G Model
ANL-1992; No. of Pages 8
ARTICLE IN PRESS
Auris Nasus Larynx xxx (2015) xxx–xxx
Contents lists available at ScienceDirect
Auris Nasus Larynx
journal homepage: www.elsevier.com/locate/anl

Nonlinear feature extraction for objective classification of complex auditory brainstem responses to diotic perceptually critical consonant-vowel syllables

Amir Salar Jafarpisheh [a], Amir Homayoun Jafari [a,b,*], Mohammadjavad Abolhassani [a], Mohammad Farhadi [c], Hamed Sadjedi [d], Akram Pourbakht [e,f], Zahra Shirzhiyan [a]

[a] Medical Physics and Biomedical Engineering Department, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran
[b] Research Center for Biomedical Technologies & Robotics (RCBTR), Tehran University of Medical Sciences, Tehran, Iran
[c] Clinical Nanomedicine Laboratory, ENT – Head & Neck Research Center, Hazrate Rasoul Akram Hospital, Iran University of Medical Sciences, Tehran, Iran
[d] Department of Engineering, Shahed University, Tehran, Iran
[e] Department of Audiology, Rehabilitation Research Center, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran
[f] Rehabilitation Research Center, Iran University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Article history:
Received 29 January 2015
Accepted 10 June 2015
Available online xxx

Keywords:
Complex auditory brainstem response (cABR)
Nonlinear analysis
RQA
Feature extraction
Recurrence time

ABSTRACT

Objective: To examine if nonlinear feature extraction method yields appropriate results in complex brainstem response classification of three different consonant vowels diotically presented in normal Persian speaking adults.
Methods: Speech-evoked auditory brainstem responses were obtained in 27 normal hearing young adults by using G.tec EEG recording system. 170 ms synthetic consonant-vowel stimuli /ba/, /da/, /ga/ were presented binaurally and the recurrence quantification analysis was performed on the responses. The recurrence time of second type was proposed as a suitable feature. ANOVA was also used for testing the significance of extracted feature. Post-comparison statistical method was used for showing which means are significantly different from each other.
Results: Dimension embedding and state space reconstruction were helpful for visualizing nonlinearity in auditory system. The proposed feature was successful in the objective classification of responses in window time 20.1–35.3 ms, which belonged to formant transition period of stimuli. Also the p value behavior of recurrence time of second type feature as a discriminant feature was close to the nature of the response that includes transient and sustained parts. On the other hand, the /ba/ and /ga/ classification period was wider than the others.
Conclusion: The extracted feature shown in this paper is helpful for the objective of distinguishing individuals with auditory processing disorders in the structurally similar voices. On the other hand, differing nonlinear feature is meaningful in a special region of response, equal to formant transition period, and this feature is related to the state space changes of brainstem response. It can be assumed that more information is within this region of signal and it is a sign of processing role of brainstem. The state changes of system are dependent on input stimuli, so the existence of top down feedback from cortex to brainstem forces the system to act differently.

1. Introduction

Animal studies have shown that auditory perceptional processing is distributed along the auditory system neurons [1,2]. A number of electrophysiological studies have recorded the complex auditory brainstem response (cABR) elicited by brief acoustic stimuli but most of them used temporal and frequency domain features like latency, amplitude, area and slope for transient peaks and magnitude of frequency following response, fundamental frequency, first formant amplitude and inter-response correlations that are linear approaches for representing differences between recorded responses [1–4]. Finding an insight about brainstem encoding of perceptually critical consonant-vowel stimuli was done through the extraction of important linear features of cABR and its relation to different acoustic stimuli containing /ba/, /da/, /ga/ [1,5].

* Corresponding author at: School of Medicine, Tehran University of Medical Sciences, Poursina Street, Tehran, Iran. Tel.: +98 2166466383.
E-mail address: h_jafari@tums.ac.ir (A.H. Jafari).

http://dx.doi.org/10.1016/j.anl.2015.06.003
0385-8146/© 2015 Published by Elsevier Ireland Ltd.

In the clinical applications, an objective tool for automatic classification of cABR signals can assist auditory professionals for better diagnosis of auditory possessing disabilities. It can be used for hearing aid fitting or cochlear implant adjustments. There are some sorts of auditory disorders that may have normal ABR or cABR in time domain linear analysis, but the use of a novel approach for the reconstruction of real signal dimension can be a tool for representing these problems. In cochlear implanted persons, this can yield to an adjustment protocol based on what a normal person really hears and what changes in his/her brainstem signal are expected. However, there is limited understanding of the neural responses evoked by various speech sounds [6].

In addition to the contribution of activity in the ascending auditory system to the speech ABR, top-down influences from higher neural centers have been shown to affect the responses. For example, auditory training and experience with a tonal language have been found to enhance responses at F0 [6,7].

Recent studies have shown that children with language based learning problems have abnormal encoding of speech at the brainstem level. Researchers tried to find an appropriate linear feature for objective classification of normal and abnormal cABR data [8]. The source filter model of speech processing showed selective deficiency in the neural encoding of acoustic features associated with the filter characteristics of speech for these subjects [5]. However, linear analysis provides poor information about system real dynamic and its transitions. Therefore, the use of nonlinear analysis may help overcome this limitation.

