Auditory steady state responses in cochlear implants

JAN WOUTERS AND MICHAEL HOFMANN

ExpORL, Dept. Neurosciences, K.U.Leuven, Belgium

Electrically Auditory Steady State Responses (EASSRs) are EEG potentials in response to periodic electrical stimuli presented through a cochlear implant (CI). Recently, for slow rate pulse trains in the 40Hz range, the electrophysiological thresholds derived from response amplitude growth functions have been demonstrated to correlate well with behavioral thresholds at these rates. In the following studies we show that auditory steady state potentials can as well be reliably evoked by amplitude-modulated or pulse-width-modulated high-rate pulse trains at clinically used carrier rates, and that stimulus artifacts can be completely removed from the electrophysiological recordings. Multichannel EEG-data have been recorded in Nucleus cochlear implant users. The properties of the resulting responses with regards to amplitude, phase and apparent latency are analyzed. The predictive value of electrophysiological thresholds derived from such responses for behavioral thresholds at these high rates is examined. This objective threshold determination method may be used in future CI fitting paradigms.

INTRODUCTION

Auditory steady state responses (ASSRs) are periodic EEG potentials as a neural response to repeated or fluctuating auditory stimuli. These responses are typically evoked by sinusoidal amplitude modulated (SAM) tones, modulated noise or pulse trains (Galambos *et al*. 1981; Picton 2003). The auditory steady state responses have some specific properties that distinguish them from event-related evoked potentials.

The ASSR is a kind of temporal integration response different from immediate responses after single-pulse or transient stimuli. Because frequency-specific stimuli can be used to elicit frequency-specific ASSRs, frequency domain analysis methodologies can be used to objectively detect a response from the EEG-background. The different rates or modulation frequencies in the stimulus give rise to ASSRs with generators at brainstem to auditory cortex, depending on the basic modulation. Large responses can be evoked at modulations of about 40 Hz, and in second stage also at 80-100 Hz. The latter responses are generated at brainstem level, the former have cortical components.

ASSRs have been an interesting topic of investigation in basic research. However, in the last few years a lot of research has been done with focus on clinical applications with ASSRs. Auditory thresholds obtained with ASSRs have been shown in several studies to be well correlated (>.90) with behavioural hearing thresholds in adults and

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.

children (Picton, 2011). In a common application stimuli typically used are 4-tone mixes with 4 carrier wave frequencies at 0.5, 1, 2 and 4 kHz, each at a different modulation frequency between 80 and 100 Hz. Both ears can be tested simultaneously with the same stimulus if and only if the 4 modulation frequencies have (slightly) different values. The ASSR thresholds are determined as the stimulation level where an ASSR response (at a specific frequency, e.g. the 8 modulation frequencies when bilaterally tested) was still significantly different (F-test) from the EEG-background. For neonates at about 3 months of age 4-frequency auditory ASSR thresholds have been determined which correlate well (.80-.90) with behavioural hearing thresholds that could only be administered on average more than 15 months later, and with an average difference and variance of about 15 and 10 dB, respectively.

At present ASSR are clinically applied in diagnostics and rehabilitation programs in the follow-up after neonatal hearing screening (Luts *et al*. 2006; Alaerts *et al*. 2010). Currently they are also applied in the investigation of objective markers for different learning disabilities, for instance dyslexia and speech-language impairments, where multi-channel ASSR recordings and different performance measures such as SNR, and phase coherence are applied (Poelmans *et al*. 2011).

Most of the ASSR results have been obtained with acoustical auditory stimulation, but the method should equally well be applicable with electrical stimulation in cochlear implants (CIs). There is a big need for objective measures to facilitate the fitting of cochlear implants. The fitting basically consists of setting thresholds and comfortable loudness levels for different stimulation channels and, up to now, this fitting has been mainly based on behavioural data obtained with the patient. Presently, thanks to the neonatal screening programs operational in many countries, many deaf children are being implanted at very young ages (below one year of age) that enhance the possibility of successful rehabilitation and integration in the hearing community. These very young ages, however, preclude reliable fitting of the cochlear implants based on behavioural data. Therefore, the need to develop objective measures for fitting cochlear implants has attracted much attention.

