STATISTICAL VARIABILITY OF AUDITORY RESPONSES TO AMPLITUDE MODULATION

Boris Gourévitch, Régine Le Bouquin Jeannès and Gérard Faucon

Laboratoire Traitement du Signal et de l'Image, INSERM, Université de Rennes 1
Campus de Beaulieu, Bâtiment 22, 35042 Rennes Cedex, FRANCE
phone: +33 2 23 23 62 20, fax: +33 2 23 23 69 17, email: boris.gourevitch@univ-rennes1.fr

ABSTRACT

The response of the human auditory system to amplitude modulation (AM) presents oscillating activity in stereoelectroencephalographic (SEEG) signals, the oscillation frequency being very close to the modulation frequency (MF). Numerical studies that may be realized on these SEEG signals mainly focus on the estimation of three parameters: the amplitude, phase and frequency of the oscillation. To improve their estimation, noise reduction techniques are discussed. The estimations errors are analyzed on simulated and real data. Results show that the parameters variability may be explained only by the signal-to-noise ratio (SNR) level in SEEG recordings.

1. INTRODUCTION

Although sound is decomposed in frequency bands in cochlea, human being perceives the temporal envelope of a sound, which plays an important role in speech intelligibility. However, psychoacoustical experiments conducted by Viemeister in 1978 showed that the higher the frequency of the AM envelope, the worse our ability to hear it [1]. This suggests that some specific post-cochlea auditory filters discriminate the speech temporal envelope. Electroencephalography analysis revealed later an oscillating electrical activity in the auditory brain areas in response to amplitude modulated noise ([2], [3]). This activity reproduces almost perfectly the AM frequency but shows an amplitude varying with the precise place where it is recorded. To lead a statistical study on these oscillations, we extract their three characteristic parameters that are the amplitude, phase and frequency. In this paper, section 2 is devoted to the database and materials. Then, we derive in section 3 a model for auditory responses. After the presentation of some estimation techniques (section 4), we will explain the parameters variability (section 5).

2. MATERIALS

Twenty epileptic patients suffering from partial refractory epilepsy participated in this study (eight males, twelve females, 18-50 years old). They were implanted with chronic SEEG electrodes ([4], [5]) in various cortical structures for surgery exploration. The anatomical positions of the leads were determined in a parallel study [6] and we only considered leads in or near auditory areas for our study. A total of 208 leads in the right (9 patients) or left auditory cortex (11 patients) were available in the whole database.

Stimuli were 1-second long white noises with a sinusoidal amplitude modulation at frequencies 4, 8, 16, 32, 64, and 128 Hz and a modulation depth of 100 %. Stimuli were shaped by rising and falling 25 ms cosine ramps/damps to avoid auditory response to sudden sound rise. The energy of stimuli is made identical. Sounds were presented binaurally via headphones to the listener by series of 50 to 100 stimuli of two randomly alternated MF (128/16 Hz ; 4/32 Hz ; 8/64 Hz). They were generated using a 16-bit D/A converter at a sampling frequency of 44.1 kHz.

White noise carriers have been used because they have flat long-term power spectrum and so they excite all afferent paths of the tonotopic decomposition. In this study, only intervals presenting some clear oscillation and free of transient response were considered. Their length was around 800 ms.

3. MODEL

For a given lead, the j-th response to the stimulus is named an epoch and may be modelized on the considered interval as a noisy sinusoidal activity

$$x^{(j)}(t) = a^{(j)}.\sin\left(2\pi\frac{f_0^{(j)}}{f_e}t+\phi^{(j)}\right)+n^{(j)}(t) \tag{1}$$

$$= s^{(j)}(t)+n^{(j)}(t) \tag{2}$$

at time t, where $a^{(j)}$, $f_0^{(j)}$ and $\phi^{(j)}$ are respectively samples of three random variables A (amplitude), F_0 (observed oscillation frequency) and Φ (phase). The quantity $n^{(j)}(t)$ is a sample of the brain activity, N, seen as a noise and $s^{(j)}(t)$ is a sample of the useful signal S, seen as a function of t, A, F_0 and Φ. In our database, 50 to 100 samples of each variable are available for a given lead. The aim of the study is to evaluate the mean amplitude $E[A]$, frequency $E[F_0]$ and phase $E[\Phi]$ of the oscillation for each lead.

