



## Paying attention to my voice or yours: An ERP study with words



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

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

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

### ARTICLE INFO

#### Article history:

Received 2 February 2015

Received in revised form 27 July 2015

Accepted 28 July 2015

Available online 30 July 2015

#### Key-words:

Self-generated voice

Non-self voice

Event-related potentials

Attention

P3

### ABSTRACT

Self-related stimuli—such as one's own face or name—seem to be processed differently from non-self stimuli and to involve greater attentional resources, as indexed by larger amplitude of the P3 event-related potential (ERP) component. Nonetheless, the differential processing of self-related vs. non-self information using voice stimuli is still poorly understood. The present study investigated the electrophysiological correlates of processing self-generated vs. non-self voice stimuli, when they are in the focus of attention.

ERP data were recorded from twenty right-handed healthy males during an oddball task comprising pre-recorded self-generated (SGV) and non-self (NSV) voice stimuli. Both voices were used as standard and deviant stimuli in distinct experimental blocks. SGV was found to elicit more negative N2 and more positive P3 in comparison with NSV. No association was found between ERP data and voice acoustic properties.

These findings demonstrated an earlier and later attentional bias to self-generated relative to non-self voice stimuli. They suggest that one's own voice representation may have a greater affective salience than an unfamiliar voice, confirming the modulatory role of salience on P3.

© 2015 Elsevier B.V. All rights reserved.

### 1. Introduction

Self-recognition represents the capacity to identify the physical and psychological aspects of ourselves, such as one's own voice, face or autobiographical memories (Gallup, 1985; Gallup et al., 2014; Gillihan & Farah, 2005). This ability emerges early in human development, around 18–24 months of age (Keenan, Gallup, & Falk, 2003; Nielsen, Dissanayake, & Kashima, 2003), and is thought to play a pivotal role in self-awareness and in a sense of agency (Gallup, 1985; Gallup et al., 2014; Jeannerod, 2003; Keenan et al., 2003; Platek et al., 2008). Impairments in this ability have been observed in some neuropsychiatric and neurodevelopmental disorders, such as schizophrenia (Waters & Badcock, 2010; Waters et al., 2012) and autism (Cygan, Tacikowski, Ostaszewski, Chojnicka, & Nowicka, 2014; Kita et al., 2011).

In the last decades, studies have suggested that self-related stimuli—such as one's own voice or face—are processed differently from non-self stimuli, involving behavioral and neural correlates that are distinct from those engaged in the processing of stimuli not related to the self. For example, behavioral studies

show that in comparison with familiar and unfamiliar stimuli, one's own face is more rapidly identified (Keenan, Freund, Hamilton, Ganis, & Pascual-Leone, 2000; Sui, Zhu, & Han, 2006; Tacikowski & Nowicka, 2010; Tong & Nakayama, 1999). Furthermore, self-related stimuli such as one's own face or name elicit greater attentional resources than non-self stimuli (Chen et al., 2008; Eichenlaub, Ruby, & Morlet, 2012; Gray, Ambady, Lowenthal, & Deldin, 2004; Miyakoshi, Nomura, & Ohira, 2007; Perrin et al., 2005; Scott et al., 2005; Sugiura et al., 2000; Folmer & Yingling, 1997; Tacikowski & Nowicka, 2010), even when they are task-irrelevant (Berlad & Pratt, 1995; Gray et al., 2004; Holeckova et al., 2006; Müller & Kutas, 1996; Ninomiya et al., 1998; Perrin et al., 1999; Scott et al., 2005; Sui et al., 2006; Tateuchi, Itoh, & Nakada, 2012).

Event-related potential (ERP) studies have demonstrated important attentional biases to self-specific stimuli, as reflected by increased N2 (Fan et al., 2011; Fan et al., 2013; Perrin et al., 1999) and P3<sup>1</sup> amplitudes (Berlad & Pratt, 1995; Cygan et al., 2014;

<sup>1</sup> Previous studies showed that the P3 component is not a unitary brain potential, and that it consists of, at least, two subcomponents that reflect distinct neural processes—the P3a and the P3b (see Polich, 2007). The P3a indexes an involuntary attentional switch elicited by an unpredictable task-irrelevant change in a regular aspect of the environment, whereas the P3b is believed to reflect the allocation of higher-order attentional resources to a task-relevant deviant event. Here, the term “P3” was consistently used throughout the manuscript to refer to the component

\* Corresponding author at: Neuropsychophysiology Lab, CIPsi, School of Psychology, University of Minho, 4710-057 Braga, Portugal. Fax: +351 25360465.  
E-mail address: [ana.pinheiro@psi.uminho.pt](mailto:ana.pinheiro@psi.uminho.pt) (A.P. Pinheiro).

Eichenlaub et al., 2012; Fan et al., 2013; Folmer & Yingling, 1997; Gray et al., 2004; Müller & Kutas, 1996; Miyakoshi et al., 2007; Perrin et al., 1999; Scott et al., 2005; Su et al., 2010; Sui et al., 2006; Tacikowski & Nowicka, 2010; Tacikowski et al., 2011, 2014; Zhao, Wu, Zimmer, & Fu, 2011) for self- compared with non-self stimuli (see Appendix A in Supplementary material for a summarized description of ERP studies on self-processing). For example, Scott et al. (2005) showed that the passive visualization of one's own face elicits increased P3 amplitude in comparison with familiar and unfamiliar faces. In the same line, enhanced mobilization of attention indexed by increased P3 amplitude was observed when participants were passively listening to one's own name in comparison with both frequent and infrequent names unrelated to the self (Berlad & Pratt, 1995). Together, these studies indicate that different categories of self-stimuli (faces, names, self-related autobiographical information) elicit larger amounts of attentional resources than stimuli unrelated to the self, suggesting a prioritized processing of self-related information at both earlier and later stages of information processing. These attentional biases may be related to the enhanced affective salience of self-related stimuli (Brosch, Scherer, Grandjean, & Sander, 2013; Fan et al., 2013; Tacikowski & Nowicka, 2010; Vuilleumier, 2005).

Nonetheless, only a few studies investigated the processing of self-generated voice cues. Voices play a special role in everyday social communication and can be considered the most important class of sounds in our social environment (Belin, Fecteau, & Bédard, 2004; Schweinberger, Kawahara, Simpson, Skuk, & Zäske, 2014). Through speech production, humans are constantly exposed to their own voice and need to constantly monitor the feedback of their own voices in order to detect potential errors and perform adjustments in vocal production to fit the challenges of the social acoustic environment (Eliades & Wang, 2008). Even though some studies suggest that individuals are less accurate in recognizing their own voice than other familiar voices (e.g., Hughes & Nicholson, 2010), probably due to differences in sound transmission when speaking vs. listening to pre-recorded self-generated speech (Maurer & Landis, 1990), people can recognize their own voice above chance level (Nakamura et al., 2001; Rosa et al., 2008). This ability is preserved even in more demanding tasks, such as when voice stimuli are acoustically transformed at the level of pitch and formant frequencies cues (Allen et al., 2005; Allen et al., 2007; Xu et al., 2013). Importantly, acoustic cues, such as fundamental frequency (F0) of phonation (i.e., the perceived pitch) and formant frequencies have been pointed out as fundamental parameters that listeners rely on to discriminate and recognize the identity of self-generated, familiar and unfamiliar voices (Baumann & Belin, 2010; Latinus & Belin, 2012; Latinus et al., 2013; Xu et al., 2013).

Behavioral and brain studies examining self-generated voice processing corroborate differences in the processing of self vs. non-self cues observed for other types of stimuli. For example, when only higher frequencies are retained in the vocal signal, the ability to recognize self-generated voices is enhanced in comparison with other familiar voices (Xu et al., 2013). Furthermore, in comparison with a non-self voice, hearing a self-generated voice elicits increased activation in the left inferior frontal and right anterior cingulate (Allen et al., 2005), right inferior frontal (Kaplan, Aziz-Zadeh, Uddin, & Iacoboni, 2008; Nakamura et al., 2001) and right parainsular brain regions (Nakamura et al., 2001). A recent ERP study suggests that the differentiation between self- and non-self voice cues at the preattentive level occurs very early in information processing, which is reflected in a centro-parietal neg-

ativity to self-generated voice deviants in the 70–100 milliseconds (ms) latency window, and in a right temporo-parietal positivity to unfamiliar voice deviants in the same time window (Graux et al., 2013). However, contrary to previous studies that used other types of self-related information (e.g., Gray et al., 2004; Sugiura et al., 2000), recent ERP evidence showed that, when participants are instructed to pay attention to a silent movie whilst ignoring vocal stimuli, the P3a amplitude to task-irrelevant self-generated voice deviants is reduced in comparison with unfamiliar and familiar voice deviants (Graux et al., 2013; Graux et al., 2014). This finding indicates a decrease in attention orienting to self-generated voice stimuli, which might reflect the prioritized processing of others' voices compared to one's own voice in a later stage of information processing (Graux et al., 2013; Graux et al., 2014). Critical differences in task demands between the studies of Graux et al., 2013; Graux et al., 2014 and studies of visual processing of self-related information may have accounted for the apparently contradictory findings. In particular, in the studies of Graux et al., 2013; Graux et al., 2014 the appearance of a deviant voice disrupted the attentional engagement on the primary task (i.e., watching a silent movie). As such, the reported P3a indexes the orienting response to unexpected deviant stimuli (Friedman, Cycowicz, & Gaeta, 2001; Knight, 1996; Spencer et al., 1999; Spencer, Dien, & Donchin, 2001). This orienting response has been described as an involuntary change of attention that is normally evoked by an unpredictable violation in an otherwise unchangeable auditory sequence (Friedman et al., 2001; Knight, 1996; Spencer et al., 1999; Spencer et al., 2001). Nonetheless, the P3 component reported in most of the abovementioned studies on visual self-related information processing (e.g., Berlad & Pratt, 1995; Cygan et al., 2014; Folmer & Yingling, 1997; Gray et al., 2004; Perrin et al., 1999; Scott et al., 2005; Su et al., 2010; Tacikowski & Nowicka, 2010; Tacikowski et al., 2014; Zhao et al., 2011) is thought to index the allocation of high-order attentional resources to a task-relevant event after the cognitive evaluation of the stimulus meaning, i.e., the P3b (Knight, 1996; Polich, 2007; Spencer et al., 1999; Spencer et al., 2001). Of note, Spencer et al., 1999; Spencer et al., 2001 demonstrated that the P3a and the P3b ERP components are dissociable.

