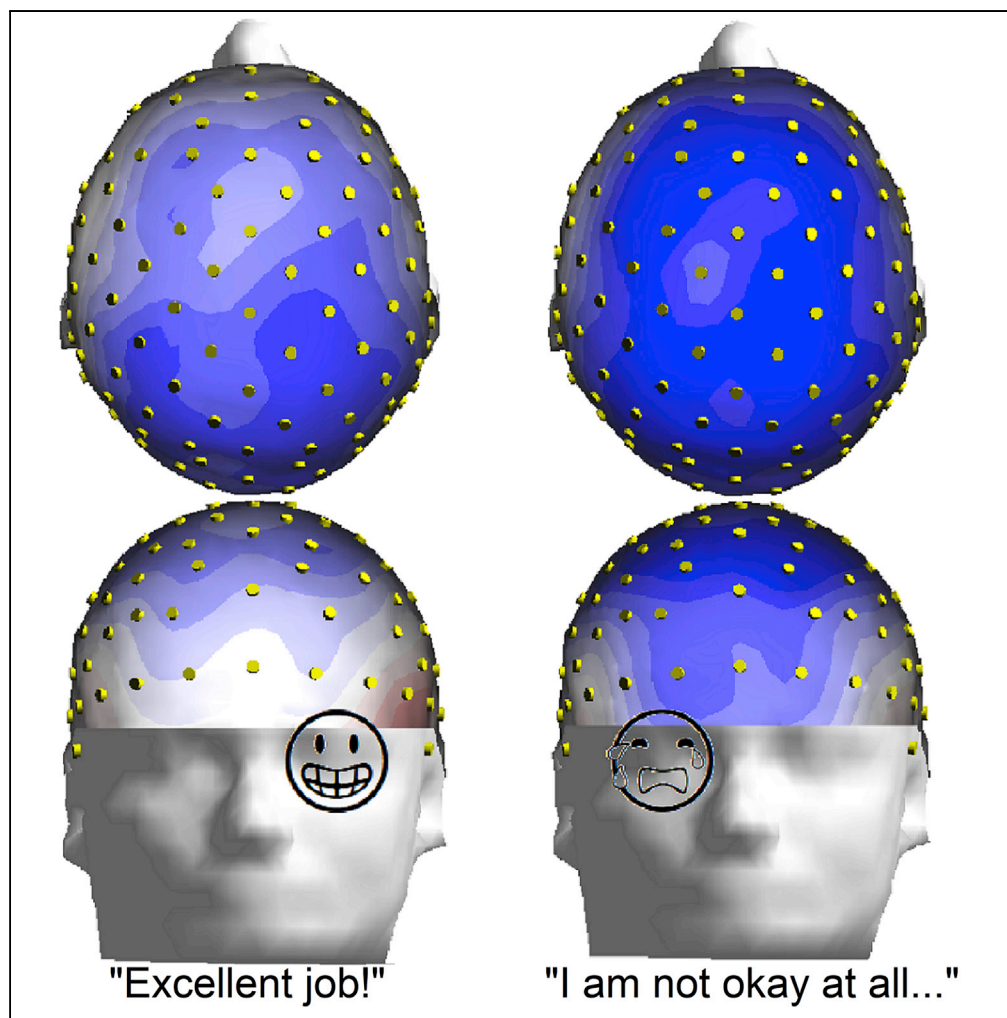


Article

ERP Markers of Valence Coding in Emotional Speech Processing



Alice Mado
Proverbio, Sacha
Santoni, Roberta
Adorni

mado.proverbio@unimib.it

HIGHLIGHTS

Specific ERP markers reflecting negative/positive speech processing were investigated

Positive valence enhanced P450 and LP responses

Negative valence enhanced N400-like anterior negativity

Similar ERP markers were found for the processing of music and affective vocalizations

Proverbio et al., iScience 23, 100933
March 27, 2020 © 2020 The Author(s).
<https://doi.org/10.1016/j.isci.2020.100933>

Article

ERP Markers of Valence Coding in Emotional Speech Processing

Alice Mado Proverbio,^{1,2,*} Sacha Santoni,¹ and Roberta Adorni¹

SUMMARY

How is auditory emotional information processed? The study's aim was to compare cerebral responses to emotionally positive or negative spoken phrases matched for structure and content. Twenty participants listened to 198 vocal stimuli while detecting filler phrases containing first names. EEG was recorded from 128 sites. Three event-related potential (ERP) components were quantified and found to be sensitive to emotional valence since 350 ms of latency. P450 and late positivity were enhanced by positive content, whereas anterior negativity was larger to negative content. A similar set of markers (P300, N400, LP) was found previously for the processing of positive versus negative affective vocalizations, prosody, and music, which suggests a common neural mechanism for extracting the emotional content of auditory information. SwLORETA applied to potentials recorded between 350 and 550 ms showed that negative speech activated the right temporo/parietal areas (BA40, BA20/21), whereas positive speech activated the left homologous and inferior frontal areas.

INTRODUCTION

It is known that speech prosody conveys highly informative cues requiring complex processing. Several event-related potential (ERP) studies have shown how the affective content of speech can modulate ERP responses as early as 200 ms for auditory stimuli very short in length (e.g., Paulmann and Kotz, 2008; Paulmann et al., 2010; Paulmann and Pell, 2010; Paulmann et al., 2013; Schirmer et al., 2013; Schirmer and Kotz, 2006) and at about 300–400 ms for more complex and longer stimuli (see the review by Kotz and Paulmann, 2011). Zinchenko et al. (2015, 2017) used an auditory processing task in dynamic multisensory stimuli and showed that negative emotions can have a modulatory effect as early as 100 ms after stimulus onset (as well as subsequently influencing P200 and N200 ERP components).

The P200 modulation is assumed to reflect enhanced attention to emotional stimuli so that they can be preferentially processed (Paulmann et al., 2013). Later ERP components (N400 and LPC) are also modulated by emotional prosody (e.g., Bostanov and Kotchoubey, 2004; Schirmer et al., 2002, 2005a; Schirmer and Kotz, 2003; Paulmann and Pell, 2010; Kanske and Kotz, 2007; Schirmer et al., 2013), but the pattern of results is strongly dependent on stimuli, tasks, and experimental designs used. Generally, ERP responses display larger amplitudes for emotional or arousing stimuli than for neutral stimuli (e.g., Schapkin et al., 2000). For example, in the pioneering psychophysiological study by Begleiter and Platz (1969) it was found that taboo words (such as swear words) elicited larger negative and positive peaks measured at three latency stages (from about 100 to 400 ms) than neutral words. More recently, Yao et al. (2016) explored the effect of word valence and arousal on auditory ERPs and reported that emotionally positive words elicited an enhancement of the late positive complex (LPC, 450–750 ms), whereas negative words elicited an enhancement of N400 (300–410 ms) response. In addition, low-arousal positive words were associated with a reduction of LPC amplitudes compared with high-arousal positive words, thus suggesting higher responses for more arousing stimuli and a positive/negative modulation of ERP components as a function of words' affective valence. Vallat et al. (2017) found enhanced late positive and late negative components (peaking between 800 and 1000 ms) at posterior and frontal areas for arousing names (but not neutral names).

Although a large literature agrees on a general effect of emotional versus neutral (affective "valence" coding, e.g., Kissler et al., 2007; Herbert et al., 2008) or arousing versus less arousing stimuli ("arousal" coding), especially for P300 (Burlinson, 1986; Cuthbert et al., 2000; Hajcak et al., 2010), it is still not very clear whether it is possible to distinguish the "polarity" (happy versus sad) of the emotion expressed, by looking at specific electrophysiological markers.

¹Milan Center for Neuroscience, Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, Milan, Italy

²Lead Contact

*Correspondence: mado.proverbio@unimib.it
<https://doi.org/10.1016/j.isci.2020.100933>



In this study it was investigated whether it was possible to identify specific ERP markers of emotional prosody, in response to ecological vocal utterances expressing a general positive or negative feeling. A few ERP studies in the literature have found greater positive potentials (P2 or P300s) to positive prosody and larger N400s or negative shifts to negative prosody and music (Grossmann et al., 2005; Proverbio et al., 2019; Proverbio and Piotti, n.d.; Yao et al., 2016), but there is no general consensus on this matter.

As for the neural circuits subtending the ability to understand speech prosody, regardless of language semantic content, the key structures seem to be the superior temporal gyrus (STG, also for being able to merge multimodal information such as body language and facial mimicry), the medial temporal cortex (MTG), and the inferior frontal cortex (IFG), in addition to the limbic system. For example, an fMRI study by Kotz et al. (2003) analyzed the brain response to words spoken with different vocal intonations: positive (joy), negative (anger), and neutral. In this study an IFG activation was observed for processing both positive and negative intonations. Comparing normal and prosodic speech across all three emotional intonations revealed bilateral, but with higher left activation of temporal and subcortical regions (putamen and thalamus) as well as activations of left inferior frontal regions for normal speech. Prosodic speech resulted in bilateral inferior frontal, prefrontal, and caudate activation (Kotz et al., 2003). Again, an fNIRS (functional near infrared spectrograph) study by Zhang et al. (2018a) recorded brain activity during passive listening of sentences consisting of non-words, i.e., lacking semantic content, but with regular syntactic structure subject-verb-object and characterized by a positive (joy), negative (fear and anger), and emotionally neutral prosody. The authors found that the superior temporal cortex, especially the right, middle, and posterior parts of the superior temporal gyrus (BA 22/42), discriminated between emotional and neutral prosodies. Furthermore, brain activation patterns were distinct when positive (happy) were contrasted to negative (fearful and angry) prosody in the left inferior frontal gyrus (BA 45) and the frontal eye field (BA8), and when angry were contrasted to neutral prosody in bilateral orbital frontal areas (BA 10/11).

