NEUROSCIENCE
INTERNATIONAL BRAIN IBRO RESEARCH ORGANIZATION

RESEARCH ARTICLE

Z. Xie et al. / Neuroscience xxx (2018) xxx–xxx

Taking Attention Away from the Auditory Modality: Context-dependent Effects on Early Sensory Representation of Speech

Zilong Xie,[a] Rachel Reetzke[a] and Bharath Chandrasekaran[a,b,c,d,e]*

[a] *Department of Communication Sciences and Disorders, The University of Texas at Austin, Austin, TX 78712, USA*
[b] *Department of Psychology, The University of Texas at Austin, Austin, TX 78712, USA*
[c] *Department of Linguistics, The University of Texas at Austin, Austin, TX 78712, USA*
[d] *Institute for Neuroscience, The University of Texas at Austin, Austin, TX 78712, USA*
[e] *Institute for Mental Health Research, The University of Texas at Austin, Austin, TX 78712, USA*

Abstract—Increasing visual perceptual load can reduce pre-attentive auditory cortical activity to sounds, a reflection of the limited and shared attentional resources for sensory processing across modalities. Here we demonstrate that modulating visual perceptual load can impact the early sensory representation of speech sounds, and that the impact of visual load is highly dependent on the predictability of the incoming speech stream. Participants (n = 20, 9 females) performed a visual search task of high (target similar to distractors) and low (target dissimilar to distractors) perceptual load, while early auditory electrophysiological responses were recorded to task-irrelevant native speech sounds. Speech sounds were presented either in a repetitive context, or a less predictable 'variable' context. Independent of auditory stimulus context, pre-attentive auditory cortical activity was reduced during high visual load, relative to low visual load. We applied a novel data-driven machine learning approach to decode speech sounds from the early auditory electrophysiological responses. Decoding performance was found to be *poorer* under conditions of high (relative to low) visual load, only when the incoming acoustic stream was predictable. When the auditory stimulus context was less predictable, decoding performance was substantially *greater* for the high (relative to low) visual load conditions. Our results provide support for shared attentional resources between visual and auditory modalities that substantially influence the early sensory representation of speech signals in a context-dependent manner. **© 2018 IBRO. Published by Elsevier Ltd. All rights reserved.**

Key words: visual perceptual load, auditory predictability, frequency-following response, machine learning.

INTRODUCTION

At any waking moment, our brains are bombarded with sensory information from multiple modalities. The central nervous system is constantly extracting meaningful patterns or regularities from the incoming stimuli (e.g., Winkler et al., 2009; Stefanics et al., 2014). Sensory systems adjust their response properties based on the stimulation context of the sensory inputs (Nelken and Ulanovsky, 2007). At the same time, we often find ourselves focusing attention on a task in one modality while ignoring information from other sensory modalities. This raises two important questions: to what extent does allocating attentional resources to one modality preclude the processing of task-irrelevant stimuli in another modality, and to what extent does such attentional modulation interact with the processing of regularities of task-irrelevant stimuli? In the current study, we addressed these questions by examining the influence of attending to a visual task on the processing of task-unrelated auditory stimuli with varying predictability.

Per load theory, visual and auditory processing share capacity-limited neural resources, and the depletion of resources by one modality diminishes available resources to the other modality (Lavie, 2005; Klemen et al., 2009). Behavioral studies have provided support for load theory, revealing the impact of modulating visual load on audition. For example, crossmodal studies show that visual tasks of high perceptual load are associated with a reduced detection sensitivity of task-irrelevant auditory stimuli, demonstrating load-induced *inattentional deafness* (Macdonald and Lavie, 2011; Raveh and

*Correspondence to: B. Chandrasekaran, The University of Texas at Austin, 2504A Whitis Ave. (A1100), Austin, TX 78712, USA. Fax: +1-512-471-2957.
E-mail address: bchandra@utexas.edu (B. Chandrasekaran).
Abbreviations: ABR, auditory brainstem response; ANOVA, analysis of variance; EEG, electroencephalography; ERP, event-related potential; F0, fundamental frequency; FFR, frequency-following response; IC, inferior colliculus; MEG, magnetoencephalography; SSA, stimulus-specific adaptation; SVM, support vector machine.

https://doi.org/10.1016/j.neuroscience.2018.05.023
0306-4522/© 2018 IBRO. Published by Elsevier Ltd. All rights reserved.

