



## Stimulus complexity matters when you hear your own voice: Attention effects on self-generated voice processing



Tatiana Conde<sup>a,b</sup>, Óscar F. Gonçalves<sup>b,c,d</sup>, Ana P. Pinheiro<sup>a,b,e,\*</sup>

<sup>a</sup> Faculdade de Psicologia, Universidade de Lisboa, Lisbon, Portugal

<sup>b</sup> Neuropsychophysiology Lab, CIPsi, School of Psychology, University of Minho, Braga, Portugal

<sup>c</sup> Spaulding Center of Neuromodulation, Department of Physical Medicine & Rehabilitation, Spaulding Rehabilitation Hospital & Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

<sup>d</sup> Bouvé College of Health Sciences, Northeastern University, Boston, MA, USA

<sup>e</sup> Cognitive Neuroscience Lab, Department of Psychiatry, Harvard Medical School, Boston, MA, USA

### ARTICLE INFO

#### Keywords:

Self-generated voice  
Non-self voice  
Stimulus type  
Event-related potentials  
Attention  
P3

### ABSTRACT

The ability to discriminate self- and non-self voice cues is a fundamental aspect of self-awareness and subserves self-monitoring during verbal communication. Nonetheless, the neurofunctional underpinnings of self-voice perception and recognition are still poorly understood. Moreover, how attention and stimulus complexity influence the processing and recognition of one's own voice remains to be clarified. Using an oddball task, the current study investigated how self-relevance and stimulus type interact during selective attention to voices, and how they affect the representation of regularity during voice perception.

Event-related potentials (ERPs) were recorded from 18 right-handed males. Pre-recorded self-generated (SGV) and non-self (NSV) voices, consisting of a nonverbal vocalization (vocalization condition) or disyllabic word (word condition), were presented as either standard or target stimuli in different experimental blocks.

The results showed increased N2 amplitude to SGV relative to NSV stimuli. Stimulus type modulated later processing stages only: P3 amplitude was increased for SGV relative to NSV words, whereas no differences between SGV and NSV were observed in the case of vocalizations. Moreover, SGV standards elicited reduced N1 and P2 amplitude relative to NSV standards.

These findings revealed that the self-voice grabs more attention when listeners are exposed to words but not vocalizations. Further, they indicate that detection of regularity in an auditory stream is facilitated for one's own voice at early processing stages. Together, they demonstrate that self-relevance affects attention to voices differently as a function of stimulus type.

### 1. Introduction

From the first instants after birth to late adulthood, human beings are exposed to their own voice more than to any other type of sound. One's own voice is a socially relevant acoustic signal through which a wealth of critical information (e.g., sex, age, health, identity, affective state) is conveyed to social partners (Kreiman and Sidtis, 2013; Sidtis and Kreiman, 2012). Despite the role played by self-voice recognition mechanisms in successful vocal communication, research in this field has been challenged by methodological issues. These include perceptual differences in self-voice perception when producing speech (due to the presence of bone-conducted sound) vs. when passively listening to pre-recorded self-generated speech (Maurer and Landis, 1990). Notwithstanding, the accurate recognition of pre-recorded self-voice stimuli

seems to occur above chance (Nakamura et al., 2001; Rosa et al., 2008a, b; Xu et al., 2013; Pinheiro et al., 2016), showing that individuals can compensate for such perceptual disparities.

#### 1.1. How special is the self-voice?

Alterations in self-voice processing mechanisms may impair verbal communication (Lane and Webster, 1991; Moeller et al., 2007), and have been implicated in psychopathological symptoms such as auditory verbal hallucinations (e.g., Waters et al., 2012; see Conde et al., 2016a for a review). However, the neurofunctional mechanisms underpinning self-voice perception remain to be clarified. The existing studies have indicated important differences in how self- and unknown voices are perceived (Allen et al., 2005; Graux et al., 2013; Graux et al., 2015;

\* Corresponding author at: Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, Portugal.

E-mail address: [appinheiro@psicologia.ulisboa.pt](mailto:appinheiro@psicologia.ulisboa.pt) (A.P. Pinheiro).

