Title: Automatic quality assessment and peak identification of auditory brainstem responses with fitted parametric peaks

Author: Joaquin T. Valderrama Angel de la Torre Isaac Alvarez Jose Carlos Segura A. Roger. D. Thornton Manuel Sainz Jose Luis Vargas

PII: S0169-2607(14)00082-0
DOI: http://dx.doi.org/doi:10.1016/j.cmpb.2014.02.015
Reference: COMM 3772

To appear in: *Computer Methods and Programs in Biomedicine*

Received date: 8-8-2013
Revised date: 22-1-2014
Accepted date: 25-2-2014

Please cite this article as: Joaquin T. ValderramaAngel de la TorreIsaac AlvarezJose Carlos SeguraA. Roger. D. ThorntonManuel SainzJose Luis Vargas Automatic quality assessment and peak identification of auditory brainstem responses with fitted parametric peaks (2014), http://dx.doi.org/10.1016/j.cmpb.2014.02.015

Automatic quality assessment and peak identification of auditory brainstem responses with fitted parametric peaks

Author Listing:

- [names] Joaquin T. [surname] Valderrama [a,*]. jvalderrama@ugr.es.
- [name] Angel [surname] de la Torre [a]. atv@ugr.es.
- [name] Isaac [surname] Alvarez [a]. isamaru@ugr.es.
- [names] Jose Carlos [surname] Segura [a]. segura@ugr.es.
- [names] A. Roger. D. [surname] Thornton [b]. ardt@soton.ac.uk.
- [name] Manuel [surname] Sainz [c,d]. msainzquevedo@gmail.com
- [names] Jose Luis [surname] Vargas [c]. jlvf472@gmail.com

Affiliations:

* Corresponding author. Tel.: +34 958 240 840; fax: +44 958 240 831. E-mail address: jvalderrama@ugr.es (J. T. Valderrama).

[a] Department of Signal Theory, Telematics and Communications, CITIC-UGR, University of Granada, C/ Periodista Daniel Saucedo Aranda s/n, 18071, Granada, Spain.

[b] MRC Institute of Hearing Research, Southampton Outstation, Royal South Hants Hospital, Brintons Terrace, Mailpoint OAU, Southampton, Hampshire SO14 OYG, United Kingdom.

[c] ENT Service, San Cecilio University Hospital, Av. Dr. Oloriz 16, 18002, Granada, Spain.

[d] Department of Surgery and its Specialties, University of Granada, Av. De Madrid 11, 18012, Granada, Spain.

Abstract:

The recording of the auditory brainstem response (ABR) is used worldwide for hearing screening purposes. In this process, a precise estimation of the most relevant components is essential for an accurate interpretation of these signals. This evaluation is usually carried out subjectively by an audiologist. However, the use of automatic methods for this purpose is being encouraged nowadays in order to reduce human evaluation biases and ensure uniformity among test conditions, patients, and screening personnel. This article describes a new method that performs automatic quality assessment and identification of the peaks, the Fitted Parametric Peaks (*FPP*). This method is based on the use of synthesized peaks that are adjusted to the ABR response. The *FPP* is validated, on one hand, by an analysis of amplitudes and latencies measured manually by an audiologist and automatically by the *FPP* method in ABR signals recorded at different stimulation rates; and on the other hand, contrasting the performance of the *FPP* method with the automatic evaluation techniques based on the correlation coefficient, F_{SP}, and cross correlation with a predefined template waveform by comparing the automatic evaluations of the quality of these methods with subjective evaluations provided by 5 experienced evaluators on a set of ABR signals of different quality. The results of this study suggest (a) that the *FPP* method can be used to provide an accurate parameterization of the peaks in terms of amplitude, latency, and width, and (b) that the *FPP* remains as the method that best approaches the averaged subjective quality evaluation, as well as provides the best results in terms of sensitivity and specificity in ABR signals validation. The significance of these findings and the clinical value of the *FPP* method are highlighted on this paper.

Keywords: Fitted Parametric Peaks (*FPP*); auditory brainstem response (ABR); automatic quality assessment; response detection; F_{SP}; subjective visual evaluation; template matching; Mexican hat wavelet.

Text body:

1. INTRODUCTION

The auditory brainstem response (ABR) is the electrical activity of the auditory nerve generated in the brainstem associated with a stimulus [1]. The recording of the ABR has been extensively used in human and animal studies for both clinical and research purposes due to its noninvasive nature. The recording of this signal is commonly used in hospitals and clinics worldwide as a hearing screening tool, to detect the hearing threshold and to detect peripheral and central lesions. Furthermore, the analysis of the ABR may help understand the underlying mechanisms of the process of hearing [2]–[8]. The ABR comprises a number of waves that occur during the first 10 ms from stimulus onset [9]. These waves are indicated by sequential Roman numerals as originally proposed by Jewett and Williston [10]. Although up to seven peaks can be identified in the ABR, the most robust are III and V.

The quality of the responses is related to the probability that a response is present, which is usually associated with the amount of noise of the recording [11], [12]. The use of automatic methods for quality assessment and response detection of ABR signals may help improve the process of automatically stopping averaging, avoiding the recording of unnecessary sweeps when there already exists an ABR of sufficient quality and consequently, making a more efficient use of the recording time [13]-[15]. Furthermore, the automated identification of the peaks, i.e., amplitudes and latencies, is also a useful tool to provide an automatic interpretation of the ABR [16]. Additionally, automated methods eliminate the need for subjective interpretations of ABR, reduce human biases, and improve uniformity among test conditions, patients, and screening assistants [17]-[22]. These advantages promote the use of automated response detection in audiology screening in order to help the operator interpretation and decision making [23].

A number of methods have been proposed in automatic evaluation of ABR [11]. Some of them include the Raileigh test, Watson's U2 test, Kuiper's test, Hodges-Ajne's test, Cochran's Q-test, and Friedman test [24], [25]; automatic computer- assisted recognition of the pattern for ABR latency/intensity functions [26]; MASTER, a Windows-based data acquisition system designed to assess human hearing by recording auditory steady-state responses [27]; zero crossing method [28]; adaptive signal enhancement [29]; multifilters and attributed automaton [30]; single-trial covariance analysis [31]; and automatic analysis methods for peak identification based on a database of ABR signals from a large (>80) number of normal hearing subjects [32], [33]. Despite the large number of automatic evaluation techniques, few of them have been implemented in commercial devices [34]. The most common reported strategies of automated ABR analysis are the correlation coefficient and the F distribution based estimation of the signal to noise ratio (SNR) using a single point of the response (F_{SP}). The correlation coefficient procedure relies on the reproducibility of two consecutive ABR signals obtained in similar conditions to determine the presence or absence of the ABR [35]. F_{SP} provides an estimation of the response SNR evaluated from the distribution of amplitudes of a single point of the response for different sweeps. The power of noise is evaluated by matching the single point distribution of amplitudes with an *F* distribution, while the power of the signal is estimated from the averaged response [36].

This article describes a new method that performs an automatic evaluation of the quality of ABR signals and identification of the peaks based on the use of templates. We have called this method Fitted Parametric Peaks (*FPP*). The *FPP* method can be useful (a) to automatically parameterize the most relevant waves of ABR signals in terms of amplitude, latency, and width, and (b) to provide an automatic estimation of the quality of ABR signals based on the individual assessment of the quality of each wave. Preliminary results of this work were presented in [37].

The rest of the paper is organized as follows. Section II describes in detail the Fitted Parametric Peaks (*FPP*) method. In section III, the performance of the described method is assessed by two experiments. Experiment 1 compares the automatic parameterization of the peaks provided by the *FPP* method with a manual procedure performed by an audiologist in a number of ABR signals obtained at different stimulation rates. Experiment 2 compares the automatic quality assessment of the *FPP* method with the automatic quality evaluation techniques based on the correlation coefficient, F_{SP}, and cross correlation with a predefined template in terms of the grade of similarity to a subjective evaluation provided by a number of experts on ABR signals of different quality. Additionally, this experiment includes a comparative study of response validation in terms of sensitivity and specificity. Section IV presents a summary and a discussion of the results. Finally, section V highlights the significance and the main contributions of this article.

2. DESCRIPTION OF THE METHOD

The most usual approach for assessing the quality of ABR signals is based in subjective evaluations provided by audiologists. However, it is well known that subjective evaluations may differ from one evaluator to another [33], [38], [39]. This bias represents a problem that could be solved using automatic quality evaluation techniques [17]-[23]. This section describes the Fitted Parametric Peaks (*FPP*) method, a new technique that provides an automatic evaluation of the quality of ABR signals and parameterization of the peaks in terms of amplitude (A), latency (L), and width (W).

2.1. Fitted Parametric Peaks

The approach of this method is based on the use of templates that fit the peaks of the ABR. The use of templates for this purpose was first proposed in [40], in which the ABR used for test is cross correlated with a template used as reference. The major disadvantage of this technique is that it requires the compilation of a database of templates corresponding to each stimulation settings (e.g., level, rate, polarity, etc.). In contrast, the *FPP* does not require the use of a database since it uses as template a parametric function. The motivation of the *FPP* quality assessment procedure relies on the subjective criterion usually applied by audiologists for the evaluation of ABR. The most persistent peaks are usually waves III and V, and therefore, an ABR response can be assumed to be valid if at least these two peaks can be identified with reasonable amplitudes at the latencies expected for these waves. Thus, the *FPP* procedure fits a parametric function modeling a peak for both waves (III and V) and evaluates the quality taking into account the similarity of the ABR signal and the fitted parametric peaks. The parametric function is given by:

$$x(t,A,L,W) = A \cdot K_0 \cdot \left(1 - \frac{(t-L)^2}{W^2}\right) \cdot \exp\left(\frac{-(t-L)^2}{2 \cdot W^2}\right)$$

This parametric function is generally known as *Mexican hat wavelet*, and corresponds (except for the sign and normalization constant) to the second derivative of a Gaussian function with mean L and standard deviation W. K_0 is a constant that makes x(t,A,L,W) have a peak-to-peak amplitude equal to A. The value of K_0 that fits this criterion is:

$$K_0 = 1 + 2 \cdot \exp(-3/2) = 1.446260320296860$$

According to the definition of the parametric function x(t,A,L,W), A is the peak-to-peak amplitude of the wave, L is the latency, and W is the semi width. Figure 1 shows an ABR signal and the parametric peak that fits wave V. The search of the parameters that define the fitted parametric peak would involve a 3 dimensional search (for A, L, and W). However, this process can be computationally optimized to a 1 dimensional search of the width. The optimal latency (L_0) and amplitude (A_0) of the fitted parametric peak can be directly estimated for each tested width parameter (W_{test}). The latency is calculated by cross correlation of the ABR signal with the parametric peak of a specific width. This step is independent of the amplitude of the parametric peak. The search of the optimal latency is performed in an interval around a referenced latency. This referenced latency can be obtained from related literature and will depend on the stimulation settings, e.g., intensity level and stimulation rate. The interval in which the optimal latency is searched must be wide enough to consider the normal variations of latencies among subjects, but at the same time, it must be narrow enough to avoid including adjacent waves. An interval of about 3 ms was found to be appropriate for this purpose. Given the width W_{test} and the latency L_0, the amplitude A_0 is directly estimated by projecting the response **y**(t) onto the parametric peak **x**(t,1,L_0,W_{test}), taking into account the properties of the scalar product of sampled signals:

$$\mathbf{x}_1(t)\cdot\mathbf{x}_2(t)=\sum_{n=0}^{N}x_1(t_n)\cdot x_2(t_n)$$

$$\left\|\mathbf{x}_1(t)\right\|^2=\mathbf{x}_1(t)\cdot\mathbf{x}_2(t)$$

With these definitions, the projection of an ABR signal **y**(t) onto the parametric function with latency L_0 and width W_{test} can be calculated using the associated unitary vector:

$$\mathbf{u}_x(t,L_0,W_{test})=\frac{\mathbf{x}(t,1,L_0,W_{test})}{\left\|\mathbf{x}(t,1,L_0,W_{test})\right\|}$$

$$\left(\mathbf{y}(t)\cdot\mathbf{u}_x(t,L_0,W_{test})\right)\cdot\mathbf{u}_x(t,L_0,W_{test})=\frac{\mathbf{y}(t)\cdot\mathbf{x}(t,1,L_0,W_{test})}{\left\|\mathbf{x}(t,1,L_0,W_{test})\right\|}\cdot\frac{\mathbf{x}(t,1,L_0,W_{test})}{\left\|\mathbf{x}(t,1,L_0,W_{test})\right\|}=\ldots$$

$$\ldots=\frac{\mathbf{y}(t)\cdot\mathbf{x}(t,1,L_0,W_{test})}{\left\|\mathbf{x}(t,1,L_0,W_{test})\right\|^2}\cdot\mathbf{x}(t,1,L_0,W_{test})=A_0\cdot\mathbf{x}(t,1,L_0,W_{test})=\mathbf{x}(t,A_0,L_0,W_{test})$$

and therefore, the amplitude can be directly computed as:

$$A_0=\frac{\mathbf{y}(t)\cdot\mathbf{x}(t,1,L_0,W_{test})}{\left\|\mathbf{x}(t,1,L_0,W_{test})\right\|^2}$$

Taking into account that the fitting is performed around each wave (i.e., around wave III or wave V), the computation of the scalar product must be restricted to an interval around the latency L_0. An interval of 2 ms can be appropriate since this interval is related to the duration of the peak. Since the latency L_0 and the amplitude A_0 are directly estimated for each width W_{test} (by cross correlation and vector projection respectively) as those providing the best fitting of the parametric function given the ABR signal and the tested width W_{test}, each width can be evaluated considering the energy of the error between the ABR signal and the parametric peak evaluated over the interval:

$$\mathbf{e} = \sum_{t_n \geq L_0 - 1ms}^{t_n \leq L_0 + 1ms} \left(y(t_n) - x(t_n, A_0, L_0, W_{test}) \right)^2 = \left\| \mathbf{y}(t) - \mathbf{x}(t, A_0, L_0, W_{test}) \right\|^2$$

and therefore, the width W_{test} of the parametric function that best fits the peak (W_{peak}) is that one minimizing the energy of the error. The optimal values of the latency L_{peak} and amplitude A_{peak} would be the corresponding L_0 and A_0 of the W_{peak}.