Overall, the STG and MTG seem to be the core cortical structures for the vocal processing of emotions (Ethofer et al., 2006; Frühholz and Grandjean, 2013; Zhang et al., 2018b; Johnstone et al., 2006). Its anterior portion has shown activation peaks for emotional expressions compared with neutral expressions in many neuroimaging studies (Bach et al., 2008; Brück et al., 2011; Witteman et al., 2012; Wildgruber et al., 2005; Sander et al., 2005; Kreifelts et al., 2010).

A marked lateralization in favor of the right superior temporal cortex has been observed, which is consistent with the greater sensitivity of the right hemisphere to the slower variations of the acoustic profiles typical of emotional expressions (Liebenthal et al., 2016; Witteman et al., 2012; Beaucousin et al., 2006; Vigneau et al., 2011). Other studies have shown a left bias for processing positive emotions (Davidson, 1992; Canli et al., 1998; Flores-Gutiérrez et al., 2007). Indeed, a large body of research indicates that approach and avoidance reactions are asymmetrically instantiated in the left and right hemisphere, respectively (e.g., Davidson, 1992). Furthermore, a considerable body of neuropsychological and psychophysiological evidence suggests that the left hemisphere is related to positive emotions, whereas the right hemisphere is related to negative emotions (e.g., Maxwell and Davidson, 2007; Roesmann et al., 2019; Wyczesany et al., 2018; Simonov, 1998).

In a previous psychophysiological work (Proverbio et al., 2019) the neural processes involved in the coding of musical stimuli and non-verbal vocalizations were investigated. Both stimulus types might be negative or positive in emotional valence. The aim was to explore the existence of common neural bases for the extraction and comprehension of the emotional content of auditory information. We wondered why a sad or happy piece of music might be universally categorized as such by people of different age, culture, education, or musical preference. Specifically, it was hypothesized that the comprehension of the emotional significance of musical stimuli (with positive and negative content) might tap at the same neural mechanism involved in the processing of speech prosodic and affective vocalizations (Panksepp and Bernatzky, 2002). It has been hypothesized (Cook et al., 2006) that minor chords share some harmonic relationships that are found in sad speech prosody. Paquette and coworkers (2018) proposed that music could involve neuronal circuits devoted to the processing of biologically relevant emotional vocalizations, and a few studies have indeed found common brain activations (including middle and superior temporal cortex, medial frontal cortex, and cingulate cortex) during listening to both vocalizations and music (Aubé et al., 2015; Escoffier et al., 2013; Belin et al., 2008). Although from these studies it can be surely argued that similar mechanisms support the comprehension of emotional content of vocalizations and music, there is not a precise

indication about the specific nature and timing of these neural processes, because of the different structure of stimuli usually compared. It is known that both types of complex stimuli are processed within the secondary auditory cortex (Whitehead and Armony, 2018). In a previous ERP study (Proverbio et al., 2019) positive and negative emotional vocalizations (e.g., laughing, crying) and digitally extracted musical tones (played on the violin) were used as stimuli. They shared the same melodic profile and main pitch/frequency characteristics. Participants listened to vocalizations or music while detecting rare auditory targets (bird tweeting or piano's arpeggios). The analyses of ERP responses showed that P200 and later late positive (LP) components were larger to positive stimuli (either voice or music based), whereas negative stimuli increased the amplitude of the anterior N400 component. Furthermore, notation obtained from acoustic signals showed how emotionally negative stimuli tended to be in a Minor key, and positive stimuli in a Major key, thus shedding some light on the brain's ability to understand music. Interestingly, Chen et al. (2008) found larger negativities (500–700 ms) to stimuli when listening to negative music and Schapkin and coworkers (2000) found larger P2 and P3 positive peaks to positive than to negative words, which is compatible with the aforementioned pattern of electrophysiological results.

The aim of the present study was therefore to further investigate whether similar markers of emotional prosody (namely positive versus negative) might be found for speech processing, similar to music and vocalizations, and whether their source reconstruction revealed distinct or shared neural circuits. We hypothesized that the mechanism supporting the comprehension of the emotional content (negative versus positive) of music, speech, and vocalizations shared common neural circuits. We also expected a possible right lateralization for the processing of negative emotions and a left lateralization for positive emotions (Davidson, 1992; Canli et al., 1998; Flores-Gutiérrez et al., 2007; Proverbio et al., 2019).

RESULTS

P450 Response

The repeated measures ANOVA analysis performed on P450 amplitude values, recorded between 350 and 550 ms at AF7-AF8, F7-F8, and FT7-FT8 electrode sites, showed the significant effect of the Emotional Valence [$F(1,19) = 7.21$, $p < 0.015$; $\eta_p^2 = 0.28$] with larger P450 responses to positive (1.29 μV , SD = 0.16) than negative stimuli (0.60, SD = 0.25).

Also significant was the Hemisphere factor [$F(1,19) = 8.83$, $p < 0.005$, $\eta_p^2 = 0.32$] with greater P450 responses over the right than left hemisphere. The significant interaction of Valence \times Hemisphere [$F(1,19) = 7.99$, $p < 0.01$, $\eta_p^2 = 0.30$] and relative post hoc comparisons indicated larger responses over right than left hemisphere ($p < 0.01$) for negative stimuli, as can be appreciated by looking at ERP waveforms of Figure 1 and topographical maps of Figure 2.

Source Reconstruction

To identify the neural generators of the surface activity within the P450 latency range a *standardized weighted low-resolution electromagnetic tomography* (swLORETA) was applied to ERP potentials recorded between 350 and 550 ms, separately for positive and negative stimuli. In this way we were able to observe neural processing of speech of different emotional valence in the same identical time window, also allowing comparison with similar ERP findings on music, vocalizations, and prosody (Proverbio et al., 2019; Proverbio and Piotti, n.d.). The inverse solution is visible in Figure 3, whereas a complete list of active electromagnetic dipoles is reported in Table 1. A further swLORETA (visible in Figure 4) was applied to the difference waves obtained by subtracting ERPs to negative stimuli from ERPs to positive stimuli in the P450 latency range. The homologous “negative minus positive” differential activity was not explored because *anterior negativity* occurred considerably later in time (1,100–1,200 ms).

During inattentive perception of negative speech the most active neural sources (ordered for magnitude) were the left superior frontal gyrus (BA8), the right medial frontal gyrus, the fusiform gyrus bilaterally, the speech-specific left superior temporal gyrus (BA42), the right medial and inferior temporal gyrus, and parietal and limbic areas.

In Table 1 a description of presumed functions associated to these areas can be also found. During inattentive perception of positive speech the most active neural sources were the left and right parahippocampal place area (BA34), the superior frontal gyrus (BA8), as well as limbic, visual, and parietal areas (please see Table 1 for details).

