Brain & Language 153–154 (2016) 38–49

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Did you or I say pretty, rude or brief? An ERP study of the effects of speaker's identity on emotional word processing

Ana P. Pinheiro [a,b,c,*], Neguine Rezaii [b], Paul G. Nestor [b,d], Andréia Rauber [e], Kevin M. Spencer [f], Margaret Niznikiewicz [b]

[a] *Neuropsychophysiology Laboratory, Psychology Research Center (CIPsi), School of Psychology, University of Minho, Braga, Portugal*
[b] *Clinical Neuroscience Division, Laboratory of Neuroscience, VA Boston Healthcare System-Brockton Division, Department of Psychiatry, Harvard Medical School, Brockton, MA, United States*
[c] *Faculty of Psychology, University of Lisbon, Lisbon, Portugal*
[d] *Department of Psychology, University of Massachusetts, Boston, MA, United States*
[e] *International Studies in Computational Linguistics, University of Tübingen, Tübingen, Germany*
[f] *Neural Dynamics Laboratory, Research Service, VA Boston Healthcare System, and Department of Psychiatry, Harvard Medical School, Boston, MA, United States*

ARTICLE INFO

Article history:
Received 24 June 2015
Revised 19 November 2015
Accepted 10 December 2015
Available online 16 February 2016

ABSTRACT

During speech comprehension, multiple cues need to be integrated at a millisecond speed, including semantic information, as well as voice identity and affect cues. A processing advantage has been demonstrated for self-related stimuli when compared with non-self stimuli, and for emotional relative to neutral stimuli. However, very few studies investigated self-other speech discrimination and, in particular, how emotional valence and voice identity interactively modulate speech processing. In the present study we probed how the processing of words' semantic valence is modulated by speaker's identity (self *vs.* non-self voice).

Sixteen healthy subjects listened to 420 prerecorded adjectives differing in voice identity (self *vs.* non-self) and semantic valence (neutral, positive and negative), while electroencephalographic data were recorded. Participants were instructed to decide whether the speech they heard was their own (self-speech condition), someone else's (non-self speech), or if they were unsure.

The ERP results demonstrated interactive effects of speaker's identity and emotional valence on both early (N1, P2) and late (Late Positive Potential – LPP) processing stages: compared with non-self speech, self-speech with neutral valence elicited more negative N1 amplitude, self-speech with positive valence elicited more positive P2 amplitude, and self-speech with both positive and negative valence elicited more positive LPP. ERP differences between self and non-self speech occurred in spite of similar accuracy in the recognition of both types of stimuli.

Together, these findings suggest that emotion and speaker's identity interact during speech processing, in line with observations of partially dependent processing of speech and speaker information.

1. Introduction

In the course of a conversation, people have to rapidly detect and integrate multiple signals in order to make sense of speech information. These signals include not only linguistic but also paralinguistic cues (Belin, Bestelmeyer, Latinus, & Watson, 2011; Schirmer & Kotz, 2006). From the perspective of a listener, it is not only important to understand "what" is being said and "how", but also to relate that information to "who" is saying it (Belin, Fecteau, & Bédard, 2004; Belin et al., 2011; Formisano, De Martino, Bonte, & Goebel, 2008). From the perspective of a speaker, it is important to distinguish between self-generated and non-self generated voices, i.e. to recognize speech as one's own.

Studies in the last two decades lend support to the idea that distinct types of information – speech, affect and identity – conveyed by the voice are processed by functionally dissociable neural pathways: the analysis of *speech* information recruits temporal (anterior and posterior superior temporal sulcus) and inferior prefrontal regions, particularly in the left hemisphere; the analysis of *vocal affect* recruits temporo-medial regions, the anterior insula, the amygdala and inferior prefrontal regions, particularly in the right hemisphere; the analysis of *vocal identity* recruits regions of the right anterior superior temporal sulcus (e.g., Belin et al., 2004, 2011). Nonetheless, the interactions between these distinct

---

* Corresponding author at: Neuropsychophysiology Lab, 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.bandl.2015.12.003

types of information remain largely unexplored. In spite of the limited number of studies probing the neuro-functional correlates of speaker's identity processing, there is evidence indicating that identity-related information is extracted and used from early stages of speech perception, within the first 200–300 milliseconds (ms) after a spoken word onset (Van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008). However, few studies have investigated the interaction between speaker's identity and speech semantic valence and, in particular, how the assessment of self-relevance (self *vs.* non-self voice) influences emotional language comprehension at the neural level.

1.1. The role of voice identity during speech perception

A considerable number of studies have demonstrated a processing advantage for self-related stimuli when compared with non-self stimuli, such as one's own name (Gray, Ambady, Lowenthal, & Deldin, 2004; Tacikowski & Nowicka, 2010; Zhao et al., 2009), self-related pronouns (Herbert, Herbert, Ethofer, & Pauli, 2011; Zhou et al., 2010), self-face (Keyes, Brady, Reilly, & Foxe, 2010; Sui, Zhu, & Han, 2006; Tacikowski & Nowicka, 2010), self-relevant objects (Miyakoshi, Nomura, & Ohira, 2007), or self-hand (Su et al., 2010).

Few studies probed the neural correlates of self-voice perception and, in particular, the time course of self-generated speech processing (Conde, Gonçalves, & Pinheiro, 2015, in press; Graux, Gomot, Roux, Bonnet-Brilhault, & Bruneau, 2015; Graux et al., 2013). More recent investigations suggest that self and non-self voices are discriminated within the first 100 ms after voice onset (e.g., Ford et al., 2001), and indicate high self-voice recognition rates (94–96% – Rosa, Lassonde, Pinard, Keenan, & Belin, 2008). ERP studies testing the corollary discharge mechanism (i.e., the expected sensory feedback resulting from one's own action) demonstrated effects on the N1 component, a negativity peaking between 80 and 120 ms after stimulus onset, and maximal at frontocentral electrodes locations (e.g., Rosburg, Boutros, & Ford, 2008). Generally, the N1 is sensitive to the physical properties of the stimuli, and has been proposed as a neurophysiological signature of automatic attention allocation to salient stimuli (Rosburg et al., 2008). Specifically, the studies that probed the corollary discharge showed that the N1 amplitude is reduced in response to the onset of unaltered self-voice auditory feedback during speech production in comparison with the passive listening to the recording of one's own voice (e.g., Baess, Widmann, Roye, Schröger, & Jacobsen, 2009; Behroozmand & Larson, 2011; Ford et al., 2001); or in response to self-triggered (i.e., button press) relative to externally-triggered tones (Behroozmand & Larson, 2011; Knolle, Schröger, & Kotz, 2013a, 2013b). As such, these studies expanded the functional significance of the N1, by demonstrating that this component is also sensitive to voice identity, and that it is specifically modulated by stimulus predictability or agency (e.g., being the author of the action of pressing a button to elicit a sound).

These experiments also reported effects on the P2 component, a positivity that is typically observed around 150–300 ms after stimulus onset, and has been associated with early stimulus categorization and attention processes (e.g., Crowley & Colrain, 2004), and more recently with the detection of the emotional salience of a stimulus (e.g., Paulmann & Kotz, 2008; Pinheiro et al., 2013, 2014). Specifically, these experiments demonstrated reduced P2 amplitude in response to self-generated sounds (Knolle et al., 2013b). These effects were interpreted as a neurophysiological signature of the conscious detection of a self-initiated auditory stimulus (Knolle et al., 2013b). However, a particular methodological feature of this type of paradigms is that they either involve active speech production or a motor condition, in which participants are asked to press a button in order to elicit a given sound (e.g., Knolle et al., 2013a, 2013b), with its disadvantages in terms of artifacts during EEG data acquisition (e.g., Ford et al., 2001).

Of relevance to the current study, evidence from experiments consisting of the passive listening to pre-recorded self-voice and non-self voice stimuli indicated that voice stimuli can be discriminated as a function of their identity in early stages of information processing. The differential sensitivity of the ERP response to self-relevance was corroborated by studies reporting effects around 100 ms. For example, Graux et al. (2013) observed that self and non-self pre-recorded voices were discriminated within the first 100 ms after voice onset: the self-voice was characterized by greater negative amplitude within this time window compared with the unfamiliar voice. Together, these studies reveal important effects of voice identity in early processing stages, i.e. within the first 200 ms after voice onset (N1, P2).

Moreover, in a recent study, we probed how self-relevance (pre-recorded self *vs.* non-self voice) modulates selective attention (Conde et al., 2015). We found that selective attention to voice stimuli was enhanced in the case of self-relevant ('my voice') compared to non-self ('someone else's voice) stimuli. This finding suggests that in experiments consisting of listening to task-relevant pre-recorded self *vs.* non-self voices, N1/P2 amplitude for the self-voice might be increased, rather than decreased as reported in experiments requiring speech production (e.g., Ford et al., 2001) or a button-press eliciting a sound (e.g., Knolle et al., 2013b). In this case, self-relevance may engender increased neural activity and larger ERP amplitude through an increase of attention (e.g., self-voice – Conde et al., 2015; self-name and self-face – Tacikowski & Nowicka, 2010). The finding of Conde et al. (2015) additionally indicated that a self-voice represents a particularly salient stimulus, suggesting that the self-relevance and the emotionality of a stimulus may produce similar effects.