A signal-to-noise ratio associated to each peak can be derived from this fitting as the ratio between the energy of the parametric peak and the energy of the error (that can be assumed to be noise):

$$SNR_{peak} = \frac{\left\| \mathbf{x}(t, A_{peak}, L_{peak}, W_{peak}) \right\|^2}{\left\| \mathbf{y}(t) - \mathbf{x}(t, A_{peak}, L_{peak}, W_{peak}) \right\|^2}$$

that can also be expressed in dB:

$$SNR_{peak}(dB) = 10 \cdot \log_{10}\left(SNR_{peak} \right)$$

The SNR can be used to evaluate the quality for each wave. Finally, a global quality parameter can be defined as the minimum SNR for waves III and V.

$$Q_{FPP}(dB) = \min\left\{ SNR_{III}(dB), SNR_V(dB) \right\}$$

The *FPP* method could implement an automated response detection paradigm considering (a) whether or not the values of amplitude, width and latency of the parametric peaks are consistent with literature, and (b) if the global quality parameter (Q_{FPP}) exceeds a given threshold. This threshold level represents the minimum quality required for considering a recording as a valid ABR signal.

The software routines that implement the *FPP* method are available in MATLAB[1] and GNU Octave[2] codes as supplementary material (section A).

3. ASSESSMENT OF THE METHOD

The *FPP* method is validated in this study with two experiments. Experiment 1 evaluates the performance of the *FPP* method through a comparison of the latencies and amplitudes of waves III and V measured manually by an audiologist and automatically by the *FPP* method in a number of ABR signals obtained at different stimulation rates. In experiment 2, the performance of the automatic quality evaluation techniques based on the *FPP*, correlation coefficient (*r*), F_{SP}, and cross correlation with a predefined template function (*Cross Corr*), is contrasted (a) with a subjective evaluation provided by 5 experts in a set of ABR signals of different quality, and (b) with a response validation study in terms of sensitivity and specificity. This section gives details about the EEG recording protocol followed on the recording process of the ABR signals and presents the results of both experiments.

---

[1] The Mathworks, Inc., Natick, MA

[2] John W. Eaton, University of Wisconsin, Madison, WI

3.1. EEG recording and signal processing

The procedure for EEG recording consisted on the presentation of auditory stimuli to the subjects and the recording of their associated electrical response (sweep). The stimulation of the auditory system was performed by 0.1 ms duration clicks in condensation polarity in order to evoke a synchronous firing of a large number of neurons [1]. The recording sessions took place in a shielded screening booth in order to minimize the effects of electromagnetic interference. The subjects were seated comfortably to reduce the myogenic noise. The intensity level 0 dBnHL was established considering the threshold level (intensity level at which stimuli are just detectable) in a group of 15 subjects (9 male and 6 female) aged from 24 to 31 yr, with no self-reported history of auditory dysfunction (normal hearing subjects). The intensity level used to obtain the ABR signals in this study was 70 dBnHL, which corresponds to 103.54 dB peak equivalent sound pressure level (dBpeSPL). The calibration of the intensity level was performed using an Artificial Ear Type 4153[3]. The EEGs were recorded by Ag/AgCl surface electrodes placed on the skin at different positions of the head. Active, ground, and reference electrodes were situated at the high forehead, low forehead, and ipsilateral mastoid respectively. The interelectrode impedances were always below 10 kΩ at the working frequencies. The recorded EEG was 70 dB amplified and bandpass filtered (100 to 3500 Hz). This signal was sampled at 25 kHz and stored using 16 bits of quantization. Digital signals were processed with algorithms implemented in MATLAB. The *FPP* method was implemented in this study using the referenced values of latency shown in table 1 and an interval of SNR assessment of 2 ms.

---

[3] Brüel & Kjær Sound & Vibration Measurement A/S, Nærum, Denmark

Table 1 shows the latency for waves III and V at different stimulation rates evoked at an intensity level of 70 dBnHL based on the data published [41]-[43]. All subjects explored in this study were volunteers and were informed in detail about the experimental protocol. A consent form was signed by the participants before the beginning of the session, which was carried out at the University of Granada (Granada, Spain) in accordance with The Code of Ethics of the World Medical Association (Helsinki Declaration of 1975, revised 2000) for experiments involving humans. This recording procedure was approved by the Human Research Ethics Committee of the University of Granada and by the Clinical Research Ethics Committee of the San Cecilio University Hospital. An expanded description of the ABR recording system used in this study can be found in [44].

3.2. Experiment 1

3.2.1. Subjects and methods

The performance of the *FPP* method to automatically parameterize the most relevant waves of ABR signals is assessed on this first experiment by a comparison of the latencies and amplitudes of waves III and V measured manually by an audiologist and automatically by the *FPP* method in a set of ABR signals obtained from 8 normal hearing subjects (7 males and 1 female; aged between 26 and 35 yr) at the stimulation rates 45, 55, 83, 100, 125, 167, and 250 Hz using the randomized stimulation and averaging technique (*RSA*). The *RSA* technique allows the recording of ABR signals at high stimulation rates using jittered stimuli [43]. The jitter of a stimulation sequence measures the amount of dispersion of the interstimulus interval in contrast to a periodical presentation of stimuli. The stimulation sequences used in this study were generated using a jitter of 4 ms. 5 recordings of 4000 sweeps were recorded from each subject at each stimulation rate, therefore, the number of ABR signals used in this study was 320 (8 subjects, 8 stimulation rates, 5 recordings). Latencies were measured manually as the difference in milliseconds between the stimulus onset and the top of the peak, and amplitudes were measured in microvolts as the difference between the top of the peak and the following trough [1], [9]. The latencies and amplitudes measured manually by an audiologist and automatically by the *FPP* method were adjusted to a 3rd order polynomial. The coefficient of determination (R^2) was calculated for each distribution. In addition to the analysis based on the polynomial fitting using the raw data (i.e., estimated amplitudes and latencies), a similar analysis was performed using normalized data. Normalization consisted of subtracting the mean value for each subject and adding the global mean in order to decrease the inter-subject variability. The values of amplitudes and latencies that did not accomplish minimum criteria to be considered as valid waves were excluded from the analysis. The criteria used in this study as threshold to detect auditory waves was $SNR_{peak} \geq 2dB$ and $A_{peak} \geq 0.05\mu V$.

3.2.2. Results

Figure 2 shows the values of the latencies and amplitudes of waves III and V measured manually (*MAN*) by an audiologist and automatically by the fitted parametric peaks method (*FPP*) in a set of 320 ABR signals from 8 normal hearing subjects at different stimulation rates. The experimental data were adjusted to a 3rd order polynomial to analyze the behavior of these parameters with the stimulation rate. This analysis shows (a) that the latency of the peaks increases as the stimulation rate increases, with a deeper shift on wave V than on wave III, and (b) that the amplitude of both peaks decreases as stimulation rate increases. These effects are a normal phenomenon, consequence of neural adaptation [8], [45], [46]. The low values of the coefficients of determination (R^2) on the parameters analyzed in this study for waves III and V (figure 2, upper panel), especially in amplitudes, are due to the great intersubject variability. The coefficients of determination increase significantly after normalization. The analysis of the data after normalization (figure 2, lower panel) points out that the coefficients of determination of the latencies and amplitudes of waves III and V are greater when the parameters are estimated automatically by the *FPP* method than when the values are measured manually. Figure 3 shows a comparative analysis of the latencies and amplitudes of waves III and V estimated manually by an audiologist and automatically by the *FPP* method with the same set of ABR signals. First of all, this figure shows that all waves III and V were correctly identified by the *FPP* method. In addition, the linear regression analysis adjusted to the experimental data points out (a) that the automatic parameterization of the peaks by *FPP* in terms of latency and amplitude is strongly related with the manual procedure ($r>0.9$ in all measures), (b) that latencies estimated by *FPP* are accurate since the linear regression curves are close to the curves *FPP=MAN* (dotted line), and (c) that a slight bias exists between the amplitudes measured manually and automatically by *FPP*, possibly as a consequence of local noise, which systematically provokes an overestimation of amplitudes by the manual method. Figure 4 shows examples of ABR signals used in this experiment from 5 subjects at different stimulation rates. The parametric peaks adjusted to the waves III and V are highlighted on this figure. In addition, this figure includes the SNR associated with each peak evaluated automatically by the *FPP* method. Table 2 presents the mean and standard deviation of the latencies, amplitudes, widths, and SNRs measured automatically by the *FPP* method on the waves III and V. This table shows the tendency of the parameters as stimulation rate increases: latencies increase, the interpeak latency between waves III and V increases because the shift of wave V is greater than in wave III, the amplitudes of both waves decrease, the widths increase in both waves, possibly as a consequence of neural desynchronization [47], and the SNRs of both waves tend to decrease due to the lower amplitude of the waves. Table 3 presents the mean and standard deviation of the latencies and amplitudes measured manually on waves III and V. The analysis of tables II and III shows that, on average, there are similarities between the values measured manually and automatically by the *FPP* method on the latencies of waves III and V, and on the amplitude of wave III. Regarding the amplitude of wave V, there is a systematic difference of a few tens of nanovolts on the values measured manually and automatically by the *FPP* method. This difference might arise because the trough that follows wave V does not fit perfectly the template. Nonetheless, the values of the latencies, amplitudes, and widths shown on both tables II and III are consistent with those reported in previous studies [4], [48]–[52].

3.3. Experiment 2

3.3.1. Subjects and methods

In this second experiment, the performance of the automatic quality assessment based on the *FPP* method is compared to the automatic quality evaluation techniques based on the correlation coefficient (*r*), the F_{SP}, and the cross correlation with a predefined template method (*Cross Corr*). The ABR signals used in this test consisted of 500 recordings from 10 normal hearing subjects (6 males and 4 females; aged between 21 and 37 yr). Each recording was obtained with auditory stimuli periodically presented at a rate of 30 Hz, at a different number of averaged sweeps (100, 300, 900, 1800, and 9500). From these 500 recordings, 40 recordings were obtained without auditory stimulation, so no ABR could be detected.

The correlation coefficient (*r*) analysis was performed on the interval [1 , 10] ms to minimize the effect of the recorded artifacts synchronized with the stimulus. The single point (SP) chosen for the implementation of the F_{SP} method was the sample 100 (corresponding to the 4th ms of the averaging window, considering f_s = 25 kHz). The template waveform used on the *Cross Corr* method was built from ABR signals recorded on 30 normal hearing subjects (17 males and 13 females; aged between 17 and 34 yr) in the same recording conditions as the test signals, using 2000 averaged sweeps. The template waveform used in the *Cross Corr* method is available as supplementary material (section B). These subjects were different from those analyzed to obtain the ABR signals used for test. Each ABR signal used to build the template waveform was normalized in amplitude according to its RMS value, cosined-tapered with a band pass window of [1 , 8] ms, and scaled in amplitude producing an RMS value equal to the mean of the RMS values of the original recordings. The mean of these signals produced the template waveform used in the *Cross Corr* method. Further details of the implementation of the methods based on the correlation coefficient (*r*), on the F_{SP}, and on the *Cross Corr* can be found, respectively, in [35], [36], and [40].

The results obtained with the automatic methods were compared to a subjective evaluation provided by 5 experts. Each expert had at least 3 years of expertise in the analysis of ABR signals. The experts were asked to rate the quality of a number of ABR signals according to the following criteria: Q = 0, no ABR is observed (no auditory response); Q = 1, wave V can be hardly detected (highly noisy ABR); Q = 2, wave V can be detected but the rest of waves are unclear (noisy ABR); Q = 3, waves III and V can be clearly detected (ABR slightly noisy); Q = 4, waves I, III, and V can be detected (good quality ABR); and Q = 5, all components of the ABR can be easily detected (excellent quality ABR). A computer application was programmed to present the test ABR signals to the evaluators and ask for the subjective quality. For each level of quality, two ABR signals were presented to the evaluator as reference. The presentation order of the ABR signals was randomized for each test. Figure 5 shows a screenshot of the computer application for subjective evaluation.