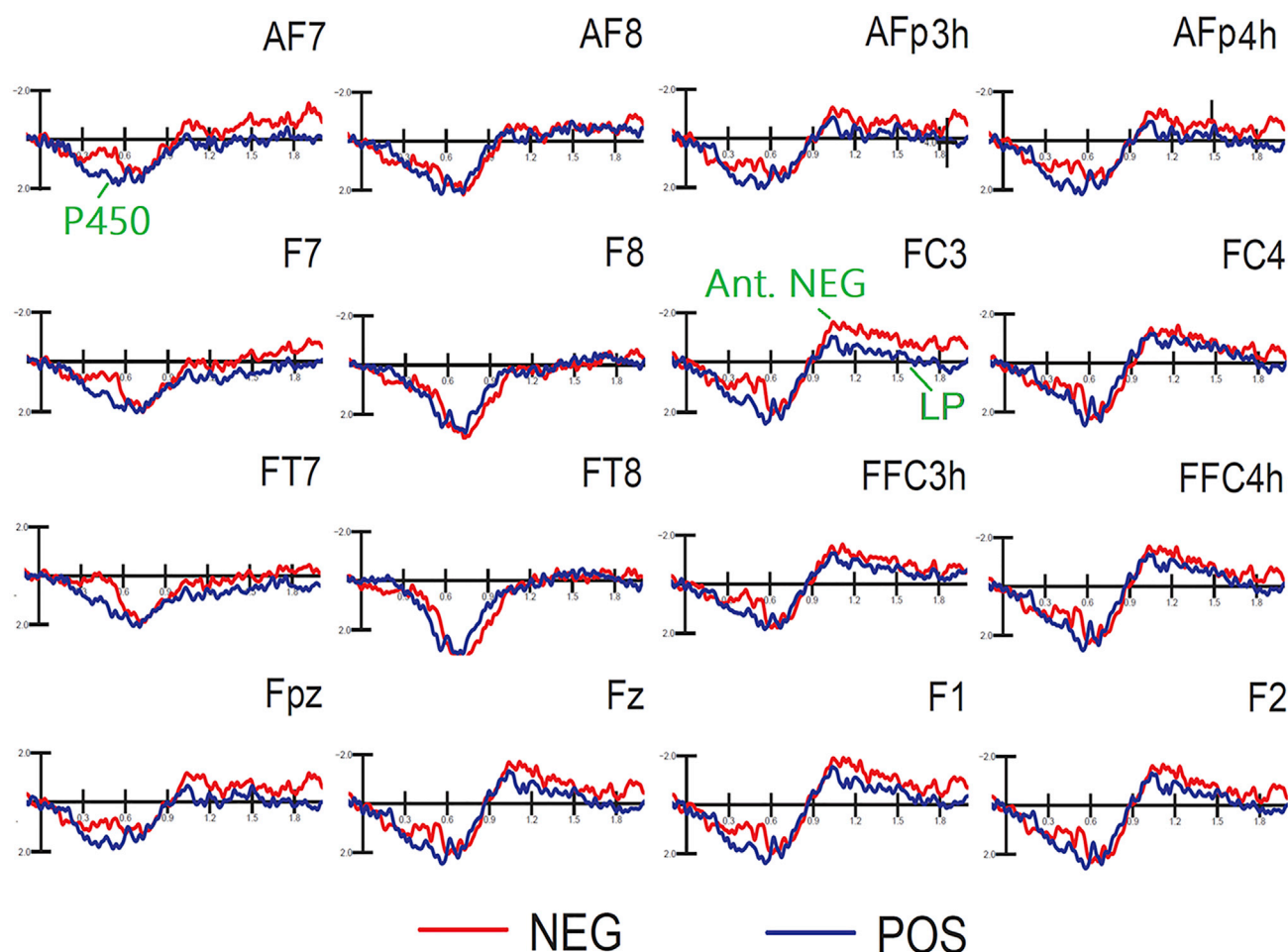


Figure 1. ERP Waveforms Recorded at Left and Right Anterior Sites in Response to Positive or Negative Sentences

The swLORETA applied to the positive – negative difference in brain activity showed a largely left hemispheric circuit for representing positive emotional hues, including the left inferior frontal area, the left inferior parietal area, and the left middle temporal area, as can be appreciated in [Figure 4](#).

Anterior Negativity

The repeated measures ANOVA analysis performed on *anterior negativity* mean area amplitude recorded between 1,100 and 1,200 ms at AFp3h-AFp4h, FC3-FC4 e FCC3h-FCC4h electrode sites yielded the significance of emotional valence [$F(1,19) = 8.92$, $p < 0.008$, $\eta_p^2 = 0.32$] with greater *anterior negativity* to negative ($-1.5 \mu V$, $SD = 0.38$) than positive speech ($0.74 \mu V$, $SD = 0.44$). The further interaction of valence \times hemisphere [$F(1,19) = 5.63$, $p < 0.03$, $\eta_p^2 = 0.23$] and relative post hoc comparisons showed that *anterior negativity* was larger to negative than positive sentences. Furthermore, it was smaller (more positive) to positive than negative sentences, over the left hemisphere ($LH-neg = -1.5 \mu V$, $SD = 0.38$; $RH-neg = -1.46 \mu V$, $SD = 0.39$; $LH-pos = -0.63 \mu V$, $SD = 0.44$; $RH-pos = -0.84 \mu V$, $SE = 0.44$). This can be appreciated by looking at ERP waveforms of [Figure 1](#) and topographical maps of [Figure 5](#).

The factor electrode was also significant [$F(2,38) = 3.53$, $p < 0.035$, $\eta_p^2 = 0.16$], showing greater anterior negativity amplitudes over frontocentral than anterior frontal sites.

Late Positivity

The repeated measures ANOVA analysis performed on LP mean area amplitude recorded between 1,300 and 1,500 ms over F1-F2, Fpz-FZ electrode sites showed a significance of emotional valence factor

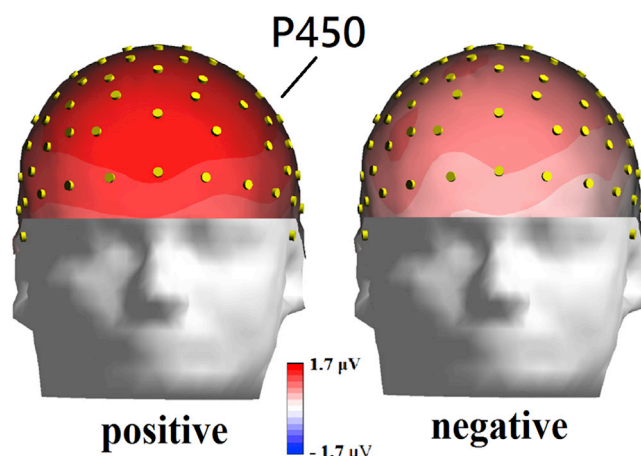


Figure 2. Isocolour Topographical Maps of P450 (Front View) Recorded between 350 and 550 ms in the Two Valence Conditions

[$F(1, 19) = 12.1$, $p < 0.0025$, $\eta_p^2 = 0.38$], with larger LP responses to positive ($-0.48 \mu\text{V}$, $SD = 0.34$) than negative stimuli ($-1.36 \mu\text{V}$, $SD = 0.36$) as can be seen in [Figures 1 and 5](#). The factor electrode was also significant [$F(3, 57) = 3.6$, $p < 0.02$, $\eta_p^2 = 0.20$], indicating greater responses over Fpz than other sites.

DISCUSSION

The aim of the present study was to investigate the neural correlates involved in the understanding of emotional language by analyzing the bioelectrical response to speech as a function of its emotional and prosodic content (positive versus negative).

The first component showing a valence-dependent modulation was P450 (350–550 ms), which was greater in amplitude to positive than negative sentences. Similarly, previous studies have shown an enhancement of earlier P200 response to positive vs. negative prosody ([Paulmann and Kotz, 2008](#)). Again, [Kanske and Kotz \(2007\)](#) found that semantically positive words elicited larger P2 responses compared with neutral words. In a recent study by [Proverbio et al. \(2019\)](#) contrasting ERP responses to positive versus negative vocalizations (for example, laughter and crying) it was found that positive stimuli elicited larger P200 responses than negative stimuli. It is possible that differences in the latency of the first modulated positive response (P200, P300, or P450) depend on stimulus duration or complexity.

ERP analyses also showed that P450 to negative stimuli was greater over the right than left frontal areas, whereas no significant hemispheric difference was observed for positive stimuli. This right hemispheric asymmetry for negatively valenced stimuli fits well with “valence hypothesis” by [Davidson et al. \(1999\)](#) according to which there would be a right hemispheric dominance for the processing of emotionally negative stimuli.

The second component considered, a negative deflection recorded between 1,100 and 1,200 ms and here named anterior negativity response, was significantly larger in response to emotionally negative verbal expressions. Since sentences did not differ in syntactic structure or word number/frequency it is reasonable to think that this difference might be ascribed to the different emotional content and valence (including prosodic inflections). Similar enhanced negativity to emotional prosody was found in an ERP study by [Schirmer et al. \(2002\)](#) and by [Grossmann et al. \(2005\)](#) reporting a greater negativity shift maximum over frontal and central sites to negative stimuli. In the latter study, 7-month-old children listened to words characterized by a happy, angry, or neutral prosody. Their ERP responses showed that the negative prosody elicited more negative potentials than the other prosodic intonations. An enhanced N400 response to negatively rather than positively valenced stimuli was also found by [Proverbio et al. \(2019\)](#) reporting larger N400s to human vocalizations, as well as to music obtained by transforming those vocalizations in instrumental music. Intriguingly, a similar N400 enhancement during processing of negative vs. positive stimuli has been reported by [Proverbio and Piotti, n.d.](#), who recently compared the processing of viola played music obtained by digitally transforming positive versus negative speech

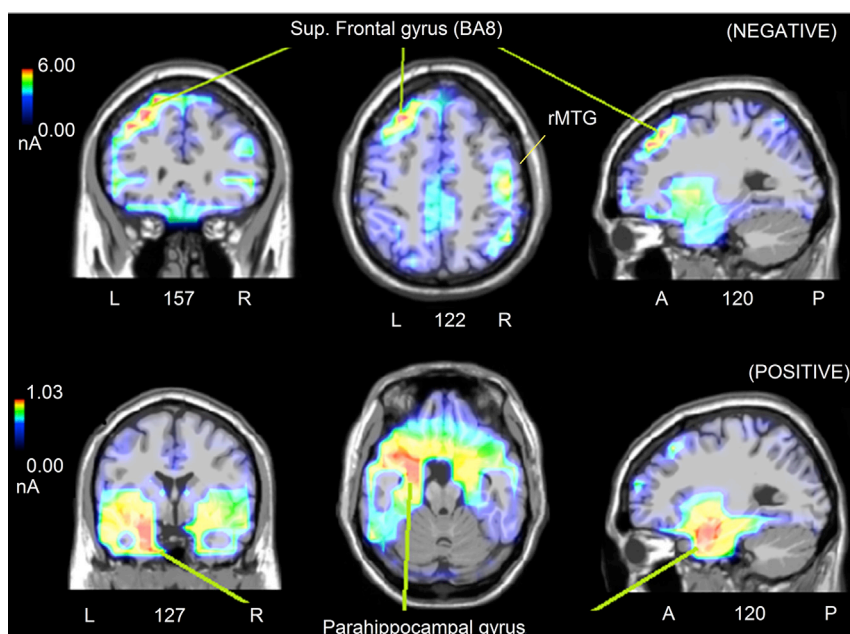


Figure 3. (Top) Coronal, Axial, and Sagittal Brain Sections Showing Some of the Sources of Activation for P450 Response (350–550 ms) to Negative Sentences According to swLORETA

