterización electrofisiológica del potencial evocado auditivo de seguimiento a la modulación del estímulo acústico

Resumen

Introducción: Comprender el lenguaje implica percibir adecuadamente los cambios temporales que se producen en la amplitud de modulación de los sonidos del habla. La detección de estos cambios disminuye en las afeciones neurológicas como la esclerosis múltiple, la neuropatía auditiva, las hipoacusias neurosensoriales, la presbiacusia y el retraso primario del lenguaje. El potencial evocado auditivo de seguimiento a la envolvente (PEA-SE) permite estudiar electrofisiológicamente la respuesta auditiva a los cambios en la modulación de un estímulo acústico.

Objetivos: Caracterizar el PEA-SE en adultos sanos y estudiar los efectos del sueño sobre este potencial.

Métodos: Se estudiaron 11 sujetos, 6 en condiciones de vigilia y 5 en sueño espontáneo. Se registró el PEA-SE obtenido con un ruido blanco al que se le hizo un barrido de frecuencias de modulación entre 20-200 Hz a 50 dB HL de intensidad.

Resultados: El PEA-SE mostró una morfología similar en todos los sujetos, caracterándose por presentar dos componentes principales. La amplitud del primer componente (localizado entre 30-50 Hz de modulación) fue significativamente mayor que la del segundo componente (entre 80-110 Hz) tanto en los sujetos despiertos como en los dormidos. Existió además un efecto significativo del estado de vigilia sobre la amplitud del PEA-SE para las frecuencias entre 88-110 Hz, 155-165 Hz y 190-200 Hz. Sin embargo, no hubo un efecto significativo del sueño sobre los componentes principales del PEA-SE.

Conclusiones: En este trabajo se corrobora el patrón morfológico del PEA-SE sustentándose su utilidad para el estudio objetivo del procesamiento temporal del sistema auditivo.

© 2011 Elsevier España, S.L. Todos los derechos reservados.

Introduction

Variations in amplitude modulation (enveloping) of speech signals are responsible for prosody, thus being essential to the understanding of language. The enveloping component of speech sounds itself contains much of the information necessary for the identification of words, sentences and phrases.\(^1,2\) That is, although proper perception of the spoken word requires a minimum amount of information frequency, the information contained in temporal or enveloping patterns is more important.\(^2\)

The processes affecting the sensitivity of the auditory system to detect changes in the envelope of an acoustic stimulus (temporal auditory acuity) cause serious language alterations. Temporal auditory acuity is affected in the elderly,\(^3\) in patients with auditory neuropathy,\(^4,5\) multiple sclerosis,\(^6\) and with sensorineural hearing losses,\(^7,8\) as well as in children with primary delay in language development.\(^9\)

The two most common methods of evaluating temporal auditory processing are the detection of a short interval of silence within a continuous acoustic stimulus\(^10-12\) and the study of the modulation transfer function (MTF). The latter represents one of the most comprehensive manners of studying temporal auditory processing and consists of determining the detection threshold of amplitude changes in a sound depending on the modulation frequency.\(^13-15\) Both methods are subjective psychoacoustic techniques, making them unreliable for studying subjects who are uncooperative towards behavioural examination.

To obtain an electrophysiological representation of the MTF, Purcell et al.\(^16\) recently adapted the methodology of auditory steady-state evoked potential response (ASSR) sinusoidally modulated in amplitude.\(^17\) With ASSR, it is possible to explore a wide range of modulating frequencies and it is also possible to obtain multiple responses consecutively in a separate manner if each signal is modulated with a different frequency.\(^18,19\) However, separate examination of different modulators is very time-consuming, so it is more efficient to carry out a continuous sweep of modulation that covers all the frequencies of interest.

Instead of using pure tones of different spectral content as stimuli while keeping modulation frequency fixed and different for each tone, as is done to obtain ASSR,\(^17\) Purcell et al.\(^16\) used a fixed carrier tone and performed a modulation frequency sweep. The responses or potentials following this sweep in the envelope of the acoustic stimulus are known as envelope-following responses (EFRs).

The literature contains few published studies on EFRs in which a continuous modulation frequency sweep is used as stimulus.\(^16,20-22\) The changes occurring in this potential with varying the frequency composition of the stimulus and the intensity have been investigated,\(^16\) as well as the effect of sleep-wakefulness.\(^16,20\) However, these studies are not sufficient, as they involved a very small number of subjects; consequently, more evidence is still required regarding the usefulness of recording EFRs to study temporal auditory processing.

In this study, we attempted to characterise EFRs electrophysiologically in a larger sample of adult subjects, while
also analysing the possible effects of the awake state on this response.