4. SPECTRAL ESTIMATION

Given the model of eq. 1, our aim is to estimate $E[A]$, $E[F_0]$ and $E[\Phi]$ with minimal influence of noise N. Then, we estimate standard deviation (STD) on real and simulated data for A, F_0 and Φ, to study physiological reproducibility of the response on all epochs.

We denote by $\gamma_{x^{(j)}}$ and $\gamma_{s^{(j)}}$ the power spectral densities (PSD) of $x^{(j)}$ and $s^{(j)}$ respectively. PSDs are implicite functions of f and for convenience, the variable f is omitted in the rest of the paper. In the following, the mean of a variable $v^{(j)}$ on j will be denoted by $\overline{v}$ and the mean of the PSD of $v^{(j)}$ will be denoted by $\overline{\gamma_{v^{(j)}}}$. For a single epoch, we may estimate the PSD of $\gamma_{s^{(j)}}$, to obtain amplitude and frequency estimation of the oscillation of each epoch, in the following way,

$$\widehat{\gamma}_{s^{(j)}} = X^{(j)*}.X^{(j)} \tag{3}$$

where $X^{(j)}$ is the Fourier transform of $x^{(j)}$ and * denotes conjugate operator.

To estimate the signal (S) PSD, γ_s, we may average epochs PSD

$$\widehat{\gamma}_s^{APSD} = \overline{\gamma_{x^{(j)}}} = \frac{1}{p}\sum_{j=1}^{p} X^{(j)*}X^{(j)} \tag{4}$$

where p is the number of epochs. Noise is not reducted but this estimator is phase independent, that may be useful if oscillations are not well synchronized in the course of epochs.

To reduce disturbances due to noise in estimating $E[A]$, $E[F_0]$ and $E[\Phi]$, we benefit from information on all epochs.

4.1 Noise reduction

To estimate γ_s, noise reduction algorithms often require the knowledge of the noise PSD. The first method consists in learning the noise PSD between stimuli (periods of spontaneous SEEG activity). Another method consists in considering that the signal can be estimated by the auditory evoked potential (AEP) itself, *i.e.* $\bar{x}(t) = \frac{1}{p}\sum_{j=1}^{p} x^{(j)}(t)$. It is based on the fact that signals $s^{(j)}$ are synchronous. In this way, the noise PSD is given by

$$\widehat{\gamma}_n = \overline{\gamma_{x^{(j)}-\bar{x}}}. \tag{5}$$

Due to the small number of epochs, there exist some local variations in the noise PSD estimated either with inter-stimuli activity or with $\widehat{\gamma}_n$ given by eq. 5. For simplicity reasons, we decide to use eq. 5 in our study. On real signals, the oscillations in epochs $x^{(j)}$ are not strictly identical in amplitude and phase and so there is a residual oscillation in $x^{(j)} - \bar{x}$. Consequently, we slightly improve the noise PSD estimation by carrying out spline interpolation around the MF. Experiments carried out show the same levels of noise PSD with or without stimuli. In the following, we describe various estimators for γ_s, some of them using this interpolated noise PSD. Let us indicate that, in numerical applications, the possible negative values in PSD estimations will be set to 0.

4.1.1 Time averaging method

The signal estimated by time averaging and corresponding to the AEP does not require noise PSD estimation and is given by

$$\widehat{s}(t) = \overline{x(t)} = \frac{1}{p}\sum_{j=1}^{p} x^{(j)}(t) \tag{6}$$

where p is the number of epochs available for a lead. It is widely used in AEP studies and minimizes $\sum_{j=1}^{p}\left(x^{(j)} - \widehat{s}\right)^2$. The PSD of this estimate $\widehat{s}$ will be noted $\widehat{\gamma}_s^{TA}$ in the following.

