Neuroscience Letters 665 (2018) 257–262

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Research article

Phase delays between tone pairs reveal interactions in scalp-recorded envelope following responses

Vijayalakshmi Easwar[a,*], Ashlee Banyard[b], Steven Aiken[c], David Purcell[a,b]

[a] National Centre for Audiology, Western University, London, Ontario, Canada N6G 1H1
[b] School of Communication Sciences and Disorders, Western University, London, Ontario, Canada N6G 1H1
[c] School of Human Communication Disorders, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2

ARTICLE INFO

Keywords:
Vowel simulation
Auditory steady-state response

ABSTRACT

Evoked potentials to envelope periodicity in sounds, such as vowels, are dependent on the stimulus spectrum. We hypothesize that phase differences between responses elicited by multiple frequencies spread tonotopically across the cochlear partition may contribute to variation in scalp-recorded amplitude. The present study evaluated this hypothesis by measuring envelope following responses (EFRs) to two concurrent tone pairs, p1 and p2, that approximated the first and second formant frequencies of a vowel, while controlling their relative envelope phase. We found that the scalp-recorded amplitude of EFRs changed significantly in phase and amplitude when the envelope phase of p2, the higher frequency tone pair, was delayed. The maximum EFR amplitude occurred at the p2 envelope phase delay of 90°, likely because the stimulus delay compensated for the average phase lead of 73.57° exhibited by p2-contributed EFRs relative to p1-contributed EFRs, owing to earlier cochlear processing of higher frequencies. Findings suggest a linear superimposition of independently generated EFRs from tonotopically separated pathways. This suggests that introducing frequency-specific delays may help to optimize EFRs to broadband stimuli like vowels.

1. Introduction

Envelope following responses (EFRs), measured using scalp electrodes, are neural responses phase-locked to the envelope periodicity of an auditory stimulus. EFRs can be elicited by narrowband stimuli such as amplitude-modulated tones as well as broadband stimuli like amplitude-modulated noise and vowels. EFRs are thought to be initiated by non-linearities in the inner ear, and are elicited at the frequency of the envelope periodicity [1]. Recent work has shown that the amplitude of EFRs elicited by broadband sounds like vowels depends on the frequency and level of vowel formants [2]. The dependence on spectral characteristics may, in part, arise from interactions between EFRs generated at the same periodicity rate from different frequency components of the broadband stimulus such as between any pair of voice harmonics [3]. Phase differences in EFRs from different stimulus frequencies may emerge from the temporal dispersion of frequency-specific processing in the cochlea. Higher frequencies are processed earlier than lower frequencies due to the traveling time incurred by the lower frequencies to reach their optimal place of excitation at the cochlear apex [4]. These frequency-dependent processing delays may persist in the tonotopically arranged inferior colliculus (IC) where EFRs to periodicity rates of ~90 Hz are predominantly generated [5,6]. The between-stimulus frequency delays may lead to differences in the onset phase of EFRs generated across stimulus frequencies, thus leading to in- and out-of-phase interactions at the same response periodicity rate when recorded at the scalp. The present study tested this hypothesis by measuring EFRs to two tone pairs, p1 and p2, mimicking consecutive harmonics at the first and second formants of the vowel /ɛ/, while controlling the relative envelope phase of p2, the higher frequency tone pair. The vowel /ɛ/ was simulated because EFRs elicited by this vowel were lower in amplitude relative to other vowels (/i/, /ɛ/, /æ/, /ɔ/ and /u/) despite being similarly audible [2], and a possible cause for lower amplitudes could be destructive phase interactions among concurrent responses. We hypothesized that: (1) the EFR amplitude will change as the relative phase between tone pair envelope changes since this will alter the phase relationship between responses elicited by each tone pair, and (2) the EFR amplitude will be maximum when the phase delay between the tone pairs compensates for the response phase difference between the EFRs elicited individually by each tone pair.