Methods

Sample

The sample consisted of 11 healthy adults with normal hearing, aged between 23 and 31 years. Of these, 6 subjects were studied in a state of relaxed wakefulness and 5 under spontaneous sleep conditions. The study was conducted at the Hearing and Speech Department of the Cuban Neuroscience Centre, after obtaining informed consent. All subjects had to fulfil the following conditions:

1. Lack of personal and family history of pathological hearing loss.
2. Normal otoscopy.
3. Normal pure tone audiometry (behavioural threshold equal or less than 20 dB HL for 125, 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz in both ears).

Electrophysiological Recording of Envelope-Following Responses

To obtain the EFRs, we used an experimental model developed at the Rotman Research Institute by the group of Purcell et al. This model was comprised by the "Multisweep" program, developed on a Labview 5.1 platform for the generation of stimuli plus collection and off-line processing of EFRs, a conventional audiometer to adjust the stimulus intensity (Madsen, Orbiter 922 model), a bioelectrical amplifier model P55C from Grass Instruments and a 16-bit precision analogue-digital conversion card model 6052E from National Instruments.

Stimulation Parameters

The stimulus used consisted of amplitude-modulated white noise (100% depth) with a frequency sweep varying from 20 to 200 Hz (in steps of 0.586 Hz). The stimulus was maintained at a fixed intensity of 50 dB HL. Each sound file recorded 30 EEG segments of 1.024 s each, with a modulating frequency sweep being performed alternately. In the first half of the stimulus (first 15 segments), the modulation frequency increased linearly from a minimum (20 Hz) to a maximum frequency (200 Hz). In the second half, the modulation decreased linearly from the maximum frequency to the minimum (from 200 to 20 Hz), symmetrically with respect to the second half. The acoustic stimuli were designed so that there was no discontinuity in the transition between the end of one and the beginning of the next. A total of 30 repetitions of this type of stimulus were averaged in each subject.

Recording Parameters

The recordings were made inside a sound-proofed chamber. We used disk electrodes (Ag/AgCl) attached to the scalp with conductive paste, placed in a region previously cleaned with alcohol. The active electrode (+) was placed in Cz (midline, vertex of the head), the reference electrode (−) in the posterior midline, below the insertion line of the hair, and the ground electrode in Fpz (midline, frontal region). The stimulus was presented in monaural form, through insert earphones (model E-A-Rtone 3A).

Bioelectrical activity was amplified with a gain of 10 000, filtered analogically between 1 and 300 Hz and digitised at a frequency of 32 kHz using a 16-bit resolution analogue–digital converter. The suppression filter of the power line was kept on during all recordings, centred at 60 Hz. Consequently, modulation frequencies between 50 and 70 Hz were excluded from analysis in this study.

Acoustic Calibration

The calibration of the stimulus and environmental noise levels were determined with a model Investigator 2260 sound level meter from Bruel & Kjaer, a type 4152 artificial ear and a 4144 microphone. The calibration was done separately for each frequency using pure tones (standard AS1591.2). Noise levels were 59, 53, 49, and 44 dB SPL for 0.5, 1, 2, and 4 kHz, respectively.

Analysis of Envelope-Following Responses

Once the recording was complete, data was processed offline and averaged synchronously in the time domain. Before obtaining the potential, we averaged an analysis window (30 segments of 1.024 s each) to determine the mean and standard deviation (SD) of the noise. The rejection threshold was set at 1.5 SD from the previously-determined noise mean.

We used Fourier analysis to estimate the amplitude and phase of the EFRs for each modulation frequency. To do this, we jointly averaged the two symmetrical halves of each analysis window. The noise was estimated using 120 spectral components that were determined through the discrete Fourier transform, with 60 components ±3.9 Hz being found on each side of the frequency of interest.

We carried out a mean comparison test to determine whether the signal amplitude differed significantly from the noise, considering a value of $P \leq 0.05$ as statistically significant.

Results

Fig. 1 shows the typical EFR morphological pattern obtained in a representative subject during wakefulness (Fig. 1A) and that of another representative subject during spontaneous sleep (Fig. 1B). The morphology of EFRs was similar in the different subjects studied; in both awake and sleeping individuals, EFRs appeared as a continuous function, not monotonous, whose amplitude decreased as the modulation frequency increased.

The EFRs were significantly different from the noise ($P < 0.05$) in all frequencies explored, except below 33 Hz and in the modulation range between 120 and 200 Hz, where there were some narrow bands of frequencies (with bandwidths of approximately 5 Hz) where no significant responses were detected in some records.
Figure 1  Typical recording of EFR obtained in two adult subjects in a state of wakefulness (A) and in a state of spontaneous sleep (B). The solid lines show the response amplitude for modulation frequencies between 20 and 200 Hz. The dashed lines show the residual noise levels calculated for these frequencies. The arrows indicate peaks of interest where the response amplitude reached its maximum values.