4.1.2 Spectral subtraction

Signal PSD estimated by amplitude spectral subtraction introduced by Boll in 1979 [7] is given by

$$\sqrt{\widehat{\gamma}_s^{ASS}} = \sqrt{\overline{\gamma_{x^{(j)}}}} - \sqrt{\widehat{\gamma}_n}. \tag{7}$$

In the same way, the power spectral subtraction estimator is

$$\widehat{\gamma}_s^{PSS} = \overline{\gamma_{x^{(j)}}} - \widehat{\gamma}_n. \tag{8}$$

This is maximum likehood estimation of PSD under gaussian assumption for N and S ([8]). Calculus shows that

$$\widehat{\gamma}_s^{PSS} = \overline{\gamma_{x^{(j)}}} - \overline{\gamma_{x^{(j)}-\bar{x}}} = \widehat{\gamma}_{\bar{x}} = \widehat{\gamma}_s^{TA}. \tag{9}$$

The power spectral subtraction estimator is equivalent to the PSD estimator of the time averaged signal. Moreover, $\widehat{\gamma}_s^{PSS}$ is the PSD estimated with the pseudo-Wiener filter ([8]) when the filtering is applied to $\overline{\gamma_{x^{(j)}}}$.

4.1.3 McAulay and Malpass estimator

An optimal amplitude estimator was proposed by McAulay and Malpass ([9]), based on a sinusoidal signal, with amplitude and phase unknown, and adapted to real-time processing by use of a priori and a posteriori information in sequential frames. In our study, SEEG epochs are not seen as sequential frames and the estimator may be written

$$\sqrt{\widehat{\gamma}_s^{MM}} = \frac{\sqrt{\overline{\gamma_{x^{(j)}}}} + \sqrt{\overline{\gamma_{x^{(j)}}} - \widehat{\gamma}_n}}{2} \tag{10}$$

with signal presence probability fixed to 1.

4.1.4 Averaged cross-spectra based method

We estimate the signal PSD using cross-products of epochs spectra, so that

$$\widehat{\gamma}_s^{ACS} = \frac{1}{p(p-1)}\sum_{k=1}^{p}\sum_{j=1, j\neq k}^{p} X^{(k)*}X^{(j)}. \tag{11}$$

We have the following relation

$$\widehat{\gamma}_s^{TA} = \overline{X}^*\overline{X} = \frac{1}{p^2}\sum_{k=1}^{p}\sum_{j=1, j\neq k}^{p} X^{(k)*}X^{(j)} + \frac{1}{p^2}\sum_{k=1}^{p} X^{(k)*}X^{(k)} \tag{12}$$

$$= \frac{(p-1)}{p}\widehat{\gamma}_s^{ACS} + \frac{1}{p}\widehat{\gamma}_s^{APSD}. \tag{13}$$

Noise is assumed to be uncorrelated between two different epochs, so $\widehat{\gamma}_s^{ACS}$ is a part of $\widehat{\gamma}_s^{TA}$ where the noise energy is in theory null contrary to $\widehat{\gamma}_s^{APSD}$. If the signal is highly correlated between epochs, it is pertinent to consider $\widehat{\gamma}_s^{ACS}$ as an estimator of the signal PSD since $\widehat{\gamma}_s^{ACS} \underset{p\to\infty}{\sim} \widehat{\gamma}_s^{TA}$.

4.2 Frequency and amplitude estimation

Given the signal PSD estimated with the previous methods, we attempt to estimate the real amplitude and frequency of the oscillation. To this end, we take the frequency and the amplitude of the nearest spectral peak to the MF (4, 8, 16, 32, 64, 128 Hz).