Electrically evoked compound action potentials (ECAPs) and electrically evoked auditory brainstem responses (EABRs) have been investigated on a large scale for these applications. However, there does not exist a one-to-one relation of behavioural thresholds with ECAP or EABR threshold measures. Correlations are apparent but do not allow the application without important additional behavioural input data (Miller *et al*. 2008; McKay *et al*. 2005). So, based on the recent acoustical ASSR investigations, the possible utility of EASSR for these purposes may lead to relevant applications because of the frequency-domain detection paradigm and because modulations at low frequencies are well preserved through the cochlear implant signal processing path. However, a considerable experimental problem is the large interference of artefacts from the electrical stimulation and power-up pulses in the surface recording electrodes. Feasibility of EASSR has been demonstrated using SAM electrical stimuli in guinea pigs (Jeng *et al*. 2007, 2008) and one study reported detection with pulsatile stimulation in CI-users but analysis

and interpretation of EASSRs was indirect (Ménard *et al*. 2004). Recently, we have shown that EASSRs can reliably be measured in CI users and good correlations have been obtained between objective and psychophysical thresholds at stimulation rates of 40 pulses per second (pps) (Hofmann and Wouters, 2010). In the present report the extension of this research to modulated pulse trains at high stimulation rates relevant for CI speech processors is described.

METHODS

A similar stimulation and seven-channel recording setup as in Hofmann and Wouters (2010) was used. Six subjects with a Cochlear Nucleus CI participated in the measurements. The stimuli from the experimental platform were directly presented to the implant of the patient using the POD programming device and the L34 research processor of Cochlear Ltd. Surface electrodes were placed on the head of the subjects following the international 10-20 system. The reference electrode was on the vertex (Cz) and the ground electrode on the clavicle contralateral to the side of the CI. Seven electrodes were placed on the high forehead (Fz), left and right of the vertex (C5 and C6), the contralateral mastoid (TP9 or TP10), and the left, middle and right back of the head (P3, Oz, P4). The electrodes are connected to an eight-channel Jaeger-Toennies low-noise differential amplifier with a gain of 50000. The amplified output was input to an external RME Multiface II sound card which was connected to the measurement and control laptop.

The electrical stimuli were trains of symmetric biphasic pulses with a phase width of 40 µs and an interphase gap of 8 µs. Three different stimulus types have been used: unmodulated pulse trains at low rates close to 40 pps (P), amplitude modulated high-rate pulse trains (AM) with a carrier rate of 900 pps and sinusoidal modulation around 40 Hz, and phase-width modulated high-rate pulse trains (PWM) with a carrier rate of 900 pps and sinusoidally modulated phase widths between 25 and 40 µs at modulation frequencies around 40 Hz. Two bipolar stimulation electrode pairs were chosen from pairs close to the apex, base and middle of the electrode array such as to maximize the differences in behavioural thresholds between the two.

EEG recordings of responses to electrical stimulation are contaminated with artefacts from the electrical stimulus pulses, RF transmission of the CI and muscle movement. These multiple artefacts are most disturbing for the detection of these steady state and low-amplitude neural responses. A multi-step process was employed to remove the artefacts. This has been a major experimental investment for this research and is described in Hofmann and Wouters (2011).

After the removal of stimulus and recording artefacts, a fast Fourier Transform (FFT) was used to calculate the complex frequency spectrum for each epoch. The mean amplitude and phase were obtained by vector-averaging the response frequency bins of the epochs. To determine the significance of a response, a two-sample Hotelling T^2 test was used on multiple measurements with two different modulation frequencies to compare the response bins for the same recording electrode and stimulus intensity. To increase statistical power, response bins for multiple recording electrodes (at the same stimulation level) and multiple measurements at different stimulation levels (and the same recording electrode) were combined. Additionally, a one-sample Hotelling T^2 test was also applied to check for the influence of stimulus artefacts on the response detection. A significance level of $p<0.05$ was used for all tests.