In the morphology of the EFRs, we noted two regions or maximum amplitude peaks that appeared well defined and stable in all subjects (Fig. 1A and B). The first peak (P1) was located between 30 and 50 Hz and the second (P2), of smaller amplitude than P1, was located between 80 and 110 Hz. We also registered a region around 27 ± 5 Hz modulation that showed a marked decrease in signal amplitude and in which no significant response was detected (it must be noted that the modulation frequency range between 50 and 70 Hz was not analysed in this study [see Methods section]).

Table 1 shows, for both groups (awake and asleep), the mean and standard deviation of the frequency at which the maximum response amplitude was reached (best modulation frequency [BMF]) for each component, as well as the amplitude and residual noise (RN) values corresponding to the BMF. Note that the amplitude values, BMF and RN were different between components but very similar between asleep and awake subjects.

To assess the possible effect of wakefulness on the P1 and P2 components of the EFRs, we performed an ANOVA analysis of observations repeated with 2 factors: group (2 levels: asleep and awake) and components. The results showed no significant effect of the group on any of the variables studied: amplitude of the component ($F_{(1,16)} = 0.85; P < .37$), BMF ($F_{(1,16)} = 3; P < .11$) and RN ($F_{(1,16)} = 1.69; P < .21$). However, we did find a significant effect of the component, showing that P1 had a greater amplitude than P2 ($F_{(1,16)} = 74.36; P < .00$). Furthermore, the noise levels were also significantly lower for P2 ($F_{(1,16)} = 48.34; P < .00$). The interaction group × component was not significant in any of the ANOVA analyses performed. The effect of sleep on the amplitude of each component is illustrated in Fig. 2.

One of the objectives of this study was to determine the possible influence of sleep on the amplitude of this response, not only in terms of the P1 and P2 components, but through

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Analysis of the Best Modulation Frequency (BMF) and its Corresponding Amplitude and Residual Noise Level Values for the P1 and P2 Components of the EFRs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P1 component (between 30 and 50 Hz)</strong></td>
<td><strong>Wakeful Subjects</strong></td>
</tr>
<tr>
<td>BMF, Hz</td>
<td>41.9 ± 2.7</td>
</tr>
<tr>
<td>BMF amplitude, nV</td>
<td>122.5 ± 38.3</td>
</tr>
<tr>
<td>RN for the BMF, nV</td>
<td>31.1 ± 9.1</td>
</tr>
<tr>
<td><strong>P2 component (between 80 and 110 Hz)</strong></td>
<td><strong>Wakeful Subjects</strong></td>
</tr>
<tr>
<td>BMF, Hz</td>
<td>94.3 ± 5.1</td>
</tr>
<tr>
<td>BMF amplitude, nV</td>
<td>53.8 ± 13.9</td>
</tr>
<tr>
<td>RN for the BMF, nV</td>
<td>15.2 ± 7.3</td>
</tr>
</tbody>
</table>

BMF: best modulation frequency; RN: residual noise.
The different modulation frequencies studied. To this end, we used the bootstrapping method of re-sampling, with which it was possible to calculate the mean, variance and SD for both groups of subjects, asleep and awake. Fig. 3 shows a graph representing the mean and SD of the EFR amplitude across the region of frequencies between 20 and 200 Hz for both sleeping and awake subjects.

ANOVA analysis revealed a significant effect of sleep–wakefulness state on the estimated EFR amplitude (F(1, 52) = 206.65; P<.00). Univariate analysis of this result showed that these differences were statistically significant for modulation frequencies between 88 and 110 Hz, 155 and 165 Hz, and 190 and 200 Hz. In these frequency bands, the response amplitude in sleeping subjects was significantly higher than in subjects who were awake (Fig. 3).

Discussion

The recording of EFRs is a valid option for the objective study of auditory system sensitivity to changes in the modulation of a continuous acoustic stimulus over time. Therefore, this response could be used as an objective index to assess temporal auditory processing. However, there are few studies to date characterising EFRs and they have used relatively small samples of adults with normal hearing.

Overall, this study used a larger sample of healthy adults with normal hearing to corroborate and expand on previous results obtained on the morphology and reproducibility of envelope-following responses to a continuous acoustic stimulus. In addition, it also analysed the possible effect of wakefulness on these responses.