1.2. The role of semantic valence in speech processing

A close relationship between the processing of self-related information and the processing of emotional valence (i.e., the overall unpleasantness/displeasure relative to pleasantness/attraction of a stimulus – Bradley & Lang, 1994) has been reported (e.g., Fossati et al., 2003). The existing evidence converges in showing that we quickly discriminate between emotionally salient and neutral stimuli, and that this differentiation occurs already at early processing stages, in the first 200 ms after speech onset (e.g., Paulmann & Kotz, 2008; Pinheiro et al., 2013, 2014). Specifically, larger P2 amplitude was observed for positive relative to neutral words (e.g., Kanske & Kotz, 2007; Kissler, Assadollahi, & Herbert, 2006), or for emotional (negative and positive) relative to neutral words (Bernat, Bunce, & Shevrin, 2001; Schapkin, Gusev, & Kühl, 2000), even when presented subliminally (Bernat et al., 2001). A putative explanation for these P2 effects is that they reflect increased automatic attentional capture by emotional stimuli (Kanske, Plitschka, & Kotz, 2011). Emotion effects were also observed in a later positive component, observed after 500 ms post-stimulus onset – the Late Positive Component (LPP). Some studies found a processing advantage for pleasant words, suggesting that they lead to increased sustained attention and deeper stimulus encoding relative to both negative and neutral words, which is reflected in increased amplitude of the LPP (e.g., Ferré, 2003; Herbert, Junghofer, & Kissler, 2008). Together, these studies show that the effects of the emotionality of verbal stimuli may be observed in both early (P2) and later (LPP) components. Thus, in addition to P2 indexing identity as described above (e.g., Knolle et al., 2013b), these findings highlight the relevance of the P2 to the investigation of both voice identity and emotion processing. We note that in case of self-relevant and emotional speech stimuli, these effects should be additive.

1.3. The interactions between self-relevance and emotion

More recent studies have demonstrated that the differential processing of positive and negative information also occurs at the personal knowledge level (Fossati et al., 2004), and that emotionally arousing words and self-relevant stimuli share some similarities, as both represent highly salient and evolutionarily relevant cues (Gray et al., 2004; Phan et al., 2004). For example, the recognition of self-relevant material (‘like me’/‘not like me’) interacts with the recognition of the emotional valence of personality trait words (e.g., “friendly”, “arrogant”). A *self-positivity bias* has been demonstrated, indicating that healthy subjects tend to process positive information as related to personal characteristics, and negative information as unrelated to personal characteristics (Heine, Lehman, Markus, & Kitayama, 1999; Herbert et al., 2008; Watson, Dritschel, Obonsawin, & Jentzsch, 2007). In the ERP study of Watson et al. (2007), participants were presented with trait words and asked to perform either an emotional judgment task (positive/negative) or a self-referential task (‘like me’/‘not like me’). An effect of valence was observed at centro-parietal electrodes between 400 and 550 ms after word onset: amplitude was less positive for negative than for positive words. Furthermore, an interaction between self-reference and valence was observed at 450–600 ms over fronto-central electrodes, and indicated less positive amplitude for positive words rated as non-self referential and for negative words rated as self-referential. These studies suggest that the interactions between self-relevance and emotionality are reflected in LPP modulations, substantiating the relevance of this ERP component to the current study.

1.4. The current study and hypotheses

The studies reviewed in the previous sections demonstrated that the effects of self-relevance are similar to those of emotion (e.g., Fields & Kuperberg, 2012; Tacikowski & Nowicka, 2010; Watson et al., 2007) and, in particular, that emotional words are distinctly processed when presented in a personal context (e.g., ‘like me’/’not like me’) *vs.* a non-personal context (e.g., Fossati et al., 2003). In the present study we sought to probe how listeners integrate ‘*what*’ is being said with ‘*who*’ is saying it, extending previous studies by using a self-voice condition. In particular, we aimed to determine whether neutral and emotional spoken words are differentially processed as function of voice identity (self *vs.* non-self). We investigated these questions by examining ERP recordings in healthy subjects who listened to a series of prerecorded words (self *vs.* non-self generated) varying in emotional valence (neutral – e.g., “brief”; positive – e.g., “pretty”; and negative – e.g., “rude”). The participants were instructed to decide whether the speech they heard was their own, if it belonged to someone else or if they were unsure.

In spite of the fact that hearing a recording of a self-generated voice (sound is perceived through air conduction alone) is different from hearing vocal feedback during speech production (sound is perceived through both air and bone conduction), previous studies have shown that self-generated vocal stimuli were successfully recognized as “self” even when no frequency filter was used to make the self-voice sound more natural[1] (Nakamura et al., 2001; Rosa et al., 2008). The ERP methodology is particularly suited to studying the time course of speech processing, allowing a careful investigation of distinct processing stages associated with voice discrimination and recognition and with emotional semantic processing.

If the processing of self *vs.* non-self distinctions occurs independently of semantic emotional valence, then we should observe no interaction between these two factors at the ERP and behavioral levels. The independent effects could be manifested in ERP differences between self and non-self speech irrespective of semantic valence type, or in differences between neutral and emotional words irrespective of speaker’s identity (self *vs.* non-self), reflected in ERP components sensitive to emotion or self-relevance modulations, i.e. N1, P2, and LPP (N1 – voice identity: Ford et al., 2001; Knolle et al., 2013b; P2 – voice identity: Knolle et al., 2013b; emotion: Bernat et al., 2001; Kanske & Kotz, 2007; Kissler et al., 2006; Schapkin et al., 2000; LPP – emotion: Ferré, 2003; Herbert et al., 2008). We should note here that studies examining the corollary discharge mechanism explore sensory predictive processes (e.g., Ford et al., 2001; Knolle et al., 2013a, 2013b), while studies examining self *vs.* other voice recognition from pre-recorded stimuli probe attentional processes and salience of the self-voice in comparison with someone else’s voice (e.g., Conde et al., 2015). In experiments that used the latter framework, larger ERP responses were recorded to self- relative to non-self voice stimuli (e.g., Conde et al., 2015). Since our experiment is the first to explicitly examine the impact of salience (both self-relevance and emotion) on early N1/P2 components in the context of participants listening to pre-recorded self- and non-self voices, we extrapolated from such studies as Conde et al. (2015) and predicted increased N1/P2 and increased LPP amplitude for self- relative to non-self speech (identity effects). Additionally, we expected increased P2 and LPP for emotional relative to neutral speech (semantic valence effects).

We also hypothesized differences in recognition accuracy as a function of identity (e.g., increased recognition of non-self relative to self-speech stimuli) or of valence (e.g., increased recognition of emotional speech compared to neutral speech). This would be consistent with some evidence showing that important information conveyed by the speaker’s voice may be perceived independently of semantic information (e.g., Belin, Zatorre, & Ahad, 2002; von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003).

If, on the other hand, semantic emotional valence interacts with voice identity in modulating speech processing, we expected to observe *valence-specific* differences between self- and non-self speech in ERP components sensitive to emotion *and* self-relevance modulations, i.e. P2 and LPP (P2 – voice identity: Knolle et al., 2013b; emotion: Bernat et al., 2001; Kanske & Kotz, 2007; Kissler et al., 2006; Schapkin et al., 2000; LPP – emotion: Ferré, 2003; Herbert et al., 2008; identity and emotion: Watson et al., 2007). Moreover, these interactions would be reflected in recognition accuracy. This would be consistent with the observation that verbal and nonverbal features interact during speech processing (e.g., von Kriegstein et al., 2003). Specifically, we predicted that these differences would be enhanced for self-speech with emotional valence relative to neutral speech, indexed by increased P2, increased LPP, and increased recognition accuracy, reflecting amplifying (additive) effects of each factor. This would be in line with the increased salience of emotional stimuli (relative to neutral stimuli) and of self-generated stimuli (relative to non-self stimuli) (e.g., Conde et al., 2015; Heine et al., 1999; Herbert et al., 2008; Watson et al., 2007). Moreover, given the self-positivity bias reported in studies using verbal material (Heine et al., 1999; Herbert et al., 2008; Watson et al., 2007), we expected increased P2 and LPP, as well as increased recognition accuracy, for self-speech with positive content relative to both neutral and negative speech.

---

[1] In order to minimize differences in sound conduction during active production of speech and passive listening to recorded self-voice stimuli, and to make a self-generated voice sound more natural, the procedure adopted by some studies was to apply an equalization filter that increases frequencies below 1000 Hz by 2 dB and decreases frequencies above 1000 Hz by 2 dB (e.g., Kaplan, Aziz-Zadeh, Uddin, & Iacoboni, 2008).