<https://doi.org/10.1016/j.ijpsycho.2018.08.007>

Received 8 June 2017; Received in revised form 5 June 2018; Accepted 10 August 2018

Available online 13 August 2018

0167-8760/© 2018 Elsevier B.V. All rights reserved.

Kaplan et al., 2008; Nakamura et al., 2001; Rosa et al., 2008a, b; Xu et al., 2013). For example, an improved recognition of the self-voice was demonstrated in acoustically demanding conditions, in which the voice signal only kept frequencies higher than the third formant (Xu et al., 2013). Also, an early discrimination between self- and non-self voice cues, occurring within 70–100 milliseconds (ms) post-stimulus onset, was revealed by event-related potential (ERP) studies (Graux et al., 2013). Furthermore, compared to unfamiliar voices, listening to the self-voice elicited increased activation in the left inferior frontal and right anterior cingulate gyri (Allen et al., 2005), right inferior frontal gyrus (Kaplan et al., 2008; Nakamura et al., 2001) and right parainsular brain regions (Nakamura et al., 2001). Self-related stimuli were also found to elicit prioritized processing. For instance, attentional enhancement by different categories of self-related stimuli (e.g., self-face/name/hand) has been consistently reported (e.g., Berlad and Pratt, 1995; Gray et al., 2004; Scott et al., 2005; Su et al., 2010; Sui et al., 2006; Tacikowski and Nowicka, 2010; Tacikowski et al., 2014). A memory advantage was additionally observed for information encoded in a self- (vs. non-self) referential manner (Symons and Johnson, 1997). Nonetheless, fewer studies examined how self-relevant stimuli modulate attentional resources during voice perception.

### 1.2. Does attention affect how the self-voice is perceived?

In dynamically changing multi-speaker contexts, vocal communication also demands a selective and flexible allocation of attentional resources to feedback generated by oneself or by others (Fritz et al., 2007a, b; Rimmele et al., 2015). Even though attention was found to modulate how one's own voice is perceived during speech production (Hu et al., 2015; Tumber et al., 2014), it remains to be clarified whether these effects are dependent on motor processes (associated with voice generation) or whether they extend to self-voice perception in general (i.e., when pre-recorded self-voice stimuli are passively presented). In a recent ERP study, we demonstrated that selective attention to voices is modulated by self-relevance, which was reflected in an increased P3 amplitude to self-compared to non-self speech (i.e., a disyllabic word) when stimuli were in the focus of attention (Conde et al., 2015). The P3 component is typically elicited by infrequent task-relevant events interspersed among frequent stimuli, and it is believed to reflect the mobilization of higher-order attentional resources after the evaluation of stimulus significance (Kok, 2001; Polich, 2007; Spencer et al., 1999, 2001). Together with previous studies focusing on other categories of self-relevant stimuli (e.g., Berlad and Pratt, 1995; Gray et al., 2004; Tacikowski and Nowicka, 2010), it is plausible that the self-voice grabs more attention due to its higher emotional salience.

Nonetheless, even when task-irrelevant (i.e., when participants are engaged in a visual distractive task), self- and unknown voices were found to distinctly modulate attention orienting in the P3 latency window (Graux et al., 2013; Graux et al., 2015). Specifically, the P3a<sup>1</sup> amplitude was increased to both familiar and unfamiliar vocalizations relative to self-generated vocalizations (Graux et al., 2013, 2015). As this ERP component is believed to reflect involuntary capture of attention towards an unexpected change in an otherwise regular environment (Friedman et al., 2001), these findings show that attention orienting is enhanced for (task-irrelevant) non-self voice cues. This is not surprising as, in a social context, a novel or totally unexpected voice

signals the presence of a conspecific who might be approached or avoided. In this specific context, the non-self voice may become more salient. Altogether, the studies mentioned above suggest that attentional demands, i.e. ignoring (Graux et al., 2013, 2015) vs. attending (Conde et al., 2015) the voice, modulate the perceived salience of one's own voice. Nevertheless, these studies do not clarify how the self-voice is perceived when in the focus of attention (as it often happens during daily conversations), as well as whether stimulus type affects how one's own voice is discriminated.