In an attempt to solve these discrepancies, we used the ERP methodology to investigate whether a self-generated voice has privileged access to attentional resources in comparison with a non-self voice, when participants are instructed to focus their attention on a sequence of vocal stimuli, presented in a modified version of the oddball task. Participants were asked to identify an infrequent vocal target stimulus interspersed with frequent vocal standards. The oddball design is a very robust and reliable paradigm for eliciting the P3 component in a short amount of time (Herrmann & Knight, 2001; Polich, 2007; Polich & Criado, 2006). Also, this design is highly suitable for controlling for differences in the physical properties between voice stimuli, since it uses a reduced set of individual stimuli which are presented both as standard and deviant stimuli in different experimental blocks. Another main advantage of using the oddball task to probe the attentive processing of SGV and NSV is that it allows the elicitation of the P3 component, with no overt response being required, as participants might be simply asked to mentally count the number of target deviant stimuli (Knight, 2001; Polich, 2007; Polich & Criado, 2006). The aims of this study were threefold: (1) to examine the role of attention in the processing of self-generated and unfamiliar voices, with the focus on the N2 and P3 ERP components; (2) to test the association between ERP correlates of voice processing and voice acoustic properties, considering previous studies demonstrating that both F0 (i.e., the perceived pitch) and formant frequencies are critical acoustic parameters that listeners rely on to process voice identity (Baumann & Belin, 2010; Latinus & Belin, 2012; Latinus et al., 2013; Xu et al., 2013); (3) to examine whether the ability to

elicited by task-relevant deviants, as our task required that participants focused their attention on the sounds and silently counted the infrequent (and task-relevant) vocal stimuli.

**Table 1.** Processing of self-generated vs. unfamiliar voices. Socio-demographic and cognitive characterization of the participants.

	M (SD)
Sociodemographic data	
Age, years	30.76 (5.39)
Years of education	15.00 (2.96)
Cognitive data*	
Full scale IQ	124.55 (12.16)
Verbal IQ	127.25 (11.35)
Performance IQ	115.90 (12.32)

Note: M: mean; SD: standard deviation; \*WAIS-III (Wechsler, 2008).

allocate attention to self-generated vs. non-self voices is associated with a more general auditory processing capacity to attentively detect pitch changes (i.e., in simple tones).

Based on previous evidence supporting the privileged processing of self-related stimuli when in the focus of attention, we hypothesized increased mobilization of attention to self-generated relative to non-self voices, which would be indexed by increased P3 amplitude. The P3 is preceded by the N2 component, which indexes the categorization of task-relevant events in an earlier attentional stage (Folstein & Van Petten, 2008; O'Donnell et al., 1993; Patel & Azzam, 2005). Even though both components are sensitive to the same manipulations (i.e., stimulus probability and stimulus' task relevance—Folstein & Van Petten, 2008; Patel & Azzam, 2005), the question of whether early discrimination and categorization of vocal distinctions indexed by the N2 are modulated by the salience of SGV and NSV stimuli remains unanswered. Following previous studies (Leitman, Sehatpour, Garidis, Gomez-Ramirez, & Javitt, 2011; Schirmer & Escoffier, 2010; Schirmer, Striano, & Friederici, 2005), self and non-self voices were presented as both standard and deviant stimuli in two distinct experimental blocks within an oddball task. Then, ERP to self and non-self vocal deviants were compared with self and non-self vocal standards, respectively, using difference waveforms. Adopting a “like from like” subtraction approach allowed us to control for the acoustic differences between stimuli, and as such, we predicted that the P3 amplitude to a self-generated voice is independent from voice acoustic properties, which would be translated in a lack of association between the acoustic parameters and the ERP correlates of self-generated voice processing. In addition, given the stability of the P3 component as a neurophysiological index of the individual processing capacity to direct attentional resources to task-relevant events (Cassidy, Robertson, & O'Connell, 2012; Kok, 2001; Polich, 2007; Sandman & Patterson, 2000), we hypothesized a positive association between P3 amplitude to pitch changes in non-vocal sounds and to changes in voice identity.

## 2. Methods

### 2.1. Participants

Twenty healthy adult males participated in the study (see Table 1). Participants were native speakers of European Portuguese and all reported normal hearing. Inclusion criteria were: right handedness (Oldfield, 1971); verbal intelligence quotient (IQ) above 90 (Wechsler Adult Intelligence Scale—WAIS-III—Wechsler, 1997); no history of neurological illness, electroconvulsive treatment, or history of drug or alcohol abuse in the past year; no current medication for medical disorders that would affect electroencephalogram (EEG) morphology. Participants were excluded if they scored  $\geq 1.7$  in the Positive Symptoms Distress Index of the Brief Symptom Inventory (BSI—Canavaro, 1999; Canavaro, 2007) (see Table 1) or if they reported a history of psychiatric disorder in oneself or in first-degree relatives (none was excluded). After

a detailed description of the study, all participants gave oral and written informed consent. The consent form was assessed by the local Institutional Review Board committee for the protection of human subjects.

### 2.2. Stimuli

#### 2.2.1. ERP experiments

**2.2.1.1. Processing of self-generated vs. unfamiliar voices.** The Portuguese word /nome/ [name] was selected from the *Affective Norms for English Words—ANEW* (Soares, Comesaña, Pinheiro, Simões, & Frade, 2011) and *P-PAL* (Soares et al., 2010) sets, according to the following criteria: neutral valence (5.33), low arousal (3.97), high frequency (>100 per million), grammatical class (noun), and short extension (2 syllables and 4 letters). In addition, the high stability of the syllabic structure of this word (i.e., consonant–vowel, consonant–vowel) aimed to reduce the variability in word pronunciation among participants.

The word was recorded by each participant (*self-generated voice condition—SGV*) and by an unknown middle-aged male (age = 42 years) who was also a native speaker of European Portuguese (*non-self voice condition—NSV*). The recording session took place at least one week before the ERP experiment, in a sound-proof room, using a portable digital recorder Roland R-26 and a Shure incorporated PG48 microphone, with a sampling rate of 44.1 kHz. In order to minimize variability in word pronunciation, the word (previously recorded by a middle-aged male voice and without regional accent—age = 47 years) was binaurally presented through a set of Sennheiser CX 300-II earphones, and participants were instructed to reproduce the word exactly as they just heard. Praat software (Boersma & Weenink, 2012) was used to match the duration and intensity of self-generated and unfamiliar voices. Audacity 2.0.2. software (<http://audacity.sourceforge.net/>) was used to remove background noise. The acoustic analysis of all vocal stimuli was performed using Praat software (see Table 2).

**2.2.1.2. Processing of simple tones.** Simple tones (duration = 100 ms; intensity = 75 dB) with frequencies of 1000 Hz and 1500 Hz were used as stimuli in a simple auditory oddball task.

#### 2.2.2. Behavioral experiment

**2.2.2.1. Self- and non-self voice discrimination task.** To ensure that all participants were able to recognize self-generated stimuli as their own voice, they performed a behavioral SGV vs. NSV discrimination task after the ERP experiments. Twenty four words, selected from the ANEW battery (Soares et al., 2012), were recorded by each participant (SGV) and by the unknown middle-aged male (NSV) from Experiment 1. All words had high frequency (>100 per million) and differed in affective valence (see Table 3): 8 negative (mean = 2.69; range = 1.68–3.74), 8 neutral (mean = 5.31; range = 4.91–5.78), 8 positive (mean = 7.20; range = 6.75–7.74). The conditions and procedure of the recording session were exactly the same as described in Experiment 1.

### 2.3. Procedure

Participants performed a total of two ERP experiments (one with vocal stimuli and another with non-vocal sounds) and one behavioral task requiring self–other voice discrimination. The order of the ERP experiments was counterbalanced across participants. The behavioral voice discrimination task took place after the ERP experiments, to avoid any prior exposure to SGV and NSV stimuli and circumvent potential learning effects associated with a previous exposure to the experimental stimuli.

**Table 2**  
Acoustic properties of the voice stimuli used in the ERP Experiment 1.

Participant	Mean F0 (Hz)	Range F0 (Hz)		Formant frequencies (Hz)		
		Min	Max	F1	F2	F3
P1	115	105	130	717	1016	2312
P2	93	87	105	581	1090	2898
P3	94	83	103	625	1185	2385
P4	103	84	133	494	1284	2125
P5	93	76	101	567	1382	2527
P6	94	78	108	719	1046	2178
P7	93	86	106	573	1138	2886
P8	108	98	120	547	1002	2260
P9	94	77	124	616	1012	2312
P10	96	80	112	557	1032	2777
P11	109	93	119	558	1071	2650
P12	85	75	91	569	1239	2615
P13	99	88	141	570	1223	2410
P14	155	127	174	613	1159	2399
P15	121	108	162	399	1273	2607
P16	124	110	154	519	1143	2557
P17	103	94	113	661	1377	2585
P18	94	87	105	520	973	2169
P19	113	106	120	374	1165	2664
P20	104	98	116	527	1093	2439
Mean (SD)	104.50 (15.80)	92.00 (13.77)	121.85 (21.65)	565.35 (85.88)	1145.13 (121.62)	2487.75 (275.56)
NSV	96	83	106	621	944	2258

Note: M: mean; SD: standard deviation.

### 2.3.1. ERP experiments

Participants were seated in an acoustically and electrically shielded room at a distance of 100 cm from the computer monitor. Participants performed two experimental blocks, each one comprising a total of 200 standard ( $P = .083$ ), and 40 target ( $P = .017$ ) stimuli. In the first block, the non-self voice was the standard stimulus and the self-voice was the target stimulus. In the second block, the self-voice was the standard stimulus and the non-self voice was the target stimulus. Stimuli were binaurally presented (at a level of 70 dB SPL) in a pseudorandom order, with the restriction that there were at least two standard stimuli between two deviant sounds, following previous studies (Háden, Németh, Török, Drávucz, & Winkler, 2013; Özgürdal et al., 2008; Schirmer et al., 2007; Simons et al., 2011). Presentation software (Neurobehavioral Systems, Albany, CA, USA) was used to control stimulus timing and presentation. Each trial began with a fixation cross (which remained on the screen during the entire experimental block), and then a vocal stimulus (duration = 483 ms) was presented, with a stimulus onset asynchrony (SOA) varying between 1283 and 1383 ms (see Fig. 1). Each experimental block lasted approximately 5 min. Order of blocks was counterbalanced across participants. Participants were instructed to silently count the number of times a target stimulus was presented.