2. Method

2.1. Participants

Twenty healthy individuals participated in the experiment. Four participants were excluded due to excessive artifacts leaving 16 subjects (6 females), aged between 30 and 52 years (*Mean age* = 44.38, *SD* = 6.50), for the analyses. All participants spoke American English as their native language, and had normal hearing as assessed by a subjective audiometric test before the ERP experiment. They were recruited from advertisements in local newspapers and in the Internet, and had, on average, 15.31 years of education (*SD* = 1.74). The inclusion criteria were: English as first language; right handedness as confirmed by the Edinburgh Inventory (Oldfield, 1971); no history of neurological or major medical illness; no history of drug or alcohol abuse (American Psychiatric Association, 2000); verbal intelligence quotient (IQ; Wechsler, 1997) above 90 (*M* = 107.22, *SD* = 21.24); no hearing, vision or upper body impairment; no history of Axis I or II disorders as determined by the *Structured Clinical Interview for DSM-IV* (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition) *for Axis I* (SCID-I – First, Spitzer, Gibbon, & Williams, 2002) *and Axis II* (SCID-II – First, Gibbon, & Spitzer, 1997) disorders; no history of Axis I disorder in first or second degree family members, as determined by the *Family History-Research Diagnostic Criteria* (FH-RDC) instrument (Andreasen, Endicott, Spitzer, & Winokur, 1977). Before participation in the study, all participants had the procedures fully explained to them and read and signed an informed consent form to confirm their willingness to participate in the study (following Harvard Medical School and Veterans Affairs Boston Healthcare System guidelines). Subjects were paid for their participation.

2.2. Stimuli

Stimuli included 70 adjectives with neutral semantic valence[2] (e.g., “raw”, “round”), 70 adjectives with positive semantic valence (e.g., “beautiful”, “bright”), and 70 adjectives with negative semantic valence (e.g., “stupid”, “sinful”) (see Appendix A and Table 1). Words were tested for frequency (Brown verbal frequency), concreteness, familiarity, imageability, number of phonemes, and number of syllables, based on the MRC Psycholinguistic Database (Coltheart, 2007; http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm). Differences between word valence categories were observed for the number of phonemes: neutral words were characterized by a higher number of phonemes than negative words ($p = .018$). Valence ratings were obtained from the norms for 13,915 English lemmas published by Warriner, Kuperman, and Brysbaert (2013), and additionally validated in a previous study of our team with a sample of 17 participants (Pinheiro, McCarley, Thompson, Gonçalves, & Niznikiewicz, 2012). As expected, valence scores of neutral words were higher than scores of negative words ($p < .001$), but lower than scores of positive words ($p < .001$; valence effect – $F(2,30) = 182.06$, $p < .001$). 210 stimuli were selected to make up two comparable lists of neutral, positive, and negative words (see Appendix A).

2.3. Procedure

Each subject participated in two experimental sessions. The first session was similar for all subjects and involved the recording of the participant’s voice.

2.3.1. Recording of the participant’s voice

Each participant was asked to read aloud a list of 210 adjectives with neutral or emotional valence (*self-speech* condition – SS). The words were shown in the center of a computer screen, one at a time. Before seeing the word, participants were instructed to listen to that same word pronounced by an unknown speaker using neutral prosody. They were instructed to match the loudness and neutral prosody of each target word as spoken by the speaker at constant voice intensity (65 dB). Occasionally, participants were instructed to repeat the word if it was not pronounced with the desired prosody and loudness. The inclusion of a “voice-model” aimed at reducing between-subjects variability in speech rate, voice loudness and pitch. Recordings were made in a quiet room with an Edirol R-09 recorder and a CS-15 cardioid-type stereo microphone, with a sampling rate of 44,100 kHz and 16-bit quantization. The electroencephalogram (EEG) was not recorded in this stage of the study.

For the *non-self speech* (NSS) condition, the same 210 words were recorded by a male (age = 43 years) or female (age = 44 years) native speaker[3] of American English unknown to the participants. The words were spoken with neutral intonation and constant voice intensity, following the same procedure as described above.

After the recording session, each word was segmented using Praat software (Boersma & Weenink, 2013). First, voice stimuli were normalized according to peak amplitude by means of a Praat script. Acoustic noise was reduced using a Fourier-based noise reduction algorithm (noise reduction = 14 dB; frequency smoothing = 150 Hz; attack/decay time = 0.15 s) implemented in Audacity 2.0.2 software (http://audacity.sourceforge.net/). Mean pitch, intensity and duration were subsequently compared across conditions (see Table 2). There were no differences between valence types in F0 and intensity ($p > .05$). However, neutral words had shorter duration than both positive ($p < .001$) and negative words ($p < .001$; $F(4,60) = 31.149$, $p < .001$).

2.3.2. ERP experiment

The ERP session took place at least two weeks after the recording session. Four hundred and twenty adjectives were presented: 210 previously recorded by the participant, and 210 previously recorded by someone unknown to the participant. The *identity* (self/non-self) and *valence* (negative/positive/neutral) of speech varied across trials, with 70 words per condition. The six combinations of speech identity and valence were ordered pseudo-randomly and presented in two lists, with the constraint of no more than three consecutive trials of the same condition. Half of the participants received the lists in AB sequence, and half in BA sequence.

The experimental procedure is outlined in Fig. 1. Each participant was seated comfortably at a distance of 100 centimeters (cm) from a computer monitor in a sound-attenuating chamber. Participants indicated if the words were spoken in their own voice, another person’s voice, or were unsure, via a button press on a Cedrus response pad (RB-830, San Pedro, USA). The availability of an “unsure” option allowed participants to make a choice between “self” and “other” with some degree of confidence, instead of a forced choice. The letters ‘S’, ‘O’, and ‘U’ (for Self, Other and Unsure) were presented in the middle of the screen at the end of each trial. Buttons of the response pad were also marked with the S, O, and U letters to minimize memory demands. Order of buttons was counterbalanced.

Before each word onset, a fixation cross was presented centrally on the screen for 1500 ms, and was kept during word presentation

---

[2] The term “semantic valence” is used throughout the manuscript to indicate neutral *vs.* emotional semantic content, as opposed to valence associated with prosodic content.

[3] For a male participant, a male control (‘non-self’) voice was used; for a female participant, a female control (‘non-self’) voice was used.

Table 1
Psycholinguistic and affective properties of the words included in the experiment, for each valence type.

		Word valence category			*F*, *p*
		Neutral	Positive	Negative	
Psycholinguistic properties	Brown verbal frequency[a]	12.61 (24.15)	17.97 (53.51)	18.44 (100.57)	0.162, .850
	Concreteness (100–700)[b]	165.34 (202.17)	106.87 (161.07)	111.04 (162.19)	2.400, .093
	Familiarity (100–700)[c]	360.71 (32.64)	329.36 (281.16)	257.70 (276.35)	2.546, .081
	Imageability[d]	263.76 (218.77)	248.33 (214.50)	200.59 (216.34)	1.619, .201
	Number of letters	5.40 (1.43)	6.50 (1.68)	5.97 (1.48)	8.982, <.001[*]
	Number of phonemes	4.43 (1.81)	3.76 (2.89)	3.30 (2.37)	3.921, .021[**]
	Number of syllables	1.67 (0.72)	1.81 (0.97)	1.61 (0.82)	1.049, .352
Affective properties	Valence[e]	5.64 (0.58)	7.34 (0.51)	2.83 (0.57)	710.108, <.001[*]

[*] $p < .001$.
[**] $p < .05$.
[a] This measure includes 14,529 entries that range from 0 to 6833 ($M = 35$; $SD = 252$).
[b] Values range from 100 to 700 ($M = 438$; $SD = 120$).
[c] Values range from 100 to 700 ($M = 488$; $SD = 99$).
[d] Values range from 100 to 700 ($M = 450$; $SD = 108$).
[e] Values range from 1 to 9 (Warriner et al., 2013). Standard deviations in parentheses.

Table 2
Acoustic properties of the voice stimuli.

	Neutral	Positive	Negative
Duration (ms)	S-M: 569.53 (23.63) NS-M: 583.5 S-F: 604.50 (50.29) NS-F: 654.75	S-M: 607.90 (37.55) NS-M: 628.79 S-F: 662.51 (57.98) NS-F: 721.7	S-M: 601.78 (31.42) NS-M: 611.65 S-F: 643.63 (48.76) NS-F: 703.02
Mean F0 (Hz)	S-M: 117.04 (13.85) NS-M: 110.24 S-F: 173.73 (16.35) NS-F: 206.58	S-O: 116.32 (15.01) NS-M: 101.18 S-F: 169.87 (12.01) NS-F: 211.97	S-M: 118.30 (13.92) NS-M: 117.89 S-F: 172.58 (11.67) NS-F: 206.59
Mean Intensity (dB)	S-M: 72.43 (3.35) NS-M: 72.66 S-F: 73.65 (2.79) NS-F: 75.86	S-M: 72.19 (3.66) NS-M: 71.97 S-F: 72.93 (2.85) NS-F: 75.34	S-M: 81.23 (26.84) NS-M: 72.19 S-F: 73.48 (2.94) NS-F: 75.92

Notes: S = self-speech; NS-M = non-self male speech; NS-F = non-self female speech; standard deviations in parentheses.

to minimize eye movements. After a 1000 ms inter-stimulus interval (ISI), a question mark signaled the beginning of the response time, i.e., for 6 s (see Fig. 1). Stimuli were presented binaurally through headphones at a sound level comfortable for each subject, and were not repeated during the experiment. Stimulus presentation and timing of events and recording of subjects' responses were controlled by Superlab Pro software package (2008; http://www.superlab.com/). Before each experimental block, participants were given a brief training with feedback.