Since negative sentences were non-targets the prefrontal activation possibly indicates increased attention and enhanced stimulus encoding. Note also the right activation of the medial temporal gyrus. (Bottom) Coronal, axial, and sagittal brain sections showing some of the sources of activation for P450 response (350–550 ms) to positive sentences according to swLORETA. The strong activity of the parahippocampal place area (PPA) compared with prefrontal cortex possibly suggests a reduced attention to positive auditory stimuli and a stronger encoding of the visual place stimulus. Note that the scale is different across conditions. A, anterior; P, posterior; L, left hemisphere; R, right hemisphere.

utterances into music. Again, Yao et al. (2016) found that emotionally positive words elicited an enhancement of the late positive complex (LPC, 450–750 ms), whereas negative words elicited an enhancement of N400 (300–410 ms) response. This similarity suggests the existence of a common mechanism for extracting and comprehending the emotional content of music, speech, and vocalizations, based on the harmonic spectrum and melodic structure of prosody, voice, and music. The unique contribution of this research compared with previous findings in the ERP literature on emotional prosody is the comparison of bioelectrical responses to speech, or vocalizations (Proverbio et al., 2019) with those elicited by musical stimuli (played by violin/viola) and the discovery of a unitary set of ERP markers of emotional valence that is not dependent on stimulus modality.

The last component modulated by emotional valence was late positivity, which was quantified at a relatively early onset (1,300–1,500 ms) but was long lasting. LP response was larger to positive than to negative sentences. Again, this effect was also found by the previously mentioned studies examining the brain response to vocalizations and music (Proverbio et al., 2019; Proverbio and Piotti, n.d.). Other ERP studies in the literature support our finding of an increased late positivity to positively valenced stimuli (Carretié et al., 2008; Hinojosa et al., 2010; Kanske and Kotz, 2007; Schacht and Sommer, 2009; Herbert et al., 2006, 2008). Bayer and Schacht (2014) observed similar results for LP in a study in which the participants were visually presented with words, faces, or images of positive, negative, or neutral valence. For what concerns a functional interpretation, the modulations of LP could reflect a prolonged evaluation of the stimulus probably caused by its intrinsic emotional relevance (Cuthbert et al., 2000; Schupp et al., 2004).

To investigate the existence of distinct neural circuits for processing different speech emotional contents a swLORETA source reconstruction was performed on the earliest modulation, corresponding to the peak of P450 response (350–550). In reviewing dipole localizations and approximate Brodmann areas (the estimated spatial accuracy of swLORETA being 5 mm), it should be noted that the technique has limited spatial resolution, as compared, for example, to MEG or fMRI. The strongest generator active for both valence

NEGATIVE

Magn	Hem	Lobe	Gyrus	BA	Presumed function
6.91	L	F	Superior frontal	8	Semantic working memory
6.81	R	F	Middle frontal	10	Attention
6.13	R	F	Medial frontal	10	
6.74	L	T	Fusiform	37	Visual perception of the living room (place, colors, shapes, etc.)
5.62	R	T	Fusiform	20	
6.03	L	T	Superior temporal gyrus	42	Speech processing
5.95	L	F	Precentral	6	
5.72	R	P	Inferior parietal lobule	40	Language-evoked imagery (NEGATIVE)
5.47	R	P	Inferior parietal lobule	39	
5.15	R	T	Middle temporal	21	Prosody (NEGATIVE)
5.13	R	T	Inferior temporal	20	
4.55	L	F	Superior frontal	9	
4.2	L	Limbic	Cingulate	31	Emotion

POSITIVE

10.35	L	Limbic	Parahippocampal	34	Parahippocampal place area (PPA)
9.58	R	Limbic	Parahippocampal	34	
9.47	L	F	Superior frontal	8	Semantic working memory
8.55	R	Limbic	Uncus	20	Emotion
6.32	R	Limbic	Cingulate	23	
8.47	L	F	Superior frontal	10	Attention
4.9	R	T	Fusiform	37	Visual perception of the living room (place, colors, shapes, etc.)
8.48	L	T	Fusiform	20	
4.9	R	O	Fusiform	37	
4.74	L	O	Lingual	18	
6.07	L	P	Postcentral	3	Language-evoked imagery (POSITIVE)
4.79	L	P	Supereior parietal lobule	7	

POSITIVE - NEGATIVE

9.39	L	F	Inferior frontal	6	Positive emotions
8.27	L	P	Inferior parietal	40	Language-evoked imagery (POSITIVE)
0.21	L	T	Middle temporal	21	Prosody (POSITIVE)
0.18	R	T	Inferior temporal	20	Prosody

Table 1. Active Electromagnetic Dipoles Explaining the P450 Surface Voltage Recorded between 350 and 550 ms according to swLORETA in Response to Negative and Positive Stimuli and the P450 Differential Activity (Positive – Negative)

Magn, magnitude in nA; Hem, hemisphere; L, left; R, right. Power RMS = 2; F, Frontal; T, Temporal; P, Parietal; O, Occipital.

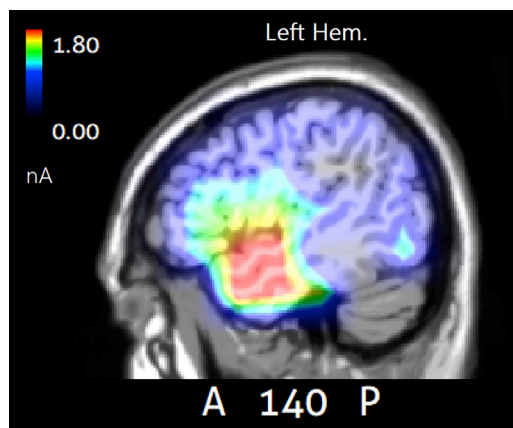


Figure 4. swLORETA Solution Relative to the Positive - Negative Difference

Left sagittal brain section (centered on the left medial temporal source) showing the electromagnetic activity explaining the differential voltage elicited by positive – negative stimuli in the P450 latency range (350–550 ms) according to swLORETA. A, anterior; P, posterior.

conditions was the superior frontal gyrus of the left hemisphere (BA 8). The left PFC seems involved in stimulus encoding and semantic working memory (Tulving et al., 1994). According to Gabrieli et al., 1998 it plays a crucial role in semantic working memory. For example, patients with left prefrontal focal lesions (du Bois-gueheneuc et al., 2006) show a clear deficit in working memory. It should be mentioned, however, that the same region (BA8) is implicated in more general memory functions, being also involved in the processing of musical stimuli (Flores-Gutiérrez et al., 2007; Platel et al., 2003; Proverbio and De Benedetto, 2018). The activation of the superior frontal gyrus (BA 8) was in fact also found in the previously mentioned ERP studies by Proverbio and coauthors (Proverbio et al., 2019; Proverbio and Piotti, n.d.) in response to music and vocalizations, with greater signals for negative stimuli. Also active during listening to negative speech was the left superior temporal gyrus (BA 42), also known as secondary auditory cortex (Wernicke area), devoted to phoneme and word recognition (e.g., Ardila et al., 2016). Heschl's gyri (BA41) were not found active at this late stage of processing. Also active was the so-called extended Wernicke area (BA20 and BA37 of the left hemisphere). According to meta-analytical studies (Ardila et al., 2016) these regions would be implicated in reception and understanding of speech. However, they might also indicate enhanced visual processing of background: objects, colors, shapes, and place (and lesser orienting toward the auditory channel). Both interpretations need further data before a definitive conclusion could be reached.

The second strongest focus of activity during listening to negative sentences was the medial/superior frontal cortex of the right hemisphere that has been characterized by the literature as involved in sustaining attention (Velanova et al., 2003), as well as in controlling attentional control and allocation (Pollmann et al., 2000; Pollmann, 2001), and context updating (Zinchenko et al., 2019). Since the task expressly required participants to ignore non-target stimuli this effect might indicate stronger attention attracting properties of negative speech. This interpretation is also supported by the much stronger activity of frontal and speech devoted areas during listening to negative speech, as opposed to the activation of visual areas and parahippocampal place areas (possibly reflecting the processing of the visual background: a place) during listening to positive speech. Other areas involved in speech processing were parietal regions, with a different hemispheric asymmetry as a function of valence. More specifically, the right inferior parietal lobule (BA 39 and BA 40) was found to be active in response to negative speech, whereas the left superior parietal lobule (BA 7) and postcentral gyrus (BA 3) were found to be active during processing of positive speech. This asymmetry might be interpreted in the light of the valence hemispheric specialization (e.g., Davidson et al., 1999), whereas their activation might be linked to their role in language-evoked imagery (Just et al., 2004).