This experiment also includes a response validation study carried out by the aforementioned automated methods in terms of sensitivity and specificity with the same set of ABR signals. The validation of responses by the automated methods was implemented considering a threshold level of quality, which varied in all methods from their lowest estimation of the quality to its greatest value. Automatic evaluations greater or equal to such threshold would be a “positive”, and they would be a “negative” otherwise. These automatic “positive” and “negative” evaluations were compared to an objective decision of response validation. This objective decision was made considering the averaged subjective evaluations of the experts greater or equal to 2, which corresponds with the detection of at least the wave V. The sensitivity and specificity parameters for each automated method were estimated at different acceptance thresholds as the true positive rate (TPR: true positives divided by all positives) and as 1 – false positive rate (FPR: false positives divided by all negatives) respectively.

3.3.2. Results

Some examples of ABR signals used for this experiment, including their associated quality evaluation provided by the automatic and subjective methods, are shown in figure 6 and table 4. In this table, *FPP* is expressed in dB, *r* is in the range [−1 , 1], F_{SP} is in absolute value, *Cross Corr* is in the range [-1 , 1], and subjective evaluations in the range [0 , 5]. Signals K and L were obtained without any auditory stimuli, thus no ABR can be detected. Figure 7.A represents the regression analysis between the subjective evaluations provided by 5 experts and the automatic quality assessment technique based on *FPP*. The linear regression analysis for each individual subjective evaluation compared to the *FPP* method is shown in the figure. The correlation coefficient for the regression analysis that considers all subjective evaluations (r = 0.72) is lower in comparison with the mean of the correlation coefficient for the individual evaluations, which suggests that there exists a bias among the evaluations of the experts. On the other hand, the correlation coefficient increases significantly on the regression analysis that considers the average of the subjective evaluations (r = 0.84, figure 7.B), which remarks that the model is better described with the averaging of a number of individual subjective evaluations. The correlation coefficient for the rest of the automatic methods compared to the averaged subjective evaluations is r = 0.78 for the evaluation based on the correlation coefficient, r = 0.77 for the evaluation based on the F_{SP} expressed in dB, and r = 0.74 for the evaluation based on the cross correlation with a predefined template waveform. The linear regression analysis between the averaged subjective evaluation and the automatic methods based on the correlation coefficient (*r*), the F_{SP}, and the cross correlation method with a predefined template (*Cross Corr*) is available as supplementary material (section C).

Figure 8 shows the receiver operating characteristics (ROC) space of a response validation study defined by the false positive rate (FPR), or 1-specificity, and the true positive rate (TPR), or sensitivity, for the automated response validation methods based on fitted parametric peaks (*FPP*), on the correlation coefficient (*r*), on the F_{SP}, and on the cross correlation with a predefined template waveform (*Cross Corr*). This figure shows that the *FPP* method presents the best results determining the existence of response for all evaluated thresholds, in exception for the thresholds corresponding to FPR evaluations lower than 0.006. The advantage of *FPP* with the other methods is especially remarkable for low FPR evaluations (lower than 0.1). The F_{SP} method presents better performance than the *r* and *Cross Corr* methods for most of the evaluated thresholds. For FPR evaluations greater than 0.55, the performances of the *r*, F_{SP}, and *Cross Corr* methods are very similar.

4. DISCUSSION

This paper describes in detail and evaluates the Fitted Parametric Peaks (*FPP*) method, a new approach of automatic quality assessment and peak parameterization based on the use of templates. The use of templates for this purpose was first proposed by C. Elberling in [40]. In his work, a cross correlation method between the ABR signal used for test and a template waveform is described. This method has the limitation of requiring a database of predefined templates for each recording condition, and while a significant match may signify a response, lack of a match do not necessarily means that no response is present, since a response could exist but not match the template [11], [40]. Another similar template-matching detection algorithm was commercially implemented in the *Algo-1 automated evoked response infant hearing screener*[4] (and successive versions). This detection algorithm is based on the weighting of a number of points in a template waveform according to their relative contribution in identifying a response, and evaluating a test signal in terms of likelihood ratio [53]. Clinical studies carried out by the *Algo-1 screener* show evidences of a high performance in screening applications [19], [53]-[55]. The approach of the *FPP* method consists of the search of the latency, width, and amplitude of a parametric peak, similar in morphology to an ABR wave that best fits the most robust waves of the ABR, waves III and V. The parametric peak waveform used as template in the *FPP* method is commonly known as *Mexican hat wavelet*, which has been successfully used in different applications of related fields, e.g., [56], [57]. The search of the parameters of the fitted peak is computationally optimized to a 1-Dimensional search on the width. The optimal latency and amplitude of the parametric peak are directly estimated for a given width. The *FPP* method described in this paper provides an automatic evaluation of the quality of ABR signals, and parameterizes the most robust waves in terms of amplitude, latency, and width.

The performance of the *FPP* method was evaluated in this study by two experiments. In the first experiment, the latencies and amplitudes of waves III and V were estimated manually by an audiologist and automatically by the *FPP*

---

[4] Natus Medical Incorporated, San Carlos, CA

method in ABR signals obtained from 8 normal hearing subjects at different stimulation rates. This analysis shows that the *FPP* method successfully identified all waves III and V. Additionally, the models for latencies and amplitudes of waves III and V as stimulation rate increases are better described when the values are estimated by the *FPP* method than manually (R^2_{FPP} > R^2_{MAN} in all parameters), which suggests that the *FPP* method provides more consistent results than the manual procedure, possibly due to the fact that the *FPP* method bases the estimation of the parameters considering an interval of the response, rather than isolated samples, which makes the *FPP* method less sensitive to noise. In addition, the results of this experiment show that, despite the difference of a few tens of nanovolts on the estimation of the amplitude of wave V, the *FPP* method provides an accurate automatic measure of the latencies, amplitudes, and widths of waves III and V, consistent with previous studies. In the second experiment, the performance of *FPP* was contrasted with the most common automatic quality evaluation procedures: the correlation coefficient (*r*) [35], the F_{SP} [36], and the cross correlation with a predefined template waveform (*Cross Corr*) [40]. These automatic quality evaluation methods were compared to a subjective evaluation provided by 5 experts. The results of this test revealed that although all automatic methods present high correlation coefficients with the averaged subjective assessment, the *FPP* remains as the method that best approaches an averaged subjective evaluation. Comparing the reliability of the visual judgments provided by the 5 experts, this test shows, on one hand, that the correlation coefficient is lower when all evaluations are considered in comparison to individual evaluations, and on the other hand, that the correlation coefficient is greater when considering an averaged subjective evaluation. These results suggest that there is an important bias among the evaluators. All individual evaluations present a similar behavior, but a different scale, which evidences that the reproducibility of visual judgments is not high. This conclusion is in accordance with previous studies [33], [38], [39], and reveals the convenience of using automatic methods. In comparison with the subjective approach, automatic quality assessment methods are uniform, consistent worldwide, and eliminate human inaccuracies. In addition to this, the objective comparison of the aforementioned automated methods in validating ABR signals (figure 8) shows that the *FPP* method presents the best results in most of the thresholds analyzed in the study.

The advantages of *FPP* in research applications are numerous. For instance, the automatic parameterization of the peaks could replace the manual labeling of waves in clinical reports, a tedious task which is usually omitted by the clinical personnel [1]. Furthermore, this functionality could be valuable to provide an automatic ABR interpretation based on response tracking (i.e., analyzing the changes on the morphology of the auditory responses according to a gradual modification of any stimulation setting, such as the intensity level or the stimulation rate). An accurate automatic ABR interpretation might have a significant clinical benefit by helping audiologists on the human decision making [17]-[23]. The online quality assessment and parameterization of the peaks carried out by *FPP* could also be appropriate in many real time clinical applications, such as the on-going evaluation of the recorded signal to automatically stopping averaging, thus eliminating unnecessary recording time [13], [14]. In addition to this, the automatic evaluation of the quality of ABR signals could be useful to carry out objective comparisons between the performances of different stimulation methods (RSA [43], QSD [58], CLAD [59], [60], etc.) and the effectiveness of different artifact rejection techniques.

The *FPP* method is not defined for clinical applications, such as screening or diagnosing. Screening and diagnosing systems, like the *Algo-1 infant hearing screener* (Natus Medical Incorporated, San Carlos, CA) [53], are designed to detect waveform abnormalities in very specific recording settings, i.e., nature of the stimuli (clicks, chirps, windowed tones...), polarity, level, rate, hardware equipment, calibration, recording procedure, etc. In screening and diagnosing applications, all parameters involved in the recording process are protocoled and closed, in exception of the subjects. Therefore, screening and diagnosing systems are useful classifying subjects as “normal” (pass) or “pathologic” (fail). The definition of the “pass” criterion requires a strictly protocoled recording procedure (recording system, stimulation and recording settings, etc.) and a clinical study with a large database of explored normal and pathologic subjects. In contrast to these systems, *FPP* can be used in a wide range of scenarios because it adapts to the normal fluctuations in amplitude, latency, width, and morphology among subjects and recording conditions. These features are appropriate in many research applications.

The automatic quality evaluation methods based on the correlation coefficient, the F_{SP}, and *FPP* present different approaches. First, the correlation coefficient bases the evaluation of the quality on the grade of reproducibility of two consecutive signals. A high positive correlation coefficient would indicate a high quality ABR if both signals are recorded in similar conditions [61]. This method presents the limitation that requires a second ABR signal to perform the test, which doubles the recording time. Additionally, a strong artifact synchronized with the stimulus would lead to an inaccurately high evaluation of the quality. The F_{SP} method bases the evaluation of the quality on the power of the averaged signal and the power of noise across sweeps. This technique requires the evaluation of all recorded sweeps, thus this method cannot be implemented offline unless the EEG is stored (or at least the single point of each sweep). In addition, this technique may present a lack of reliability when evaluating a signal that could not be a response. For instance, this technique would provide a high evaluation index when the ABR is affected by a strong artifact synchronized with the stimulus. Finally, the *FPP* method approaches the perspective of expert subjective evaluators, rating the grade of identification and quality of the most important waves, does not require the access to the EEG, and provides information regarding the parameterization of the peaks. We believe that since the correlation coefficient method measures the reproducibility of the response, the F_{SP} method measures the level of noise of the recording, and the *FPP* method evaluates the existence of ABR waves, the use of a combination of all these automatic methods could improve significantly the accuracy in automatic evaluations and provide a better automatic interpretation of ABR signals.

Future research could include the search of appropriate template functions that fit the waves of other auditory evoked potentials, such as compound action potentials (CAPs), middle latency responses (MLRs), or late latency responses (LLRs) using the approach of *FPP*.

5. CONCLUSION AND SIGNIFICANCE

A novel automatic method for quality assessment and peak identification of ABR signals, the Fitted Parametric Peaks (*FPP*), is described and evaluated in this article. The approach of *FPP* opens a new paradigm in template-matching algorithms, avoiding the need of a database of templates and including additional information regarding the most relevant components of ABR signals. The computational efficiency of the *FPP* method could be appropriate for its implementation in real time processing applications. The results presented in this article suggest that *FPP* method presents a high level of accuracy identifying the most important waves of the ABR, and estimating their latency, amplitude, and width. The measure of these parameters with the *FPP* method seems to be less sensitive to noise than the manual procedure because it considers an interval of the response rather than isolated samples. The automatic identification of the peaks could facilitate the wave labeling process and could be useful to provide an automatic ABR interpretation, with a significant clinical value by helping the operator with the decision making. In comparison with the automatic evaluation techniques based on the correlation coefficient (*r*), on F_{SP}, and on the cross correlation with a predefined template waveform (*Cross Corr*), the *FPP* remains as the method (a) that best approaches a subjective evaluation of the quality, and (b) that provides the best results in the validation of ABR signals in most of the analyzed thresholds. This study has also shown that the subjective evaluations provided by different experts were biased among evaluators, i.e., all evaluators had the same criteria but their scales of assessment were different. This bias can be a problem for the reliability of a subjective evaluation, especially when the evaluator is not an expert. The use of the automatic *FPP* method described in this paper could be valuable in this context.

6. CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this research work.

ACKNOWLEDGMENT

The authors gratefully acknowledge the participation of the voluntary subjects and evaluators involved in this study. This research is granted by the project "Design, implementation and evaluation of an advanced system for recording Auditory Brainstem Response (ABR) based in encoded signalling" (TEC2009-14245), R&D National Plan (2008-2011), Ministry of Economy and Competivity (Government of Spain) and "European Regional Development fund Programme" (2007-2013); and by the grant "University Professor Training Program" (FPU, AP2009-3150), Ministry of Education, Culture, and Sports (Government of Spain).