**2.3.1.2. Processing of simple tones.** The 1000 Hz standard tone ( $n = 200$ ;  $P = .83$ ) and the 1500 Hz target tone ( $n = 40$ ;  $P = .17$ ) were binaurally presented in a pseudo-random order, with the restriction that at least 2 standard stimuli preceded a deviant sound. Each trial included a fixation cross (remaining on the screen during the

entire trial) and a tone (duration = 100 ms), with a SOA ranging between 900 and 1000 ms (see Fig. 2). Participants were instructed to silently count the number of times a target tone was presented.

### 2.3.2. Behavioral experiment

**2.3.2.1. Self- and non-self voice discrimination task.** Twenty four words uttered by the SGV or the NSV (total = 48) were randomly presented. Each trial comprised a warning slide (duration = 1000 ms) signaling the beginning of a trial, the vocal stimulus (duration = 1000 ms), and a question mark (duration = 2000 ms or until a response was provided by the subjects) (see Fig. 3). Participants were instructed to judge if the voice uttering each word was their own voice (self) or the voice of somebody else (non-self), indicating their response by pressing one of two buttons. The order of buttons was counterbalanced across participants.

### 2.4. EEG data recording and analysis

The EEG was recorded using a 64-channel BioSemi Active Two System (<http://www.biosemi.com/products.htm>). The EEG was continuously sampled at 512 Hz, with a bandpass of 0.01–100 Hz and stored for later analysis. Horizontal and vertical eye movements were monitored by placing electrodes at left and right temples (horizontal electrooculogram-EOG) and one below the left eye (vertical EOG). Electrodes were also placed on left and right mastoids for offline referencing.

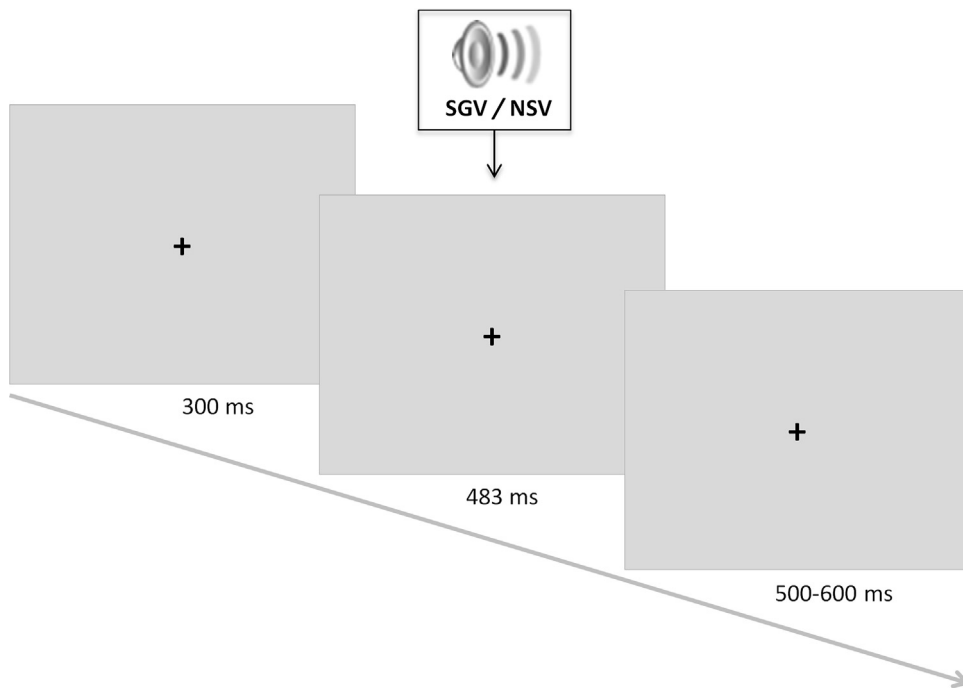
The EEG data were analyzed using the Brain Vision Analyzer 2.0.4 software ([www.brainproducts.com](http://www.brainproducts.com)). A high-pass filter of .01 Hz was applied to the EEG data. Individual EEG epochs of 1100 ms, time-locked to the onset of the voice stimuli and with a 100 ms prestimulus baseline, were computed. Ocular artifacts were corrected using the Gratton et al.'s method (Gratton, Coles, & Donchin, 1983). Single epochs containing excessive motor artifacts (threshold of  $\pm 100 \mu\text{V}$ ) were excluded from the analysis. For each participant, ERP averages were based on at least 75% of the trials in each condition (standard SGV =  $179.82 \pm 15.38$ ; deviant SGV =  $35.06 \pm 3.19$ ; standard NSV =  $175.65 \pm 14.43$ ; deviant NSV =  $36.59 \pm 3.26$ ). In order to control for the acoustic differences between SGV and NSV, difference waveforms were calculated by

**Table 3**  
Psycholinguistic and affective properties of the words used in the voice discrimination task.

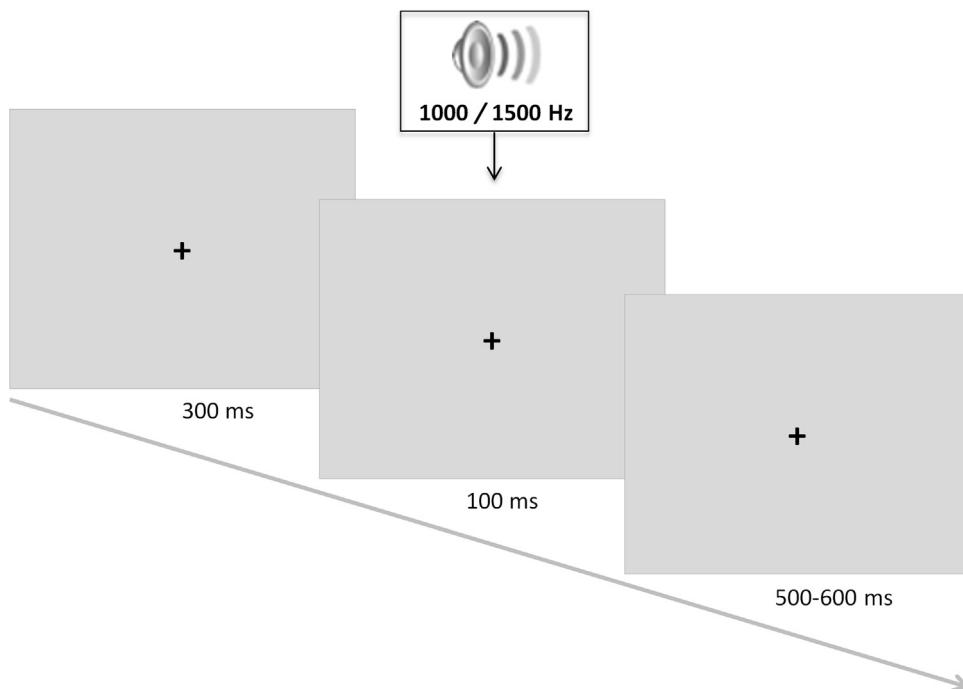
Word type	Valence M (SD)	Arousal M (SD)	Frequency (per million) M (SD)
Negative	2.69 (0.64)	6.32 (0.41)	199.24 (126.19)
Neutral	5.31 (0.34)	3.81 (0.30)	294.21 (243.21)
Positive	7.20 (0.39)	5.61 (0.45)	294.49 (288.91)

Note: psycholinguistic and affective properties were retrieved from the ANEW database (Soares et al., 2010). M: mean; SD: standard deviation.





**Fig.1.** Schematic illustration of a trial in Experiment 1.



**Fig. 2.** Schematic illustration of a trial in Experiment 2.

subtracting SGV standards from SGV deviants, and NSV standards from NSV deviants.

After a careful visual inspection of grand average waveforms, two auditory ERP components with maximal effect at centro-parietal (CPz/1/2) and parietal (Pz/1/2) electrodes were identified: N2 and P3. For both SGV and NSV difference waveforms, P3 amplitude was measured as the mean amplitude between 320 and 500 ms post-stimulus onset, whereas N2 amplitude was measured as the mean amplitude between 200 and 300 ms post-stimulus onset. N2 and P3 peak latency was measured in the same intervals.

We followed the same analysis procedure in the oddball task with simple tones. In this case, EEG epochs of 800 ms were segmented (including a –100 ms pre-stimulus baseline) and difference waveforms were calculated by subtracting the 1000 Hz standard stimulus to the 1500 Hz deviant stimulus. The N2 and P3 to simple tones had also maximal effect at centro-parietal (CPz/1/2) and parietal (Pz/1/2) electrodes. N2 and P3 were computed as the mean amplitude considering the following latency intervals: 180–280 ms (N2), and 300–450 ms post-stimulus onset (P3).

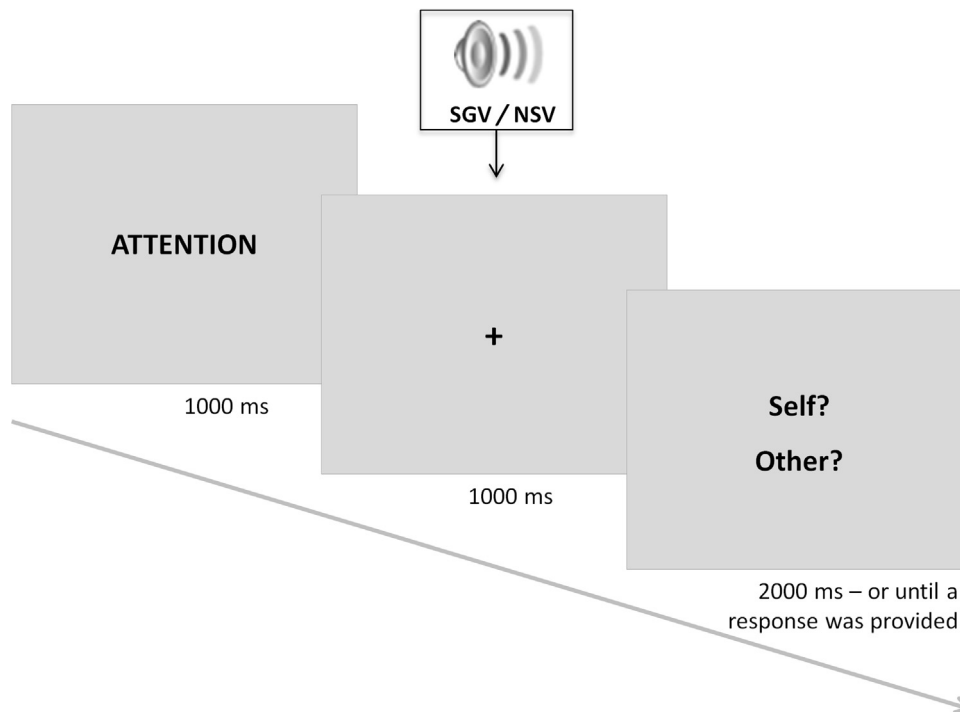


Fig. 3. Schematic illustration of a trial in Experiment 3.