Specifically active for processing negative speech were also the right medial and inferior temporal gyri (BA21 and BA20). These areas seem to respond to both human and animal negative vocalizations (Belin et al., 2008), as well as to music extracted from human negative vocalizations (Proverbio et al., 2019). A similar finding was reported by Altenmüller et al. (2002) who found that listening to negative melodies engaged the right superior temporal gyrus (BA 42 and BA 22) and the right anterior cingulate cortex (BA 24), whereas listening to positive music engaged the left middle temporal gyrus (BA 21) and left inferior temporal gyrus (BA 20). This pattern of lateralization fits with the hypothesis proposed by various authors according to which positive emotions and stimuli that induce this type of emotion would be

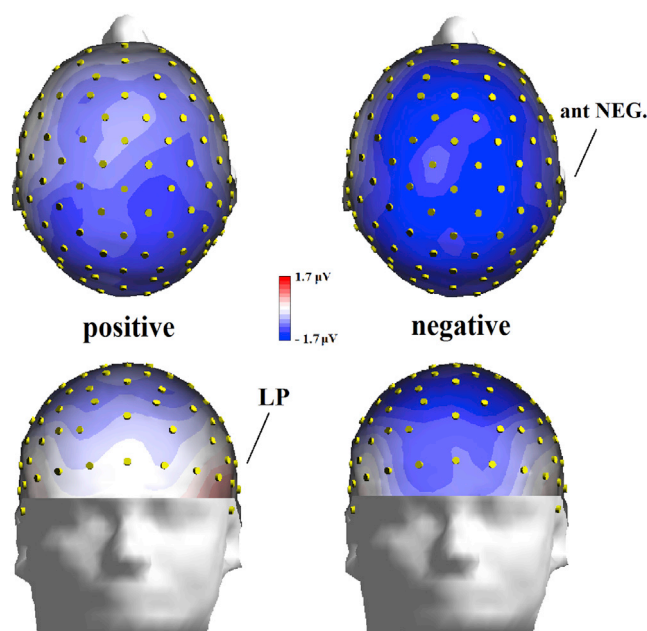


Figure 5. Isocolour Topographical Maps of Anterior Negativity Voltages (Top and Front Views) Recorded between 1,100 and 1,200 ms in the Two Valence Conditions

The LP focus over the left inferior frontal area can also be observed.

processed mainly by the left hemisphere, whereas negative emotions would be mainly processed by the right hemisphere (Davidson, 1992; Canli et al., 1998). Quite similarly, Flores-Gutiérrez et al. (2007) had participants listening to music with positive or negative emotional value and found that "unpleasant" music was mainly elaborated by the right hemisphere, whereas the "pleasant" music mainly activated the right hemisphere (Flores-Gutiérrez et al., 2007).

Conversely, specifically active for processing positive speech was the parahippocampal gyrus (PHG), bilaterally but with somewhat greater left hemisphere asymmetry. Since positive speech was task irrelevant (and less biologically relevant than negative speech) and participants were required to visually fixate on the center of a picture displaying a place (namely, a living room) this activation might be linked to the role of the parahippocampal gyrus in place processing (parahippocampal place area, PPA; Epstein, 1998). However, the role of PHG in the semantic processing of verbal stimuli with positive emotional value cannot be disregarded. Indeed, in an fMRI study in which participants were required to silently read positive, negative, and neutrals adjectives it was shown that adjectives with positive emotional value elicited a more pronounced amygdaloid and PHG activation than emotionally neutral and negative adjectives (Herbert et al., 2009). Furthermore, an fMRI study (Johnstone et al., 2006) has identified the neural circuits (also in the parahippocampal area) involved in processing voices connoted by an emotionally positive prosody. The further swLORETA applied to the differential positive – negative response (in the P450 latency range) allowed us to investigate the neural circuits specifically involved in the processing of positive emotional connotations of auditory information, regardless of other factors, such as the processing of visual background, task-related factors (working memory, attention), and speech processing. Source reconstruction data supported the existence of a strongly left-lateralized neural network including the left inferior frontal area (which agrees with the available literature on positive emotions carried by auditory stimuli, such as Proverbio et al., 2019; Belin et al., 2008; Kotz et al., 2003; Koelsch et al., 2006; and Zhang et al., 2018a, 2018b), the left inferior parietal area (e.g., Flores-Gutiérrez et al., 2007), and left middle temporal gyrus (e.g., Proverbio and Piotti, n.d.; Altenmüller et al., 2002).

In conclusion, the data outlined above have identified three ERP markers indexing speech valence processing, namely, enhanced early and late anterior positivities (P450 and LP) for processing speech that conveyed positive emotions and enhanced anterior negativity response for speech that conveyed negative

emotions. It should be noted that a somewhat similar set of ERP components (P300, N400, and LP) was also seen in a series of ERP studies involving the perception of non-verbal vocalizations and music, as well as a study investigating the brain response to music extracted from the same speech stimuli of the present study. It is possible that differences in the latency of the above ERP responses (across studies) is dependent on differences in stimulus nature, duration, and complexity.

The similarity hints at a common mechanism for extracting emotional cues from auditory information. Furthermore, according to source reconstruction, and in line with recent fMRI studies, the key role of right medial temporal cortex for processing negative prosodic/melodic information was supported (Davidson, 1992; Canli et al., 1998; Flores-Gutiérrez et al., 2007; Proverbio et al., 2019; Proverbio and Piotti, n.d.; Belin et al., 2008). In addition, a later activation of the left inferior frontal area (at LP level) for processing positive prosodic/melodic information was found, which was also predicted on the basis of previous research (Proverbio et al., 2019; Belin et al., 2008; Kotz et al., 2003; Koelsch et al., 2006; Zhang et al., 2018a, 2018b).

In fact, according to the literature (Ethofer et al., 2006; Bach et al., 2008; Lattner et al., 2005; Brück et al., 2011; Zhang et al., 2018a, 2018b), (after sensory processing and the extraction of acoustic parameters) the right superior temporal cortex would identify the affective nature of prosody by means of multimodal integration, whereas the bilateral inferior frontal gyrus would perform a sophisticated evaluation and semantic understanding of the vocally expressed emotions, at a later stage. A general right hemispheric dominance for processing negative emotional content of auditory information was also found in the present study.

Limitations of the Study

The study is somewhat limited by the reduced spatial resolution of EEG technique as compared with other neuroimaging techniques such as fMRI or MEG. Another potential limit might be the uncommon choice of time windows for ERP quantifications. However, the relatively late time windows were justified by the stimulation prolonged over time.

METHODS

All methods can be found in the accompanying [Transparent Methods supplemental file](#).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.100933>.

ACKNOWLEDGMENTS

This study was funded by 2017-ATE-0058 grant no. 16940 by University of Milano-Bicocca to A.M.P., entitled "The processing of emotional valence in language and music: an ERP investigation." The authors are deeply grateful to Francesco de Benedetto for his technical support. We are also grateful to Dr. Patricia Klaas for the final English editing.

AUTHOR CONTRIBUTIONS

A.M.P. designed the study, interpreted the data, and wrote the paper. S.S. created and validated the stimuli. S.S. and R.A. recorded and analyzed EEG/ERP data. All authors have approved the submitted version of the manuscript.

DECLARATION OF INTERESTS

The authors declare that the research was conducted in the absence of any real or perceived conflict of interest.

