



Research paper

Auditory event-related responses to diphthongs in different attention conditions

David J. Morris^{a,*}, Kurt Steinmetzger^b, John Tøndering^a^a Department of Nordic Studies and Linguistics, University of Copenhagen, Copenhagen-S, Denmark^b Department of Speech, Hearing and Phonetic Sciences, University College London, UK

HIGHLIGHTS

- The direction of a diphthong yields differences in the P1-N1-P2 responses.
- N1-P2 responses to diphthongs occur even when attention is on a visual task.
- P2 vertex amplitude evoked with a diphthong varies with attention condition.

ARTICLE INFO

Article history:

Received 8 December 2015

Received in revised form 4 May 2016

Accepted 5 May 2016

Available online 6 May 2016

Keywords:

ERP

Attention

N1

P2

Acoustic change complex

Vowel perception

ABSTRACT

The modulation of auditory event-related potentials (ERP) by attention generally results in larger amplitudes when stimuli are attended. We measured the P1-N1-P2 acoustic change complex elicited with synthetic overt (second formant, $F2\Delta = 1000$ Hz) and subtle ($F2\Delta = 100$ Hz) diphthongs, while subjects (i) attended to the auditory stimuli, (ii) ignored the auditory stimuli and watched a film, and (iii) diverted their attention to a visual discrimination task. Responses elicited by diphthongs where $F2$ values rose and fell were found to be different and this precluded their combined analysis. Multivariate analysis of ERP components from the rising $F2$ changes showed main effects of attention on P2 amplitude and latency, and N1-P2 amplitude. P2 amplitude decreased by 40% between the attend and ignore conditions, and by 60% between the attend and divert conditions. The effect of diphthong magnitude was significant for components from a broader temporal window which included P1 latency and N1 amplitude. N1 latency did not vary between attention conditions, a finding that may be related to stimulation with a continuous vowel. These data show that a discernible P1-N1-P2 response can be observed to subtle vowel quality transitions, even when the attention of a subject is diverted to an unrelated visual task.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

As Simons and Chabris' invisible gorilla convincingly demonstrates [23], when we direct our attention to a specific occurrence, other events can go unnoticed. Functional descriptions of the modulation of perception by attention attribute a central role to salience filters that monitor perceptual input, and bias signals that regulate sensitivity control by priming neural populations in sensory cortices, see [13] for a review. Sensitivity and neural responsiveness is upregulated when a perceptual decision must be rendered,

Abbreviations: dB HL, deciBel hearing level; EEG, electroencephalography; ERP, event-related potential; $F2$, second formant; MMN, mismatch negativity; RT, response time.

* Corresponding author at: Department of Nordic Studies and Linguistics, University of Copenhagen, Njalsgade 120, Copenhagen-S 2300, Denmark.

E-mail address: dmorris@hum.ku.dk (D.J. Morris).

<http://dx.doi.org/10.1016/j.neulet.2016.05.002>

0304-3940/© 2016 Elsevier Ireland Ltd. All rights reserved.

and this may co-occur with downregulation of sensitivity for neural populations tuned to non-task relevant stimuli [10,27]. This complimentary co-regulation of top-down sensitivity is evident in cross-modal electrophysiology studies, where the introduction of visual stimuli typically results in a decrease in the amplitude of late auditory event-related potentials (ERPs). For example, the effect of simultaneously-performed tasks on auditory ERPs was investigated by Lavoie et al. [14] in presentation conditions which were: eyes open; eyes closed; reading a book; watching a movie, and; adding 3 digit numbers. N1-P2 amplitudes to click stimuli from 10 normal hearing listeners were reported to be lowest when arithmetic was performed concurrent with the presentation of auditory stimulus, and highest when subjects performed no task and had their eyes closed. Oray et al. [20] presented loud white noise tone bursts alone and also asynchronously paired with visual stimuli which were black and white vertical bars. They found mean decreases in amplitude between auditory alone and auditory-visual conditions

that, when measured at the vertex, were 40% for N1 and 25% for P2. They attributed this attention-related reduction to a suppressive mechanism that benefited auditory perception when presented in the absence of visual stimuli.

The advantage that selective attention provides auditory perception has also been attributed to a temporal boost that promotes the neural processing of attended stimuli, an increase that may be indexed by N1 latency [21,26]. Folyi et al. [6] investigated this N1 latency acceleration, with ERP measurements to near-threshold tonebursts, in conditions where participants ignored the stimuli and watched a silent movie, or responded to stimulus onset. They found that N1 latencies decreased and amplitudes increased in the attend condition relative to the ignore condition. Effects of attention on auditory ERPs have also been investigated in mismatch negativity (MMN) studies. While the auditory MMN, elicited with fine frequency changes between standard and deviant tones, was found to decrease in amplitude when subjects sat passively or read a book [18], it was not affected by the level of difficulty of a simultaneously performed visual discrimination task [19].