APPENDIX

Supplementary data associated with this article can be found, in the online version, at [URL].

REFERENCES

[1] J. W. Hall, New Handbook of Auditory Evoked Responses, (Pearson: Allyn and Bacon, Boston MA, 2007).

[2] C. D'Avanzo, A. Goljahani, G. Pillonetto, G. Nicolao, G. Sparacino, A multi-task learning approach for the extraction of single-trial evoked

potentials, Computer Methods and Programs in Biomedicine 110 (2) (2013) 125-136.

[3] A. R. D. Thornton, A. Slaven, Auditory brainstem responses recorded at fast stimulation rates using maximum length sequences, British Journal of Audiology 27 (3) (1993) 205-210.

[4] S. M. Leung, A. Slaven, A. R. D. Thornton, G. J. Brickley, The use of high stimulus rate auditory brainstem responses in the estimation of hearing threshold, Hearing Research 123 (1-2) (1998) 201-205.

[5] H. Zabala-Fernandez, R. Orglmeister, L. Trahms, T. H. Sander, Identification enhancement of auditory evoked potentials in EEG by epoch concatenation and temporal decorrelation, Computer Methods and Programs in Biomedicine 108 (3) (2012) 1097-1105.

[6] A. Erenberg, J. Lemons, C. Sia, D. Tunkel, P. Ziring, M. Adams, J. Holstrum, M. McPherson, N. Paneth, B. Strickland, Newborn and infant hearing loss: Detection and intervention, Pediatrics 103 (2) (1999) 527-530.

[7] Z. D. Jiang, D. M. Brosi, X. M. Shao, A. R. Wilkinson, Maximum length sequence brainstem auditory evoked responses in term neonates who have perinatal hypoxia ischemia, Pediatric Research 48 (5) (2000) 639-645.

[8] J. T. Valderrama, A. de la Torre, I. Alvarez, J. C. Segura, A. R. D. Thornton, M. Sainz, J. L. Vargas, A study of adaptation mechanisms based on ABR recorded at high stimulation rate, Clinical Neurophysiology (2013), http://dx.doi.org/10.1016/j.clinph.2013.06.190.

[9] R. F. Burkard, M. Don, The Auditory Brainstem Response, in Auditory Evoked Potentials. Basic principles and clinical application, Chap. 11 (Lippincott Williams & Wilkins, Baltimore MD, 2007).

[10] D. L. Jewett, J. S. Williston, Auditory-evoked far fields averaged from the scalp of humans, Brain 94 (4) (1971) 681-696.

[11] C. Elberling, M. Don, Detecting and assessing synchronous neural activity in the temporal domain (SNR, response detection), in Auditory Evoked Potentials. Basic principles and clinical application, Chap. 5 (Lippincott Williams & Wilkins, Baltimore MD, 2007).

[12] E. Sejdic, L. A. Lipsitz, Necessity of noise in physiology and medicine, Computer Methods and Programs in Biomedicine 111 (2) (2013) 459-470.

[13] M. Don, C. Elberling, Use of quantitative measures of auditory brain-stem response peak amplitude and residual background noise in the decision to stop averaging, Journal of the Acoustical Society of America 99 (1) (1996) 491-499.

[14] O. Ozdamar, R. E. Delgado, Measurement of signal and noise characteristics in ongoing auditory brainstem response averaging, Annals of Biomedical Engineering 24 (6) (1996) 702-715.

[15] G. Sparacino, S. Milani, E. Arslan, C. Cobelli, A Bayesian approach to estimate evoked potentials, Computer Methods and Programs in Biomedicine 68 (3) (2002) 233-248.

[16] R. E. Delgado, O. Ozdamar, Automated auditory brainstem response interpretation, IEEE Engineering in Medicine and Biology Magazine 13 (2) (1994) 227-237.

[17] G. G. Gentiletti-Faenze, O. Yañez-Suarez, J. M. Cornejo-Cruz, Evaluation of automatic identification algorithms for auditory brainstem response used in universal hearing loss screening, in Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings 3, Cancun, Mexico, 2003, pp. 2857-2860.

[18] R. E. Eilers, E. Miskiel, O. Ozdamar, R. Urbano, J. E. Widen, Optimization of automated hearing test algorithms: Simulations using an infant response model, Ear and Hearing 12 (3) (1991) 191-198.

[19] B. S. Hermmann, A. R. Thornton, J. M. Joseph, Automated infant hearing screening using the ABR: development and validation, American Journal of Audiology 4 (2) (1995) 6-14.

[20] W. H. McFarland, F. B. Simmons, F. R. Jones, An automated hearing screening technique for newborns, Journal of Speech and Hearing Disorders 45 (4) (1980) 495-503.

[21] O. Ozdamar, R. E. Delgado, R. E. Eilers, R. C. Urbano, Automated electrophysiologic hearing testing using a threshold-seeking algorithm, Journal of the American Academy of Audiology 5 (2) (1994) 77-88.

[22] K. D. Pool, T. Finitzo, Evaluation of a computer-automated program for clinical assessment of the auditory brain stem response, Ear and Hearing 10 (5) (1989) 304-310.

[23] Joint Committee on Infant Hearing, Year 2007 position statement: principles and guidelines for early hearing detection and intervention programs, Pediatrics 120 (4) (2007) 898-921.

[24] M. Cebulla, E. Stürzebecher, K. Wernecke, Objective detection of auditory brainstem potentials. Comparison of statistical tests in the time and frequency domains, Scandinavian Audiology 29 (1) (2000) 44-51.

[25] E. Stürzebecher, M. Cebulla, Objective detection of auditory evoked potentials. Comparison of several statistical tests in the frequency domain on the basis of near-threshold ABR data, Scandinavian Audiology 26 (1) (1997) 7-14.

[26] E. Vannier, O. Adam, P. Karasinski, M. Ohresser, J. Motsch, Computer-assisted ABR interpretation using the automatic construction of the latency-intensity curve, Audiology 40 (4) (2001) 191-201.

[27] M. S. John, T. W. Picton, MASTER: a Windows program for recording multiple auditory steady-state responses, Computer Methods and Programs in Biomedicine 61 (2) (2000) 125-150.

[28] J. Fridman, E. R. John, M. Bergelson, J. B. Kaiser, H. W. Baird, Application of digital filtering and automatic peak detection to brain stem auditory evoked potential, Electroencephalography and Clinical Neurophysiology 53 (4) (1982) 405-416.

[29] F. Chan, F. Lam, P. Poon, W. Qiu, Detection of brainstem auditory evoked potential by adaptive filtering, Medical and Biological Engineering and Computing 33 (1) (1995) 69-75.

[30] T. Gronfors, Peak identification of auditory brainstem responses with multifilters and attributed automaton, Computer Methods and Programs in Biomedicine 40 (2) (1993) 83-88.

[31] G. C. Galbraith, Enhanced brainstem and cortical evoked response amplitudes: Single-trial covariance analysis, Perceptual and Motor Skills 92 (3 part 1) (2001) 659-672.

[32] V. Sundaramoorthy, M. J. Pont, C. Degg, J. A. Cook, A computerized database of ‘normal’ auditory brainstem responses, British Journal of Audiology 34 (3) (2000) 197-201.

[33] E. Vannier, O. Adam, J. Motsch, Objective detection of brainstem auditory evoked potentials with a priori information from higher presentation levels, Artificial Intelligence in Medicine 25 (3) (2002) 283-301.

[34] J. W. Hall, K. A. Rupp, Auditory brainstem response: Recent developments in recording and analysis, Advances in Oto-rhino-laryngology 53 (1997) 21-45.

[35] B. A. Weber, G. L. Fletcher, A computerized scoring procedure for auditory brainstem response audiometry, Ear and Hearing 1 (5) (1980) 233-236.

[36] C. Elberling, M. Don, Quality estimation of averaged auditory brainstem responses, Scandinavian Audiology 13 (3) (1984) 187-197.

[37] J. T. Valderrama, I. Alvarez, A. de la Torre, J. C. Segura, M. Sainz, J. L. Vargas, A portable, modular, and low cost auditory brainstem response recording system including an algorithm for automatic identification of responses suitable for hearing screening, in Proceedings of the

IEEE/EMBS Special Topic Conference on Point-of-Care HealthCare Technologies (PoCHT), Bangalore, India, 2013, art. no. 6461314, pp. 180-183.

[38] S. A. Arnold, Objective versus visual detection of the auditory brain stem response, Ear and Hearing 6 (3) (1985) 144-150.

[39] T. Kakiashvili, W. W. Koczkodaj, M. Woodbury-Smith, Improving the medical scale predictability by the pairwise comparisons method: Evidence from a clinical data study, Computer Methods and Programs in Biomedicine 105 (3) (2012) 210-216.

[40] C. Elberling, Auditory electrophysiology. The use of templates and cross correlation functions in the analysis of brain stem potentials, Scandinavian Audiology 8 (3) (1979) 187-190.

[41] T. W. Picton, D. R. Stapells, K. B. Campbell, Auditory evoked potentials from the human cochlea and brainstem, The Journal of Otolaryngology 10 (Supplement 9) (1981) 1-41.

[42] J. L. Stone, M. Calderon-Arnulphi, K. S. Watson, K. Patel, N. S. Mander, N. Suss, J. Fino, J. R. Hughest, Brainstem auditory evoked potentials-A review and modified studies in healthy subjects, Journal of Clinical Neurophysiology 26 (3) (2009) 167-175.

[43] J. T. Valderrama, I. Alvarez, A. de la Torre, J. C. Segura, M. Sainz, J. L. Vargas, Recording of auditory brainstem response at high stimulation rates using randomized stimulation and averaging, Journal of the Acoustical Society of America 132 (6) (2012) 3856-3865.

[44] J. T. Valderrama, I. Alvarez, A. de la Torre, J. C. Segura, M. Sainz, J. L. Vargas, Educational approach of a BAER recording system based on experiential learning, Technics Technologies Education Management 6 (4) (2011) 876-889.

[45] A. R. D. Thornton, M. J. Coleman, The adaptation of cochlear and brainstem auditory evoked potentials in humans, Electroencephalography and Clinical Neurophysiology 39 (4) (1975) 399-406.

[46] P. G. Gillespie, U. Muller, Mechanotransduction by hair cells: Models, molecules, and mechanisms, Cell 139 (1) (2009) 33-44.

[47] M. A. Petoe, A. P. Bradley, W. J. Wilson, Spectral and synchrony differences in auditory brainstem responses evoked by chirps of varying durations, Journal of the Acoustical Society of America 128 (4) (2010) 1896-1907.

[48] H. Pratt, H. Sohmer, Intensity and rate functions of cochlear and brainstem evoked responses to click stimuli in man, Archives of Oto-Rhino-Laryngology 212 (2) (1976) 85-92.

[49] T. Yagi, K. Kaga, The effect of the click repetition rate on the latency of the auditory evoked brain stem response and its clinical use for a neurological diagnosis, Archives of Oto-Rhino-Laryngology 222 (2) (1979) 91-96.

[50] R. E. Lasky, A developmental study on the effect of stimulus rate on the auditory evoked brain-stem response, Electroencephalography and Clinical Neurophysiology 59 (5) (1984) 411-419.

[51] G. Lina-Granade, L. Collet, A. Morgon, B. Salle, Maturation and effect of stimulus rate on brainstem auditory evoked potentials, Brain and Development 15 (4) (1993) 263-269.

[52] Z. D. Jiang, Y. Y. Wu, A. R. Wilkinson, Age-related changes in BAER at different click rates from neonates to adults, Acta Paediatrica, International Journal of Paediatrics 98 (8) (2009) 1284-1287.

[53] J. G. Peters, An automated infant screener using advanced evoked response technology, The Hearing Journal 39 (1986) 25-30.

[54] K. J. Doyle, S. Fujikawa, P. Rogers, E. Newman, Comparison of newborn hearing screening by transient otoacoustic emissions and auditory brainstem response using ALGO-2, International Journal of Pediatric Otorhinolaryngology 43 (3) (1998) 207-211.

[55] K. J. Doyle, B. Burggraaff, S. Fujikawa, J. Kim, Newborn hearing screening by otoacoustic emissions and automated auditory brainstem response, International Journal of Pediatric Otorhinolaryngology 41 (2) (1997) 111-119.

[56] V. Bajaj, R. B. Pachori, Automatic classification of sleep stages based on the time-frequency image of EEG signals, Computer Methods and Programs in Biomedicine 112 (3) (2013) 320-328.

[57] J. Widjaja, U. Suripon, Retinal blood vessel detection using wavelet-matched filter, Optical Engineering 52 (3) (2013) 037204.

[58] D. L. Jewett, G. Caplovitz, B. Baird, M. Trumpis, M. P. Olson, L. J. Larson-Prior, The use of QSD (q-sequence deconvolution) to recover superposed, transient evoked-responses, Clinical Neurophysiology 115 (12) (2004) 2754-2775.

[59] R. E. Delgado, O. Ozdamar, Deconvolution of evoked responses obtained at high stimulus rates, Journal of the Acoustical Society of America 115 (3) (2004) 1242-1251.