Please cite this article in press as: Xie Z et al. Taking Attention Away from the Auditory Modality: Context-dependent Effects on Early Sensory Representation of Speech. Neuroscience (2018), https://doi.org/10.1016/j.neuroscience.2018.05.023

UNCORRECTED PROOF

Lavie, 2015). The neural mechanisms underlying these modulatory influences are still under debate. From a central processing account, while incoming sensory information is fully processed by both modalities at perceptual levels, the sensory information that gains prominence is determined at a more central level of processing (Deutsch and Deutsch, 1963; Dux et al., 2006). However, a recent magnetoencephalography (MEG) study revealed substantial reduction in the early auditory cortical evoked activity to task-irrelevant auditory stimuli during high, relative to low, visual load (Molloy et al., 2015). This suggests that crossmodal influences can potentially be discerned even during the *encoding* stage of auditory processing. We directly test the hypothesis that visual load can influence the early sensory representation of auditory signals by examining the frequency-following response (FFR), a scalp-recorded 'neurophonic' component that faithfully captures phase-locked activity to periodic stimuli (Smith et al., 1975; Chandrasekaran and Kraus, 2010).

FFRs have been extensively used to index the fidelity of early, pre-attentive speech encoding in humans (Chandrasekaran and Kraus, 2010). The scalp-recorded FFRs derived from electroencephalography (EEG) are hypothesized to reflect activity primarily from subcortical auditory ensembles (Smith et al., 1975; Chandrasekaran and Kraus, 2010; Bidelman, 2015). However, there may also be cortical contributions to the FFR as well (Coffey et al., 2016). A recent proposal suggests that the FFRs reflect an integrated, dynamic interplay between pre-attentive cortical and subcortical circuitry (Kraus and White-Schwoch, 2015). This interplay is facilitated by ascending, as well as corticofugal pathways (Suga, 2008; Bajo and King, 2015; Malmierca et al., 2015).

Prior work has demonstrated that FFRs to speech stimuli are highly sensitive to stimulus context. Specifically, evidence has shown that sensory fidelity is enhanced for speech sounds presented in predictive contexts relative to less predictable contexts (Chandrasekaran et al., 2009; Slabu et al., 2012; Lau et al., 2016; Lehmann et al., 2016). The neural mechanism underlying context-dependent modulation of the FFRs is unclear. Two distinct mechanisms may be at play: (1.) corticofugal modulation that selectively enhances the representation of regularities in the signal via predictive processing; and (2.) novelty detection, which reflects predominantly local stimulus-specific adaptation (SSA) (Chandrasekaran et al., 2014).

Here we examined the impact of manipulating visual perceptual load on the FFRs to speech signals under predictable and less predictable contexts. Native Chinese speakers performed a visual search task of high or low perceptual load (Fig. 1B). On a random subset of trials, task-irrelevant Mandarin lexical tones (Fig. 1A) were presented in a predictable context or a variable context (Fig. 1C). We recorded early cortical evoked activity, along with FFRs to the tones, and utilized a novel machine learning approach to decode speech category (Mandarin tone) information from the FFRs. We evaluate the extent to which decoding performance, which reflects the fidelity of stimulus