Using nonlinear methods with applicability to short and noisy data with the aim of representing non-evident changes in physiological behavior of human body system are novel approaches in neurosciences [9]. Detection of weak transition in signal dynamics was done using recurrence time statistics and especially for transitions with very low energy [10]. Recurrence quantification analysis (RQA) is a tool for the representation of similarities and dissimilarities of signals that cannot easily be seen in time domain data. This analysis is based on dimension reconstruction according to Takens' theory. According to this theory, by considering a time series that is sampled from system behavior in one observable dimension, the reconstruction of a scope about multidimensional behavior is possible.

The base of the field of nonlinear dynamics is the representation of trajectories in their phase space [11]. Recurrence plots are new methods based on the nonlinear analysis that have been developed in the last decade. Recurrence plot (RP) represents the times at which states in a phase space recur. It enables us to investigate the *m*-dimensional phase space trajectory through a two dimensional representation of its recurrence [11]. Most of the RP related methods are based on quantifying nearest neighbors in phase space. It has been shown that two types of recurrence points exist: true and sojourn. Therefore two types of the recurrence time can be defined T1 and T2 respectively [10,12]. The use of the sojourn point's concept and its relating recurrence times increased the hope for detecting system's state transition points.

A novel study has shown that the Fuzzy nonlinear model can represent input–output behavior of brainstem in generating cABR to /da/ [13]. Therefore, this can be acceptable to find nonlinear features that can be used for cABR signal classification.

In this paper, we try to use RQA nonlinear method for representing differences between cABR elicited by three different perceptually critical diotic stimuli /ba/, /da/ and /ga/. Exploring this feature can be a tool for the objective classification of these responses and a proof for the hypothesis of processing the role of brainstem.

2. Material and methods

2.1. Participants

Twenty-seven volunteer students from Tehran University of Medical Sciences (13 women and 14 men), aged from 22 to 29 years (mean ± SD = 24.34 ± 1.95), participated in this study. None of the subjects had a history of auditory, learning or neurologic problems. All students were monolingual Persian speakers by self-report and pure tone hearing thresholds for both ears were equal to or better than 20 dB HL for octave frequencies 250–8000 Hz. Subjects gave written consent to participate intensively in the study. All procedures were approved by the deputy of research review board and ethics community of Tehran University of Medical Sciences.

2.2. Stimuli

In this research, three different 170 ms one syllable consonant-vowel synthesized stimuli including /ba/, /da/ and /ga/ were presented to each person at a sampling rate of 20 kHz. These three stimuli were obtained from Kraus and colleagues auditory neuroscience laboratory at Northwestern University and are the same with previous research undertaken by them [1]. Fig. 1 shows time–frequency specifications of these three stimuli schematically.

Stimuli durations are 170 ms with voicing (100 Hz F0) onset at 10 ms. The formant transition durations are 50 ms and comprise a linearly rising F1 (400–720 Hz), flat F4 (3300 Hz), F5 (3750 Hz) and F6 (4900 Hz). Ten milliseconds of initial frication are centered at frequencies around F4 and F5. After the 50 ms formant transition period, F2 and F3 remain constant at their transition end point frequencies of 1240 and 2500 Hz, respectively, for the remainder of the syllable. The stimuli differ only in the starting points of F2 and F3. For [ba], F2 and F3 rise from 900 Hz and 2400 Hz, respectively. For [da], F2 and F3 fall from 1700 and 2580, respectively. For [ga], F2 and F3 decrease from 3000 and 3100, respectively. These synthesized stimuli have an identical and constant F0 throughout their entire duration [1].

The diotic stimuli were delivered with a high precision synchronized stimuli delivery system that was designed and manufactured for simultaneous presentation of the same sound to each ear. The rate of presentation was 4.65/s. Both stimuli polarities (condensation and rarefaction) were presented. The test stimuli were presented to both ears through Etymotic ER-30 earphone (Etymotic Research, Elk Grove Village, IL) at an intensity of 83 dB SPL. To ensure subject cooperation and promote stillness, all subjects watched videotaped programs such as movies or cartoons of their choice. They were instructed to attend to the video rather than to the stimuli.

2.3. EEG acquisition and analysis

2.3.1. cABR signal extraction

G.tec EEG recording system was used to record evoked potentials synchronized with auditory stimuli from Cz-to-ipsilateral earlobe, with forehead as ground and digitized at 19,200 Hz.

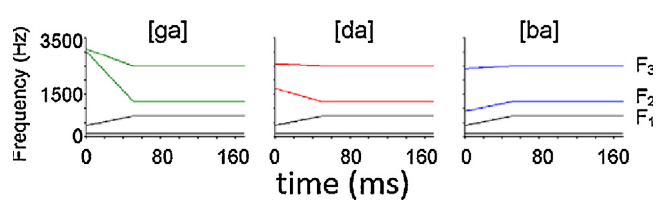

Fig. 1. Schematic representations of the spectral composition of [ga] (left), [da] (center), and [ba] (right) stimuli as a function of time (in ms). The first 50 ms are the formant transition period, followed by the steady-state [a] portion [1].

Please cite this article in press as: Jafarpisheh AS, et al. Nonlinear feature extraction for objective classification of complex auditory brainstem responses to diotic perceptually critical consonant-vowel syllables. Auris Nasus Larynx (2015), http://dx.doi.org/10.1016/j.anl.2015.06.003