4.3 Simulations

We tested noise reduction methods by simulation with $p = 50$ epochs, $f_0^{(j)} = 4$ Hz, signal length $q = 800$ ms and sampling frequency $f_e = 1$ kHz. Amplitude was fixed to $a^{(j)} = 10$ and phase was set to $\phi^{(j)} = 0$. For noise $n^{(j)}$ modelling, we derive an AR model from all epochs recorded in the primary auditory cortex, where the averaged signal has been subtracted for each lead. Experiments showed the same shape of noise PSD between all MF. Consequently, the AR modelling is performed at MF equal to 128 Hz, due to the low level of auditory evoked response at this high MF. The AR coefficients are given in Table 1. For this simulation,

$a_0 = 1$	$a_1 = -1.6471$	$a_2 = 0.6041$	$a_3 = 0.1676$
$a_4 = -0.0801$	$a_5 = -0.0429$	$a_6 = 0.0075$	

Table 1: AR coefficients for the SEEG generating model

the signal-to-noise ratio (SNR) is defined by

$$SNR_1 = 10\log_{10}\left(\frac{V[s]}{V[n]}\right) \tag{14}$$

for a lead. Due to some equivalences related in section 4.1, we tested the estimators $\widehat{\gamma}_s^{ASS}$, $\widehat{\gamma}_s^{ACS}$, $\widehat{\gamma}_s^{TA}$, $\widehat{\gamma}_s^{MM}$ and $\widehat{\gamma}_s^{APSD}$. We included also an estimation of the PSD using a Hamming windowed version of the time averaged signal (γ^{HWTA}) to see the window influence. Results presented in figure 1 concern the mean amplitude, the root mean square (RMS) error on the MF estimate and the RMS error on exact amplitude.

The averaged cross-spectra based method seems to be the best estimator for amplitude estimation (fig. 1a and 1b). It slightly outperforms the time averaging method. Nevertheless, the interesting performances of the time averaging method justifies the use of averaged responses in studies based on event related potentials. Windowing brings no real advantages. Using averaged PSD (eq. 4) leads to erroneous values in amplitude and frequency estimation for low SNR. Amplitude spectral subtraction leads to important bias for amplitude estimation for low SNR levels. Figure 1c indicates that the proposed methods cannot identify the oscillation peak below a SNR of -15dB.