Received: October 11, 2019

Revised: December 20, 2019

Accepted: February 19, 2020

Published: March 27, 2020

REFERENCES

- Altenmüller, E., Schürmann, K., Lim, V., and Parlit, D. (2002). Hits to the left - flops to the right. Different emotions during music listening are reflected in cortical lateralisation patterns. *Neuropsychologia* 40, 1017–1025.
- Ardila, A., Bernal, B., and Rosselli, M. (2016). How localized are language brain areas? A review of Brodmann areas involvement in oral language. *Arch. Clin. Neuropsychol.* 31, 112–122.
- Aubé, W., Angulo-Perkins, A., Peretz, I., Concha, L., and Armony, J.L. (2015). Fear across the senses: brain responses to music, vocalizations and facial expressions. *Soc. Cogn. Affect. Neurosci.* 10, 399–407.
- Bach, D.R., Grandjean, D., Sander, D., Herdener, M., Strik, W.K., and Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *NeuroImage* 42, 919–927.
- Bayer, M., and Schacht, A. (2014). Event-related brain responses to emotional words, pictures, and faces—a cross-domain comparison. *Front. Psychol.* 5, 1106.
- Beaucousin, V., Lacheret, A., Turbelin, M.R., Morel, M., Mazoyer, B., and Tzourio-Mazoyer, N. (2006). fMRI study of emotional speech comprehension. *Cereb. Cortex* 17, 339–352.
- Begleiter, H., and Platz, A. (1969). Cortical evoked potentials to semantic stimuli. *Psychophysiology* 6, 91–100.
- Belin, P., Fecteau, S., Charest, I., Nicastro, N., Hauser, M.D., and Armony, J.L. (2008). Human cerebral response to animal affective vocalizations. *Proc. R. Soc. B Biol. Sci.* 275, 473–481.
- Bostanov, V., and Kotchoubey, B. (2004). Recognition of affective prosody: continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology* 41, 259–268.
- Brück, C., Kreifelts, B., and Wildgruber, D. (2011). Emotional voices in context: a neurobiological model of multimodal affective information processing. *Phys. Life Rev.* 8, 383–403.
- du Boisgueheneuc, F., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., Samson, Y., Zhang, S., and Dubois, B. (2006). Functions of the left superior frontal gyrus in humans: a lesion study. *Brain* 129 (Pt 12), 3315–3328.
- Burleson, M.H. (1986). Multiple P3s to emotional stimuli and their theoretical significance. *Psychophysiology* 23, 684–694.
- Canli, T., Desmond, J.E., Zhao, Z., Glover, G., and Gabrieli, J.D.E. (1998). Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport* 9, 3233–3239.
- Carretié, L., Hinojosa, J.A., Albert, J., López-Martín, S., De La Gándara, B.S., Igoa, J.M., and Sotillo, M. (2008). Modulation of ongoing cognitive processes by emotionally intense words. *Psychophysiology* 45, 188–196.
- Chen, J., Yuan, J., Huang, H., Chen, C., and Li, H. (2008). Music-induced mood modulates the strength of emotional negativity bias: an ERP study. *Neurosci. Lett.* 445, 135–139.
- Cook, N.D., Fujisawa, T.X., and Takami, K. (2006). Evaluation of the affective valence of speech using pitch substructure. *IEEE Trans. Audio Speech Lang. Process.* 14, 142–151.
- Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., and Lang, P.J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biol. Psychol.* 52, 95–111.
- Davidson, R.J. (1992). Anterior Cerebral Asymmetry and the Nature of Emotion. *Brain and Cognition* 20, 125–151.
- Davidson, R.J., Abercrombie, H., Nitschke, J.B., and Putnam, K. (1999). Regional brain function, emotion and disorders of emotion. *Curr. Opin. Neurobiol.* 9, 228–234.
- Epstein, R. (1998). A cortical representation of the local visual environment. *Nature* 392, 598–601.
- Escoffier, N., Zhong, J., and Schirmer, A. (2013). Emotional expressions in voice and music: same code, same effect? *Hum. Brain. Mapp.* 34, 1796–1810.
- Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., and Wildgruber, D. (2006). Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. *NeuroImage* 30, 580–587.
- Flores-Gutiérrez, E.O., Díaz, J.L., Barrios, F.A., Favila-Humara, R., Guevara, M.A., del Río-Portilla, Y., and Corsi-Cabrera, M. (2007). Metabolic and electric brain patterns during pleasant and unpleasant emotions induced by music masterpieces. *Int. J. Psychophysiol.* 65, 69–84.
- Frühholz, S., and Grandjean, D. (2013). Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: a quantitative meta-analysis. *Neurosci. Biobehav. Rev.* 37, 24–35.
- Gabrieli, J.D., Poldrack, R.A., and Desmond, J.E. (1998). The role of left prefrontal cortex in language and memory. *Proc. Natl. Acad. Sci. U S A* 95, 906–913.
- Grossmann, T., Striano, T., and Friederici, A.D. (2005). Infants' electric brain responses to emotional prosody. *Neuroreport* 16, 1825–1828.
- Hajcak, G., MacNamara, A., and Olvet, D.M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Dev. Neuropsychol.* 35 (2), 129–155.
- Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., and Rockstroh, B. (2006). Processing of emotional adjectives: evidence from startle EMG and ERPs. *Psychophysiology* 43, 197–206.
- Herbert, C., Junghofer, M., and Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology* 45, 487–498.
- Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., and Kissler, J. (2009). Amygdala activation during reading of emotional adjectives - an advantage for pleasant content. *Social Cogn. Affect. Neurosci.* 4, 35–49.
- Hinojosa, J.A., Méndez-Bértolo, C., and Pozo, M.A. (2010). Looking at emotional words is not the same as reading emotional words: behavioral and neural correlates. *Psychophysiology* 47, 748–757.
- Johnstone, T., Van Reekum, C.M., Oakes, T.R., and Davidson, R.J. (2006). The voice of emotion: an fMRI study of neural responses to angry and happy vocal expressions. *Social Cogn. Affect. Neurosci.* 1, 242–249.
- Just, M.A., Newman, S.D., Keller, T.A., McEleney, A., and Carpenter, P.A. (2004). Imagery in sentence comprehension: an fMRI study. *NeuroImage* 21, 112–124.
- Kanske, P., and Kotz, S.A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* 1148, 138–148.
- Kissler, J., Herbert, C., Peyk, P., and Junghofer, M. (2007). Buzzwords: early cortical responses to emotional words during reading. *Psychol. Sci.* 18, 475–480.
- Koelsch, S., V Cramon, D.Y., Müller, K., and Friederici, A.D. (2006). Investigating emotion with music: an fMRI study. *Hum. Brain Mapp.* 27, 239–250.
- Kotz, S.A., Meyer, M., Alter, K., Besson, M., von Cramon, D.Y., and Friederici, A.D. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Lang.* 86, 366–376.
- Kotz, S.A., and Paulmann, S. (2011). Emotion, language, and the brain. *Lang. Linguist. Compass* 5, 108–125.
- Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., and Wildgruber, D. (2010). Association of trait emotional intelligence and individual fMRI-activation patterns during the perception of social signals from voice and face. *Hum. Brain Mapp.* 31, 979–991.
- Lattner, S., Meyer, M., and Friederici, A.D. (2005). Voice perception: sex, pitch, and the right hemisphere. *Hum. Brain Mapp.* 24, 11–20.
- Liebenthal, E., Silbersweig, D.A., and Stern, E. (2016). The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Front. Neurosci.* 10, 506.
- Maxwell, J.S., and Davidson, R.J. (2007). Emotion as motion: asymmetries in approach and avoidant actions. *Psychol. Sci.* 18, 1113–1119.
- Panksepp, J., and Bernatzky, G. (2002). Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behav. Process.* 60, 133–155.
- Paulmann, S., and Kotz, S.A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport* 19, 209–213.
- Paulmann, S., and Pell, M.D. (2010). Contextual influences of emotional speech prosody on face

processing: how much is enough? *Cogn. Affect. Behav. Neurosci.* 10, 230–242.

Paulmann, S., Seifert, S., and Kotz, S.A. (2010). Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. *Soc. Neurosci.* 5, 59–75.

Paulmann, S., Bleichner, M., and Kotz, S.A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Front. Psychol.* 4, 345.

Platel, H., Baron, J.C., Desgranges, B., Bernard, F., and Eustache, F. (2003). Semantic and episodic memory of music are subserved by distinct neural networks. *Neuroimage* 20, 244–256.

Pollmann, S., Weidner, R., Müller, H.J., and von Cramon, D.Y. (2000). A fronto-posterior network involved in visual dimension changes. *J. Cogn. Neurosci.* 12, 480–494.

Pollmann, S. (2001). Switching between dimensions, locations, and responses: the role of the left frontopolar cortex. *Neuroimage* 14, 118–124.

Proverbio, A.M., and De Benedetto, F. (2018). Auditory enhancement of visual memory encoding is driven by emotional content of the auditory material and mediated by superior frontal cortex. *Biol. Psychol.* 132, 164–175.

Proverbio, A.M., De Benedetto, F., and Guazzzone, M. (2019). Shared neural mechanisms for processing emotions in music and vocalizations. *Eur. J. Neurosci.* <https://doi.org/10.1111/ejn.14650>.

Proverbio, A.M., Piotti, E. (n.d.). Common neural bases for processing speech prosody and music: An integrated model. *Psychol. Sci.*

Roesmann, K., Dellert, T., Junghofer, M., Kissler, J., Zwitserlood, P., Zwanzger, P., and Dobel, C. (2019). The causal role of prefrontal hemispheric asymmetry in valence processing of words - insights from a combined cTBS-MEG study. *Neuroimage* 191, 367–379.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.R., and Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage* 28, 848–858.

Schacht, A., and Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cogn. Affect. Behav. Neurosci.* 9, 28–43.

Schapkin, S.A., Gusev, A.N., and Kuhl, J. (2000). Categorization of unilaterally presented emotional words: an ERP analysis. *Acta Neurobiol. Exp. (Wars)* 60, 17–28.

Schirmer, A., Kotz, S.A., and Friederici, A.D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cogn. Brain Res.* 14, 228–233.

Schirmer, A., and Kotz, S.A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *J. Cogn. Neurosci.* 15, 1135–1148.

Schirmer, A., Striano, T., and Friederici, A.D. (2005a). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport* 16, 635–639.

Schirmer, A., and Kotz, S.A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10, 24–30.

Schirmer, A., Chen, C.B., Ching, A., Tan, L., and Hong, R.Y. (2013). Vocal emotions influence verbal memory: neural correlates and interindividual differences. *Cogn. Affective, Behav. Neurosci.* 13, 80–93.

Schupp, H.T., Junghöfer, M., Weike, A.I., and Hamm, A.O. (2004). The selective processing of briefly presented affective pictures: an ERP analysis. *Psychophysiology* 41, 441–449.

Simonov, P.V. (1998). The functional asymmetry of the emotions. *Zh. Vyssh. Nerv. Deiat. Im. I P Pavlova* 48, 375–380.

Tulving, E., Kapur, S., Craik, F.I.M., Moscovitch, M., and Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc. Natl. Acad. Sci. U S A* 91, 2016–2020.