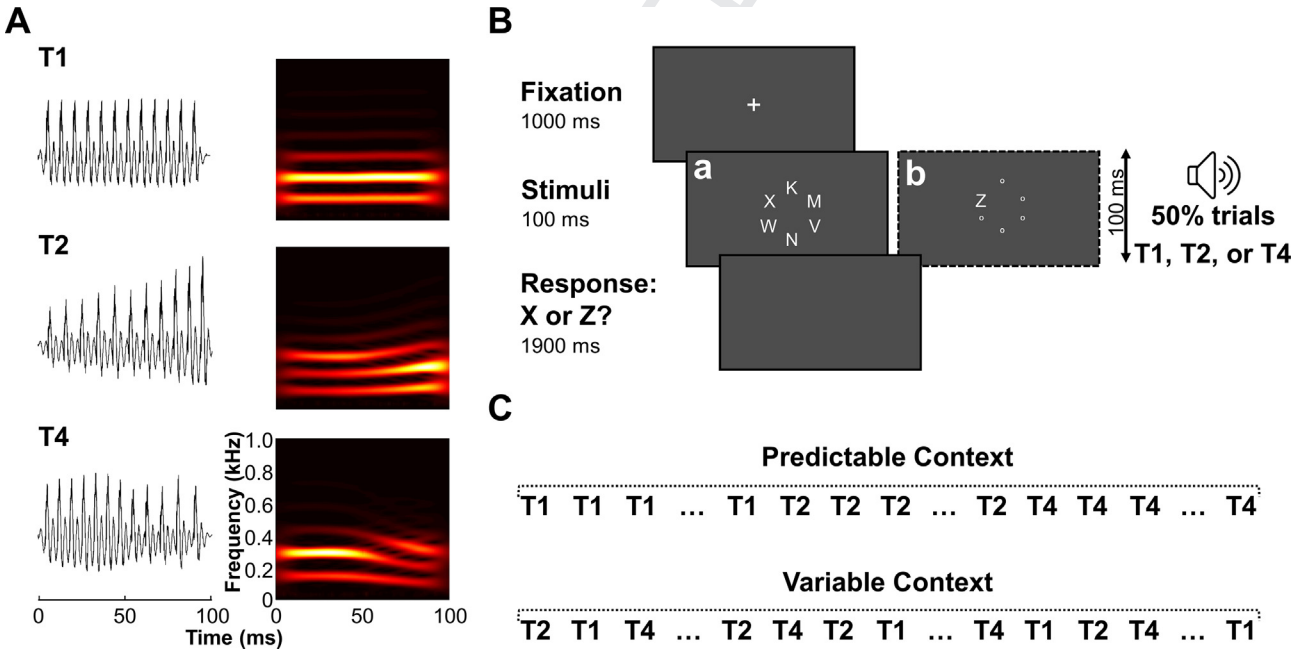

Fig. 1. Stimuli and task design. (A) Waveforms and spectrograms of the auditory stimuli, i.e., 100 ms Mandarin tones T1, T2, and T4. (B) Trial structure of the visual search task adapted from Molloy et al. (2015). Each trial began with a 1000-ms fixation cross at the center of the screen. Immediately after, a visual letter array of either high (a) or low (b) load was presented for 100 ms. On a random 50% of the trials, a 100-ms auditory stimulus (Mandarin tone T1, T2, or T4) was presented concurrently with the visual stimuli. In the remaining 50% trials, only the visual letter array was presented. After stimulus representation, a blank screen was presented for a maximum of 1900 ms, during which participants were instructed to identify the visual target (X or Z) as quickly and accurately as possible. Once participants made a response, the task moved to the next trial. (C) The auditory stimuli are presented in either predictable (top) or variable (bottom) contexts. In the predictable auditory contexts, the tones are presented in blocks within which each tone is presented repetitively. In the variable auditory contexts, the tones are presented in a random order. In both contexts, an equal number of each of the three Mandarin tones was used.

Please cite this article in press as: Xie Z et al. Taking Attention Away from the Auditory Modality: Context-dependent Effects on Early Sensory Representation of Speech. Neuroscience (2018), https://doi.org/10.1016/j.neuroscience.2018.05.023

Download English Version:

https://daneshyari.com/en/article/8840616

Download Persian Version:

https://daneshyari.com/article/8840616

Daneshyari.com