2.4. EEG data acquisition and analysis

2.4.1. EEG recording procedure

The EEG was recorded with 64 electrodes mounted on a custom-made cap (Electro-cap International, USA), according to the expanded 10–20 system (American Electroencephalographic Society, 1991), using Biosemi Active 2 system (Biosemi B.V., Amsterdam, Netherlands). The electrode offset was kept below 40 mV. The EEG was acquired in a continuous mode at a digitization rate of 512 Hz, with a bandpass filter of 0.01–100 Hz, and stored on hard disk for later analysis. Horizontal and vertical ocular movements were recorded for eye movement and blink detection and rejection, via electrodes placed on the external canthus of both eyes (horizontal electrooculogram) and one below the left eye (vertical electrooculogram).

2.4.2. EEG data analysis

The EEG data were processed using BrainVision Analyzer 2 software (Brain Products, Munich, Germany). The EEG channels were referenced offline to the average of the left and right mastoids. EEG data were high-pass filtered with a 0.1 Hz filter. Individual EEG epochs associated with correct responses were created for each stimulus type (SS-neutral; SS-positive; SS-negative; NSS-neutral; NSS-positive; NSS-negative), with −200 ms pre-stimulus baseline and 1000 ms post-stimulus epoch. The EEG was baseline corrected using the −200 to 0 ms prestimulus interval. The EEG channels were corrected for vertical and horizontal eye movements using the method of Gratton, Coles, and Donchin (1983). Segments were also semiautomatically screened for eye movements, muscle artifacts, electrode drifting and amplifier blocking. EEG epochs exceeding ±100 μV were excluded from further EEG analysis. After artifact rejection, at least 70% of the trials in individual ERPs per condition per subject entered the analyses.

Grand average waveforms to the six conditions are shown in Fig. 2. Voice stimuli evoked clearly identifiable frontocentral N1 and P2 components, peaking at 180 ms and 240 ms, respectively, and followed by a centroparietal LPP component with an onset of approximately 500 ms. N1, P2 and LPP components were computed by determining the mean activity on averaged waveforms for each subject, voice identity type, semantic valence type, and electrode sites. Mean amplitudes for N1 and P2 were computed in time windows of 130–210 ms and 215–380 ms respectively, centered on the component peaks (e.g., Pinheiro et al., 2014). For the LPP component, the time window was 500–700 ms (e.g., Proverbio, Adorni, Zani, & Trestianu, 2009). Amplitude measurements were based on the average amplitude within the specified time windows.

2.5. Statistical analyses

The SPSS statistical software package (Version 22.0, SPSS Inc., Chicago, IL, USA) was used for the statistical analyses. Only significant results are presented (alpha level was set at .05).

2.5.1. ERP data

Based on a careful inspection of grand average waveforms and topographical maps (see Fig. 2), the following electrodes were selected for the statistical analyses: left medial (FC3, C3, CP3),

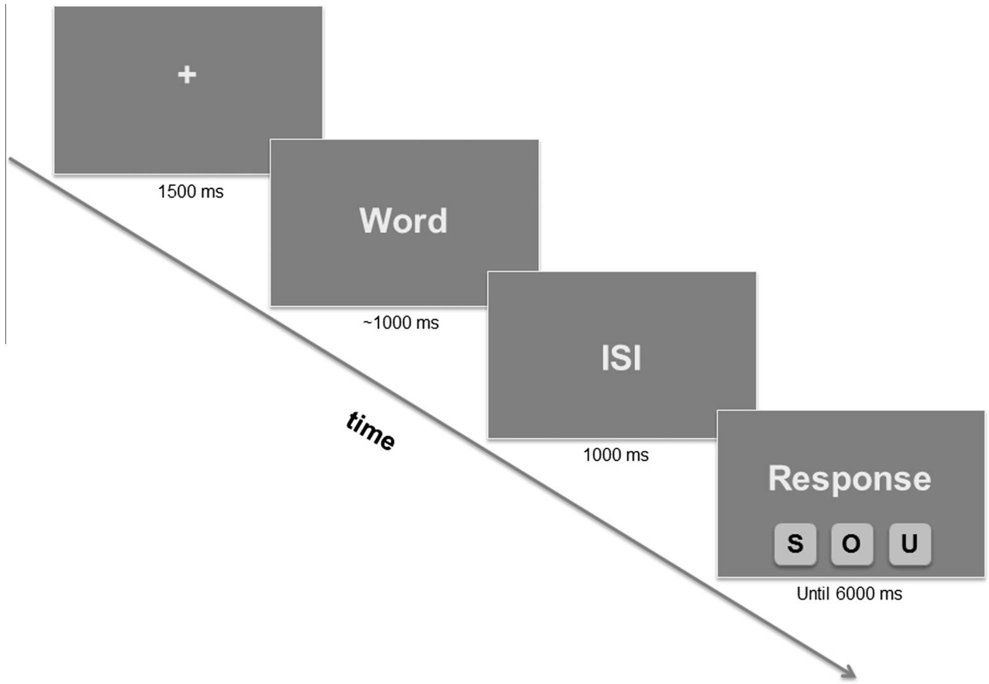

Fig. 1. Illustration of an experimental trial.

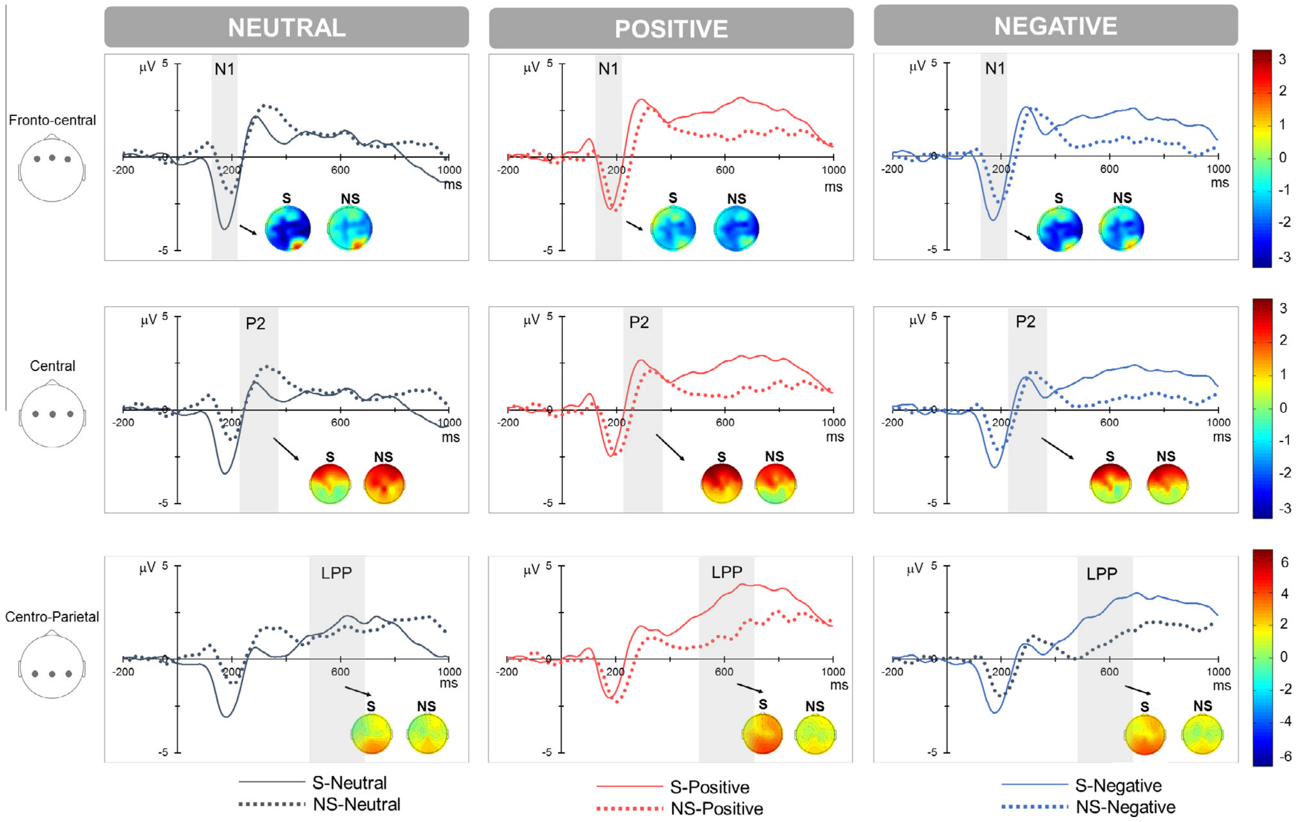

Fig. 2. Contrasts between self and non-self speech with neutral, positive and negative content, and grand average scalp maps showing the spatial distribution of the N1, P2 and LPP effects.

midline (FCz, Cz, CPz), and right medial (FC4, C4, CP4). Repeated measures analyses of variance (ANOVAs) were separately computed for N1, P2, and LPP mean amplitude. The analysis included the within-subjects factors of identity (SS, NSS), valence (neutral, positive, negative), and region of interest (left medial, midline, right medial). In a separate analysis, and in order to rule out the effects of differences in voice acoustic properties, voice mean F0[4] was added as a covariate. Significant interactions between identity and valence were followed with repeated-measures ANOVA for self and non-self voice stimuli separately, with the within-subjects factors of valence and region of interest.