Vallat, R., Lajnef, T., Eichenlaub, J.B., Berthomier, C., Jerbi, K., Morlet, D., and Ruby, M. (2017). Increased evoked potentials to arousing auditory stimuli during sleep: implication for the understanding of dream recall. *Front. Hum. Neurosci.* 21, 132.

Velanova, K., Jacoby, L.L., Wheeler, M.E., McAvoy, M.P., Petersen, S.E., and Buckner, R.L. (2003). Functional-anatomic correlates of sustained and transient processing components engaged during controlled retrieval. *J. Neurosci.* 17, 8460–8470.

Vigneau, M., Beaucois, V., Hervé, P.Y., Jobard, G., Petit, L., Crivello, F., Mellet, E., Zago, L., Mazoyer, B., and Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to

phonological, lexico-semantic, and sentence processing?: insights from a meta-analysis. *Neuroimage* 54, 577–593.

Whitehead, J.C., and Armony, J.L. (2018). Singing in the brain: neural representation of music and voice as revealed by fMRI. *Hum. Brain Mapp.* 39, 4913–4924.

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., and Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage* 24, 1233–1241.

Witteman, J., Van Heuven, V.J., and Schiller, N.O. (2012). Hearing feelings: a quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia* 50, 2752–2763.

Wyczesany, M., Capotosto, P., Zappasodi, F., and Prete, G. (2018). Hemispheric asymmetries and emotions: evidence from effective connectivity. *Neuropsychologia* 121, 98–105.

Yao, Z., Yu, D., Wang, L., Zhu, X., Guo, J., and Wang, Z. (2016). Effects of valence and arousal on emotional word processing are modulated by concreteness: behavioral and ERP evidence from a lexical decision task. *Int. J. Psychophysiology* 110, 231–242.

Zhang, D., Zhou, Y., and Yuan, J. (2018a). Speech prosodies of different emotional categories activate different brain regions in adult cortex: an fNIRS study. *Sci. Rep.* 8, 218.

Zhang, M., Ge, Y., Kang, C., Guo, T., and Peng, D. (2018b). ERP evidence for the contribution of meaning complexity underlying emotional word processing. *J. Neurolinguist.* 45, 110–118.

Zinchenko, A., Kanske, P., Obermeier, C., Schröger, E., and Kotz, S.A. (2015). Emotion and goal-directed behavior: ERP evidence on cognitive and emotional conflict. *Social Cogn. Affect. Neurosci.* 10, 1577–1587.

Zinchenko, A., Obermeier, C., Kanske, P., Schröger, E., and Kotz, S.A. (2017). Positive emotion impedes emotional but not cognitive conflict processing. *Cogn. Affect. Behav. Neurosci.* 17, 665–677.

Zinchenko, A., Conci, M., Taylor, P., Müller, H.J., and Geyer, T. (2019). Taking attention out of context: frontopolar transcranial magnetic stimulation abolishes the formation of new context memories in visual search. *J. Cogn. Neurosci.* 31, 442–452.

iScience, Volume 23

Supplemental Information

ERP Markers of Valence Coding in Emotional Speech Processing

Alice Mado Proverbio, Sacha Santoni, and Roberta Adorni

Transparent Methods

Material and Methods

Participants

Twenty healthy right-handed university students (9 men and 11 women) participated in this study as unpaid volunteers. They earned academic credits for their participation. Their mean age was 22.75 years ($SD = 1.8$). All had normal vision and hearing and reported no history of neurological illness or drug abuse. Their right-handedness and right ocular dominance were assessed through the Italian version of the Edinburgh Handedness Inventory, a laterality preference questionnaire (group score = 0.87, $SD = 0.09$; min = -1, max = +1). All experiments were conducted with the understanding and written consent of each participant. No participant was excluded for technical reasons. The experimental protocol was approved by the Ethical Committee of the University of Milano-Bicocca (RM-2019-176).

Stimuli

Stimuli were 223 verbal expressions characterized by a positive (99) or negative semantic (99) and prosodic content, and 25 neutral sentences pronounced in a neutral emotional tone. Positive categories included: euphoria, surprise, encouragement, gratitude, appreciation, hope and relief. Negative categories included: self-pity, boredom, frustration, depression, despair, discouragement and reaction to physical pain. Stimulus sentences were pronounced by two professional speakers, one female and male. More precisely: 112 utterances, of which 50 were negative, 49 were positive and 13 neutral ones (targets) were spoken by a woman, while the other 111 utterances, of which 49 were negative, 50 positives and 12 neutral ones were spoken by a man. Speakers were asked to pronounce the sentences spontaneously but with feeling, identifying with the characters involved in the specific context, the way actors do.

A microphone (*Tonor brand*) and audio file editor software *Audacity* were used to record audio tracks. The files were all in WAV format. The two categories of stimuli, positive and negative, were balanced for number of words and letters, as proved by two one-way ANOVAs carried out across categories (number of words: $[F(1,98) = 0.71795, p = 0.39]$; number of letters $[F(1,98) = 1.3923, p = 0.24]$). The average number of words per sentence was 3.28, the average number of letters per sentence was 15.89.

Audio tracks lasted exactly 2200 ms and were normalized and levelled (frequency of 44100 Hz, resolution of 16 bits). Their acoustic intensity was measured through a PCE-999 sound level meter. Measures were 59.2 dB (SD= 3.8) for positive utterances and 56.5 dB (SD= 4.3) for negative utterances, which did not statistically differ, as proofed by means of a one-way ANOVA ($p= 0.99$). Auditory stimulation was pseudo-randomized and consisted in 6 runs of 37 or 38 stimuli. Each run included an equal number of stimuli pronounced by the male or female speaker, and of positive vs. negative stimuli. Furthermore, each block contained a minimum of 3 and maximum of 5 rare targets (sentences containing a proper name). The Inter Stimulus Interval (ISI) was 600 ms and randomly varied for additional 200 ms.

Stimulus validation

To assess the actual emotional valence of positive and negative stimuli, a much larger initial set of 450 stimuli (200 positive, 200 negative, 50 neutral) was administered to a group of native Italian Psychology graduated scholars ranging in age from 24 to 31 years ($M= 26.4$, $SD= 2.73$) including 3 females and 2 males. They were presented stimuli in binaural mode, randomly mixed and in short sequences, through of set of headphones. Subjects had to evaluate the emotional content of each audio track by means of a 3-point Likert scale, where 0 corresponded to negative content, 1 to neutral content and 2 to positive content. On the basis of the validation results 99 utterances obtaining a “2” evaluation from 100% of judges were selected as positive stimuli; likewise, 99 utterances obtaining a “0” evaluation from 100% of judges were selected as negative stimuli. A repeated measures ANOVA performed on the data showed a significant difference in the valence attributions across positive and negative classes [$F(1.98)= 45400$, $p= 0.0000$]. Further 25 male- and female-voiced stimuli that contained a proper name were selected as stimuli. See Table S1 for some examples of stimuli.

Procedure

Participants were *comfortably seated* inside an electrically shielded anechoic cabin (a Faraday cage), in front of a PC screen located outside the cabin, and visible through a mirror. Auditory stimuli were played *through headphones* (Sennheiser HD 202, Sennheiser electronic GmbH & Co. KG.). In order to avoid the occurrence of the occipito-parietal alpha rhythm, indexing the inactivity of the visual cortex and potentially contaminating the EEG signals as artifacts, a visual stimulus was also provided (the image of a domestic

environment, namely a living room, visible in Fig. S1) which was semantically compatible with the auditory context (i.e., the voice of a person who tells something, like in a conversation between friends). The image was printed on transparency and placed on the monitor placed in front of the subject at a distance of about 114 cm. The screen was connected to an IBM-PC computer, outside the cabin. Participants rested their head against the back of their armchair and had to stay relaxed and perfectly still (while fixating the foveal point).

They were also asked to avoid any body or eye movements. Each participant was provided with written experimental instructions. Prior to EEG recording, a training session occurred in which participants familiarized with setting and stimuli that were all in all similar to those used during the experiment. During the training session two 2 runs of 25 stimuli each were presented. The task consisted in responding, as quickly and as accurately as possible, by pressing a key on a joystick with the index finger if the sentence containing a person name. The task was not relevant for the ERP study, its actual purpose was only to keep the subject's attention over the auditory stimulation. The hand used was alternated during the recording session. Hand order and experimental runs were counterbalanced across subjects. At the beginning of each session, participants were reminded which response hand would be used. 6 runs of 25 stimuli each were administered.

All the sequences began with the warning stimuli "Attention", "Ready", "Go" (in Italian), presented acoustically and in sequence, and ended with the word "Thank You". At the end of the experimental session, the experimenter was available to offer explanations and to answer about the research's aim.

EEG recordings

The EEG was continuously recorded using the software package *EEvoke* (ANT, Enschede) from 128 scalp sites arranged according to the international 10–5 system (Oostenveld & Praamstra, 2001) at a sampling rate of 512 Hz. Horizontal and vertical eye movements were also recorded. Average mastoids served as the reference lead. The EEG and electrooculogram (EOG) were amplified with a half-amplitude bandpass of 0.016–100 Hz. A 50 Hz notch filter was also applied. Electrode impedance was maintained below 5 k Ω . EEG epochs were synchronized with stimulus onset. A computerized artifact rejection criterion was applied before averaging to discard epochs in which eye movements, blinks, excessive muscle potential or amplifier blocking occurred. The artifact rejection criterion was a peak-to-peak amplitude exceeding 50 μ V,

and the rejection rate was ~5%. ERPs were averaged off-line from -100 to 2 sec after stimulus onset and analyzed via software *EEProbe* (ANT, Enschede). ERP time epoch length was 2 seconds.