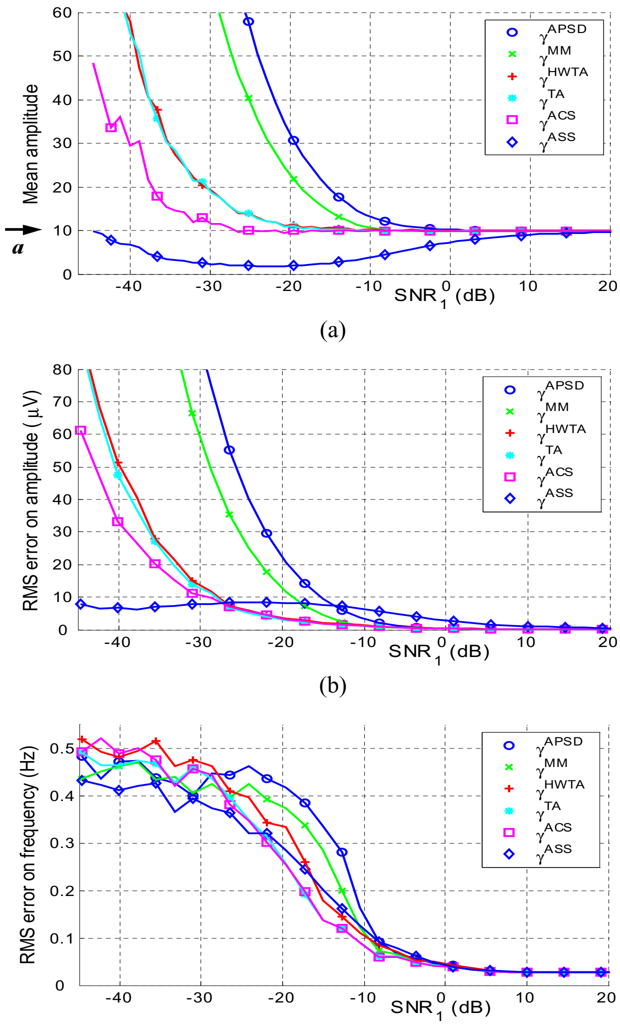

Figure 1: (a) Estimation of amplitude (b) RMS error on amplitude (c) RMS error on frequency of the oscillation, function of the SNR level

4.4 Phase estimation

To keep a sense to phase comparisons, the phase of an oscillation is evaluated at the true MF ($f_m = 4, 8, 16, 32, 64, 128$ Hz) even if the oscillation frequency is not exactly the same. Given an epoch, the phase can be estimated using the discrete Fourier transform of the observation

$$\widehat{\theta}^{(j)} = \arg\left(\frac{1}{q}\sum_{t=1}^{q} x^{(j)}(t)e^{-2i\pi\frac{f_m}{f_e}t}\right). \qquad (15)$$

The averaged phase may be deduced from all epochs by

$$\widehat{\theta}_E = \arg\left(\sum_{j=1}^{p} e^{i\widehat{\theta}^{(j)}}\right). \qquad (16)$$

In this expression, we assume that the amplitude of the oscillation does not vary between individual responses. Using the averaged signal (eq. 6) instead of $x^{(j)}(t)$ in eq. 15, we obtain another estimator

$$\widehat{\theta}_{TA} = \arg\left(\frac{1}{q}\sum_{t=1}^{q} \overline{x(t)}e^{-2i\pi\frac{f_m}{f_e}t}\right). \qquad (17)$$

With parameters $a^{(j)} = 10$, $f_0^{(j)} = 8$, $p = 50$, and 500 trials for each SNR level, the lowest STD is obtained for $\widehat{\theta}_{TA}$ (figure 2).

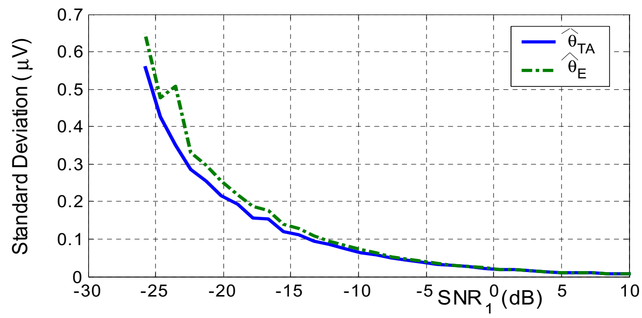

Figure 2: Phase estimation using eq. 17 (solid line) and eq. 16 (dash-dotted line), versus the SNR.

5. PHYSIOLOGICAL REPRODUCIBILITY OF THE RESPONSE

Considering the model in eq. 1, we analyze now the STD on the parameters A, F_0, and Φ observed on the epochs for each lead. This STD may have a physiological origin such as a reproducibility problem, and/or may be also caused by noise disturbances. Having now a PSD estimator for S ($\widehat{\gamma}_s^{ACS}$), and a noise PSD estimator (interpolated $\widehat{\gamma}_n$ PSD), we estimate the SNR for the SEEG oscillating activity for each lead

$$SNR_2 = 10\log_{10}\left(\frac{\widehat{\gamma}_s^{ACS}}{\widehat{\gamma}_n}\right). \qquad (18)$$

We use the methodology described in section 4.2 applied to eq. 3 to obtain the oscillation amplitude and frequency for an epoch. We obtain the phase for an epoch using eq. 15. Then, we compare the STD observed on simulated data and on real data for each parameter.