To detect stationary EEG activity synchronous with the stimulation, different methods are available (Picton *et al*. 2003). All of these methods are based on a comparison of the assumed response, the signal, with the spontaneous EEG activity not linked to the stimulation, the noise. Time-domain and frequency domain techniques can both be applied. Different statistical tests give identical results. However, for the analysis of EASSRs in the frequency and complex domains, any remaining artefacts after stimulus artefact removal will also contain components that will appear in the frequency bin corresponding to the modulation frequency and that may be erroneously interpreted as a neural response by one-sample tests. This problem is more prominent for electrical stimulation at high rates (Hofmann and Wouters, 2011). One way to distinguish between stimulus artefacts and neural responses is to evaluate the behaviour with changing modulation frequency. This leads to the application of two-sample tests (by using data for two different but nearby modulation frequencies). Response phase delay changes linearly with modulation frequency, which is caused by the constant latency of the evoked neural response per frequency range (Picton *et al*. 2003). For the potentially remaining artefacts of the high-rate pulse trains, the spectral component at the modulation frequency is independent of modulation frequency with a phase that depends on the initial phase of the stimulus. Therefore, stimulus artefacts have a phase delay of 0°. To use this phase dependency on modulation frequency for response detection, two measurements with different modulation frequencies are required for each stimulus condition. A suitable test for these purposes is a two-sample Hotelling T^2 test.

The EASSR properties response amplitude, phase delay, apparent latency and electrophysiological thresholds were evaluated in six subjects with each two stimulation electrode pairs. Measurements were conducted at modulation frequencies of 35 and 44 Hz, with the three stimulus types as described above, and on eight evenly spaced decreasing current levels between comfort level for modulated 900 pps pulse trains and below T level for unmodulated 900 pps pulse trains. The distribution of significant response amplitudes was analyzed per recording electrode, and the median response amplitude was compared to the noise level in the FFT response frequency bin. Electrophysiological thresholds were derived from the recorded responses for the different analysis methods and the ability to determine thresholds for all recordings was evaluated. In this paper we focus on response amplitudes and threshold values.

RESULTS

Artefacts resulting from the bipolar stimulus and power-up pulses could be completely removed in all subjects and on all electrodes even at pulse rates up to 900 pps with the developed methods. No recording electrodes had to be excluded from the analysis.

The mean response amplitudes to stimulation at comfort levels per recording electrode for all subjects are shown in figure 1 (n=219 from six subjects on two stimulation electrode pairs, three stimulus types, two modulation frequencies of 35 and 44 Hz, seven recording electrodes, significant ASSR responses in about 45% of the measurements). Reliable responses could be obtained from the recording electrodes at the contralateral mastoid and the back of the head, with median response amplitudes of 99 nV and 65 nV, respectively. Noise levels were at 1160 nV and 690 nV, respectively, which were reduced to 35 and 22 nV after time-domain averaging. Each of the recording electrode sites at the contralateral mastoid or the back of the head lead to similarly good SNRs and thus seem to be equally suited for recording EASSRs.

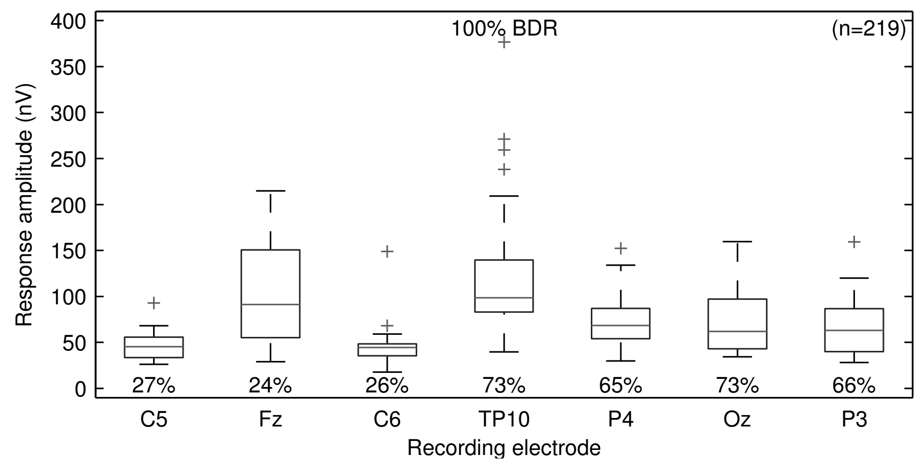

Fig. 1: Response amplitudes for different recording electrodes. Box plots are shown for the amplitudes of significant responses at comfort level stimulation per recording electrode for all subjects (n=219), percentages show the number of significant responses (different from noise) relative to all measurements. For subjects with a CI in the right ear, electrodes C5, P4 and TP10 were swapped with electrodes C6, P3 and TP9, respectively.