ERP components were measured with reference to the average baseline voltage over the interval of -100 to 0 ms and identified at the sites and latencies at which they reached their maximum amplitudes. Topographical voltage maps of the ERPs were made by plotting color-coded isopotentials obtained by interpolating voltage values between surface electrodes at specific latencies. Standardized low-resolution electromagnetic tomography (sLORETA; Pascual-Marqui, 2002) was performed on surface ERP voltages. LORETA was applied to the first observed modulation of ERPs, due to emotional valence, occurring between 350 and 550 ms (P450), as well as on the difference wave obtained by subtracting the largest from the smallest P450 wave, elicited by positive and negative stimuli, respectively. LORETA is a discrete linear solution to the inverse EEG problem, and it corresponds to the 3D distribution of neural electric activity that maximizes similarity (i.e. maximizes synchronization) in terms of orientation and strength between neighboring neuronal populations (represented by adjacent voxels). In this study, an improved version of standardized weighted low-resolution brain electromagnetic tomography was used; this version incorporates a singular value decomposition-based lead field weighting method (i.e. swLORETA, Palmero-Soler *et al.*, 2007). The source space properties included a grid spacing (the distance between two calculation points) of 5 points and an estimated signal-to-noise ratio of 3, which defines the regularization (higher values indicating less regularization and therefore less blurred results). The use of a value of three to four for the computation of SNR in the Tikhonov's regularization (1963) produces superior accuracy of the solutions for any inverse problems assessed. swLORETA was performed on the group data (grand-averaged data) to identify statistically significant electromagnetic dipoles ($p < 0.05$) in which larger magnitudes correlated with more significant activation. The data were automatically re-referenced to the average reference as part of the LORETA analysis. A realistic boundary element model (BEM) was derived from a T1-weighted 3D MRI dataset through segmentation of the brain tissue. This BEM model consisted of one homogeneous compartment comprising 3446 vertices and 6888 triangles. Advanced source analysis (ASA) employs a realistic head model of three layers (scalp, skull and brain) created using the BEM. This realistic head model comprises a set of irregularly shaped boundaries and the conductivity values for the compartments between them. Each boundary is approximated by a number of points, which are

interconnected by plane triangles. The triangulation leads to a more or less evenly distributed mesh of triangles as a function of the chosen grid value. A smaller value for the grid spacing results in finer meshes and vice versa. With the aforementioned realistic head model of three layers, the segmentation is assumed to include current generators of brain volume, including both gray and white matter. Scalp, skull, and brain region conductivities were assumed to be 0.33, 0.0042 and 0.33, respectively (Zanow and Knösche, 2004). The source reconstruction solutions were projected onto the 3D MRI of the Collins brain, which was provided by the Montreal Neurological Institute. The probabilities of source activation based on Fisher's F -test were provided for each independent EEG source, the values of which are indicated in a so-called 'unit' scale (the larger, the more significant). Both the segmentation and generation of the head model were performed using the ASA software program (*Advanced Neuro Technology, ANT*). ERP components were measured when and where they reached their maximum amplitude at scalp sites, and according to the previous literature on emotional prosody. The mean area amplitude of P450 component was measured at AF7-AF8, F7-F8 and FT7-FT8 electrode sites in the time window between 350 and 550 ms. The mean area amplitude of Anterior Negativity was measured at Fp1 Afp3h, Afz, Fpz, Afp4h and Fp2 electrode sites in the time window between 1100 and 1200 ms. This time window is uncommon but the latency of enhanced negativity to negative material has been shown to vary from 500-700 ms (Chen et al., 2008) to 1100-1200 (Proverbio and Piotti in press) depending on stimulus duration and complexity.

The mean area amplitude of Late Positivity was measured at fronto-central electrodes F1, F2, Fz, Fpz, in the time window between 1300 and 1500 post-stimulus ms. ERP data were subjected to multifactorial repeated-measures ANOVAs with three within-subjects factors: emotional valence (positive, negatives), electrode (dependent on the ERP component of interest) and hemisphere (left, right). Multiple comparisons among means were performed with Tukey's *post hoc* test. The effect size for the statistically significant factors was estimated using partial eta-squared (η_p^2).

The mean reaction time to rare target names was 1577.54 ms (after voice onset). Accuracy was very high: 97.64%

Figure S1. Related to Figure 1. Visual background, permanently present on the screen. The red central circle was the fixation point.

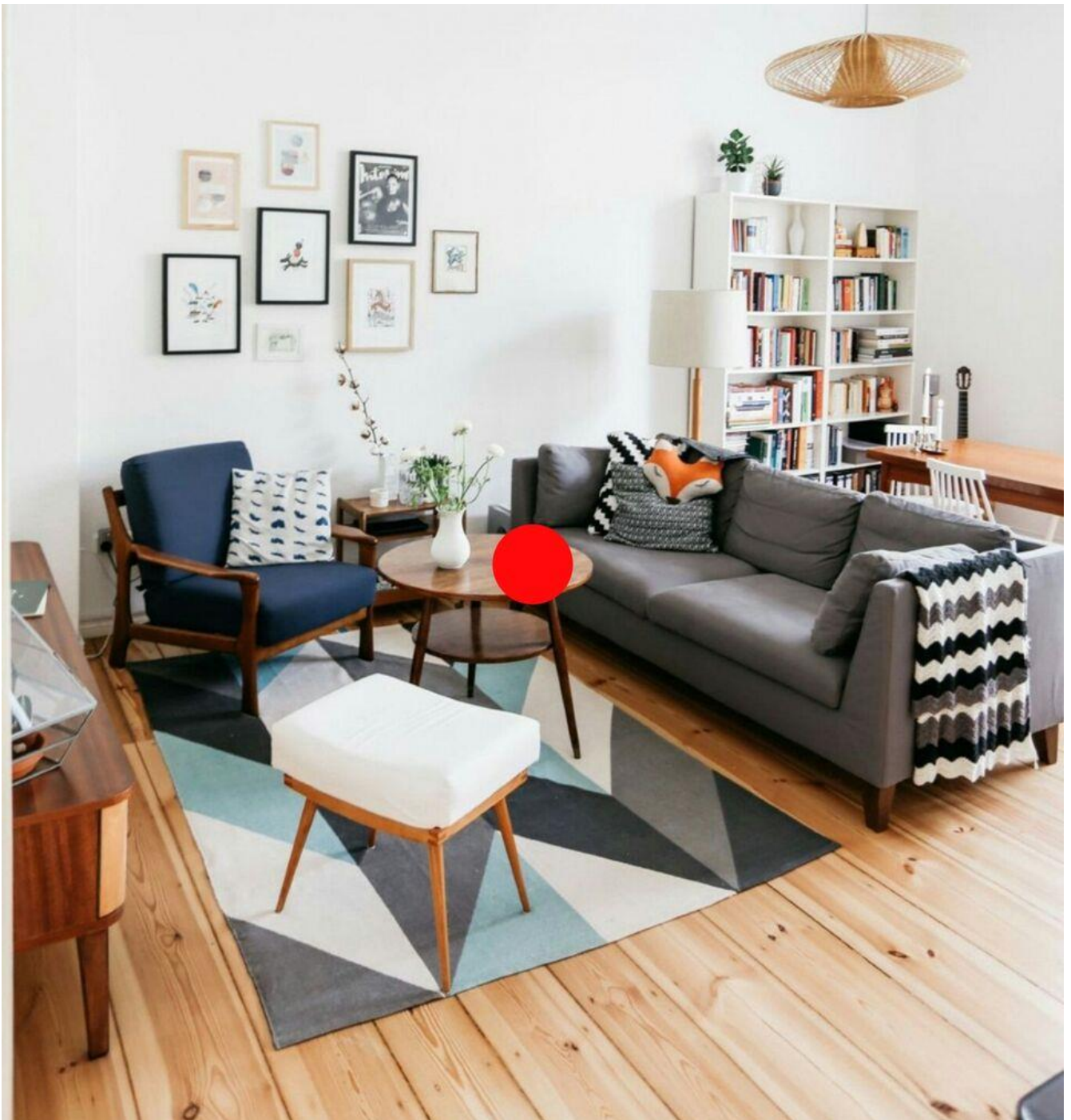


Table S1. Related to Figure 1. Some examples of stimuli belonging to the three stimulus classes. Emotionally neutral stimuli acted as targets. The original stimuli were in Italian language.

NEGATIVE	NEUTRAL	POSITIVE
Oh poor me!	Say hi to Ugo.	How nice to see you!
I am a failure...	I hear Carlo singing.	Amazing!
Aaahi, that's bad!	Now Bruno is better.	Excellent job!
It's all useless...	I am calling Eva now.	Victory!
I'm really screwed.	Paola and I are friends.	Aaah, what a relief...
I feel very sad.	Marta is arrived.	Seriously, you are the best.
There is nothing left to do...	I am having dinner with Lara tonight.	You are always the number one!
I'm not okay at all...	Here is Luca finally.	Hey, congratulations!