Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Is laughter a better vocal change detector than a growl?

Ana P. Pinheiro [a,b,*], Carla Barros [b], Margarida Vasconcelos [b], Christian Obermeier [c] and Sonja A. Kotz [c,d]

[a] Faculty of Psychology, University of Lisbon, Lisbon, Portugal
[b] Neuropsychophysiology Laboratory, School of Psychology, University of Minho, Braga, Portugal
[c] Dept. of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[d] Faculty of Psychology and Neuroscience, Dept. of Neuropsychology & Psychopharmacology, Maastricht University, The Netherlands

ARTICLE INFO

Article history:
Received 19 September 2016
Reviewed 18 November 2016
Revised 26 January 2017
Accepted 27 March 2017
Action editor Ed Wilding
Published online 11 April 2017

Keywords:
Emotion
Voice
Prediction error
Mismatch negativity
Beta oscillations

ABSTRACT

The capacity to predict what should happen next and to minimize any discrepancy between an expected and an actual sensory input (prediction error) is a central aspect of perception. Particularly in vocal communication, the effective prediction of an auditory input that informs the listener about the emotionality of a speaker is critical. What is currently unknown is how the perceived valence of an emotional vocalization affects the capacity to predict and detect a change in the auditory input. This question was probed in a combined event-related potential (ERP) and time-frequency analysis approach. Specifically, we examined the brain response to standards (Repetition Positivity) and to deviants (Mismatch Negativity – MMN), as well as the anticipatory response to the vocal sounds (pre-stimulus beta oscillatory power). Short neutral, happy (laughter), and angry (growls) vocalizations were presented both as standard and deviant stimuli in a passive oddball listening task while participants watched a silent movie and were instructed to ignore the vocalizations. MMN amplitude was increased for happy compared to neutral and angry vocalizations. The Repetition Positivity was enhanced for happy standard vocalizations. Induced pre-stimulus upper beta power was increased for happy vocalizations, and predicted the modulation of the standard Repetition Positivity. These findings indicate enhanced sensory prediction for positive vocalizations such as laughter. Together, the results suggest that positive vocalizations are more effective predictors in social communication than angry and neutral ones, possibly due to their high social significance.

* *Corresponding author*. Neuropsychophysiology Laboratory, CIPsi, School of Psychology, University of Minho, Campus de Gualtar, 4710-957, Braga, Portugal.
E-mail address: ana.pinheiro@psi.uminho.pt (A.P. Pinheiro).

http://dx.doi.org/10.1016/j.cortex.2017.03.018

1. Introduction

In a constantly changing environment humans face the challenge of having to prioritize sensations that compete for attention. Perception becomes more effective when sensory predictions are formed and updated based on the comparison of predicted and actual sensory feedback to minimize a prediction error (e.g., Arnal & Giraud, 2012). The automatic nature of such a mechanism plays a critical role in social communication: as much of the sensory input in our daily life has an affective tone, our capacity to effectively respond to unpredicted changes based on their emotional salience significantly contributes to effective social interactions (Jessen & Kotz, 2011; Jessen, Obleser, & Kotz, 2012).

In social communication, the voice represents one of the most relevant sound categories (Belin, Fecteau, & Bedard, 2004): it plays a pivotal role in conveying not only verbal information, but also important cues about the identity, age, and emotional state of a speaker (Belin, Bestelmeyer, Latinus, & Watson, 2011). However, when compared to the study of facial emotion expressions, fewer studies have investigated the neural basis of vocal emotion processing. The existing studies support a multi-stage model of vocal emotional perception and recognition (Paulmann & Kotz, 2008; Paulmann, Seifert, & Kotz, 2010; Schirmer & Kotz, 2006; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006). An open question is how human listeners automatically detect saliency in vocalizations that may signal a change from an expected vocalization, and how the valence expressed by the voice (i.e., its perceived pleasantness *vs*. unpleasantness – e.g., Bradley & Lang, 2000) influences this process. For instance, consider a mismatch between an angry vocalization and an utterance describing a happy event. The utterance will predict that the accompanying vocalization should be happy as well, but what you will hear is the opposite. The difference between how and when the vocal input occurs, and how it was expected to be is referred to as a prediction error leading to surprise, and a likely behavioral adaptation of a listener (e.g., Friston, 2012). As vocal information unfolds dynamically over time, the high temporal resolution of the electroencephalogram (EEG) is ideal to tackle these types of conflict in two ways: with a phase-locked evoked response and a non-phase locked oscillatory response. Specifically, pre-stimulus oscillatory activity may be better suited to probe how future auditory events are anticipated (e.g., Bernasconi, Manuel, Murray, & Spierer, 2011; van Ede, Jensen, & Maris, 2010), and therefore to shed light on the neurofunctional processes underlying the formation of a prediction.