2.5.2. Accuracy data

Behavioral measures included recognition accuracy (correct identification of self-speech as 'self' and correct identification of non-self speech as 'non-self') and number of unsure responses. The effects of identity and valence on voice recognition accuracy and number of 'unsure' responses were tested separately for each behavioral measure using a repeated-measures ANOVA with identity (self, non-self) and valence (neutral, positive, negative) as within-subject factors. In a separate analysis, voice mean F0 was added as a covariate.

All the analyses were corrected for non-sphericity using the Greenhouse–Geisser method (the original *df* is reported). Main effects and interactions were followed up with pairwise comparisons using Bonferroni correction for multiple comparisons. Effect sizes for significant effects are reported using the partial η-square method (*partial* η^2).

3. Results

3.1. ERP results

Results are presented according to the following structure: first, we present significant main effects (Hypothesis 1), and then we describe the interaction between the two factors under study (voice identity and semantic valence; Hypothesis 2) for each component separately.

3.1.1. N1

3.1.1.1. Main effect. Voice identity manipulation did not yield a significant effect ($F(1,15) = 3.03$, $p > .05$).

3.1.1.2. Interaction. The interaction between voice identity and valence reached statistical significance ($F(2,30) = 8.28$, $p = .001$, *partial* $\eta^2 = .36$). We followed this interaction by computing separate repeated-measures ANOVA that tested the effects of voice identity for each semantic valence type, i.e. neutral, positive and negative speech (the within-subjects factors were voice identity and region of interest). N1 was significantly more negative for SS than for NSS with neutral valence (identity effect – $F(1,15) = 13.23$, $p = .002$, *partial* $\eta^2 = .47$).

The covariate was not significantly related to N1 amplitude ($F(1,14) = .30$, $p > .05$).

3.1.2. P2

3.1.2.1. Main effects. The effects of identity ($F(1,15) = 1.82$, $p > .05$), and valence ($F(2,30) = .31$, $p > .05$) were not significant.

3.1.2.2. Interaction. The interaction between identity and valence reached statistical significance ($F(2,30) = 3.79$, $p = .034$, *partial* $\eta^2 = .20$). We followed this interaction by computing separate repeated-measures ANOVA that tested the effects of voice identity for each semantic valence type, i.e. neutral, positive and negative speech (the within-subjects factors were voice identity and region of interest). P2 amplitude was significantly more positive for SS relative to NSS with positive valence only (identity effect – $F(1,15) = 4.79$, $p = .045$, *partial* $\eta^2 = .24$).

The covariate was not significantly related to P2 amplitude ($F(1,14) = .22$, $p > .05$).

3.1.3. LPP

3.1.3.1. Main effects. LPP amplitude was significantly more positive for SS relative to NSS (identity effect – $F(1,15) = 7.34$, $p = .016$, *partial* $\eta^2 = .33$).

3.1.3.2. Interaction. The interaction between identity and valence was significant ($F(2,30) = 3.36$, $p = .048$, *partial* $\eta^2 = .18$). We followed this interaction by computing separate repeated-measures ANOVA that tested the effects of voice identity for each semantic valence type, i.e. neutral, positive and negative speech (the within-subjects factors were voice identity and region of interest). Increased LPP for SS relative to NSS occurred only for speech with positive (identity effect – $F(1,15) = 16.83$, $p = .001$, *partial* $\eta^2 = .53$) and with negative (identity effect – $F(1,15) = 6.74$, $p = .020$, *partial* $\eta^2 = .31$) valence.

The covariate was not significantly related to LPP amplitude ($F(1,14) = .03$, $p > .05$).

3.2. Behavioral results

3.2.1. Number of correct responses

3.2.1.1. Main effects. Participants were similarly accurate in the recognition of self and non-self speech (identity effect – $F(1,15) = .43$, $p > .05$), and of neutral and emotional speech (valence effect – $F(2,30) = .17$, $p > .05$; see Table 3).

3.2.1.2. Interaction. The interaction between identity and valence was not significant ($F(2,30) = .78$, $p > .05$).

The covariate was not significantly related to the number of correct responses ($F(1,14) = 4.00$, $p > .05$).

3.2.2. Number of unsure responses

3.2.2.1. Main effects. No differences were found in the number of unsure responses based on identity ($F(1,15) = .64$, $p > .05$) or valence types ($F(2,30) = .18$, $p > .05$) (see Table 3).

3.2.2.2. Interaction. The interaction between identity and valence was not significant ($F(2,30) = .53$, $p > .05$).

The covariate was not significantly related to the number of unsure responses ($F(1,14) = .11$, $p > .05$).

4. Discussion

The comprehension of spoken language is a complex process since many linguistic and non-linguistic features are embedded in the speech signal and need to be integrated within a fraction of a second. This study probed how speaker's identity influenced the processing of the semantic valence of adjectives. Our ERP results lend support to our second hypothesis, by demonstrating differences in the processing of speech stimuli that were dependent both on identity and valence, and that occurred in spite of similar accuracy in the recognition of both types of stimuli. The interaction between both factors indicated that implicit verbal processing occurred even if not explicitly required.

---

[4] Considering that the processes under study involve self-other voice discrimination, the mean difference in F0 between each participant's vocal stimuli and the control voice was calculated and included in the analysis. F0 has been pointed out as a fundamental parameter that listeners rely on to discriminate and recognize the identity of self, familiar and unfamiliar voices (Baumann & Belin, 2010; Latinus & Belin, 2012; Latinus, McAleer, Bestelmeyer, & Belin, 2013; Xu, Homae, Hashimoto, & Hagiwara, 2013).

Table 3
Percentage of correct and unsure responses in the recognition of self and non-self speech with neutral, positive and negative semantic content.

Identity	Emotion	Correct responses	Unsure responses
Self	Neutral	91.61 (8.52)	1.88 (2.26)
	Positive	93.21 (7.48)	1.52 (1.98)
	Negative	92.32 (7.84)	1.70 (2.67)
Non-self	Neutral	93.04 (8.70)	2.05 (3.00)
	Positive	94.02 (8.12)	0.98 (1.25)
	Negative	94.11 (6.55)	1.07 (1.52)

Note: Standard deviations in parentheses.

4.1. ERP interactive effects of voice identity and semantic valence on distinct stages of speech processing

N1 differed as a function of voice identity, but only in the case of neutral speech: N1 amplitude was increased for self-speech. This finding agrees with previous studies showing that pre-recorded self-speech (e.g., the neutral word "table") elicited increased attentional resources compared to non-self speech (e.g., Conde et al., 2015). The significant interaction between voice identity and valence at this early processing stage suggests that neutral words automatically attract more attention if spoken by one's own voice than by an unknown voice. However, since this effect was observed before the words' recognition point (i.e., the point in the speech signal in which full access to the meaning of the word takes place), caution is required in the interpretation of the current results.

Nonetheless, it is worth noting that more recent studies suggest that semantic processing is initiated simultaneously with perceptual processing (e.g., Penolazzi, Hauk, & Pulvermüller, 2007), that the differentiation of voice signals as a function of speaker's identity may occur before full access to the word meaning (e.g., Conde et al., in press), and that the accurate identification of the speaker is facilitated by linguistic knowledge (Perrachione & Wong, 2007). These studies agree that both linguistic and nonlinguistic (e.g., speaker's identity) cues are simultaneously extracted from the same signal (e.g., Belin et al., 2011) and interact during speech processing (e.g., von Kriegstein et al., 2003). The contribution of valence to speaker's identity discrimination in such an early processing stage is in line with observations of partially dependent processing of speech and speaker information (e.g., Perrachione & Wong, 2007; Schweinberger, Walther, Zäske, & Kovács, 2011). Since no voice identity differences were observed for positive and negative speech (in spite of the fact that the same word stimuli were presented in both identity conditions), it is plausible that the semantic operations that start at the same time as vocal features analysis may interfere with the discrimination of voice identity. While we cannot exclude the possibility of a contribution from low-level acoustic cues, we note that the analysis of covariance showed that the N1 effects were not significantly predicted by voice F0 variation, which plays a critical role in voice identity recognition and discrimination (e.g., Baumann & Belin, 2010; Latinus & Belin, 2012; Latinus et al., 2013; Xu et al., 2013). However, the role of phonetic variation in the explanation of these findings should be considered. Indeed, the processing of speaker's identity information seems to occur independently of the particular acoustic structure of the stimuli around 300–600 ms, while within the first 250 ms voice identity processing is dependent of speech (Schweinberger et al., 2011). This finding suggests that 250–300 ms of continuous speech may be necessary for above-chance identification of familiar voices (Schweinberger, Herholz, & Sommer, 1997). Over the course of the voice signal, more non-linguistic (e.g., speaking rate, loudness, mean F0, F0 variability) and linguistic information is available to the listeners, which may significantly improve both voice identity and emotion recognition (e.g., Schweinberger et al., 1997). Therefore, later effects indexed by the P2 and LPP may provide more relevant information about the processes under analysis.