* Corresponding author at: National Centre for Audiology, 2262 Elborn College, Western University, London, Canada N6G 1H1.
E-mail address: veaswar@nca.uwo.ca (V. Easwar).

https://doi.org/10.1016/j.neulet.2017.12.014
Received 7 July 2017; Received in revised form 2 December 2017; Accepted 5 December 2017
Available online 06 December 2017
0304-3940/

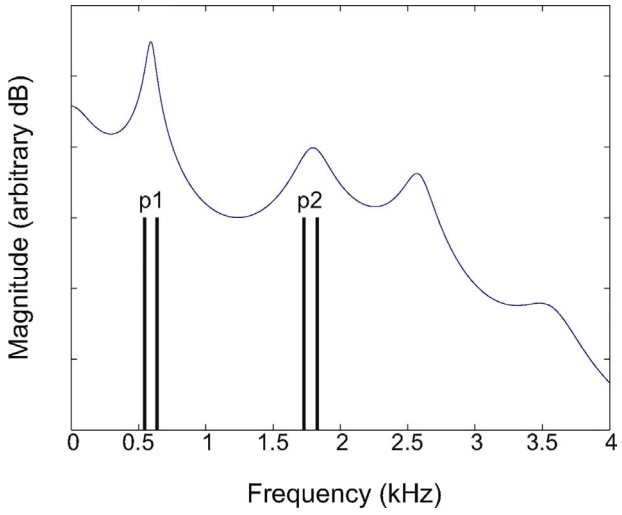

Fig. 1. Stimulus tones are shown in the frequency domain as equal height, black, vertical lines placed at frequencies indicated on the horizontal axis. The vertical axis is an arbitrary dB scale for illustration purposes. Tones were all equal magnitude for this initial investigation of potential EFR interactions. The blue line above the tone pairs represents the linear predictive coding (LPC) spectral envelope of the vowel /ɛ/ from the word "pet" in Choi et al. [2]. The tone pairs p1 and p2 were positioned as simple analogues of the first and second vowel formants, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2. Methods

2.1. Participants

Nineteen normal hearing adults (18–27 years) with hearing thresholds ≤20 dB HL participated in the study after providing written consent. The study protocol was approved by the Health Sciences Research Ethics Board of Western University.

2.2. Stimuli

Eight stimuli were used; each stimulus consisted of two equal-amplitude tone pairs, p1 and p2, that were intended to be simple analogues of the first two formants of /ɛ/ vowels from [2]. The formant frequencies were the median values from the three instances of /ɛ/ in the words "pet", "bed" and "said" used as stimuli in [2]. The first and second formant frequency estimates were 575 and 1814 Hz, respectively.

The first four stimuli (stimuli i–iv) consisted of p1 and p2 tone pairs at different beat frequencies (bf) whereas the final four stimuli (stimuli v–viii) consisted of p1 and p2 tone pairs at the same bf (see Fig. 1 for spectrum). Therefore, the first four stimuli elicited two EFRs simultaneously whereas the final four stimuli elicited one composite EFR. Tones in the p1 tone pair in stimuli i–iv, henceforth referred to as p1 bf 91 Hz, were 545.90 and 636.72 Hz. These tones were separated by 90.82 Hz, rounded to 91 Hz in this paper for brevity in reporting, to simulate the average fundamental frequency (f_0) of the /ɛ/ vowels in [2] (i.e., vowel harmonics separated by 91 Hz). The exact frequency of the tones had an integer number of cycles in each analysis epoch (1.024 s window). Tones in the p2 tone pair of these four stimuli, henceforth referred to as p2 bf 99 Hz, were 1728.52 and 1827.15 Hz. These tones were separated by 98.63 Hz, rounded to 99 Hz in this paper for brevity in reporting, to simulate the median f_0 of 98 Hz in the stimulus "pet". While the cosine onset envelope phases of the p1 tone pairs remained constant across stimuli i–iv, the envelope phases of p2 tone pairs were 0°, 90°, 180°, and 270° in the four stimuli i–iv, respectively (see Fig. 2 for stimulus envelope phases). This was achieved by progressively delaying tone onsets by $\frac{1}{4}$ cycles of the stimulus envelope (i.e., increments of ~2.5 ms). For stimuli i–iv, the p2 bf 99 Hz was designed to demonstrate that the EFR follows stimulus envelope phase without any significant changes in response amplitude. The difference in EFR frequency between the responses elicited by the p1 and p2 tone pairs was exactly eight fast Fourier transform (FFT) bins.