Overall, response amplitudes from AM and PWM high-rate pulse trains had a shallower response-growth function and resulted in larger response amplitudes at lower stimulus intensities than for the low-rate pulse trains.

In figure 2 box plots of the electrophysiological thresholds derived from the recorded responses for the two different analysis methods mentioned above are presented. There was no difference between the thresholds of the two-sample Hotelling T^2 test for the modulated high-rate pulse trains applied to recordings either with or without stimulus artefact removal. However, for recordings without stimulus artefact removal, it was not possible to deduce all electrophysiological thresholds from the two-sample Hotelling T^2 test for low-rate pulse trains (failing in 25% of the cases). As a comparison, a one-sample Hotelling T^2 test applied to recordings without stimulus artefact removal did not yield thresholds in 33, 8 and 83% of the measurements with low-rate, AM and PWM high rate pulse trains, respectively. This clearly demonstrates the importance of the two-sample approach of the combination of data at two nearby modulation frequencies for artefact effect reduction and thus the determination of reliable electrophysiological thresholds from EASSRs.

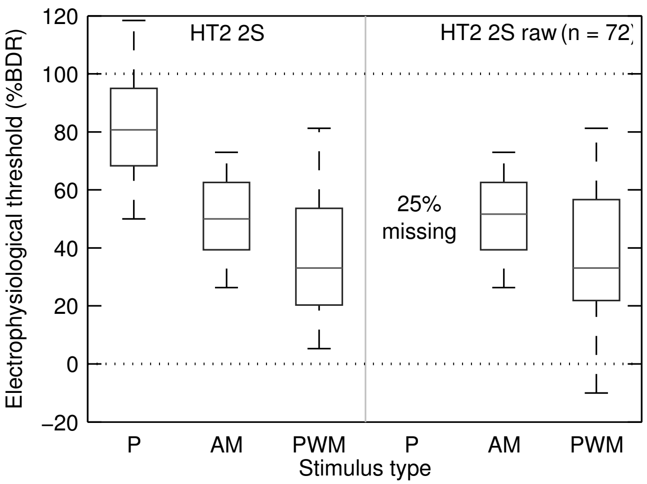

Fig. 2: Electrophysiological thresholds derived from response growth functions. Box plots are shown of the thresholds across subjects for the two different analysis methods relative to the behavioural dynamic range at pulse rates of 900 pps (n=132); HT2 2S: thresholds derived with two-sample Hotelling T2 test; raw: thresholds obtained from recordings without stimulus artefact removal; percentage missing: some electrophysiological thresholds could not be determined.

The thresholds obtained for the three stimuli were significantly different from each other (figure 2). Mean differences between objective thresholds and T levels at 900 pps were 77%, 49% and 31% of the behavioural dynamic range (BDR) for low-rate, AM and PWM high-rate pulse trains, respectively. For low-rate pulse trains, thresholds were more correlated with T levels at 40 pps than with T levels at 900 pps

(r_{40}=0.96 and r_{900}=0.80). Both AM (r_{40}=0.80 and r_{900}=0.96) and PWM (r_{40}=0.72 and r_{900}=0.96) high-rate pulse trains had higher correlations with T levels at 900 pps than at the low-rate pulse trains. All correlations were significant with $p<0.01$.

CONCLUSIONS

The results of this study show that it is possible to use modulated high-rate pulse trains to evoke EASSRs in CI users. An improved artefact removal procedure has been used that can completely remove stimulus artefacts from recordings of EASSRs to stimuli at clinical pulse rates. Electrophysiological thresholds derived from EASSRs excited by modulated high-rate pulse trains were lower than thresholds for low-rate pulse trains and correlated very well (r=0.96) with behavioural T levels at clinically used pulse rates in CI speech processors. The median of the electrophysiological thresholds for the PWM-stimuli was only about 30% of the dynamic range higher than the behavioural T levels.