## 2.5. Statistical analyses

All statistical analyses were conducted using IBM SPSS 21.0 software (SPSS, Corp., USA).

### 2.5.1. ERP data

All statistical analyses were based on unfiltered ERP data. The mean amplitude and latency of the N2 and P3 components were analyzed using four separate repeated-measures analyses of variance (ANOVA), with voice identity (SGV, NSV), region (centro-parietal, parietal electrodes) and electrode site (CPz/1/2, Pz/1/2) as within-subjects factors. Main effects and interactions were followed with pairwise comparisons between conditions, using the Bonferroni adjustment for multiple comparisons. Analyses were corrected for non-sphericity using the Greenhouse–Geisser correction method, when appropriate.

### 2.5.2. Accuracy in voice recognition

A repeated-measures ANOVA tested the effects of emotional valence (affective dimension indexing the pleasantness and unpleasantness of a stimulus—Bradley & Lang, 1994) on the recognition accuracy (hit rates) of both SGV and NSV, with valence (negative, neutral, positive) and voice identity (SGV, NSV) as within-subjects factors. Main effects and interactions were followed with pairwise comparisons between conditions, using the Bonferroni adjustment. The analysis was corrected for non-sphericity using the Greenhouse–Geisser correction method when appropriate.

### 2.5.3. Correlation between ERP and acoustic data

Pearson's correlation coefficient was calculated to test the association between: (a) ERP responses to simple tones and the ERP responses to both SGV and NSV conditions, at CPz and Pz electrodes (N2 and P3 amplitudes); (b) voices' acoustic properties from each participant (i.e., F0 and formant frequencies—F1–F3) and the P3 amplitude elicited by SGV (at CPz and Pz). Levels of significance

were adjusted for multiple comparisons using the Bonferroni correction.

## 3. Results

### 3.1. ERP data

In Experiment 1, the accuracy in target detection was 98.13% (range = 92.5–100%) for SGV deviants and 98.75% (range = 95–100%) for NSV deviants. In Experiment 2, accuracy of target detection was 98.25% (range = 90–100%), indicating that participants were able to successfully detect deviance in auditory stimulation. Therefore, data from all participants were included in the ERP analyses.

Fig. 4 illustrates grand average difference waveforms for SGV and NSV conditions at the 6 electrodes included in the statistical analyses, with topographical maps for N2 and P3 components. Fig. 5 illustrates grand average difference waveforms for simple tones at the 6 electrodes included in the statistical analyses. A high cutoff filter at 15 Hz was applied to the grand average difference waveforms presented in Figs. 4 and 5 for illustration purposes only.

**N2:** The repeated measures ANOVA revealed a main effect of identity,  $F(1, 19) = 10.42, p = .004, \eta_p^2 = .354$ , and region,  $F(1, 19) = 8.84, p = .002, \eta_p^2 = .402$ , on N2 amplitude. Pairwise comparisons showed that N2 was more negative to the SGV relative to the NSV condition ( $p = .004$ ). Also, pairwise comparisons revealed that N2 was more negative at centro-parietal than parietal electrodes. No significant main effects or interactions were observed for the N2 latency ( $p > .05$ ).

**P3:** A significant main effect of identity,  $F(1, 19) = 4.42, p = .049, \eta_p^2 = .189$ , revealed more positive P3 amplitude to the SGV than to the NSV condition ( $p = .049$ ). No significant main effects or interactions were observed for the P3 latency ( $p > .05$ ).

### 3.2. Behavioral data: accuracy in voice recognition

Participants recognized their own voice with a mean accuracy of 92.92% (SD = 10.30; range = 66.67–100%) and the NSV with a mean

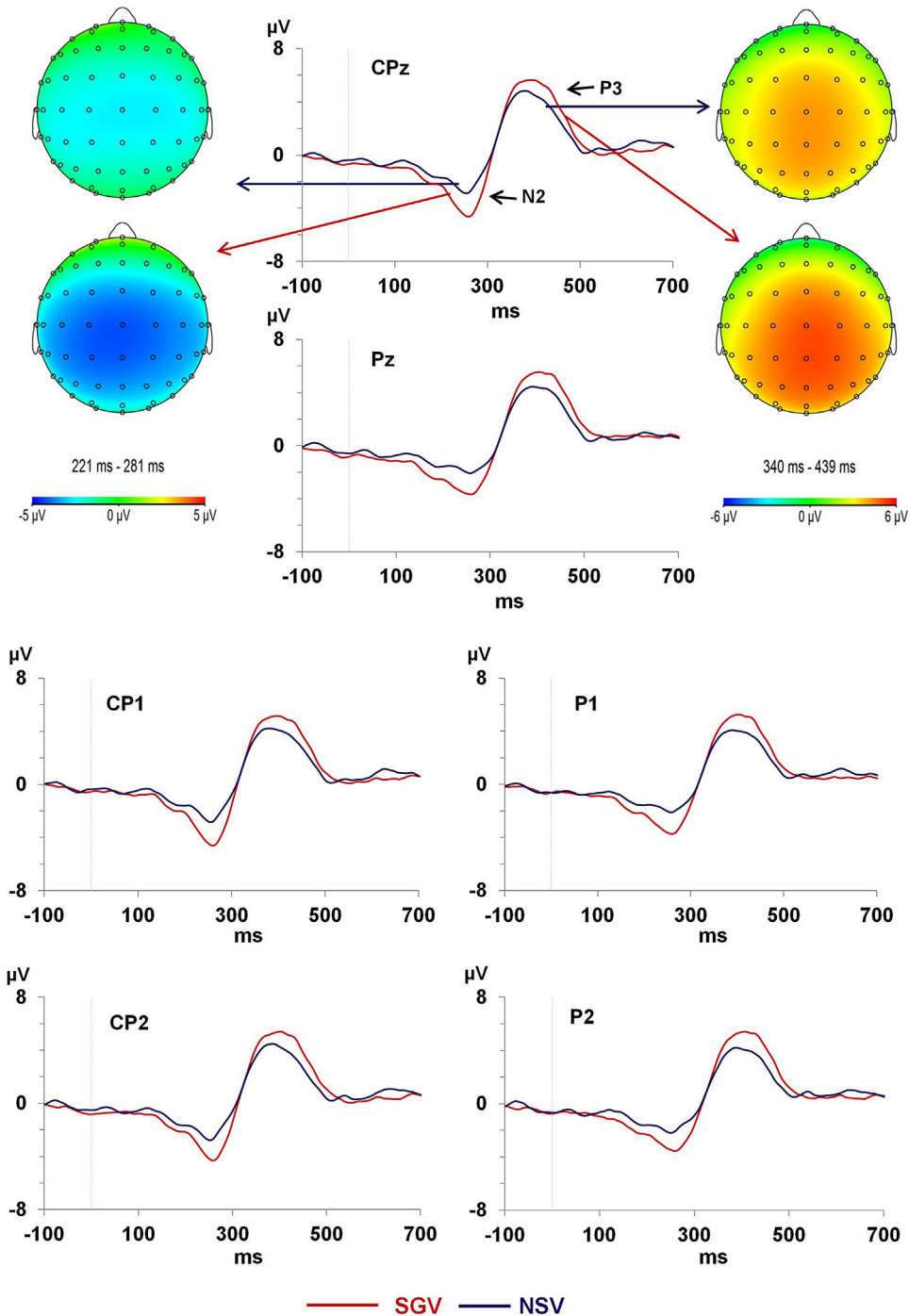


Fig. 4. Illustration of grand average difference waveforms for the SGV and NSV conditions.

accuracy of 95.20% (SD = 8.03; range = 70.83–100%). A significant main effect of identity,  $F(1, 19) = 6.07$ ,  $p = .024$ ,  $\eta_p^2 = .242$ , showed that recognition accuracy was higher for the NSV than for the SGV condition ( $p = .024$ ).

### 3.3. Correlation between ERP and acoustic data

The analysis revealed a significant association between the N2 amplitude to simple tones and to vocal sounds at the CPz channel only (Fig. 6): the more negative the N2 to pitch changes in simple tones, the more negative the N2 to SGV ( $r = .630$ ,  $p = .003$ ) and to the

NSV ( $r = .657$ ,  $p = .002$ ). No significant associations were found for the P3 amplitude ( $p > .004$ ).

## 4. Discussion

This study investigated the effects of attention on the ERP correlates of self-generated and non-self voice processing. By using a modified oddball task, we found that both early and high-order attentional processing stages are modulated by voice identity. Confirming our initial hypothesis, we observed that a self-generated voice engages more attentional resources than the voice of