[60] O. Ozdamar, J. Bohorquez, Signal-to-noise ratio and frequency analysis of continuous loop averaging deconvolution (CLAD) of overlapping evoked potentials, Journal of the Acoustical Society of America 119 (1) (2006) 429-438.

[61] S. M. Mason, A. P. Su, R. A. Hayes, Simple online detector of auditory evoked cortical potentials, Medical and Biological Engineering and Computing 15 (6) (1977) 641-647.

Figure Legends:

- Figure 1. Parameters involved in the automatic quality evaluation technique based on Fitted Parametric Peaks (*FPP*). The parametric peak fitted to the wave V of an ABR test signal is highlighted.
- Figure 2. Latencies (L) and amplitudes (A) of waves III and V measured manually (*MAN*) and automatically by the *FPP* method in a set of 320 ABR signals obtained from 8 normal hearing subjects at different stimulation rates. Normalized data in terms of the mean value are also presented on this

figure to decrease the intersubject variability. The coefficients of determination (R^2) obtained in this study on each parameter suggest that the model of amplitudes and latencies is better described with the *FPP* method.

- Figure 3. Comparative analysis of the latencies and amplitudes of waves III and V estimated manually by an audiologist and automatically by the *FPP* method. The linear regression model of the experimental data is compared with the curve *FPP = MAN* (dotted line).
- Figure 4. Examples of ABR signals from 5 normal hearing subjects obtained at different stimulation rates using the randomized stimulation and averaging (*RSA*) technique [43]. The parametric peaks adjusted to waves III and V are highlighted on this figure and the automatic quality evaluation provided by the *FPP* method for each wave is presented.
- Figure 5. Computer application screenshot used on the subjective evaluation of the quality. Two ABR signals are shown as reference for each quality level. The subjective evaluator is asked to rate the quality for each test ABR between 0 (no ABR) to 5 (excellent quality ABR).
- Figure 6. Examples of ABR signals of different quality used for test. The signals K and L are obtained without auditory stimulation. The quality evaluation provided for each signal by both automatic and subjective methods is provided in table 4.
- Figure 7. (A) Linear regression analysis for each individual subjective evaluation compared to the automatic evaluation provided by the *FPP* method. (B) Linear regression analysis for the averaged subjective evaluation. This figure highlights the existing bias among evaluators. The

model is better described when an averaged subjective evaluation is considered (r = 0.84).

- Figure 8. ROC space of a response validation study defined by the false positive rate (FPR), or 1-specificity, and the true positive rate (TPR), or sensitivity, for the automated response validation methods based on fitted parametric peaks (*FPP*), on the correlation coefficient (*r*), on the F_{SP}, and on the cross correlation with a predefined template waveform (*Cross Corr*).

Table Legends:

- Table 1. Referenced latencies (in ms) for waves III and V at different stimulation rates.
- Table 2. Mean (and standard deviation in parentheses) of the latencies (L), amplitudes (A), widths (W), and SNRs of waves III and V measured automatically by the *FPP* on a set of 320 ABR signals obtained from 8 normal hearing subjects at different stimulation rates. Latencies and widths are measured in ms, amplitudes in µV, and SNR in dB.
- Table 3. Mean (and standard deviation in parentheses) of the latencies (L), amplitudes (A), widths (W), and SNRs of waves III and V measured manually on a set of 320 ABR signals obtained from 8 normal hearing subjects at different stimulation rates. Latencies and widths are measured in ms and amplitudes in µV.
- Table 4. Evaluation of the quality provided by the automatic evaluation techniques based on *FPP*, *r*, F_{SP}, and *Cross Corr*, by the individual subjective evaluation of the experts (*Ev1-Ev5*), and by the averaged

subjective evaluation (*All Ev*) for the ABR signals shown in figure 6 as examples.

Table 1.

Stimulation rate	L_{III}	L_{V}
30 Hz	3.72	5.68
45 Hz	3.74	5.69
55 Hz	3.80	5.79
72 Hz	3.86	5.90
83 Hz	3.90	5.97
100 Hz	3.94	6.07
125 Hz	4.00	6.21
167 Hz	4.03	6.40
250 Hz	3.99	6.72

Table 2.

Stimulation Rate	L_{III}	L_{V}	L_{V} - L_{III}	A_{III}	A_{V}	W_{III}	W_{V}	SNR_{III}	SNR_{V}
45 Hz	3.74 (0.13)	5.71 (0.20)	1.97 (0.15)	0.25 (0.08)	0.26 (0.07)	0.37 (0.05)	0.46 (0.05)	9.19 (2.83)	12.41 (3.03)
55 Hz	3.79 (0.10)	5.80 (0.20)	1.99 (0.15)	0.23 (0.07)	0.25 (0.08)	0.38 (0.04)	0.47 (0.06)	8.30 (3.37)	12.78 (3.55)
72 Hz	3.86 (0.13)	5.91 (0.18)	2.04 (0.15)	0.21 (0.07)	0.22 (0.07)	0.37 (0.05)	0.49 (0.06)	8.58 (3.34)	12.38 (2.68)
83 Hz	3.91 (0.12)	5.98 (0.19)	2.06 (0.14)	0.20 (0.08)	0.21 (0.07)	0.38 (0.05)	0.53 (0.08)	8.05 (3.96)	12.67 (3.37)
100 Hz	3.92 (0.15)	6.09 (0.22)	2.12 (0.15)	0.17 (0.07)	0.19 (0.06)	0.38 (0.06)	0.50 (0.07)	7.62 (3.67)	13.16 (2.91)
125 Hz	4.01 (0.17)	6.21 (0.20)	2.21 (0.16)	0.14 (0.06)	0.18 (0.05)	0.40 (0.08)	0.50 (0.07)	7.64 (3.22)	12.57 (3.14)
167 Hz	4.18 (0.15)	6.41 (0.26)	2.19 (0.23)	0.15 (0.07)	0.15 (0.04)	0.52 (0.14)	0.52 (0.08)	6.59 (4.02)	11.25 (3.05)
250 Hz	4.33 (0.25)	6.77 (0.25)	2.42 (0.15)	0.11 (0.05)	0.13 (0.04)	0.54 (0.12)	0.58 (0.10)	5.42 (2.23)	11.39 (3.06)

Table 3.

Stimulation Rate	L_{III}	L_V	L_V - L_{III}	A_{III}	A_V
45 Hz	3.73 (0.15)	5.70 (0.23)	1.97 (0.20)	0.25 (0.10)	0.29 (0.07)
55 Hz	3.78 (0.11)	5.78 (0.22)	1.98 (0.18)	0.23 (0.10)	0.28 (0.09)
72 Hz	3.87 (0.14)	5.90 (0.22)	2.03 (0.19)	0.21 (0.08)	0.25 (0.07)
83 Hz	3.91 (0.16)	5.91 (0.40)	1.97 (0.43)	0.20 (0.09)	0.24 (0.07)
100 Hz	3.90 (0.14)	6.07 (0.24)	2.11 (0.20)	0.17 (0.08)	0.22 (0.07)
125 Hz	4.00 (0.20)	6.21 (0.23)	2.20 (0.24)	0.15 (0.07)	0.21 (0.05)
167 Hz	4.16 (0.19)	6.40 (0.29)	2.22 (0.27)	0.15 (0.06)	0.20 (0.08)
250 Hz	4.36 (0.37)	6.77 (0.30)	2.41 (0.21)	0.12 (0.06)	0.17 (0.04)

Table 4.

ABR	FPP	r	F_{SP}	Cross Corr	Ev1	Ev2	Ev3	Ev4	Ev5	All Ev
A	8.8	0.97	54.1	0.84	5	5	4	5	5	4.8
B	10.6	0.99	113.8	0.77	5	5	3	5	5	4.6
C	7.6	0.95	12.5	0.86	5	4	3	4	5	4.2
D	14.2	0.54	3.6	0.80	4	4	4	5	4	4.2
E	7.1	0.70	5.6	0.58	4	3	3	5	4	3.8
F	5.8	0.42	2.5	0.61	3	4	3	3	3	3.2
G	6.5	0.53	3.7	0.65	3	1	1	3	4	2.4
H	4.8	0.61	2.1	0.71	4	3	2	4	3	3.2
I	1.4	0.10	1.6	0.64	0	2	1	1	2	1.2
J	1.9	0.27	2.1	0.59	1	3	1	3	2	2.0
K	1.9	0.40	1.7	0.62	0	1	0	0	0	0.2
L	-1.7	-0.17	0.6	0.36	0	0	0	0	0	0.0

Figure 1

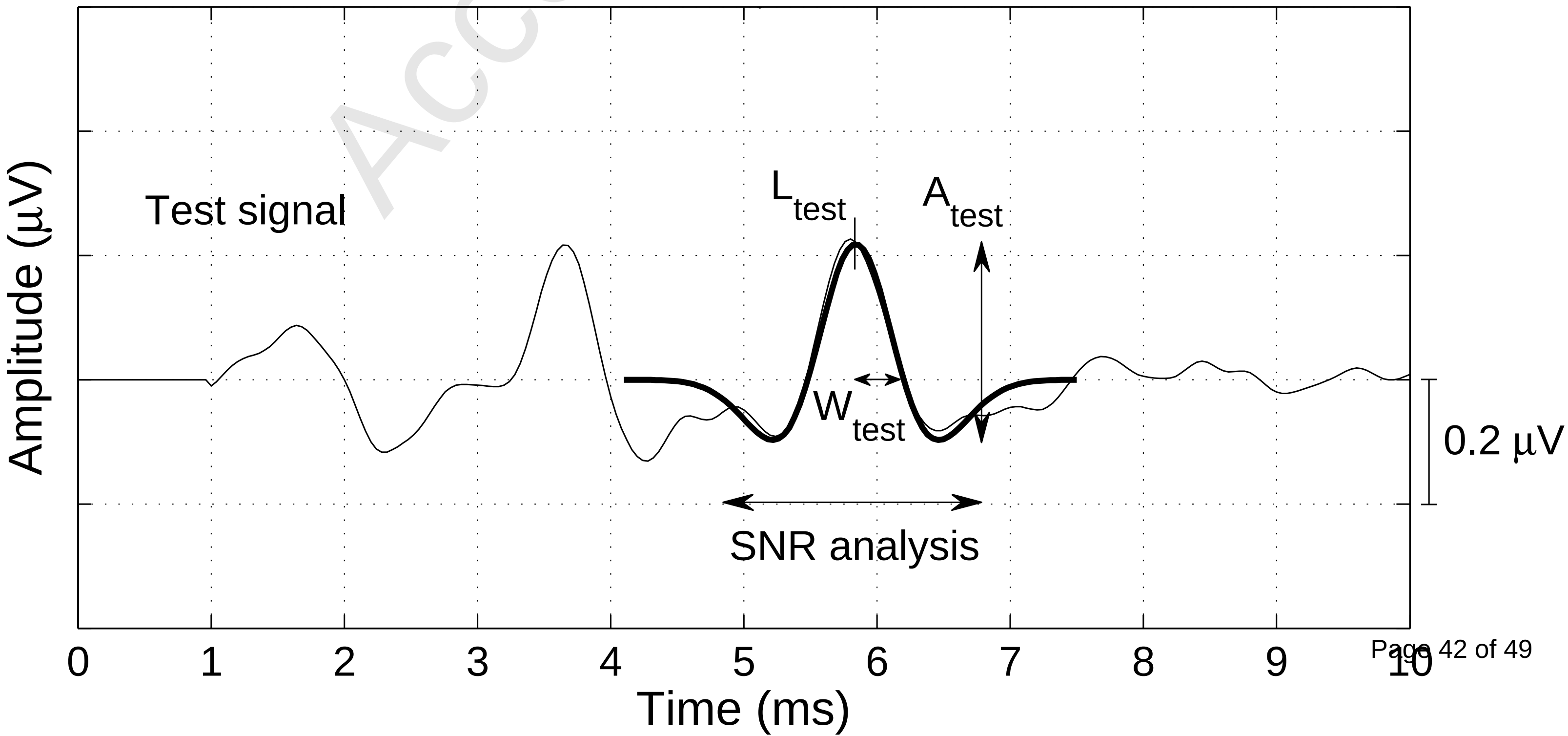
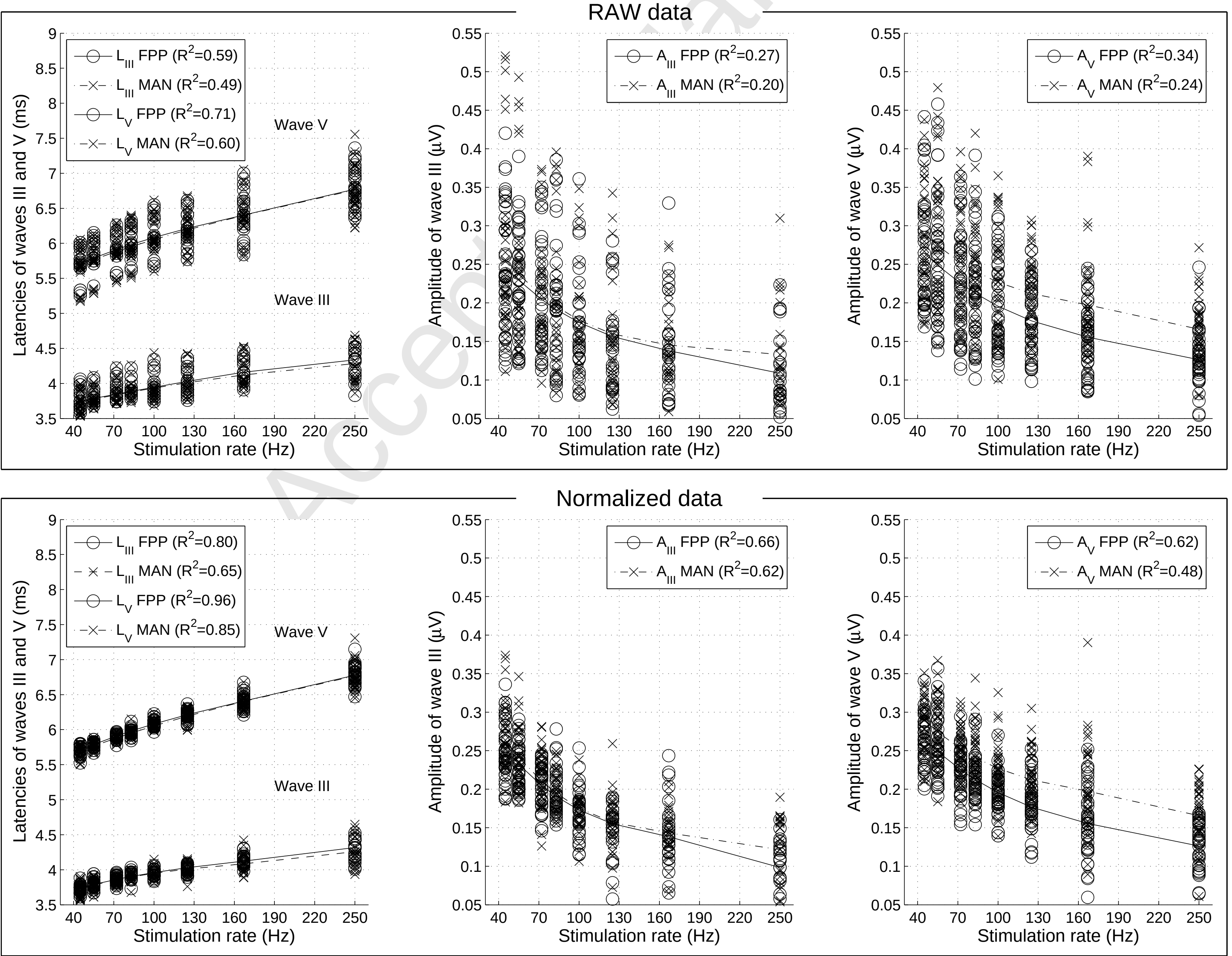