In the final four stimuli (stimuli v–viii), henceforth referred to as p1 + p2 bf 99 Hz, tones in the p2 tone pair were identical to those in stimuli i–iv. That is, the tone frequencies were 1728.52 and 1827.15 Hz, and the cosine onset envelope phase varied between 0° (stimulus v) and 270° (stimulus viii). In stimuli v–viii, the tone frequencies of the p1 tone pair were fixed at 545.90 and 644.53 Hz to match the bf of the p2 tone pair, and the envelope phase of the p1 tone pair was held constant. This was designed to evaluate variations in the amplitude and phase of the composite EFR when EFRs at the same bf are initiated from distinct stimulus frequencies (i.e., approximately the first and second formants of /ɛ/).

The eight stimuli formed one stimulus sweep of duration 8.864 s. Tones had a steady-state portion of ~1.024 s with additional linear onset and offset ramps of 36 ms.

2.3. Stimulus presentation and response recording

Experiments were controlled using software developed in LabVIEW (v8.5; National Instruments, TX) with a National Instruments PCI-6289 M-series acquisition card. Stimuli were generated with 16-bit resolution at 32,000 samples/second, and responses were recorded with 18-bit resolution at 8000 samples/second. Acoustic signals were produced by an Etymotic ER2 earphone shielded with Mu metal. Stimulus levels were set at 70 dB SPL using a Tucker-Davis Technologies PA5 attenuator and SA1 power amplifier. A Brüel and Kjær 2250 sound level meter and Type 4157 ear simulator, (Nærum, Denmark) were used for calibration in flat(Z)-weighted Leq mode.

Stimuli were presented without gaps for 400 sweeps (59 min) to a single ear chosen randomly for each participant (right ear tested in 10 participants), with the earphone sealed in the canal with a foam tip. Three disposable MEDI-TRACE Ag/AgCl electrodes were located at the vertex (non-inverting) and just below the hairline at the posterior midline of the neck (inverting) with a ground (or common) on the collarbone. Electrode sites were cleaned with Nuprep to ensure electrode impedances were ≤5 kΩ at 30 Hz with inter-electrode differences ≤2 kΩ, which were measured with an F-EZM5 GRASS impedance meter. Brain activity was conditioned using a GRASS LP511 EEG amplifier with a gain of 50 k and bandpass filtering from 3 to 3000 Hz. An additional gain of 2 was applied by the PCI-6289 card for a total gain of 100 k. Participants were encouraged to sleep while reclined in a comfortable chair located in an electromagnetically-shielded sound booth. The possibility of stimulus artifact was investigated by delivering the stimulus to a Zwislocki ear simulator while positioning all electrodes, leads, and cables in as similar position as possible to standard data collection. One participant and one phantom (electrodes suspended in water with impedances ~1.5 kΩ) showed no large signals at the response frequencies. In the phantom head, the maximum amplitudes recorded at 91 and 99 Hz were 3.54 and 1.84 nV, respectively. In the human participant, the maximum amplitudes recorded at 91 and 99 Hz were 14.84 and 8.41 nV, respectively. It is therefore unlikely that stimulus artifact contributed substantially to the reported response amplitudes.

2.4. Response analysis

Analysis was completed offline using an FFT implemented in MATLAB (MathWorks, Natick, MA). For the purposes of reducing myogenic artifacts, a noise metric was calculated for each epoch as the average amplitude in a frequency band that included the response frequencies between 80 and 240 Hz. Epochs with noise metrics smaller than the mean noise metric +2 SDs (called the artifact rejection