Interactive effects of self-relevance and valence were also observed approximately 200 ms after speech onset, reflected in increased P2 amplitude for self-compared to non-self speech with positive content. This finding lends new support to the sensitivity of the P2 component to both voice identity (Knolle et al., 2013b) and emotion (Bernat et al., 2001; Kanske & Kotz, 2007; Kissler et al., 2006; Schapkin et al., 2000), and corroborates the implicit semantic activation evoked by voice processing (e.g., von Kriegstein et al., 2003). Studies using an auditory-to-motor prediction task (testing the differential processing of self-initiated *vs.* externally presented sounds – e.g., Knolle et al., 2013a, 2013b) suggested that the P2 component indexes the conscious detection of a self-initiated sound (Knolle et al., 2013b). However, this type of task differs from passively listening to self *vs.* non-self voice stimuli, such as in the current study. Studies probing the processing of neutral and emotional words highlighted the modulatory role of stimulus valence on P2 amplitude, with some indicating a processing advantage for positive compared to neutral words, reflected in increased P2 amplitude (e.g., Kanske & Kotz, 2007; Kissler et al., 2006). The absence of significant main effects on the P2 component suggest that the processing of voice identity is not independent of semantic valence, and is compatible with previous studies that demonstrated that the analysis of non-verbal vocal features does not occur independently of verbal (e.g., phonological, semantic) analysis during speech processing (e.g., von Kriegstein et al., 2003). Again, we should note that even though these effects occurred before full access to word meaning took place, recent evidence has accumulated suggesting that semantic effects may occur earlier than what traditional language models would hypothesize (e.g., Kryuchkova, Tucker, Wurm, & Baayen, 2012; Pulvermüller, Shtyrov, & Hauk, 2009; Scott, O'Donnell, Leuthold, & Sereno, 2009). These studies indicate that semantic operations start before the word identification is complete (Penolazzi et al., 2007; Van Berkum, Brown, Hagoort, & Zwitserlood, 2003; van den Brink, Brown, & Hagoort, 2006; Van Petten, Coulson, Rubin, Plante, & Parks, 1999). For example, more recently, Kryuchkova et al. (2012) reported effects of spoken words' valence (e.g., "poison", "hospital") already 150 ms after stimulus onset, before the words' uniqueness points. Memory effects could have also conceivably contributed to the current pattern of findings: participants may have remembered the words they read in the recording session and, as a result, the words were activated with minimal phonological information. Considering the positivity bias reported in previous studies (Kanske & Kotz, 2007; Kissler et al., 2006), it is possible that positive words (e.g., "beautiful", "pretty") were better remembered and, as a result, elicit increased attentional capture when spoken in one's own voice than in another person's voice. An alternative explanation is that differences in the acoustic properties of the voice stimuli accounted for these effects. Nonetheless, the analysis of covariance suggested that this is not the case: the F0 difference between self- and non-self speech stimuli was not significantly related to P2 amplitude.

ERP differences driven by voice identity and speech content were also observed in a later processing stage (500–700 ms) indexed by the LPP. As the mean words' duration was 632.77 ms, these effects coincided with full access to the word's meaning. A main effect of voice identity was reflected in more positive LPP amplitude for self- than for non-self speech. Semantic valence – in terms of positive *vs.* negative attributes – did not yield significant effects, suggesting that words did not independently elicit different amounts of sustained attention allocation as a function of their emotional salience. It is plausible that the lack of differences between neutral and emotional words, irrespective of voice identity, at this post-lexical processing stage is related to the nature of the behavioral task: the task was focused on perceptual discrimination ('is this my voice or somebody else's voice?') and did not require explicit semantic evaluation.

However, we observed interactive effects of self-relevance and emotion, reflected in more positive LPP for self-speech with emotional valence compared with emotional non-self speech stimuli. This ERP component has been consistently described in emotion research (indeed, it represents the most consistent finding in ERP emotion research – e.g., Olofsson, Nordin, Sequeira, & Polich, 2008), reflecting sustained attention that is enhanced for emotionally salient compared to neutral stimuli (e.g., Hajcak, MacNamara, Foti, Ferri, & Keil, 2013; Schupp et al., 2000). The effects of self-relevance were found in the LPP as well. For example, larger amplitudes were observed for self-generated or self-related stimuli, such as one's own name (e.g., Tacikowski & Nowicka, 2010), one's own handwritting (Chen et al., 2008), or pictures of one's own face (e.g., Keyes et al., 2010; Tacikowski & Nowicka, 2010). Our study extended these observations by showing more positive LPP for *emotional self*-generated words compared with the same words spoken by an unfamiliar voice. This finding suggests that the motivational value of both *self*- and *emotional* speech may impact upon sustained attention and result in increased elaborative processing. Together with the P2 findings, this observation is consistent with a synergistic interaction between self-relevance and emotion. Since enhanced LPP amplitude for emotional or self-relevant stimuli has been taken as an index of increased attention and cognitive processing related to stimuli with higher emotional relevance, the larger LPP for emotional self-speech observed in our study suggests additive effects of salience associated either with emotional words and self-related stimuli. In other words, a self-relevant context ("my" voice) seems to enhance the processing of emotional speech: words such as "rude" or "pretty" trigger deeper processing if spoken in one's own voice at a later post-lexical stage. It is plausible that this reflects an attempt to determine how the words fit with the participant's self-concept (e.g., Watson et al., 2007). This effect occurred in spite of the fact that the same words were presented in both identity conditions, and after controlling for the effects of voice F0 differences.

The LPP results substantiate the claim that self-related processing is somewhat 'special' in the brain, as previously demonstrated by studies probing the differential processing of self- and non-self stimuli and reporting enhanced amplitude (e.g., P300) to self-related cues, such as one's face (e.g., Keyes et al., 2010; Tacikowski & Nowicka, 2010), one's own name (e.g., Tacikowski & Nowicka, 2010), or one's own voice (Conde et al., 2015). These differences have been taken as evidence for the existence of unique mechanisms underpinning self-recognition (e.g., Sui et al., 2006). The increased salience of emotional self-speech fits well with the observation that the representation of self-related cues involves neurofunctional processes that are distinct from those activated in more general cognitive processing (e.g., Northoff et al., 2006).

Together, the ERP results lend new support to the interaction and integration of speech and voice perception processes. Consistent with our ERP findings, Perrachione and Wong (2007) proposed that the voice perception model of Belin et al. (2004) – that defends the existence of functionally dissociable brain pathways dedicated to the processing of voice identity, speech and affect information – should consider a bi-directional integration of processes underlying speech and voice perception.

4.2. Similarly high recognition of self and non-self speech

In spite of the ERP differences between self *vs.* non-self speech processing, participants recognized both types of stimuli with equally high accuracy. For all conditions, the accuracy rate was over 90%, in line with more recent studies (Hughes & Nicholson, 2010; Rosa et al., 2008). This suggests that, even though hearing a recording of our own voice is distinct from hearing the sound of our voice when speaking, it can still be accurately recognized (e.g., Graux et al., 2013, 2015; Hughes & Nicholson, 2010; Rosa et al., 2008). It is possible that the higher exposure to one's own played-back voice due to new technologies (e.g., cell phones, answering machines, video recordings, internet phone chatting) has resulted in an easier recognition of self-voice (e.g., Rosa et al., 2008). As participants were simply required to discriminate self and non-self speech stimuli, we should note that the high performance may represent a ceiling effect due to low task difficulty.

4.3. Caveats and future directions

One methodological issue of this study is related to the fact that the acoustic characteristics of one's own recorded voice are different from those of one's own voice as it is internally experienced, due to differences in the means of transmission of the sound (e.g., Maurer & Landis, 1990). Unfortunately, there is not yet a way to account for this distortion. Furthermore, because we did not match stimuli on familiarity (i.e., the self-voice was more familiar than the non-self voice), it is possible that the ERP differences reflect familiarity effects rather than self-specific processing, even though a more recent study (Graux et al., 2015) provided support for the existence of distinct brain processes underlying the discrimination between a self-voice and a familiar voice. Future studies should address these issues. Future studies should also address whether similar results are obtained if a semantic task (in which participants are instructed to judge the valence of speech stimuli) is used instead of a perceptual task emphasizing the distinction between self and non-self speech stimuli.

Our knowledge about the processes involved in self-voice recognition is still poorly understood. Therefore, more experiments need to be carried out in order to increase our understanding of the neural computations involved in sound-to-meaning transformations underlying the recognition of self *vs.* non-self speech with high or low emotional salience. Ultimately, this knowledge may contribute to our understanding of the altered neuro-functional processes in disorders characterized by self-voice recognition and emotional impairments, such as schizophrenia (e.g., Costafreda, Brébion, Allen, McGuire, & Fu, 2008; Ford et al., 2001; Johns et al., 2001) and, specifically, in elucidating clinical symptoms such as auditory verbal hallucinations (AVH). AVH carry a rich amount of linguistic and paralinguistic cues that convey not only speech, but also identity and affect information (Nayani & David, 1996). More recent evidence has suggested a link between voice processing abnormalities and AVH (e.g., Heinks-Maldonado et al., 2007). Probing which dimensions of voice processing (e.g., identity, speech and affect information) are impaired in AVH may shed light on the pathological mechanisms underlying specific phenomenological features of hearing voices in the absence of corresponding external stimulation.