For simulated data, we evaluated the STD on 1000 epochs of 800 ms length using eq. 1 where $a^{(j)} = 10$, $\phi^{(j)} = 0$ and for $f_0^{(j)} = 4, 8, 16, 32, 64, 128$ Hz. In figures 3, 4, 5 results are shown for MF equal to 16 Hz and 64 Hz, but they are similar for other MF.

We note that the STD observed on real data may be explained entirely by the SNR level, for each parameter (amplitude, frequency and phase). This suggests that the theoretical model really observed for the oscillations produced by amplitude modulated noises is

$$x^{(j)}(t) = a.\sin\left(2\pi\frac{f_0}{f_e}t + \phi\right) + n^{(j)}(t) \qquad (19)$$

where a, f_0, and ϕ are practically constants.

6. CONCLUDING REMARKS

This preliminary study is a key-point before any analysis of the oscillation. The study on signal estimators validates the use of the averaged signal and shows that the technique based on cross-spectra is well adapted to SEEG data. The reproducibility of the physiological response (amplitude, frequency, phase) is acceptable, since the distribution STD of these three parameters only depends on the SNR level in SEEG activity.

ACKNOWLEDGMENTS

The authors wish to thank the Laboratory of Neurophysiology and Neuropsychology of the University of Aix-Marseille (France) for recording the database.

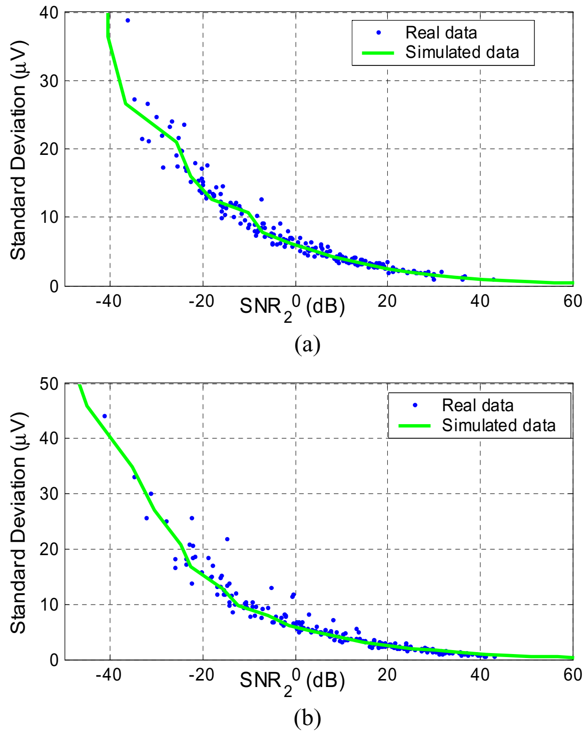

Figure 3: STD of amplitude estimation obtained on real data and simulated data for MF equal to (a) 16 Hz and (b) 64 Hz

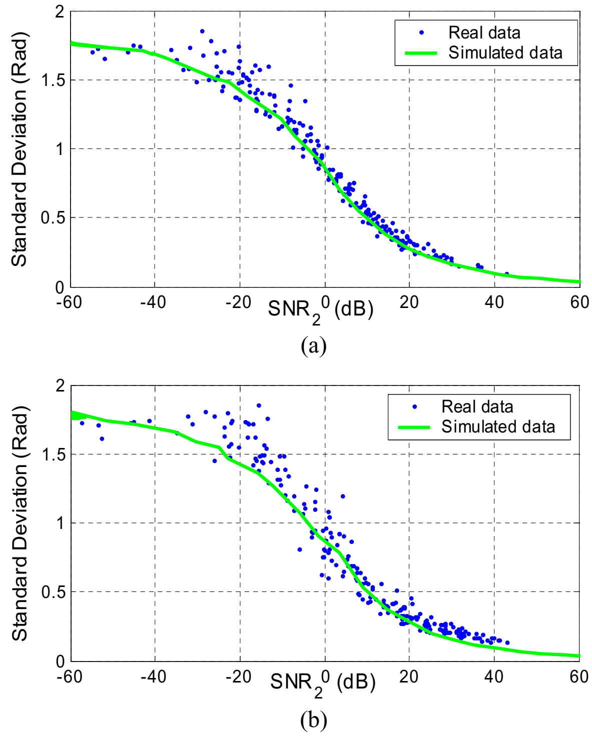

Figure 4: STD of phase estimation obtained on real data and simulated data for MF equal to (a) 16 Hz and (b) 64 Hz