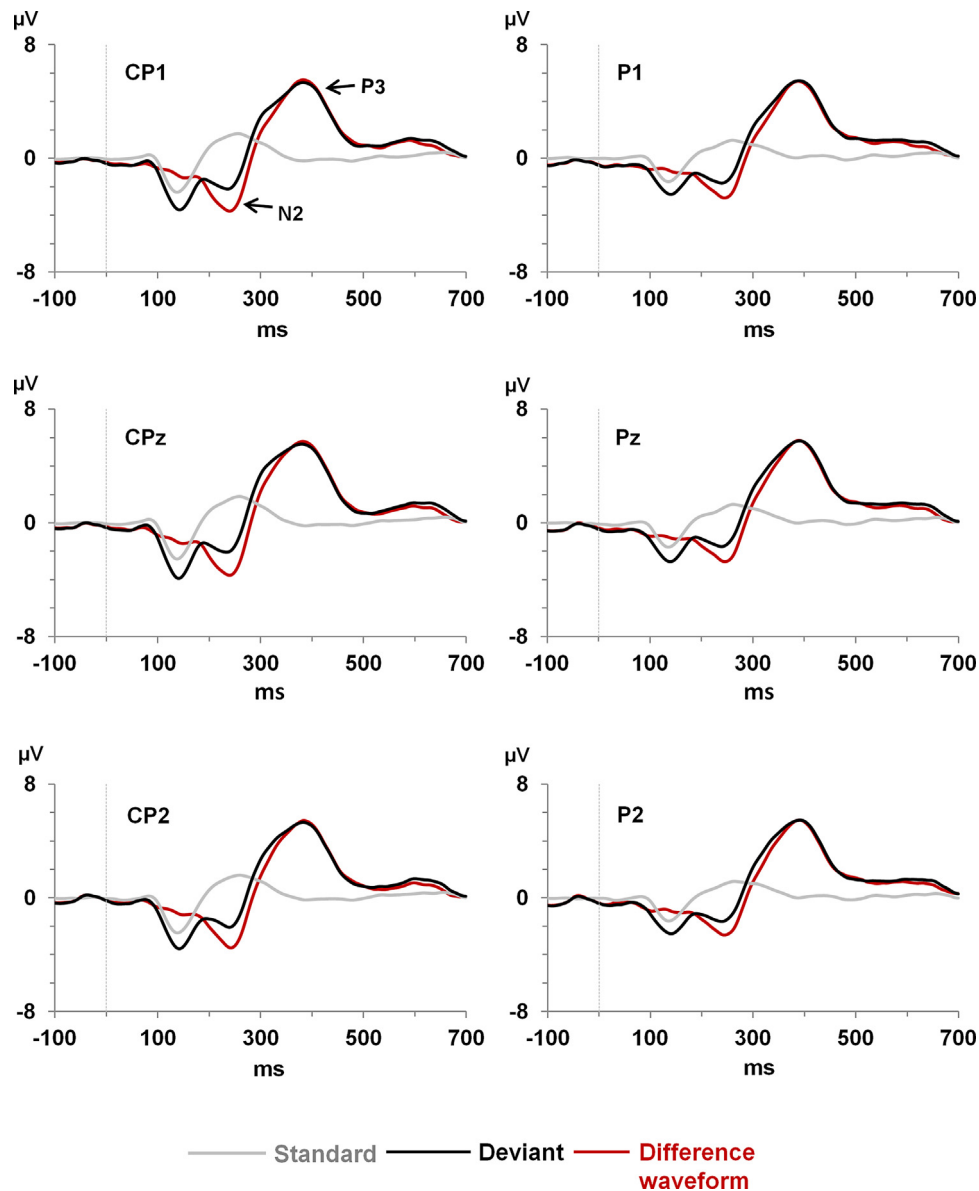


Fig. 5. Illustration of grand average waveforms for simple tones.

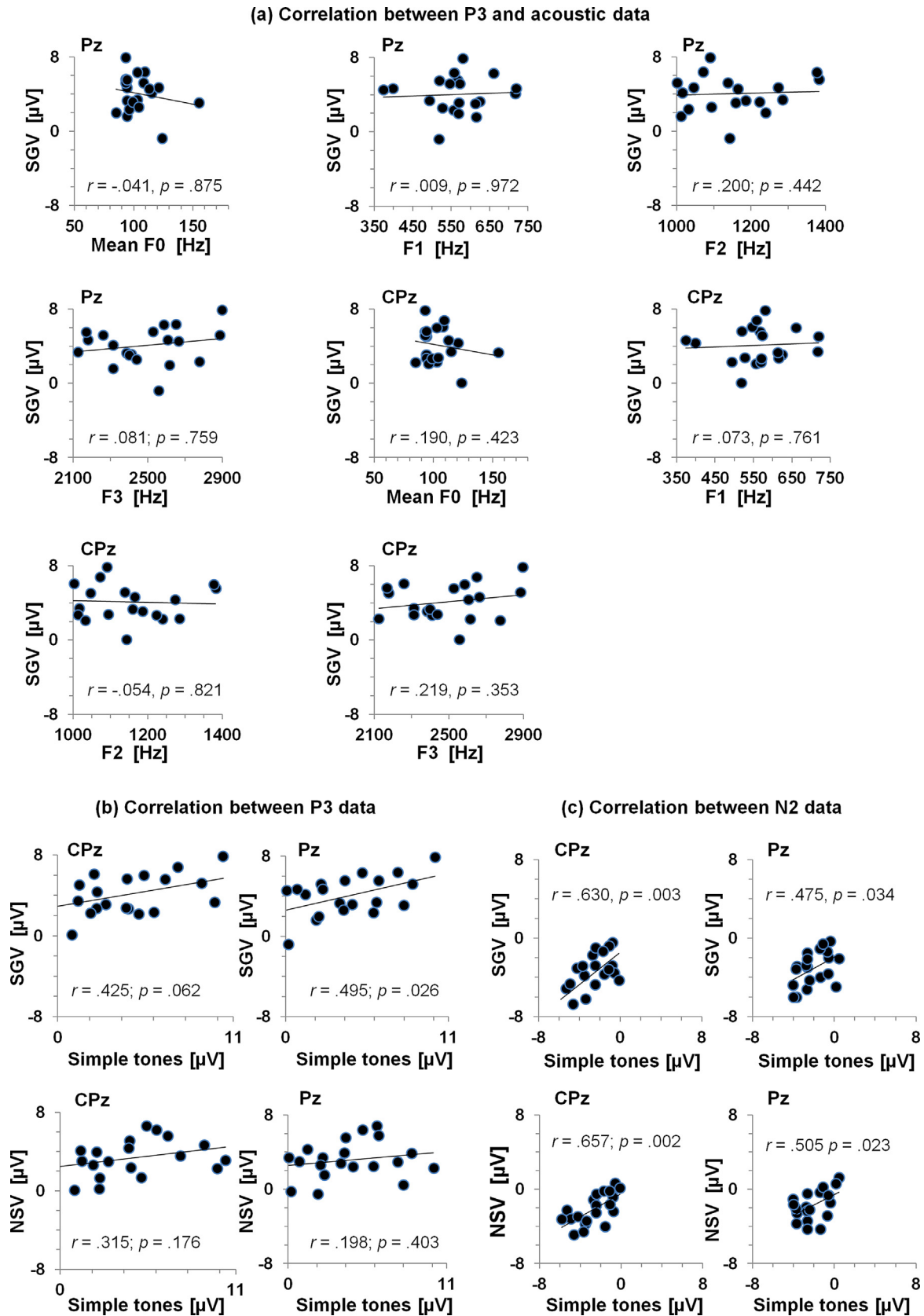
somebody else, demonstrated by increased N2 and P3 amplitude to the former.

Considering the functional significance of the P3 component as an index of attentional resources allocation and cognitive evaluation of a stimulus meaning (Knight, 1996; Spencer et al., 1999; Spencer et al., 2001), our result of increased P3 amplitude to self- relative to non-self voice deviants suggests that more attentional resources were directed to the self-generated voice. Moreover, this finding is consistent with previous studies that used other categories of self-related stimuli, such as one's own face (e.g., Scott et al., 2005; Sui et al., 2006; Tacikowski & Nowicka, 2010), name (e.g., Berlad & Pratt, 1995; Perrin et al., 1999; Perrin et al., 2005; Tacikowski et al., 2014), autobiographical information (Gray et al., 2004) or self-hand (Su et al., 2010). A plausible explanation for the increased allocation of attentional resources to self-related vs. non-self information is that such stimuli are assessed as more emotionally and personally significant than non-self stimuli, thus attracting more attention. Support for this explanation comes from studies demonstrating that events with higher emotional salience engage more

attentional resources when compared with neutral stimuli (e.g., Carretié, Martín-Loeches, Hinojosa, & Mercado, 2001; Delplanque, Silvert, Hot, & Sequeira, 2005; Vuilleumier, 2005), and that P3 amplitude is modulated by the affective salience of the eliciting stimulus, with larger amplitude being devoted, for instance, to emotional vs. neutral pictures (Delplanque et al., 2005; Delplanque et al., 2006; Keil et al., 2002; Olofsson et al., 2008) or sounds (Czigler, Cox, Gyimesi, & Horváth, 2007; Thierry & Roberts, 2007). Psychophysiological support for increased affective significance of self-related vs. non-self stimuli was also provided by Sugiura et al. (2000) who reported larger electrodermal response (an index of increased emotional reactivity) during the visualization of one's own face than when seeing a familiar face. It is well known that emotion and attention are strikingly related: given the limited capacity of our information processing system, attentional resources tend to be preferentially captured by events with greater affective salience in comparison with neutral events (Carretié et al., 2001, 2008; Olofsson et al., 2008; Vuilleumier, 2005).

Importantly, when considering the apparent contradiction between our findings of increased P3 to self-voice and the results





**Fig. 6.** Results of the correlation analysis using Pearson's coefficient between the P3 amplitude to SGV and the acoustic properties, and between the P3 amplitude to simple tones and to voice stimuli. Levels of significance were adjusted for multiple comparisons using Bonferroni correction and thus correlations are significant if  $p \leq 0.003$ .

reported by Graux et al., 2013; Graux et al., 2014 of smaller P3a to self vs. non-self familiar and unfamiliar voices, it is important to note crucial differences in task demands between both experimental tasks. In our study, the subject's attention was focused on the target (deviant) voice stimuli and, thus, the functional meaning of the P3 component most likely reflects the cognitive evaluation of the stimulus significance as well as the mobilization of high-order attentional resources to a task-relevant target event, i.e., the P3b (Knight, 1996; Polich, 2007; Spencer et al., 1999; Spencer et al., 2001). However, in the studies of Graux and colleagues, participants directed attention to a silent movie, whilst ignoring the voice stimuli. In the latter task, the P3 more prominently reflects an orienting response, that is, an involuntary change of attention elicited by an unexpected violation in an otherwise unchangeable auditory stream, i.e., a P3a (Cycowicz & Friedman, 1999; Friedman et al., 2001; Knight, 1996; Spencer et al., 1999; Spencer et al., 2001). This involuntary capture of attention is thought to be a crucial biological process with implications for survival (Friedman et al., 2001). In contrast with the P3 elicited by our task-relevant voice deviants, the P3a to unexpected task-irrelevant deviants typically has a more frontal scalp distribution that is thought to reflect the contribution of the frontal cortex (Friedman et al., 2001; Schröger et al., 2000; Spencer et al., 1999, 2001). Thus, in our view, it seems plausible that in the studies of Graux et al. (2013, 2014) the P3a amplitude was increased to non-self voice stimuli (i.e., familiar/unknown) since the detection of a novel and completely unexpected voice stimulus might be more important for survival than the detection of one's own voice, but specifically when the participant is not paying voluntary attention to the auditory sequence.

Moreover, we found larger N2 amplitude to SGV in comparison with NSV stimuli. The N2 is thought to index early high-order operations of discrimination and categorization of stimuli that signal change in an otherwise invariant sequence of events (O'Donnell et al., 1993; Patel & Azzam, 2005; Salisbury et al., 1994). Thus, this component is at the interface between more automatic and controlled processes of the orienting response (Li, Yuan, & Lin, 2008; Yuan et al., 2007). As such, the increased amplitude observed for SGV indicates that self and non-self voices were discriminated at an early attentional processing stage and might suggest an easier detection of the SGV relative to the NSV. More negative N2 amplitude was previously reported by other studies investigating self-stimuli processing (Fan et al., 2011; Fan et al., 2013; Perrin et al., 1999). Since attentional modulation by emotional salience was shown to take place in early stages of information processing (Alexandrov, Klucharev, & Sams, 2007; Carretié et al., 2001; Wang et al., 2009), the heightened affective salience of SGV might have contributed to its facilitated detection. It is plausible that the enhanced brain response to SGV in this earlier attentional stage contributes to increased higher-order attention to self-voice in later cognitive stages of voice information processing, as indexed by the P3.