REFERENCES

Alaerts, J., Luts, H., Van Dun, B., Desloovere, C., and Wouters, J. (**2010**). "Latencies of auditory steady state responses recorded in early infancy" Audiology and Neuro-Otology, **15** (2), 116-127

Galambos, R., Makeig, S., and Talmachoff, P. J. (**1981**). "A 40-Hz auditory potential recorded from the human scalp" Proceedings of the National Academy of Sciences of the United States of America, **78** (4), 2643-7

Hofmann, M., and Wouters, J. (**2010**). "Electrically evoked auditory steady state responses in cochlear implant users" Journal of the Association of Research in Otolaryngology, **11** (2), 267-282

Hofmann, M., and Wouters, J. (**2011**). "Improved detection of auditory steady state responses with electrical stimuli in cochlear implant users" Submitted

Jeng, F.-C., Abbas, P. J., Brown, C. J., Miller, C. A., Nourski, K. V., and Robinson, B. K. (**2007**) "Electrically evoked auditory steady-state responses in guinea pigs" Audiology and Neuro-Otology, **12** (2), 101-12

Jeng, F.-C., Abbas, P. J., Brown, C. J., Miller, C. A., Nourski, K. V., and Robinson, B. K. (**2008**) "Electrically evoked auditory steady-state responses in a guinea pig model" Audiology and Neuro-Otology, **13** (3), 161-71

Luts, H., Desloovere, C., and Wouters, J. (**2006**). "Clinical application of dichotic multiple-stimulus auditory stead-state responses in high-risk newborns and young children" Audiology and Neuro-Otology, **11**, 1, 24-37

McKay, C. M., Fewster, L., and Dawson, P. (**2005**), "A different approach to using neural response telemetry for automated cochlear implant processor programming" Ear and Hearing, **26** (4 Suppl), 38S-44S

Ménard, M., Gallégo, S., Truy, E., Berger-Vachon, C., Durrant, J. D., and Collet, L. (2004), "Auditory steady-state response evaluation of auditory thresholds in cochlear implant patients" International Journal of Audiology **43**, 39-43

Miller, C. A., Brown, C. J., Abbas, P. J., and Chi, S. L. (**2008**), "The clinical application of potentials evoked from the peripheral auditory system" Hearing Research **242** (1-2), 184-97

Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. W. (**2003**). "Human auditory steady state responses" International Journal of Audiology **42** (2), 177-219

Picton, T.W. (**2011**). *Human auditory evoked potentials*, Plural Publishing, San Diego

Poelmans, H., Luts, H., Vandermosten, M., Boets, B., Ghesquière, P., and Wouters, J. (**2011**). "Auditory steady-state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia" Ear and Hearing, in press

Neurocortical mechanisms of comprehension in degraded speech

JONAS OBLESER[1], ANTJE STRAUß[1] AND ANNA WILSCH[1]

[1] *Max Planck Research Group "Auditory Cognition", Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*

Comprehending speech is an astonishing faculty of the human brain, especially so under adverse listening conditions. How and by which neural mechanisms do we cope so well with the fleeting percepts of speech? In addition to "facilitating" influences such as semantic context, listeners also cope with challenging listening situation by fully exploiting their sensory and cognitive resources, e.g. their working memory capacities ("compensation"). We will propose a framework for studying neural signatures of facilitation and compensation, and discuss recent data from functional MRI (fMRI), magneto- and electroencephalography studies (M/EEG; with an emphasis on neural oscillations) that utilize acoustically degraded speech stimuli to study the neural underpinnings of these facilitation and compensation mechanisms in detail.

INTRODUCTION

The so-called "bottom-up" flow of auditory information from the sensory periphery to the central auditory areas in the superior temporal plane (and further on in the cerebral cortex) is not as well understood as comparable processes in the visual domain. Nevertheless, great progress has been made over the last decades. For recent reviews (from a neuroscience perspective) of suggested mechanisms and pathways see e.g. Hackett (2008), Rauschecker and Scott (2009), and Schreiner and Winer (2007).

Our starting point in this paper is the well accepted and intuitively clear assumption that these so-called bottom-up or afferent processes in audition alone cannot explain the astonishing human ability to cope with substantial degradation of the auditory input and still achieve comprehension – this is arguably most relevant in comprehending speech. In principle, however, it applies to any sound of potential behavioural relevance (e.g., we are also able to recognise known music pieces or known voices of music in adverse listening situations).

Two notes of caution: First, for simplicity, the term "top-down" shall be assigned to all these processes that help facilitate or compensate in the process of comprehension in adverse listening. This does not imply that all such top-down processes are assumed to feed back to basic auditory processing levels, as top-down in the strict sense would indicate. In contrast, it has been convincingly argued before that this is not very likely to be the case, and that top-down adjustments to the

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.