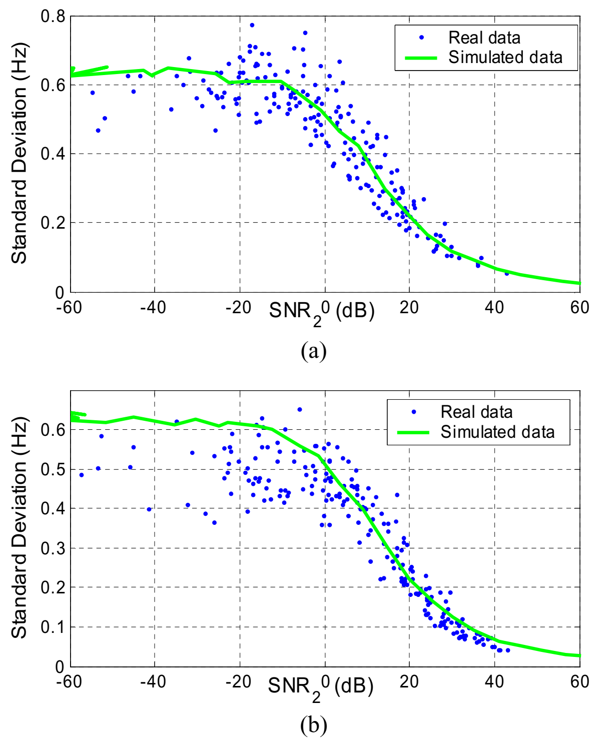

Figure 5: STD of frequency estimation obtained on real data and simulated data for MF equal to (a) 16 Hz and (b) 64 Hz

REFERENCES

[1] N.F. Viemeister, "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am., vol. 66, pp. 1364-1380, 1979.

[2] C. Pantev, M. Hoke, K. Lehnertz, B. Lutkenhoner, G. Anogianakis, and W. Wittkowski, "Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields," Electroencephalogr. Clin. Neurophysiol., vol. 69, pp. 160-170, 1988.

[3] T.W. Picton, "Human auditory steady state responses," in Evoked Potentials III, C. B. T. Barber, Ed. Boston : Butterworth, pp. 117-124, 1987.

[4] J. Bancaud, J. Talairach, A. Bonis, C. Schaub, G. Szikla, P. More, M. Bordas-Ferrer, "La stéréoélectroencéphalographie dans l'épilepsie: informations neurophysiopathologiques apportées par l'investigation fonctionnelle stéréotaxique," Ed. Paris: Masson, 1965.

[5] J. Talairach and P. Tournoux, "Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: an approach to cerebral imaging," Georg Thieme Verlag, 1988.

[6] C. Liégeois-Chauvel, J.B. de Graaf, V. Laguitton, P. Chauvel, "Temporal envelope processing in the human left and right auditory cortices," Cereb. Cortex, in press.

[7] S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. on ASSP, vol. 27, pp. 113-120, 1979.

[8] O. Cappé, "Techniques de réduction de bruit pour la restauration d'enregistrements musicaux," Ph.D. Thesis, Paris: Ecole Nationale Supérieure des Télécommunications, 1994.

[9] R.J. McAulay, M.L. Malpass, "Speech enhancement using a soft-decision noise suppression filter", IEEE Trans. on ASSP, vol. 28, pp. 137-145, 1980.