4.4. Conclusions

This study demonstrated interactive effects of speaker's identity and emotional valence during speech processing, indexed by the N1, P2 and LPP components. The LPP effects, observed after the words' recognition point, indicate that the processing of spoken emotional words is modulated by self-relevance, corroborating the non-independent effects of emotion and self-relevance reported by previous ERP studies (e.g., Fields & Kuperberg, 2012; Watson et al., 2007). In other words, emotional language comprehension seems to be intrinsically contextualized, i.e. it depends on the speaker ("me" *vs.* "not me"). These findings add to existing studies on voice and speech perception suggesting that these two abilities are more closely integrated than previously thought (e.g., Perrachione & Wong, 2007).

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 (European Regional Development Fund, Portugal) (*Fundo Europeu de Desenvolvimento Regional*) through the European programs QREN (*Quadro de Referência Estratégico Nacional*), and COMPETE (Operational Programme 'Thematic Factors of Competitiveness', Portugal) (*Programa Operacional Factores de Competitividade*), awarded to A.P.P.

We are grateful to all the participants of this study for their contribution to science.

Appendix A

List of words used in the experiment:

Negative	Neutral	Positive
Abnormal	Actual	Adorable
Afraid	Airy	Alive
Angry	Aloof	Beautiful
Awkward	Ample	Blessed
Bad	Annual	Brave
Beaten	Automatic	Bright
Blind	Average	Brilliant
Bloody	Basic	Calm
Confused	Blank	Careful
Crazy	Blond	Caring
Creepy	Blue	Charming
Cruel	Bold	Clean
Damaged	Brief	Confident
Dead	Broad	Cute
Dirty	Brown	Divine
Dreadful	Casual	Elegant
Dumb	Central	Fabulous
Enraged	Civil	Faithful
Failed	Classic	Famous
Faulty	Close	Fantastic
Fearful	Collected	Free
Foolish	Common	Friendly
Furious	Compact	Funny
Guilty	Constant	Gentle
Helpless	Cubic	Gifted
Horrid	Curly	Glad
Hostile	Daily	Good
Ill	Deep	Gorgeous
Infected	Dry	Gracious
Inferior	Familiar	Grateful
Insane	Flat	Handsome
Jealous	Full	Happy
Lazy	Herbal	Healthy
Lonely	High	Honest
Lost	Informal	Hopeful
Mad	Involved	Incredible
Malign	Large	Inspired
Mean	Lay	Joyful
Messy	Local	Kind
Morbid	Long	Loved
Nasty	Main	Lovely
Nervous	Mild	Loyal
Painful	Mutual	Lucky
Pathetic	Narrow	Magical
Poor	Near	Merry
Punished	Neutral	Nice
Rejected	Open	Perfect
Rude	Overt	Playful
Sad	Plain	Precious
Scabby	Plural	Pretty
Scared	Private	Protected
Selfish	Purple	Proud
Shabby	Quiet	Pure
Shamed	Red	Relaxed
Sick	Regular	Romantic
Sinful	Related	Safe
Sneaky	Round	Satisfied
Stinking	Sharp	Secure
Stupid	Slim	Sexy
Terrible	Small	Slender
Terrified	Square	Special
Tragic	Straight	Splendid
Ugly	Subtle	Strong
Unhappy	Thick	Super
Upset	Tiny	Terrific
Useless	Usual	Thoughtful
Violent	Wet	Truthful
Weak	White	Useful
Wicked	Wild	Wealthy
Wrong	Yellow	Wise

References

American Electroencephalographic Society (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology, 8*(2), 200–202.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders. Diagnostic and statistical manual of mental disorders 4th edition TR*. Washington, DC: American Psychiatric Association Press. http://dx.doi.org/10.1016/B978-1-4377-2242-0.00016-X.

Andreasen, N. C., Endicott, J., Spitzer, R. L., & Winokur, G. (1977). The family history method using diagnostic criteria. Reliability and validity. *Archives of General Psychiatry, 34*(10), 1229–1235. http://dx.doi.org/10.1001/archpsyc.1977.01770220111013.

Baess, P., Widmann, A., Roye, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *European Journal of Neuroscience, 29*(January), 1514–1521. http://dx.doi.org/10.1111/j.1460-9568.2009.06683.x.

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.

Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience, 12*(1), 54. http://dx.doi.org/10.1186/1471-2202-12-54.

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology, 102*(4), 711–725. http://dx.doi.org/10.1111/j.2044-8295.2011.02041.x.

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.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research, 13*(1), 17–26. http://dx.doi.org/10.1016/S0926-6410(01)00084-2.

Bernat, E., Bunce, S., & Shevrin, H. (2001). Event-related brain potentials differentiate positive and negative mood adjectives during both supraliminal and subliminal visual processing. *International Journal of Psychophysiology, 42*(1), 11–34. http://dx.doi.org/10.1016/S0167-8760(01)00133-7.

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer [computer program]*. Version 5.3.53. <http://www.praat.org/>.

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. http://dx.doi.org/10.1016/0005-7916(94)90063-9.

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.

Coltheart, M. (2007). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A, 33*(4), 497–505. http://dx.doi.org/10.1080/14640748108400805.

Conde, T., Gonçalves, O. F., & Pinheiro, A. P. (in press). The effects of stimulus complexity on the preattentive processing of self-generated and non-self voices: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience*, http://dx.doi.org/10.3758/s13415-015-0376-1.

Conde, T., Gonçalves, O. F., & Pinheiro, A. P. (2015). Paying attention to my voice or yours: An ERP study with words. *Biological Psychology, 111*, 40–52. http://dx.doi.org/10.1016/j.biopsycho.2015.07.014.

Costafreda, S. G., Brébion, G., Allen, P., McGuire, P. K., & Fu, C. H. Y. (2008). Affective modulation of external misattribution bias in source monitoring in schizophrenia. *Psychological Medicine, 38*(6), 821–824. http://dx.doi.org/10.1017/S0033291708003243.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology, 115*, 732–744. http://dx.doi.org/10.1016/j.clinph.2003.11.021.

Ferré, P. (2003). Effects of level of processing on memory for affectively valenced words. *Cognition & Emotion, 17*(6), 859–880. http://dx.doi.org/10.1080/02699930302309.

Fields, E. C., & Kuperberg, G. R. (2012). It's All About You: An ERP study of emotion and self-relevance in discourse. *NeuroImage, 62*(1), 562–574. http://dx.doi.org/10.1016/j.neuroimage.2012.05.003.

First, M. B., Gibbon, M., & Spitzer, R. (1997). *Structured clinical interview for DSM-IV axis II personality disorders (SCID-II, version 2.0)*. New York, NY: Biometrics Research Department, New York State Psychiatric Institute.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV axis I diagnosis-patient edition (SCID-I/P, version 2.0)*. New York, NY: Biometric Research Department, New York State Psychiatric Institute.

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry, 158*(12), 2069–2071. http://dx.doi.org/10.1176/appi.ajp.158.12.2069.

Formisano, E., De Martino, F., Bonte, M., & Goebel, R. (2008). "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science (New York, N. Y.), 322*, 970–973. http://dx.doi.org/10.1126/science.1164318.

Fossati, P., Hevenor, S. J., Graham, S. J., Grady, C., Keightley, M. L., Craik, F., & Mayberg, H. (2003). In search of the emotional self: An fMRI study using positive and negative emotional words. *The American Journal of Psychiatry, 160*(11), 1938–1945. http://dx.doi.org/10.1176/appi.ajp.160.11.1938.

Fossati, P., Hevenor, S. J., Lepage, M., Graham, S. J., Grady, C., Keightley, M. L., ... Mayberg, H. (2004). Distributed self in episodic memory: Neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage, 22*(4), 1596–1604. http://dx.doi.org/10.1016/j.neuroimage.2004.03.034.

Gratton, G., Coles, M. G. H., & 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-Brilhault, F., & Bruneau, N. (2015). 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.

Graux, J., Gomot, M., Roux, S., Bonnet-Brilhault, F., Camus, V., & Bruneau, N. (2013). My voice or yours? An electrophysiological study. *Brain Topography, 26*, 72–82. http://dx.doi.org/10.1007/s10548-012-0233-2.

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*(2), 216–224. http://dx.doi.org/10.1016/S0022-1031(03)00092-1.

Hajcak, G., MacNamara, A., Foti, D., Ferri, J., & Keil, A. (2013). The dynamic allocation of attention to emotion: Simultaneous and independent evidence from the late positive potential and steady state visual evoked potentials. *Biological Psychology, 92*(3), 447–455. http://dx.doi.org/10.1016/j.biopsycho.2011.11.012.