### 1.3. Does stimulus complexity matter when the self-voice is perceived?

Differences in stimulus complexity should be considered when interpreting findings of self-voice perception studies. For example, the multidimensional model of voice perception (Belin et al., 2004; Belin et al., 2011) predicts that both linguistic and nonlinguistic (i.e., identity and affective) vocal cues are processed by partially dissociated but interacting brain pathways. Consistent with this model, critical bi-directional interactions between verbal and nonverbal cues were demonstrated during speaker's recognition (Fleming et al., 2014; Nygaard and Pisoni, 1998; Remez et al., 1997; Schweinberger et al., 1997; Zarate et al., 2015). Specifically, speaker recognition was found to be improved with increased stimulus duration (Cook and Wilding, 1997; Schweinberger et al., 1997), as well as with increased phonetic variability (Roebuck and Wilding, 1993).

Evidence for the effects of signal complexity on how one's own voice is processed is still scarce and limited to the realm of speech production. In one of such studies, Ventura et al. (2009) used magnetoencephalography (MEG) to investigate the effects of stimulus complexity on auditory cortical responsiveness (indexed by the M100<sup>2</sup> component) to self-generated sensory feedback during voice production (Ventura et al., 2009). Importantly, Ventura et al. (2009) demonstrated that the magnitude of auditory cortical suppression to one's own voice feedback (reflected in diminished M/N100 amplitude) depends on stimulus complexity: less complex vocal sounds (/a/) elicited larger M100 amplitude attenuation than more complex and dynamic self-voice stimuli (/a-a-a/ and /a-a-a-a/). Other studies revealed that when participants are presented with experimentally induced changes in voice feedback during speech production (e.g., increased voice F0), they tend to vocally compensate in the direction opposed to the changes introduced, i.e., they lower their voice F0 (e.g., Burnett et al., 1998; Chen et al., 2013). The magnitude of the compensatory vocal responses is modulated both by word content (Patel and Schell, 2008) and by language experience (participants' native language — Liu et al., 2010).

Even though the role of acoustic complexity on self-voice processing has been highlighted by the studies mentioned above, evidence is limited to experimental designs involving voice generation and short vocalizations. However, two important limitations should be noted. First, such designs are typically concerned with how one's own voice is perceived when vocalizing vs. when passively listening to a recording of the same voice (i.e., a self- vs. self-voice comparison), and not with how the self-voice is distinguished from non-self voice cues. Second, neuroscience techniques (e.g., EEG/ERP, MRI/fMRI) are highly sensitive to physiological artifact noise (e.g., muscle activity, eye movements), which limits the development of online voice production tasks with more complex speech stimuli beyond the steady vowel /a/ used in a

<sup>1</sup> The P3a and the P3b components are dissociable brain potentials that reflect distinct neurocognitive processes (see Polich, 2007). The P3a is thought to reflect the involuntary capture of attention by an unpredictable violation of a regular and invariant aspect of the environment. The P3b indexes the mobilization of higher-order attentional resources to a task-relevant deviant (target) event. As in the current experiment participants were asked to focus their attention on the vocal sounds and to silently count the infrequent (and task-relevant) vocal stimuli, henceforth we used the term “P3” to refer to the sub-component elicited by the task-relevant (target) stimuli.

<sup>2</sup> In voice perturbation tasks, the M100 component (magnetic counterpart of the N1 ERP component) is thought to reflect the operation of an internal predictive mechanism: when sensory feedback matches the prediction, auditory cortical suppression (reflected in reduced M100/N1 amplitude to self-compared to non-self or to altered self-voice feedback) is observed; however, an error signal is generated when the incoming self-voice feedback deviates from the prediction (Behroozmand and Larson, 2011; Heinks-Maldonado et al., 2005; Sitek et al., 2013; Hickok et al., 2011).

Download English Version:

<https://daneshyari.com/en/article/11023447>

Download Persian Version:

<https://daneshyari.com/article/11023447>

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