It is worth noting that the differential ERP response to task-relevant self vs. unfamiliar voice deviants does not seem to be due to the contribution of specific acoustic properties of the stimuli, as our subtraction method involved the direct comparison between the same stimuli playing different roles (i.e., SGV standards were subtracted from SGV deviants, and vice-versa). Furthermore, F0 and formant frequencies were not associated with the electrophysiological responses analyzed in this study, even though previous studies have found that both acoustic properties are critical cues underlying self and non-self voice discrimination and recognition (e.g., Latinus et al., 2013; Xu et al., 2013). Moreover, we found an association between the N2 amplitude to vocal and non-vocal sounds, showing that the greater the ability to discriminate pitch changes in simple tones, the easier the discrimination and categorization of vocal deviants. Since the N2 is thought to reflect

classification processes and is at the interface between sensory and more cognitive stages of task-relevant change detection (Li et al., 2008; Yuan et al., 2007), this association suggests that the early categorization of self- and non-self voice deviants is modulated by the capacity to categorize acoustic changes, dependent on pitch cues. This keeps with previous evidence showing that the perception of voice identity involves the integration of pitch cues (Baumann & Belin, 2010; Belin et al., 2004; Belin et al., 2011; Latinus & Belin, 2012; Latinus et al., 2013; Schweinberger et al., 2014; Xu et al., 2013). However, at a higher-order (cognitive) stage of information processing, the ability to direct attentional resources to voice stimuli was not associated with the more general ability to direct attention to auditory stimuli representing a pitch change. Together, these findings provide additional support to the proposal that the salience of the voice stimulus (and not the physical differences between SGV and NSV) seems to play the most critical role in modulating attention during voice processing at a post-perceptual stage. In the same line, Tacikowski and Nowicka (2010) proposed that, after the recognition of a particular stimulus is completed, it is the meaning assigned to a given stimulus (e.g., face or name) that determines the engagement of attention. Supporting this explanation, the accuracy rates from the voice discrimination task in our study demonstrate that participants were able to successfully identify their voices as self-generated, in spite of differences in the perception of the voice signal when actively producing speech vs. passively listening to pre-recorded self-generated speech (Maurer & Landis, 1990).

Together, our results provide further support for the 'special' role of self-related stimuli. Corroborating previous evidence for increased attentional resources to self-related stimuli (e.g., one's own face, name or voice), our findings indicate that there is something "special" on how the brain processes information about the self, with one's own voice receiving a privileged access to attentional resources. The earlier and later attentional biases towards self-generated vs. non-self voice might explain the advantage of SGV recognition in acoustically demanding conditions, as reported by Xu et al., (2013). Since self-awareness refers to the capacity of subjects being the object of one's own attention (Jeannerod, 2003; Keenan et al., 2003; Morin, 2011), the ability to allocate attention to self-relevant stimuli is likely to have a critical social and adaptive value. In particular, the capacity to direct attention to one's own voice is fundamental for efficient online speech monitoring (Brébion, Smith, Gorman, & Amador, 1996; Ilankovic et al., 2011; Kambeitz-Ilankovic et al., 2013), and thus, plays a critical role in everyday social communication. Impairments in this capacity were found to contribute to the external misattribution of self-generated voice stimuli commonly observed in schizophrenia patients with auditory verbal hallucinations (Allen et al., 2004; Allen et al., 2007; Hashimoto & Sakai, 2003; Ilankovic et al., 2011; Kambeitz-Ilankovic et al., 2013; Kumari et al., 2010), which highlights the need to investigate these processes in this disorder.

In contrast with some previous studies (e.g., Graux et al., 2014; Sugiura et al., 2000), we did not control for familiarity effects. However, in a recent study, Graux et al. (2014) provided evidence for a differential attention orienting to self-generated vs. familiar voices. This suggests that familiarity is not the main factor accounting for increased attention to self-generated relative to an unknown voice. Moreover, our experimental stimuli consisted of a dissyllabic word and, as variables such as duration and phonetic variability were found to have a critical impact on speaker's recognition accuracy (e.g., Cook & Wilding, 1997; Schweinberger et al., 1997), future studies should explore whether stimulus complexity modulates how attentional resources are directed to voice stimuli. In order to account for a better understanding of how attention shapes self-processing, it is also critical to explore how self-generated voices are processed when participants are engaged in a primary task

and are not paying attention to voice stimuli. These studies were already initiated by our research team and the analysis of results is currently under way.

## 5. Conclusion

In conclusion, our study provided, for the first time, electrophysiological evidence for increased allocation of attentional resources to a self-generated voice in comparison with a non-self voice, irrespective of voice acoustic properties, extending previous studies that focused on the processing of visual self-related information. This attentional bias seems to be driven by the greater affective salience of one's own voice representation.

## Acknowledgments

This work was supported by Grants IF/00334/2012, PTDC/PSI-PCL/116626/2010 and PTDC/MHN-PCN/3606/2012 funded by Fundação para a Ciência e a Tecnologia (FCT, Portugal) and FEDER (Fundo Europeu de Desenvolvimento Regional) through the European programs QREN (Quadro de Referência Estratégico Nacional), and COMPETE (Programa Operacional Factores de Competitividade), awarded to A.P.P., and a by a FCT Doctoral Grant (SFRH/BD/77681/2011) awarded to T.C.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.biopsycho.2015.07.014>.

## References

- Alexandrov, Y. I., Klucharev, V., & Sams, M. (2007). Effect of emotional context in auditory-cortex processing. *International Journal of Psychophysiology*, 65(3), 261–271. <http://dx.doi.org/10.1016/j.ijpsycho.2007.05.004>
- Allen, P. P., Johns, L. C., Fu, C. H., Broome, M. R., Vythelingum, G. N., & McGuire, P. K. (2004). Misattribution of external speech in patients with hallucinations and delusions. *Schizophrenia Research*, 69, 277–287. <http://dx.doi.org/10.1016/j.schres.2003.09.008>
- Allen, P. P., Amaro, E., Fu, C. H. Y., Williams, S. C. R., Brammer, M., Johns, L. C., & McGuire, P. K. (2005). Neural correlates of the misattribution of self-generated speech. *Human Brain Mapping*, 26(1), 44–53. <http://dx.doi.org/10.1002/hbm.20120>
- Allen, P., Amaro, E., Fu, C. H. Y., Williams, S. C. R., Brammer, M. J., Johns, L. C., & McGuire, P. K. (2007). Neural correlates of the misattribution of speech in schizophrenia. *The British Journal of Psychiatry*, 190, 162–169. <http://dx.doi.org/10.1192/bjp.bp.106.025700>
- Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: common dimensions for different vowels and speakers. *Psychological Research*, 74(1), 110–120. <http://dx.doi.org/10.1007/s00426-008-0185-z>
- Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*, 8(3), 129–135. <http://dx.doi.org/10.1016/j.tics.2004.01.008>
- Berlad, I., & Pratt, H. (1995). P300 in response to the subject's own name. *Electroencephalography and Clinical Neurophysiology*, 96, 472–474.
- Boersma, P., & Weenink, D., 2012. Praat, Version 5.3. <http://www.fon.hum.uva.nl/praat/>
- Brébion, G., Smith, M. J., Gorman, J. M., & Amador, X. (1996). Reality monitoring failure in schizophrenia: the role of selective attention. *Schizophrenia Research*, 22(2), 173–180.
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59.
- Brosch, T., Scherer, K. R., Grandjean, D., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, 143, 1–10. <http://dx.doi.org/10.4414/smw.2013.13786>
- Canavarro, M. C. (1999). Inventário de Sintomas Psicopatológicos—BSI [Psychopathological Symptoms Inventory—BSI]. In R. Mário Simões, M. Gonçalves, & L. S. Almeida (Eds.), *Testes e provas psicológicas em Portugal* (2) (pp. 96–109). Braga: APPORT/SHO.
- Canavarro, M. C. (2007). Inventário de sintomas psicopatológicos: uma revisão crítica dos estudos realizados em Portugal [Psychopathological symptoms inventory: a critical revision of the Portuguese studies]. In L. Almeida, M. Simões, C. Machado e, & M. Gonçalves (Eds.), *Avaliação psicológica. Instrumentos validados para a população Portuguesa [Psychological assessment: tests validated for the Portuguese population]* (3) (pp. 305–331). Coimbra: Quarteto Editora.
- Carretié, L., Hinojosa, J. A., Albert, J., López-Martín, S., De La Gándara, B. S., Igoa, J. M., & Sotillo, M. (2008). Modulation of ongoing cognitive processes by emotionally intense words. *Psychophysiology*, 45(2), 188–196. <http://dx.doi.org/10.1111/j.1469-8986.2007.00617.x>
- Carretié, L., Martín-Loeches, M., Hinojosa, J. A., & Mercado, F. (2001). Emotion and attention interaction studied through event-related potentials. *Journal of Cognitive Neuroscience*, 13(8), 1109–1128. <http://dx.doi.org/10.1162/089892901753294400>
- Cassidy, S. M., Robertson, I. H., & O'Connell, R. G. (2012). Retest reliability of event-related potentials: evidence from a variety of paradigms. *Psychophysiology*, 49, 659–664. <http://dx.doi.org/10.1111/j.1469-8986.2011.01349.x>
- Chen, A., Weng, X., Yuan, J., Lei, X., Qiu, J., Yao, D., & Li, H. (2008). The temporal features of self-referential processing evoked by Chinese handwriting. *Journal of Cognitive Neuroscience*, 20(5), 816–827. <http://dx.doi.org/10.1162/jocn.2008.20505>
- Cook, S., & Wilding, J. (1997). Earwitness testimony: never mind the variety, hear the length. *Applied Cognitive Psychology*, 11(2), 95–111. [10.1002/\(SICI\)1099-0720\(199704\)11:2<95::AID-ACP429>3.0.CO;2-O](http://dx.doi.org/10.1002/(SICI)1099-0720(199704)11:2<95::AID-ACP429>3.0.CO;2-O)
- Cycowicz, Y. M., & Friedman, D. (1999). The effect of intention to learn novel, environmental sounds on the novelty P3 and old/new recognition memory. *Biological Psychology*, 50(1), 35–60. [http://dx.doi.org/10.1016/S0301-0511\(99\)00004-6](http://dx.doi.org/10.1016/S0301-0511(99)00004-6)
- Cygan, H. B., Tacikowski, P., Ostaszewski, P., Chojnicka, I., & Nowicka, A. (2014). Neural correlates of own name and own face detection in autism spectrum disorder. *PUBLIC LIBRARY OF SCIENCE*, 9(1), e86020. <http://dx.doi.org/10.1371/journal.pone.0086020>
- Czinger, I., Cox, T. J., Gyimesi, K., & Horváth, J. (2007). Event-related potential study to aversive auditory stimuli. *Neuroscience Letters*, 420, 251–256. <http://dx.doi.org/10.1016/j.neulet.2007.05.007>
- Delplanque, S., Silvert, L., Hot, P., Rigoulot, S., & Sequeira, H. (2006). Arousal and valence effects on event-related P3a and P3b during emotional categorization. *International Journal of Psychophysiology*, 60(3), 315–322. <http://dx.doi.org/10.1016/j.ijpsycho.2005.06.006>
- Delplanque, S., Silvert, L., Hot, P., & Sequeira, H. (2005). Event-related P3a and P3b in response to unpredictable emotional stimuli. *Biological Psychology*, 68(2), 107–120. <http://dx.doi.org/10.1016/j.biopsycho.2004.04.006>
- Eichenlaub, J.-B., Ruby, P., & Morlet, D. (2012). What is the specificity of the response to the own first-name when presented as a novel in a passive oddball paradigm? An ERP study. *Brain Research*, 1447, 65–78. <http://dx.doi.org/10.1016/j.brainres.2012.01.072>
- Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, 453, 1102–1106. <http://dx.doi.org/10.1038/nature06910>
- Fan, W., Zhang, Y., Wang, X., Wang, X., Zhang, X., & Zhong, Y. (2011). The temporal features of self-referential processing evoked by national flag. *Neuroscience Letters*, 505(3), 233–237. <http://dx.doi.org/10.1016/j.neulet.2011.10.017>
- Fan, W., Chen, J., Wang, X. Y., Cai, R., Tan, Q., Chen, Y., & Zhong, Y. (2013). Electrophysiological correlation of the degree of self-reference effect. *PUBLIC LIBRARY OF SCIENCE*, 8(12), e80289. <http://dx.doi.org/10.1371/journal.pone.0080289>
- Folmer, R. L., & Yingling, C. D. (1997). Auditory P3 responses to name stimuli. *Brain and Language*, 31(1), 306–311. <http://dx.doi.org/10.1006/brln.1997.1828>
- Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews*, 25(4), 355–373. [http://dx.doi.org/10.1016/S0149-7634\(01\)00019-7](http://dx.doi.org/10.1016/S0149-7634(01)00019-7)
- Gallup, G. G. (1985). Do minds exist in species other than our own. *Neuroscience and Biobehavioral Reviews*, 9(4), 631–641. [http://dx.doi.org/10.1016/0149-7634\(85\)90010-7](http://dx.doi.org/10.1016/0149-7634(85)90010-7)
- Gallup, G. G., Platek, S. M., & Spaulding, K. N. (2014). The nature of visual self-recognition revisited. *Trends in Cognitive Sciences*, 18(2), 57–58. <http://dx.doi.org/10.1016/j.tics.2013.10.012>
- Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*, 131(1), 76–97. <http://dx.doi.org/10.1037/0033-2909.131.1.76>
- Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55(4), 468–484.
- Graux, J., Gomot, M., Roux, S., Bonnet-Brihault, F., Camus, V., & Bruneau, N. (2013). My voice or yours? An electrophysiological study. *Brain Topography*, 26(1), 72–82. <http://dx.doi.org/10.1007/s10548-012-0233-2>
- Graux, J., Gomot, M., Roux, S., Bonnet-Brihault, F., & Bruneau, N. (2014). Is my voice just a familiar voice? An electrophysiological study. *Social Cognitive and Affective Neuroscience*, 10(1), 101–105. <http://dx.doi.org/10.1093/scan/nsu031>
- Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, 40, 216–224. [http://dx.doi.org/10.1016/S0022-1031\(03\)00092-1](http://dx.doi.org/10.1016/S0022-1031(03)00092-1)
- Håden, G. P., Németh, R., Török, M., Drávucz, S., & Winkler, I. (2013). Context effects on processing widely deviant sounds in newborn infants. *Frontiers in Psychology*, 4(674), 1–8. <http://dx.doi.org/10.3389/fpsyg.2013.00674>
- Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI



- study. *Human Brain Mapping*, 20(1), 22–28. <http://dx.doi.org/10.1002/hbm.10119>
- Herrmann, C. S., & Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience and Biobehavioral Reviews*, 25(6), 465–476, 239150.
- Holeckova, I., Fischer, C., Giard, M. H., Delpuech, C., & Morlet, D. (2006). Brain responses to a subject's own name uttered by a familiar voice. *Brain Research*, 1082(1), 142–152. <http://dx.doi.org/10.1016/j.brainres.2006.01.089>
- Hughes, S. M., & Nicholson, S. E. (2010). The processing of auditory and visual recognition of self-stimuli. *Consciousness and Cognition*, 19(4), 1124–1134. <http://dx.doi.org/10.1016/j.concog.2010.03.001>
- Illankovic, L. M., Allen, P. P., Engel, R., Kambeitz, J., Riedel, M., Müller, N., & Hennig-Fast, K. (2011). Attentional modulation of external speech attribution in patients with hallucinations and delusions. *Neuropsychologia*, 49(5), 805–812. <http://dx.doi.org/10.1016/j.neuropsychologia.2011.01.016>
- Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behavioural Brain Research*, 142(1–2), 1–15. [http://dx.doi.org/10.1016/S0166-4328\(02\)00384-4](http://dx.doi.org/10.1016/S0166-4328(02)00384-4)
- Kambeitz-Illankovic, L., Hennig-Fast, K., Benetti, S., Kambeitz, J., Pettersson-Yeo, W., O'Daly, O., & Allen, P. (2013). Attentional modulation of source attribution in first-episode psychosis: a functional magnetic resonance imaging study. *Schizophrenia Bulletin*, 39(5), 1027–1036. <http://dx.doi.org/10.1093/schbul/sbs101>
- Kaplan, J. T., Aziz-Zadeh, L., Uddin, L. Q., & Iacoboni, M. (2008). The self across the senses: an fMRI study of self-face and self-voice recognition. *Social Cognitive and Affective Neuroscience*, 3, 218–223. <http://dx.doi.org/10.1093/scan/nsn014>
- Keenan, J. P., Gallup, G. G., & Falk, D. (2003). *The face in the mirror: the search for the origins of consciousness*. New York, NY: Harper Collins Publishers.
- Keenan, J. P., Freund, S., Hamilton, R. H., Ganis, G., & Pascual-Leone, A. (2000). Hand response differences in a self-face identification task. *Neuropsychologia*, 38, 1047–1053. [http://dx.doi.org/10.1016/S0028-3932\(99\)00145-1](http://dx.doi.org/10.1016/S0028-3932(99)00145-1)
- Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, 39, 641–649. <http://dx.doi.org/10.1017/S0048577202394162>
- Kita, Y., Gunji, A., Inoue, Y., Goto, T., Sakihara, K., Kaga, M., & Hosokawa, T. (2011). Self-face recognition in children with autism spectrum disorders: a near-infrared spectroscopy study. *Brain and Development*, 33(6), 494–503. <http://dx.doi.org/10.1016/j.braindev.2010.11.007>
- Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature*, 383, 256–259.
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3), 557–577. <http://dx.doi.org/10.1017/S0048577201990559>
- Kumari, V., Fannon, D., Ffytche, D. H., Raveendran, V., Antonova, E., Premkumar, P., & Kuipers, E. (2010). Functional MRI of verbal self-monitoring in schizophrenia: performance and illness-specific effects. *Schizophrenia Bulletin*, 36(4), 740–755. <http://dx.doi.org/10.1093/schbul/sbn148>
- Latinus, M., & Belin, P. (2012). Perceptual auditory aftereffects on voice identity using brief vowel stimuli. *PUBLIC LIBRARY OF SCIENCE*, 7(7), e41384. <http://dx.doi.org/10.1371/journal.pone.0041384>
- Latinus, M., McAleer, P., Bestelmeyer, P. E. G., & Belin, P. (2013). Norm-based coding of voice identity in human auditory cortex. *Current Biology*, 23(12), 1075–1080. <http://dx.doi.org/10.1016/j.cub.2013.04.055>
- Leitman, D. I., Sehatpour, P., Garidis, C., Gomez-Ramirez, M., & Javitt, D. C. (2011). Preliminary evidence of pre-attentive distinctions of frequency-modulated tones that convey affect. *Frontiers in Human Neuroscience*, 5(96), 1–8. <http://dx.doi.org/10.3389/fnhum.2011.00096>
- Li, H., Yuan, J., & Lin, C. (2008). The neural mechanism underlying the female advantage in identifying negative emotions: an event-related potential study. *Neuroimage*, 40, 1921–1929. <http://dx.doi.org/10.1016/j.neuroimage.2008.01.033>
- Müller, H. M., & Kutas, M. (1996). What's in a name? Electrophysiological differences between spoken nouns, proper names and one's own name. *Neuroreport*, 8(1), 221–225. <http://dx.doi.org/10.1097/00001756-199612200-00045>
- Maurer, D., & Landis, T. (1990). Role of bone conduction in the self-perception of speech. *Folia Phoniatrica et Logopaedica*, 42(5), 226–229.
- Miyakoshi, M., Nomura, M., & Ohira, H. (2007). An ERP study on self-relevant object recognition. *Brain and Cognition*, 63(2), 182–189. <http://dx.doi.org/10.1016/j.bandc.2006.12.001>
- Morin, A. (2011). Self-awareness part 1: definition, measures, effects, functions, and antecedents. *Social and Personality Psychology Compass*, 5(10), 807–823. <http://dx.doi.org/10.1111/j.1751-9004.2011.00387.x>
- Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., & Kojima, S. (2001). Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia*, 39(10), 1047–1054. [http://dx.doi.org/10.1016/S0028-3932\(01\)00037-9](http://dx.doi.org/10.1016/S0028-3932(01)00037-9)
- Nielsen, M., Dissanayake, C., & Kashima, Y. (2003). A longitudinal investigation of self-other discrimination and the emergence of mirror self-recognition. *Infant Behavior and Development*, 26(2), 213–226. [http://dx.doi.org/10.1016/S0163-6383\(03\)00018-3](http://dx.doi.org/10.1016/S0163-6383(03)00018-3)
- Ninomiya, H., Onitsuka, T., Chen, C. H., Sato, E., & Tashiro, N. (1998). P300 in response to the subject's own face. *Psychiatry and Clinical Neurosciences*, 52, 519–522. <http://dx.doi.org/10.1046/j.1440-1819.1998.00445.x>
- O'Donnell, B. F., Shenton, M. E., McCarley, R. W., Faux, S. F., Smith, R. S., Salisbury, D. F., & Jolesz, F. (1993). The auditory N2 component in schizophrenia: relationship to MRI temporal lobe gray matter and to other ERP abnormalities. *Biological Psychiatry*, 34(1–2), 26–40. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8373937>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97–113. [http://dx.doi.org/10.1016/0028-3932\(71\)90067-4](http://dx.doi.org/10.1016/0028-3932(71)90067-4)
- Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological Psychology*, 77(3), 247–265. <http://dx.doi.org/10.1016/j.biopsycho.2007.11.006>
- Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: selected studies of the event-related Potential. *International Journal of Medical Sciences*, 2(4), 147–154. <http://dx.doi.org/10.7150/ijms.2.147>
- Perrin, F., García-Larrea, L., Mauguière, F., & Bastuji, H. (1999). A differential brain response to the subject's own name persists during sleep. *Clinical Neurophysiology*, 110, 2153–2164. [http://dx.doi.org/10.1016/S1388-2457\(99\)00177-7](http://dx.doi.org/10.1016/S1388-2457(99)00177-7)
- Perrin, F., Maquet, P., Peigneux, P., Ruby, P., Degueldre, C., Baletau, E., & Laureys, S. (2005). Neural mechanisms involved in the detection of our first name: a combined ERPs and PET study. *Neuropsychologia*, 43(1), 12–19. <http://dx.doi.org/10.1016/j.neuropsychologia.2004.07.002>
- Platek, S. M., Wathne, K., Tierney, N. G., & Thomson, J. W. (2008). Neural correlates of self-face recognition: an effect-location meta-analysis. *Brain Research*, 1232, 173–184. <http://dx.doi.org/10.1016/j.brainres.2008.07.010>
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical EEG and Neuroscience*, 118(10), 2128–2148. <http://dx.doi.org/10.1016/j.clinph.2007.04.019>
- Polich, J., & Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology*, 60(2), 172–185. <http://dx.doi.org/10.1016/j.ijpsycho.2005.12.012>
- Rosa, C., Lassonde, M., Pinaud, C., Keenan, J. P., & Belin, P. (2008). Investigations of hemispheric specialization of self-voice recognition. *Brain and Cognition*, 68, 204–214. <http://dx.doi.org/10.1016/j.bandc.2008.04.007>
- Salisbury, D. F., O'Donnell, B. F., McCarley, R. W., Shenton, M. E., & Benavente, A. (1994). The N2 event-related potential reflects attention in schizophrenia deficit. *Biological Psychiatry*, 39(1), 1–13. [http://dx.doi.org/10.1016/0301-0511\(94\)90053-1](http://dx.doi.org/10.1016/0301-0511(94)90053-1)
- Sandman, C., & Patterson, J. V. (2000). The auditory event-related potential is a stable and reliable measure in elderly subjects over a 3 year period. *Clinical Neurophysiology*, 111, 1427–1437. [http://dx.doi.org/10.1016/S1388-2457\(00\)00320-5](http://dx.doi.org/10.1016/S1388-2457(00)00320-5)
- Schirmer, A., & Escoffier, N. (2010). Emotional MMN: anxiety and heart rate correlate with the ERP signature for auditory change detection. *Clinical Neurophysiology*, 121(1), 53–59. <http://dx.doi.org/10.1016/j.clinph.2009.09.029>
- Schirmer, A., Simpson, E., & Escoffier, N. (2007). Listen up! Processing of intensity change differs for vocal and nonvocal sounds. *Brain Research*, 1176, 103–112. <http://dx.doi.org/10.1016/j.brainres.2007.08.008>
- Schirmer, A., Striano, C. A. T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport*, 16(6), 635–639. <http://dx.doi.org/10.1097/00001756-200504250-00024>
- Schröger, E., Giard, M. H., & Wolff, C. (2000). Auditory distraction: event-related potential and behavioral indices. *Clinical Neurophysiology*, 111(8), 1450–1460. [http://dx.doi.org/10.1016/S1388-2457\(00\)00337-0](http://dx.doi.org/10.1016/S1388-2457(00)00337-0)
- Schweinberger, S. R., Herholz, A., & Sommer, W. (1997). Recognizing famous voices influence of stimulus duration and different types of retrieval cues. *Journal of Speech, Language, and Hearing Research*, 40(2), 453–463.
- Schweinberger, S. R., Kawahara, H., Simpson, A. P., Skuk, V. G., & Zäske, R. (2014). Speaker perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1), 15–25. <http://dx.doi.org/10.1002/wics.1261>
- Scott, L. S., Luciana, M., & Wewerka, S. (2005). Electrophysiological correlates of facial self-recognition in adults and children. *Cognition, Brain, Behavior*, 9(3), 211–238.
- Simons, C. J. P., Sambeth, A., Krabbendam, L., Pfeifer, S., van Os, J., & Riedel, W. J. (2011). Auditory P300 and N100 components as intermediate phenotypes for psychotic disorder: familial liability and reliability. *Clinical Neurophysiology*, 122(10), 1984–1990. <http://dx.doi.org/10.1016/j.clinph.2011.02.033>
- Soares, A. P., Comesaña, M., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., et al. (2010). P-PAL: uma base lexical com índices psicolinguísticos do Português Europeu [P-PAL: a European Portuguese lexical database]. *Linguística*, 2(3), 67–72.
- Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods*, 44(1), 256–269. <http://dx.doi.org/10.3758/s13428-011-0131-7>
- Spencer, K. M., Dien, J., & Donchin, E. (1999). A componential analysis of the ERP elicited by novel events using a dense electrode array. *Psychophysiology*, 36(3), 409–414. <http://dx.doi.org/10.1017/S0048577299981180>
- Spencer, K. M., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology*, 38(2), 343–358. <http://dx.doi.org/10.1017/S0048577201000324>
- Su, Y., Chen, A., Yin, H., Qiu, J., Lv, J., Wei, D., & Wang, T. (2010). Spatiotemporal cortical activation underlying self-referential processing evoked by self-hand. *Biological Psychology*, 85(2), 219–225. <http://dx.doi.org/10.1016/j.biopsycho.2010.07.004>