Test signal
L_{test}
A_{test}
W_{test}
SNR analysis
0.2 μV
Amplitude (μV)
Time (ms)
0
1
2
3
4
5
6
7
8
9
10

Figure 2


RAW data
Normalized data
Latencies of waves III and V (ms)
Wave V
Wave III
L_{III} FPP (R^2=0.59)
L_{III} MAN (R^2=0.49)
L_V FPP (R^2=0.71)
L_V MAN (R^2=0.60)
Amplitude of wave III (μV)
A_{III} FPP (R^2=0.27)
A_{III} MAN (R^2=0.20)
Amplitude of wave V (μV)
A_V FPP (R^2=0.34)
A_V MAN (R^2=0.24)
L_{III} FPP (R^2=0.80)
L_{III} MAN (R^2=0.65)
L_V FPP (R^2=0.96)
L_V MAN (R^2=0.85)
A_{III} FPP (R^2=0.66)
A_{III} MAN (R^2=0.62)
A_V FPP (R^2=0.62)
A_V MAN (R^2=0.48)
Stimulation rate (Hz)
40 70 100 130 160 190 220 250
3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5 0.55

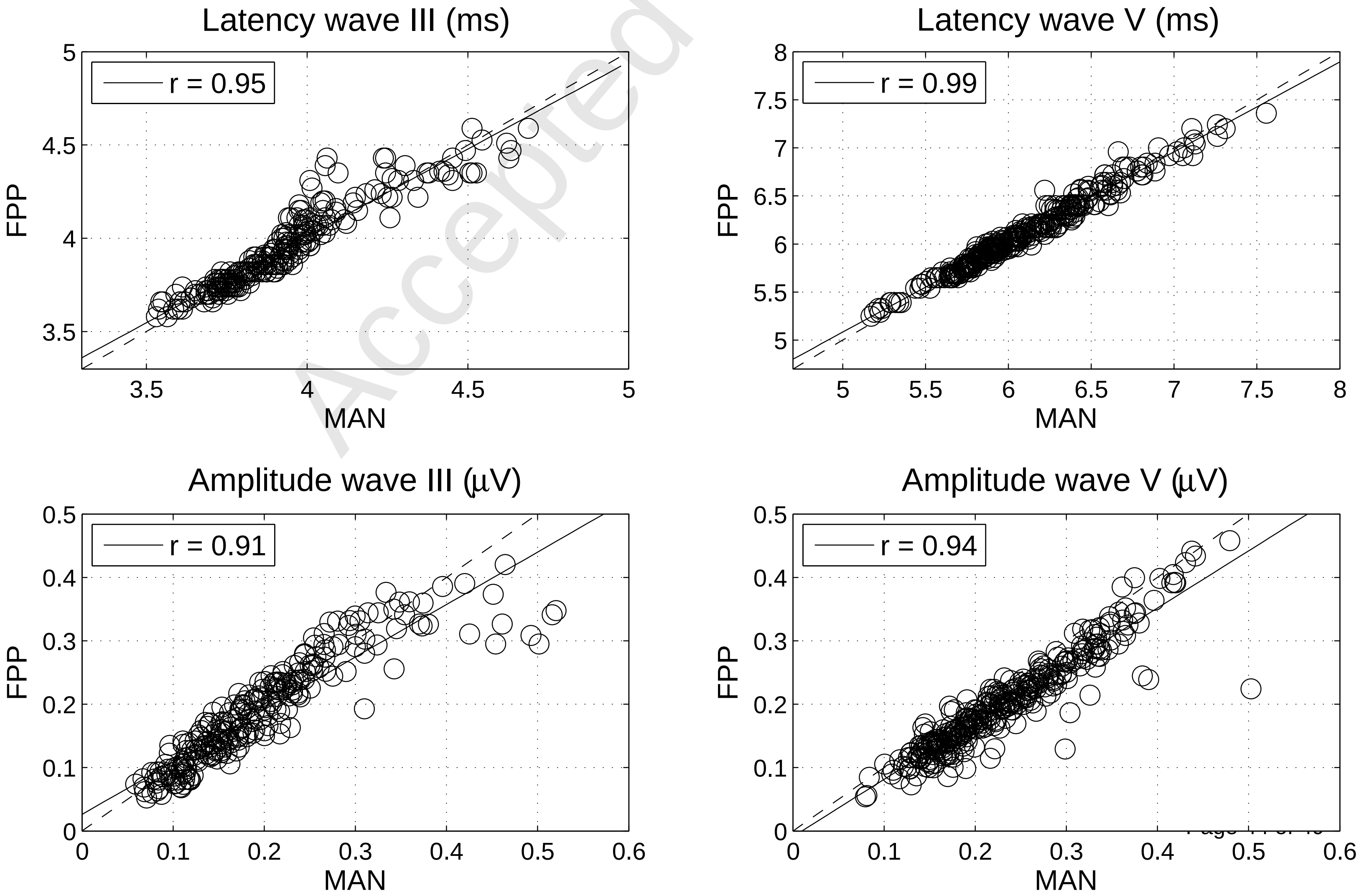

Latency wave III (ms)
r = 0.95
FPP
MAN
5
4.5
4
3.5
3.5
4
4.5
5
Latency wave V (ms)
r = 0.99
FPP
MAN
8
7.5
7
6.5
6
5.5
5
5
5.5
6
6.5
7
7.5
8
Amplitude wave III (μV)
r = 0.91
FPP
MAN
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
Amplitude wave V (μV)
r = 0.94
FPP
MAN
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6

Figure 4

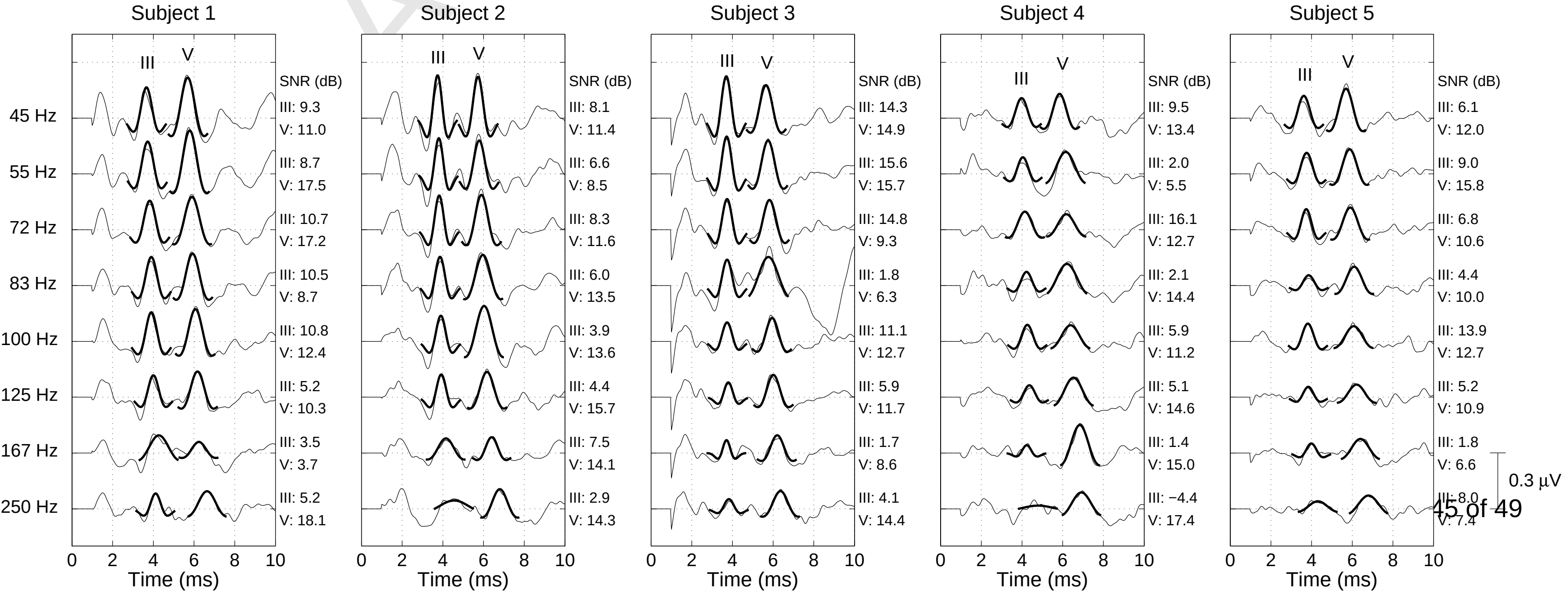
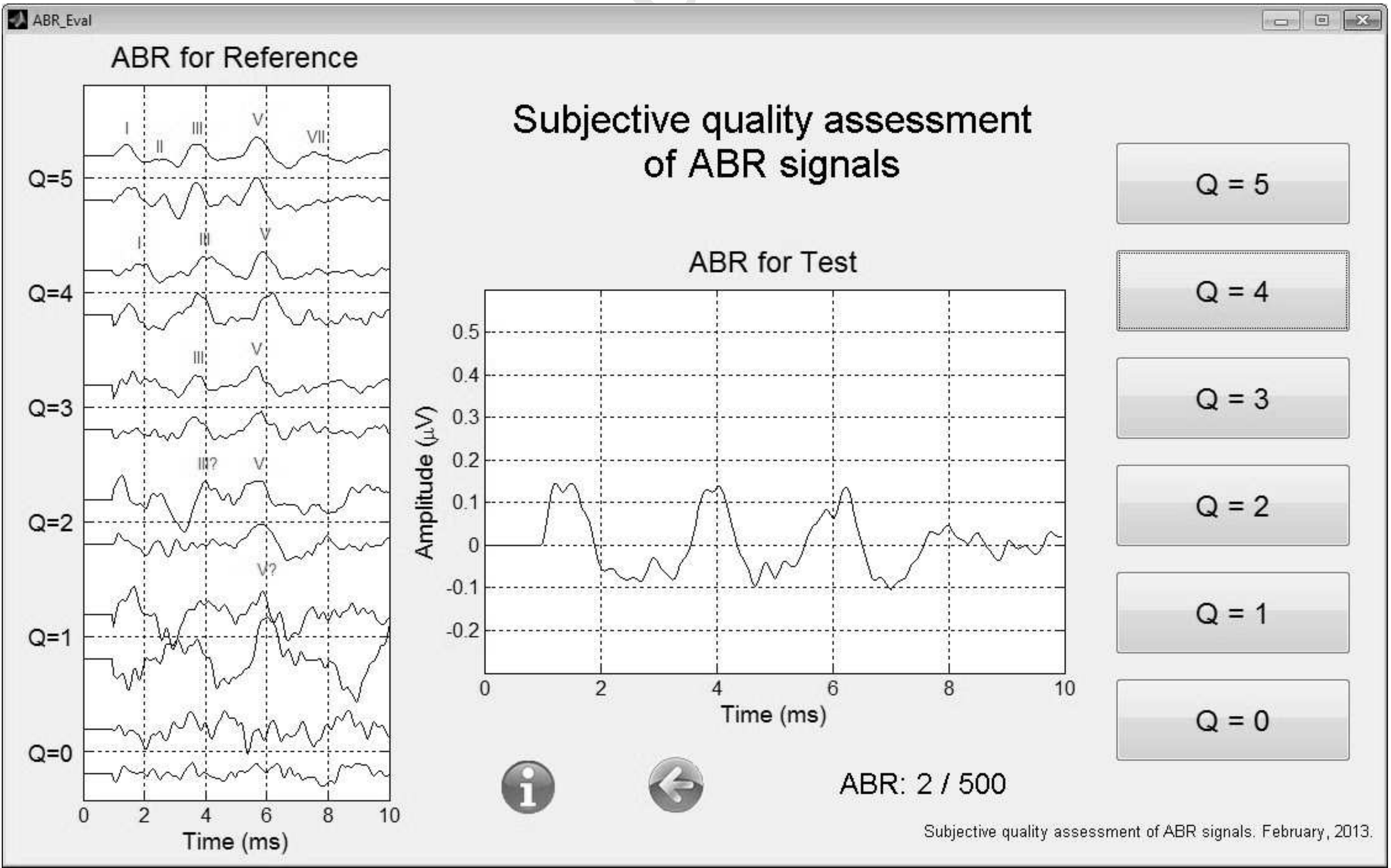

Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
III
V
SNR (dB)
45 Hz
55 Hz
72 Hz
83 Hz
100 Hz
125 Hz
167 Hz
250 Hz
Subject 1: III: 9.3 V: 11.0; III: 8.7 V: 17.5; III: 10.7 V: 17.2; III: 10.5 V: 8.7; III: 10.8 V: 12.4; III: 5.2 V: 10.3; III: 3.5 V: 3.7; III: 5.2 V: 18.1
Subject 2: III: 8.1 V: 11.4; III: 6.6 V: 8.5; III: 8.3 V: 11.6; III: 6.0 V: 13.5; III: 3.9 V: 13.6; III: 4.4 V: 15.7; III: 7.5 V: 14.1; III: 2.9 V: 14.3
Subject 3: III: 14.3 V: 14.9; III: 15.6 V: 15.7; III: 14.8 V: 9.3; III: 1.8 V: 6.3; III: 11.1 V: 12.7; III: 5.9 V: 11.7; III: 1.7 V: 8.6; III: 4.1 V: 14.4
Subject 4: III: 9.5 V: 13.4; III: 2.0 V: 5.5; III: 16.1 V: 12.7; III: 2.1 V: 14.4; III: 5.9 V: 11.2; III: 5.1 V: 14.6; III: 1.4 V: 15.0; III: −4.4 V: 17.4
Subject 5: III: 6.1 V: 12.0; III: 9.0 V: 15.8; III: 6.8 V: 10.6; III: 4.4 V: 10.0; III: 13.9 V: 12.7; III: 5.2 V: 10.9; III: 1.8 V: 6.6; III: 8.0 V: 7.4
0.3 µV
0
2
4
6
8
10
Time (ms)

Figure 5


ABR_Eval
ABR for Reference
I
II
III
V
VII
I
III
V
III
V
III?
V
V?
Q=5
Q=4
Q=3
Q=2
Q=1
Q=0
0
2
4
6
8
10
Time (ms)
Subjective quality assessment of ABR signals
ABR for Test
Amplitude (μV)
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
0
2
4
6
8
10
Time (ms)
Q = 5
Q = 4
Q = 3
Q = 2
Q = 1
Q = 0
ABR: 2 / 500
Subjective quality assessment of ABR signals. February, 2013.

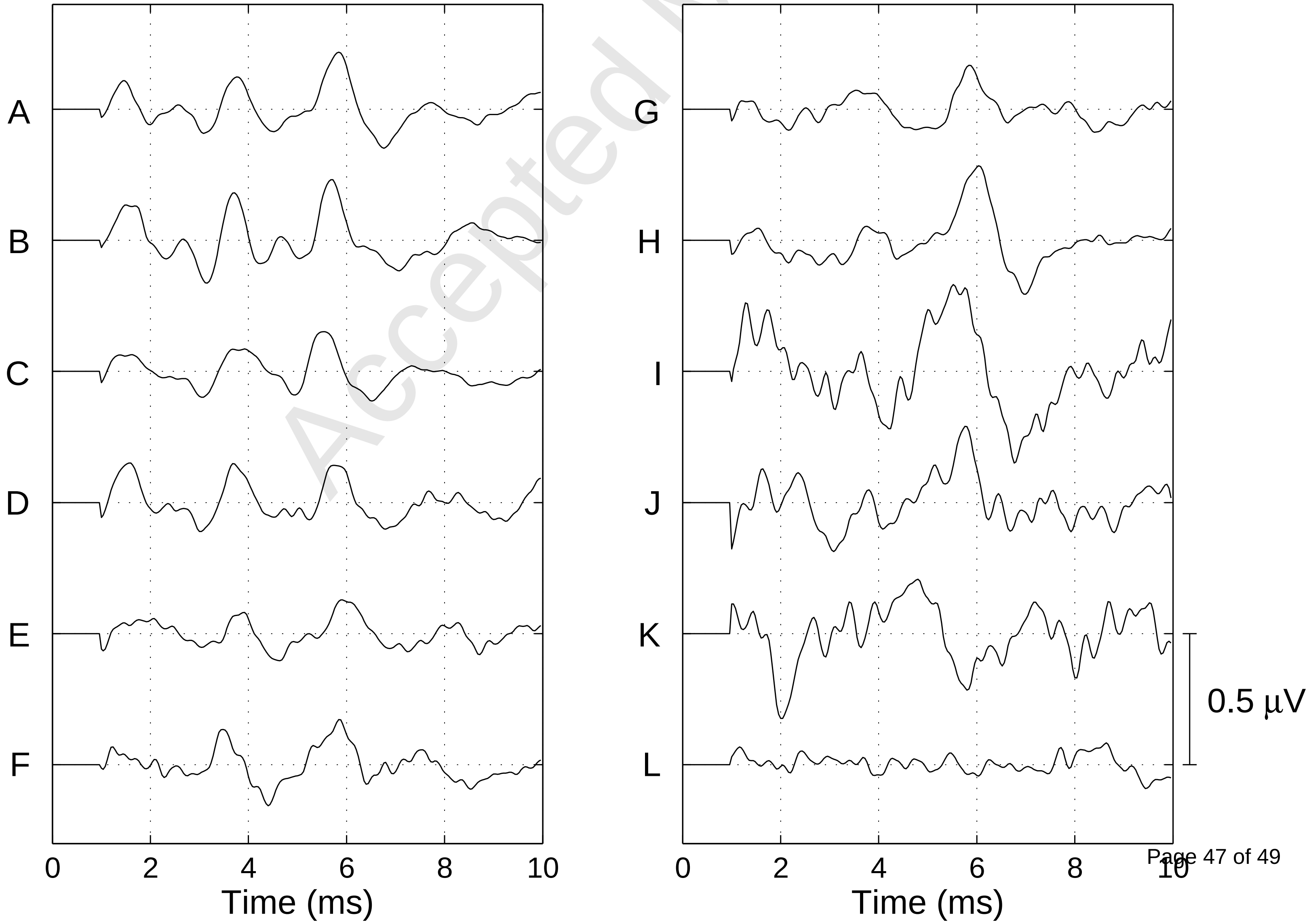
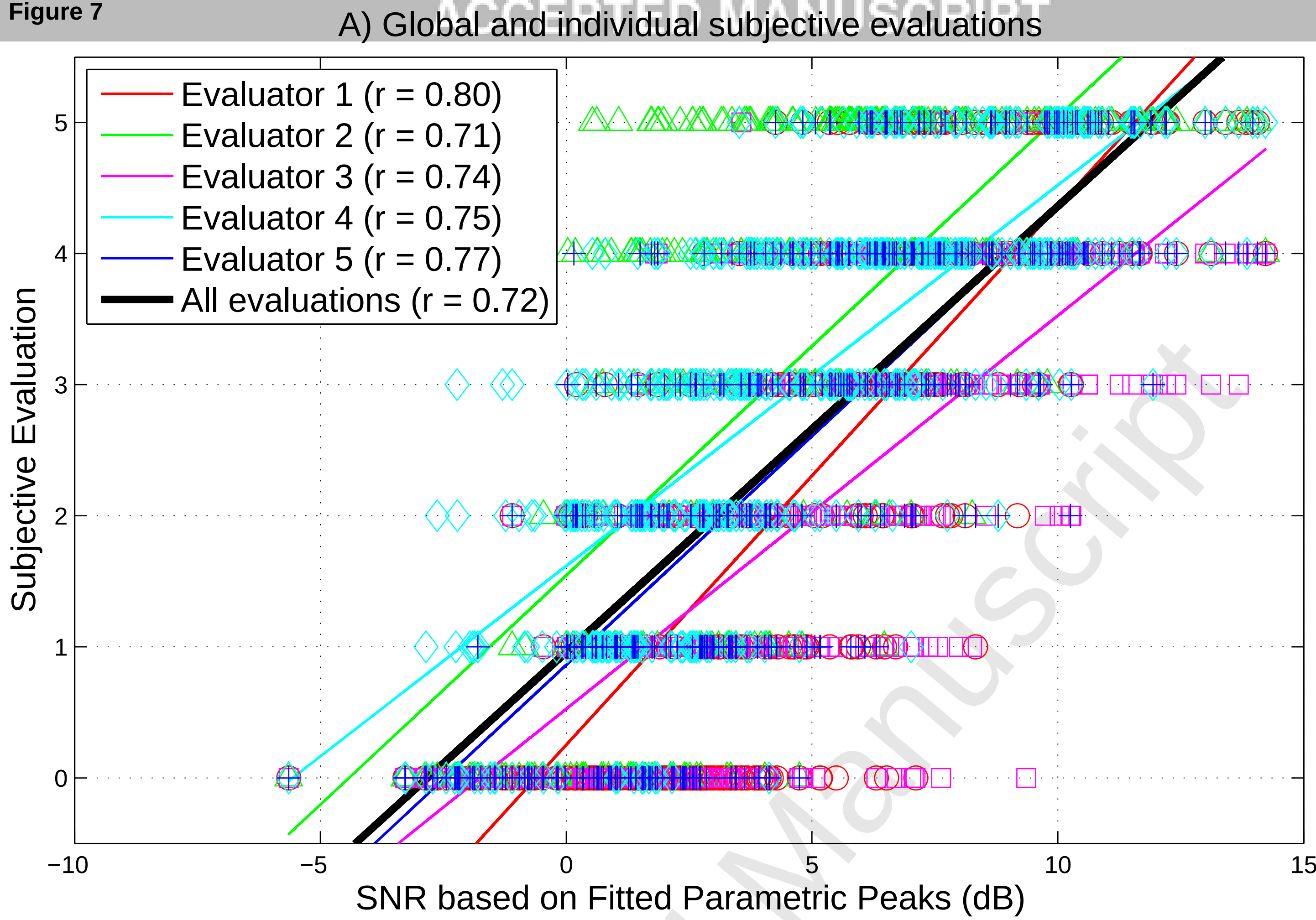
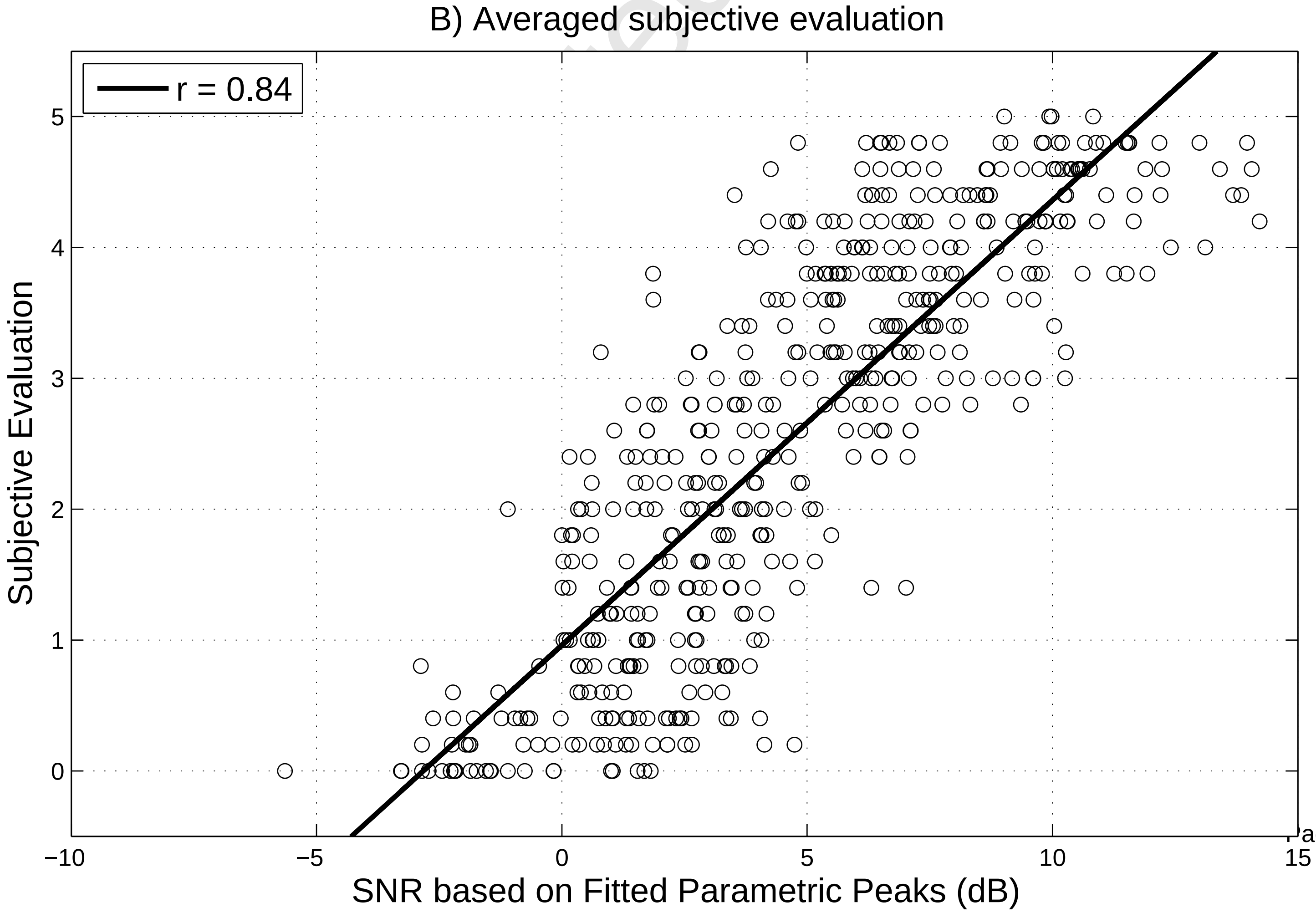
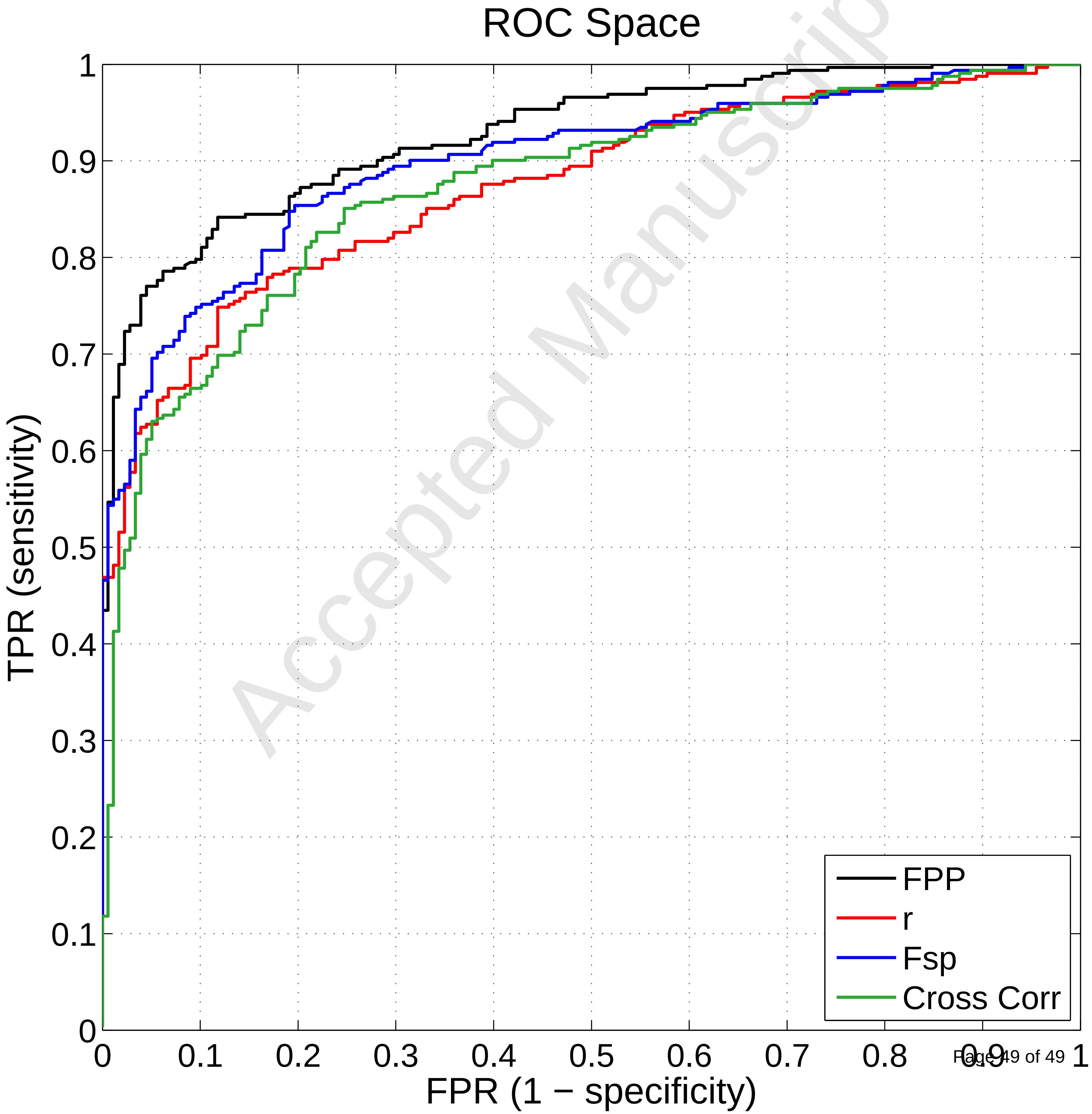