1.1. *Detecting emotional change – insights from ERPs*

ERPs offer a unique glimpse into the temporal window of predictive effects in emotional voice processing. A commonly used electrophysiological event-related measure to estimate predictive processes is the Mismatch Negativity (MMN). The MMN is a negative ERP component that peaks at 100–250 msec after sound onset, and signals the preattentive change detection in the sound environment (e.g., Näätänen, 1995, 2001). In MMN experiments participants are instructed to ignore a stream of sounds that differ in probability (high-probability or *standard* sounds *vs*. low-probability or *deviant* sounds), and to focus their attention on a concurrent task such as watching a movie. Recent accounts of the functional significance of the MMN suggest that this component is a neurophysiological signature of predictive processing and, in particular, of a prediction error (e.g., Garrido, Kilner, Stephan, & Friston, 2009). Two important processes seem to be at play. On the one hand, the detection of regularity in an auditory scene is required: the automatic extraction of statistical regularities (i.e., a frequently presented stimulus or *standard* sound) leads to increased top-down expectations, thereby resulting in suppressed neural responsiveness to the expected sound. In other words, the information about a frequently occurring stimulus is stored in a memory representation that then can facilitate predictions about what will happen next in an auditory environment. On the other hand, in the case of change detection, the mismatch between the top-down expectation and the perceived sensory input (i.e., a low-probability stimulus or *deviant* sound) leads to a prediction error that enhances neural responsiveness to the unexpected sound. As such the MMN reflects the difference between top-down expectation and incoming bottom-up sensory signals, and represents a prediction error signal (Baldeweg, 2007; Garrido et al., 2009; Todd, Michie, Schall, Ward, & Catts, 2012; Wacongne, Changeux, & Dehaene, 2012; Winkler & Czigler, 2012).

It is worth noting that some of the studies that used a passive roving standard stimulation to probe predictive processing have also revealed repetition effects to standard sounds that predicted the MMN elicitation (Baldeweg, 2007; Costa-Faidella, Baldeweg, Grimm, & Escera, 2011; Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg, 2005). They showed that an increase in the number of stimulus repetitions resulted in an increase of the P50 and P2 amplitudes, which was termed 'Repetition Positivity'. These effects are typically observed in response to standard stimuli at frontocentral electrode sites from 50 to 250 msec post-stimulus onset (Baldeweg, 2007; Costa-Faidella et al., 2011; Haenschel et al., 2005). They are proposed to reflect a neurophysiological correlate of a suppressed prediction error due to more efficient top-down predictions (Baldeweg, 2007).

The MMN may indicate how a change in emotional voice quality is detected preattentively. However, only a few MMN studies have investigated vocal emotional perception. The existing evidence confirms a rapid categorization of vocalizations based on their emotional relevance. Automatic distinctions of emotional vocalizations indexed by the MMN may be based on a minimal amount of acoustic information, such as mean F0 and its variation over time (Leitman, Sehatpour, Garidis, Gomez-Ramirez, & Javitt, 2011). Schirmer and colleagues reported an earlier MMN peak latency for happy than for neutrally intoned pseudowords (Schirmer, Striano, & Friederici, 2005), and a larger MMN amplitude for angry relative to neutral meaningless syllables that was positively correlated to state anxiety (Schirmer & Escoffier, 2010). Chen and collaborators (Chen, Lee, & Cheng, 2014) described a MMN amplitude increase for pseudowords expressing disgust compared to happiness. Using magnetoencephalography (MEG), Thonnessen et al. (2010) observed increased activation