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review, 106*(4), 766–794. http://dx.doi.org/10.1037/0033-295X.106.4.766.

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Archives of General Psychiatry, 64*(3), 286–296. http://dx.doi.org/10.1001/archpsyc.64.3.286.

Herbert, C., Herbert, B. M., Ethofer, T., & Pauli, P. (2011). His or mine? The time course of self-other discrimination in emotion processing. *Social Neuroscience, 6*(3), 277–288. http://dx.doi.org/10.1080/17470919.2010.523543.

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology, 45*(3), 487–498. http://dx.doi.org/10.1111/j.1469-8986.2007.00638.x.

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.

Johns, L. C., Rossell, S., Frith, C., Ahmad, F., Hemsley, D., Kuipers, E., & McGuire, P. K. (2001). Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychological Medicine, 31*(04), 705–715. http://dx.doi.org/10.1017/S0033291701003774.

Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research, 1148*, 138–148. http://dx.doi.org/10.1016/j.brainres.2007.02.044.

Kanske, P., Plitschka, J., & Kotz, S. A. (2011). Attentional orienting towards emotion: P2 and N400 ERP effects. *Neuropsychologia, 49*(11), 3121–3129. http://dx.doi.org/10.1016/j.neuropsychologia.2011.07.022.

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*(3), 218–223. http://dx.doi.org/10.1093/scan/nsn014.

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain and Cognition, 72*(2), 244–254. http://dx.doi.org/10.1016/j.bandc.2009.09.006.

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: Insights from ERP studies. *Progress in Brain Research, 156*(06), 147–183. http://dx.doi.org/10.1016/S0079-6123(06)56008-X.

Knolle, F., Schröger, E., & Kotz, S. A. (2013a). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex, 49*(9), 2449–2461. http://dx.doi.org/10.1016/j.cortex.2012.12.012.

Knolle, F., Schröger, E., & Kotz, S. A. (2013b). Prediction errors in self- and externally-generated deviants. *Biological Psychology, 92*(2), 410–416. http://dx.doi.org/10.1016/j.biopsycho.2012.11.017.

Kryuchkova, T., Tucker, B. V., Wurm, L. H., & Baayen, R. H. (2012). Danger and usefulness are detected early in auditory lexical processing: Evidence from electroencephalography. *Brain and Language, 122*, 81–91. http://dx.doi.org/10.1016/j.bandl.2012.05.005.

Latinus, M., & Belin, P. (2012). Perceptual auditory aftereffects on voice identity using brief vowel stimuli. *PLoS ONE, 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: CB, 23*(12), 1075–1080. http://dx.doi.org/10.1016/j.cub.2013.04.055.

Maurer, D., & Landis, T. (1990). Role of bone conduction in the self-perception of speech. *Folia Phoniatrica et Logopaedica, 42*(5), 226–229. http://dx.doi.org/10.1159/000266070.

Miyakoshi, M., Nomura, M., & Ohira, H. (2007). An ERP study on self-relevant object recognition. *Brain and Cognition, 63*, 182–189. http://dx.doi.org/10.1016/j.bandc.2006.12.001.

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.

Nayani, T. H., & David, A. S. (1996). The auditory hallucination: A phenomenological survey. *Psychological Medicine, 26*(1), 177–189. http://dx.doi.org/10.1017/S003329170003381X.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain – A meta-analysis of imaging studies on the self. *NeuroImage, 31*(1), 440–457. http://dx.doi.org/10.1016/j.neuroimage.2005.12.002.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. 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.

Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *NeuroReport, 19*(2), 209–213. http://dx.doi.org/10.1097/WNR.0b013e3282f454db.

Penolazzi, B., Hauk, O., & Pulvermüller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology, 74*(3), 374–388. http://dx.doi.org/10.1016/j.biopsycho.2006.09.008.

Perrachione, T. K., & Wong, P. C. M. (2007). Learning to recognize speakers of a non-native language: Implications for the functional organization of human auditory cortex. *Neuropsychologia, 45*(8), 1899–1910. http://dx.doi.org/10.1016/j.neuropsychologia.2006.11.015.

Phan, K. L., Taylor, S. F., Welsh, R. C., Ho, S. H., Britton, J. C., & Liberzon, I. (2004). Neural correlates of individual ratings of emotional salience: A trial-related fMRI study. *NeuroImage, 21*(2), 768–780. http://dx.doi.org/10.1016/j.neuroimage.2003.09.072.

Pinheiro, A. P., del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., ... Niznikiewicz, M. A. (2013). Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: An electrophysiological investigation. *Psychological Medicine, 43*(03), 603–618. http://dx.doi.org/10.1017/S003329171200133X.

Pinheiro, A. P., McCarley, R. W., Thompson, E., Gonçalves, O. F., & Niznikiewicz, M. (2012). From semantics to feelings: How do individuals with schizophrenia rate the emotional valence of words? *Schizophrenia Research and Treatment*. http://dx.doi.org/10.1155/2012/431823 431823.

Pinheiro, A. P., Rezaii, M., Rauber, A., Liu, T., Nestor, P. G., McCarley, R. W., ... Niznikiewicz, M. A. (2014). Abnormalities in the processing of emotional prosody from single words in schizophrenia. *Schizophrenia Research, 152*(1), 235–241. http://dx.doi.org/10.1016/j.schres.2013.10.042.

Proverbio, A. M., Adorni, R., Zani, A., & Trestianu, L. (2009). Sex differences in the brain response to affective scenes with or without humans. *Neuropsychologia, 47*(12), 2374–2388.

Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language, 110*(2), 81–94. http://dx.doi.org/10.1016/j.bandl.2008.12.001.

Rosa, C., Lassonde, M., Pinard, C., Keenan, J. P., & Belin, P. (2008). Investigations of hemispheric specialization of self-voice recognition. *Brain and Cognition, 68*(2), 204–214. http://dx.doi.org/10.1016/j.bandc.2008.04.007.

Rosburg, T., Boutros, N. N., & Ford, J. M. (2008). Reduced auditory evoked potential component N100 in schizophrenia – A critical review. *Psychiatry Research, 161*(3), 259–274. http://dx.doi.org/10.1016/j.psychres.2008.03.017.

Schapkin, S. A., Gusev, A. N., & Kühl, J. (2000). Categorization of unilaterally presented emotional words: An ERP analysis. *Acta Neurobiologiae Experimentalis, 60*(1), 17–28.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*, 24–30. http://dx.doi.org/10.1016/j.tics.2005.11.009.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology, 37*(2), 257–261. http://dx.doi.org/10.1111/1469-8986.3720257.

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. <http://jslhr.asha.org/cgi/reprint/40/2/453/npapers2://publication/uuid/56AA9A0A-B16C-4E29-A077-7DED959E7D84>.

Schweinberger, S. R., Walther, C., Zäske, R., & Kovács, G. (2011). Neural correlates of adaptation to voice identity. *British Journal of Psychology, 102*, 748–764. http://dx.doi.org/10.1111/j.2044-8295.2011.02048.x.

Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological Psychology, 80*(1), 95–104. http://dx.doi.org/10.1016/j.biopsycho.2008.03.010.

Su, Y., Chen, A., Yin, H., Qiu, J., Lv, J., Wei, D., ... Wang, T. (2010). Spatiotemporal cortical activation underlying self-referencial processing evoked by self-hand. *Biological Psychology, 85*(2), 219–225. http://dx.doi.org/10.1016/j.biopsycho.2010.07.004.

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*(2), 318–324. http://dx.doi.org/10.1016/j.biopsycho.2010.03.009.

Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology, 40*, 235–248. http://dx.doi.org/10.1111/1469-8986.00025.

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*, 580–591. http://dx.doi.org/10.1162/jocn.2008.20054.

van den Brink, D., Brown, C. M., & Hagoort, P. (2006). The cascaded nature of lexical selection and integration in auditory sentence processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32*(2), 364–372. http://dx.doi.org/10.1037/0278-7393.32.3.364.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25*(2), 394–417.

von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research, 17*(1), 48–55. http://dx.doi.org/10.1016/S0926-6410(03)00079-X.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191–1207. http://dx.doi.org/10.3758/s13428-012-0314-x.

Watson, L. A., Dritschel, B., Obonsawin, M. C., & Jentzsch, I. (2007). Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Research, 1152*(1), 106–110. http://dx.doi.org/10.1016/j.brainres.2007.03.049.

Wechsler, D. (1997). *WAIS-III: Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Xu, M., Homae, F., Hashimoto, R.-I., & 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.

Zhao, K., Yuan, J., Zhong, Y., Peng, Y., Chen, J., Zhou, L., ... Ding, D. (2009). Event-related potential correlates of the collective self-relevant effect. *Neuroscience Letters, 464*(1), 57–61. http://dx.doi.org/10.1016/j.neulet.2009.07.017.

Zhou, A., Shi, Z., Zhang, P., Liu, P., Han, W., Wu, H., ... Xia, R. (2010). An ERP study on the effect of self-relevant possessive pronoun. *Neuroscience Letters, 480*(2), 162–166. http://dx.doi.org/10.1016/j.neulet.2010.06.033.