- Sugiura, M., Kawashima, R., Nakamura, K., Okada, K., Kato, T., Nakamura, A., & Fukuda, H. (2000). Passive and active recognition of one's own face. *Neuroimage*, 11(1), 36–48. <http://dx.doi.org/10.1006/nimg.1999.0519>
- Sui, J., Zhu, Y., & Han, S. (2006). Self-face recognition in attended and unattended conditions: an event-related brain potential study. *Neuroreport*, 17(4), 423–427. <http://dx.doi.org/10.1097/01.wnr.0000203357.65190.61>
- Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: an ERP study. *Biological Psychology*, 84, 318–324. <http://dx.doi.org/10.1016/j.biopsycho.2010.03.009>
- Tacikowski, P., Jednoróg, K., Marchewka, A., & Nowicka, A. (2011). How multiple repetitions influence the processing of self-, famous and unknown names and faces: an ERP study. *International Journal of Psychophysiology*, 79(2), 219–230. <http://dx.doi.org/10.1016/j.ijpsycho.2010.10.010>
- Tacikowski, P., Cygan, H. B., & Nowicka, A. (2014). Neural correlates of own and close-other's name recognition: ERP evidence. *Frontiers in Human Neuroscience*, 8(194), 1–10. <http://dx.doi.org/10.3389/fnhum.2014.00194>
- Tateuchi, T., Itoh, K., & Nakada, T. (2012). Neural mechanisms underlying the orienting response to subject's own name: an event-related potential study. *Psychophysiology*, 49(6), 786–791. <http://dx.doi.org/10.1111/j.1469-8986.2012.01363.x>
- Thierry, G., & Roberts, M. V. (2007). Event-related potential study of attention capture by affective sounds. *Neuroreport*, 18(3), 245–248. <http://dx.doi.org/10.1097/WNR.0b013e328011dc95>
- Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. *Journal of Experimental Psychology*, 25(4), 1016–1035.
- Vüilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9(12), 585–594. <http://dx.doi.org/10.1016/j.tics.2005.10.011>
- Wang, J., Nicol, T., Skoe, E., Sams, M., & Kraus, N. (2009). Emotion modulates early auditory response to speech. *Journal of Cognitive Neuroscience*, 21(11), 2121–2128. <http://dx.doi.org/10.1162/jocn.2008.21147>
- Waters, F., & Badcock, J. C. (2010). First-rank symptoms in schizophrenia: reexamining mechanisms of self-recognition. *Schizophrenia Bulletin*, 36(3), 510–517. <http://dx.doi.org/10.1093/schbul/sbn112>
- Waters, F., Woodward, T., Allen, P., Aleman, A., & Sommer, I. (2012). Self-recognition deficits in schizophrenia patients with auditory hallucinations: a meta-analysis of the literature. *Schizophrenia Bulletin*, 38(4), 741–750. <http://dx.doi.org/10.1093/schbul/sbq144>
- Wechsler, D. (1997). Escala de inteligência de Wechsler para adultos. In *Instruções para a administração e cotação* (3ª Edição). Lisboa: Cegoc-Tea [Wechsler Adult Intelligence Scale—Third Edition: Administration and scoring manual]. (trad. e adapt. Portuguesa Cegoc-Tea, 2008).
- Xu, M., Homae, F., Hashimoto, R., & Hagiwara, H. (2013). Acoustic cues for the recognition of self-voice and other-voice. *Frontiers in Psychology*, 4(735), 1–7. <http://dx.doi.org/10.3389/fpsyg.2013.00735>
- Yuan, J., Zhang, Q., Chen, A., Li, H., Wang, Q., Zhuang, Z., & Jia, S. (2007). Are we sensitive to valence differences in emotionally negative stimuli? Electrophysiological evidence from an ERP study. *Neuropsychologia*, 45, 2764–2771. <http://dx.doi.org/10.1016/j.neuropsychologia.2007.04.018>
- Zhao, K., Wu, Q., Zimmer, H. D., & Fu, X. (2011). Electrophysiological correlates of visually processing subject's own name. *Neuroscience Letters*, 491, 143–147. <http://dx.doi.org/10.1016/j.neulet.2011.01.025>