In the present study we examined the P1-N1-P2 complex recorded in response to synthesized diphthongs, defined here as spectral changes that alter the quality of the vowel. It has been demonstrated that when recorded with speech stimuli, the P1-N1-P2 is stable over test sessions [28] and it can be elicited with synthetic speech [1,32], although component peaks can have shorter latencies relative to elicitation with natural speech stimuli [25]. The synthetic vowels that we used were presented to subjects in a stimulus paradigm akin to the ‘alternating strategy’ used in Martin et al. [17]. It has been suggested that measurement paradigms involving continuous acoustic stimulation with alternating changes in a-b-a-b form have the potential to benefit the electrical signal-noise ratio of the response, as the number of samples in the average ERP may be doubled by combining a-b and b-a transitions. We also included a subtle diphthong, in order to examine how the P1-N1-P2 response to fine-grained vowel quality transitions were affected by different attention conditions. In an auditory-visual presentation condition we instructed subjects to divert their attention from the concurrently presented diphthongs, to a cognitive task which was visual discrimination of Japanese symbols. The visual discrimination task required subjects to attend to the spatial detail of the symbols, in this way competing for attentional resources with the vowel formant transitions, a competition that we hypothesised may override ERP markers of subtle diphthongs.

2. Material and methods

2.1. Subjects

Sixteen healthy subjects who were university students and staff participated in this study. Data from four subjects were eliminated from analysis due to excessive noise in the responses. The remaining 12 subjects (6 female) had a mean age of 29 years 6 mos (SD 9 years 7 mos). All subjects reported that their right hand was dominant and they used this hand to provide responses during the test session. They all had normal vision and normal hearing as revealed by audiometric screening (pure tone thresholds less than 25 dB HL at 250–4000 Hz, in both ears) and all provided informed consent prior to the experiment. All subjects were native Danish speakers and none of the subjects had any prior knowledge of Japanese Kanji orthography, which was used in the visual task.

2.2. Stimuli

The auditory stimuli were two one-second and two two-second audio files that were vowels synthesized from a Klatt grid in PRAAT [3]. Overt diphthongs, where second formant (F2) values increased



Fig 1. Example of stimuli from the visual task. The three Japanese Kanji ideograms were presented on a computer screen where two were the same and one was different.

by 1000 Hz and subtle diphthongs, where F2 increased by 100 Hz occurred in the middle portion of the audio files, while all other acoustic parameters remained constant. The onset and offset F2 value was 1000 Hz which corresponded approximately to cardinal vowel 15 as described in [8, pp. 12–13] or [ɜ:]. The vowel quality after the subtle transition was slightly fronted, while after the overt transition, it corresponded approximately to cardinal vowel two. This vowel quality is closest to an [ɛ:] in Danish. A compressed version of the one second stimuli are available supplemental to this article as audio files 1 and 2, for stimuli with F2 transitions to 1100 and 2000 Hz, respectively. Fig. 2 shows the spectral details of the vowel and the transitions, which were 40 ms in duration. The four audio files were presented randomly so that they formed a continuous triphthongal chain (interstimulus interval = 0) within which vowel quality transitions were followed by steady-state portions of 460 and 960 ms. The sound field presentation level of the continuous stimuli was 70 dB (A) and this varied by less than 1 dB across the 3 vowel qualities. In total, 1440 diphthongs were presented to each participant and presentation conditions were counterbalanced across subjects.

2.3. Presentation conditions

Subjects underwent EEG testing in three conditions: attend, ignore and divert. During the attend condition they were instructed to respond to the auditory stimuli when F2 increased by 1000 Hz, that is, when the transition was from an [ɜ:] to an [ɛ:]. These rising overt diphthongs occurred irregularly due to the randomization of the one and two second stimulus files. In the ignore condition subjects watched a silent film. During the divert condition subjects performed a visual discrimination task involving three Japanese symbols, where two were the same and one was different (see Fig. 1). This was similar to the visual search task that was used in [2], where subjects were required to identify words from a grid of letters. However, we were hesitant about the possibility of interference between lexical items in combination with synthetic speech, and therefore created the stimuli for the visual task from five pairs of Kanji ideograms. Further motivation for the selection of this task in the divert condition is based on the assumption that it would require subjects to devote cognitive, particularly visuo-spatial, processing resources to the features of the Kanji symbols, but would only minimally task the subjects’ episodic working memory. Presentation of the visual trials was temporally independent of the diphthongal change, as we were interested in unmixing the crosstalk between the auditory ERP and the response from the visual task, an undertaking that would not be able to be accomplished if we had used sequential presentation of auditory and visual stimuli as, for instance, employed in [15].

During the divert condition subjects were instructed to devote their attention to the visual task and select as quickly and accurately as possible the symbol that was different by pressing the corresponding button on a response box. Prior to testing, subjects completed a training round with 12 visual trials in the absence of auditory stimuli, during which onscreen feedback to responses was provided as correct or incorrect. Feedback was not provided dur-

Download English Version:

<https://daneshyari.com/en/article/4343195>

Download Persian Version:

<https://daneshyari.com/article/4343195>

[Daneshyari.com](https://daneshyari.com)