The first aspect to discuss with respect to the electrophysiological characterisation of EFRs is their morphological pattern. EFRs have been described as a continuous, non-monotone function whose amplitude decreases with increasing modulation frequency. This behaviour has been explained by the low pass filter effect created by the cranium. It has also been demonstrated that, despite the fact that potential amplitude decreases with increasing modulation frequency, the signal-to-noise ratio increases, since RN decreases even more than the EFR amplitude. The data obtained in this study about the morphology of EFRs match this description.

Another aspect to analyse is the structure of the components. As observed in previous studies, EFRs were characterised by the presence of two major components (P1 and P2). It has been considered that these may be similar to the ASSR caused by continuous tones modulated in amplitude at a frequency of 40 or 80 Hz, respectively.

We corroborated the reports about the amplitude of these components. The P1 amplitude was significantly greater than that of P2 in both wakeful and sleeping subjects. Once again, this presents analogies with the ASSR at 40 and 80 Hz, and could be attributed to different sources for both components. We still discuss whether the ASSR at 40 and 80 Hz are generated only by overlap of transient mid-latency auditory evoked potentials (MLAEP) and short-latency potentials (SLAEP), respectively, or if there could also be a contribution from neural populations responding specifically to oscillatory activity. In any case, the 40 Hz response generated by overlay of the MLAEP at higher levels of the auditory pathway (thalamocortical radiations and primary cortex) was recorded with greater amplitude at the surface than the far-field 80 Hz response generated at the brainstem by SLAEP overlap was recorded. It has also been suggested that there is some resonance in the auditory system for these two stimulation frequencies (40 and 80 Hz). This would explain the increase in amplitude of responses and the fact that this effect of amplitude increase by resonance is greater at 40 than at 80 Hz.

As in previous studies, we found in this analysis a marked decrease in the amplitude of the EFRs around 27 Hz modulation. A likely explanation for this finding is that, for those modulation frequencies, there may be a mismatch of the MLAEP and SLAEP components that would cause this decrease in amplitude. However, this should be clarified through further research.

Taking into account the influence that sleep has on the electrical activity of the brain and on the recording of AEPs, it is important to know how sleep affects EFR morphology. We maintain the analogy raised earlier between the P1 and P2 components of the EFRs and the ASSR obtained at 40 and 80 Hz modulation, where it is known that both responses have different generation sites. Sleep affects 40 Hz responses, generated at higher levels of the central nervous system, more than 80 Hz responses. Bearing this information in mind, it would be reasonable to expect sleep to have a greater effect on the amplitude of the response detected to frequencies below 50 Hz than on the response to frequencies above 70 Hz. However, we found no significant difference between sleep and wakefulness for either of the two main EFR components, although we did note a lack of influence of sleep on response amplitude in some bands of frequencies above 88 Hz, sometimes even reaching even greater potential amplitude during sleep.

Previous studies have found different effects of sleep on EFRs. Some authors have reported that sleep
decreased the amplitude of P1 but not of P2,\textsuperscript{16} while others found a significant decrease in the amplitude of both components.\textsuperscript{20} One possible explanation for this contradiction between different studies regarding the effects of sleep on EFRs is that such effects have only been investigated previously in a very small number of subjects (2–3 individuals). In our case, there was an additional limiting factor, since we used different samples of subjects for the records obtained during sleep and wakefulness. The wide variability among individuals, determined by the size of the head and the thickness of the skull,\textsuperscript{37} could mask possible differences caused by the effect of sleep on the EFR components.

Another limitation of this study is the use of Fourier analysis to estimate the EFRs. While this method provides a reliable estimate of the energy spectrum of stationary signals (such as ASSR), it is not the most appropriate method for estimating non-stationary responses, where the modulation frequency is constantly changing over time, as is the case of EFRs. For this reason, some authors\textsuperscript{20,21} have proposed the use of time-frequency analysis methods for the study of EFRs. This should be investigated in depth in future works.

**Conclusions**

This work confirms and extends previous results on the morphological pattern, main components and reproducibility between adult individuals of the EFRs obtained with a continuous sweep of modulation frequencies (between 20 and 200 Hz). It corroborates and supports the usefulness of this potential for objective evaluation of temporal auditory processing. In this study, the morphology of EFRs in sleeping, normal hearing adults was very similar to that of wakeful subjects. The response amplitude in some bands with modulation frequencies above 88 Hz was significantly affected by sleep. This effect was not observed in any of the two main components of the potential (components P1 and P2) or their corresponding values for amplitude and residual noise level. Nevertheless, these results have limitations related mainly to the use of suppression filters on the power line, the sample selection and the use of Fourier analysis for signal processing. These limitations should be redressed in future research.

**Conflict of Interest**

The authors have no conflicts of interest to declare.

**Acknowledgements**

We wish to thank Professor Terence W. Picton for his cooperation, in providing the experimental model with which electrophysiological recordings were made.

**References**