A
B
C
D
E
F
G
H
I
J
K
L
0
2
4
6
8
10
Time (ms)
0
2
4
6
8
10
Time (ms)
0.5 µV

Figure 7


A) Global and individual subjective evaluations
Evaluator 1 (r = 0.80)
Evaluator 2 (r = 0.71)
Evaluator 3 (r = 0.74)
Evaluator 4 (r = 0.75)
Evaluator 5 (r = 0.77)
All evaluations (r = 0.72)
Subjective Evaluation
0
1
2
3
4
5
−10
−5
0
5
10
15
SNR based on Fitted Parametric Peaks (dB)


B) Averaged subjective evaluation
r = 0.84
Subjective Evaluation
0
1
2
3
4
5
−10
−5
0
5
10
15
SNR based on Fitted Parametric Peaks (dB)


ROC Space
TPR (sensitivity)
FPR (1 − specificity)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
FPP
r
Fsp
Cross Corr

Supplementary Material

Section A: MATLAB & GNU Octave ROUTINES (R).

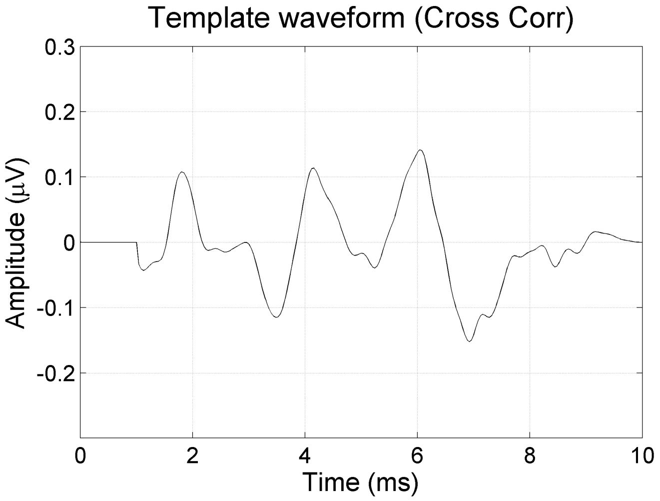
R1: MATLAB & GNU Octave routine that implements the FPP methodology

```matlab
function [x,L0,W0,A0,SNR] = FPP(Linit,ABR,fs)
% Input parameters:  Linit (Latency used for initialization in ms)
%                    y (ABR signal used for test)
%                    fs (sampling frequency in Hz)
% Output parameters: x (Peak fitted to the ABR signal)
%                    L0 (Latency of the fitted peak in ms)
%                    W0 (Width of the fitted peak in ms)
%                    A0 (Amplitude of the fitted peak in uV)
%                    SNR (Quality evaluation provided by FPP)

% Initialization
t = (0:length(y)-1)/fs*1e3;  % Time axis in ms
W_test = 0.1:0.01:1;           % Axis of evaluated widths in ms
Ainit = 1;                     % Amplitude used for initialization in uV
maxlag = ceil(1.5e-3*fs);      % Maximum time displacement allowed = 3 ms
Int_time = 1;                  % Time interval of 2 ms around the latency
PN_i = 1e10;                   % Power of noise parameter initialization

% 1 Dimensional search on the width parameter
for i=1:length(W_test)
    % Step 1 - Peak initialization
    x_test = Peak_Generation(Linit,W_test(i),Ainit,fs,y);

    % Step 2 - Search of the optimal latency for the analyzed width
    Interval = t>Linit-Int_time & t<Linit+Int_time;
    [Corr,lag] = xcorr(y(Interval),x_test(Interval),maxlag,'coeff');
    [~,idx] = max(Corr);
    lag_max = lag(idx)/fs*1e3;  % Time displacement in ms of L for best fit
    L0 = Linit+lag_max;         % Best latency (L0) of the peak for W_test
    x_test = Peak_Generation(L0,W_test(i),Ainit,fs,y);    % Updated peak

    % Step 3 - Search of the optimal amplitude for the analyzed width
    Interval = t>L0-Int_time & t<L0+Int_time;
    A0 = dot(y(Interval),x_test(Interval))/...        % Optimal amplitude
        dot(x_test(Interval),x_test(Interval));
    x_test = Peak_Generation(L0,W_test(i),A0,fs,y);   % Updated peak

    % Step 4 - Evaluation of the error and power of noise
    e = y(Interval)-x_test(Interval);
    PN = dot(e,e);                              % Power of noise estimation

    % Step 5 - Optimal approximation of the peak to the ABR signal
    if (PN<PN_i)
        PN_i = PN;                              % PN of reference updated
        L_peak = L0;                            % Optimal latency
        A_peak = A0;                            % Optimal amplitude
        W_peak = W_test(i);                     % Optimal width
        x = Peak_Generation(L_peak,W_peak,A_peak,fs,y);     % Fitted peak
        Px = dot(x(Interval),x(Interval));      % Power of the Peak
        SNR = 10*log10(Px/PN);                  % Signal to Noise Ratio
    end
end
```

R2: MATLAB & GNU Octave routine that implements Peak Generation

```
function [x] = Peak_Generation(L,W,A,fs,ABR)
% Input parameters:  L (Latency in ms), W (Width in ms),
%                    A (Amplitude in µV), fs (sampling frequency in Hz)
%                    ABR (ABR signal used for test)
% Output parameters: x (Peak generated by the function)

t = (0:length(ABR)-1)/fs*1e3;          % Time axis in ms
K0 = 1+2*exp(-3/2);                    % Normalization constant
E = exp(-(t-L).^2/(2*W^2));            % Exponential term of the peak
x = E.*(1-(t-L).^2/(W^2));             % Peak, no amplitude adjusted
x = x/K0;                              % Normalization of the amplitude
x = x-mean(x);                         % Normalization of the amplitude
x = A*x';                              % Peak with amplitude adjusted
```

Section B: Template waveform used in the *Cross Corr* method.



Section C: Linear regression analysis between the averaged subjective evaluations of the quality provided by 5 experienced audiologists and the automatic methods based on the FPP, on the correlation coefficient (r), on the F_{SP}, and on the cross correlation (Cross Corr) in a set of ABR signals of different quality.