



The Language, Tone and Prosody of Emotions: Neural Substrates and Dynamics of Spoken-Word Emotion Perception

Einat Liebenthal^{1*}, David A. Silbersweig¹ and Emily Stern^{1,2}

¹ Department of Psychiatry, Brigham and Women's Hospital, Boston, MA, USA, ² Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA

OPEN ACCESS

Edited by:

Jonathan B. Fritz,
The University of Maryland, College
Park, USA

Reviewed by:

Dan Zhang,
Tsinghua University, China
Iain DeWitt,
National Institute of Deafness and
Communication Disorders, USA

*Correspondence:

Einat Liebenthal
eliebenthal@partners.org

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 26 January 2016

Accepted: 24 October 2016

Published: 08 November 2016

Citation:

Liebenthal E, Silbersweig DA and
Stern E (2016) The Language, Tone
and Prosody of Emotions: Neural
Substrates and Dynamics of
Spoken-Word Emotion Perception.
Front. Neurosci. 10:506.
doi: 10.3389/fnins.2016.00506

Rapid assessment of emotions is important for detecting and prioritizing salient input. Emotions are conveyed in spoken words via verbal and non-verbal channels that are mutually informative and unveil in parallel over time, but the neural dynamics and interactions of these processes are not well understood. In this paper, we review the literature on emotion perception in faces, written words, and voices, as a basis for understanding the functional organization of emotion perception in spoken words. The characteristics of visual and auditory routes to the amygdala—a subcortical center for emotion perception—are compared across these stimulus classes in terms of neural dynamics, hemispheric lateralization, and functionality. Converging results from neuroimaging, electrophysiological, and lesion studies suggest the existence of an afferent route to the amygdala and primary visual cortex for fast and subliminal processing of coarse emotional face cues. We suggest that a fast route to the amygdala may also function for brief non-verbal vocalizations (e.g., laugh, cry), in which emotional category is conveyed effectively by voice tone and intensity. However, emotional prosody which evolves on longer time scales and is conveyed by fine-grained spectral cues appears to be processed via a slower, indirect cortical route. For verbal emotional content, the bulk of current evidence, indicating predominant left lateralization of the amygdala response and timing of emotional effects attributable to speeded lexical access, is more consistent with an indirect cortical route to the amygdala. Top-down linguistic modulation may play an important role for prioritized perception of emotions in words. Understanding the neural dynamics and interactions of emotion and language perception is important for selecting potent stimuli and devising effective training and/or treatment approaches for the alleviation of emotional dysfunction across a range of neuropsychiatric states.

Keywords: emotions, semantics, amygdala, word processing, fMRI, ERPs (event-related potentials), speech perception, voice perception

INTRODUCTION

Spoken words naturally contain linguistic and paralinguistic elements that are both important and mutually informative for communication. The linguistic information consists of the literal, symbolic meaning of the word, whereas the paralinguistic information consists of the physical, contextual form of the word. For example, the meaning of the word “mad,” whether spoken in

the sense of “mentally disturbed,” “furious,” or “wildly excited,” can be disambiguated based on evaluation of contextual paralinguistic information such as the speaker’s current emotional status, as disclosed by their voice tone and facial expression. The linguistic and paralinguistic bits of information unveil in parallel as the spoken word unfolds over time. However, the neural dynamics of each process and the nature of neural interactions between linguistic and paralinguistic processes in spoken word perception are not well understood.

In this paper, we review the literature on perception of emotion in faces, written words, and voices, as a basis for understanding the neural architecture of emotion perception in spoken words. In particular, we critically consider evidence from animal, and human lesion and neuroimaging, studies for the existence of a fast route for emotion perception in spoken words that is analogous to the route described for facial expressions. We compare the characteristics of auditory and visual routes to the amygdala, in terms of neural dynamics, hemispheric lateralization, and functionality, across these stimulus classes. The comparison of neural substrates and neural dynamics of emotion perception across sensory modalities (auditory, visual) and stimulus types (non-verbal, verbal) informs the issue of whether certain aspects of the neural processing of emotions can be considered supramodal and universal, and therefore broadly applicable to linguistic input. We base the initial inquiry on the perception of emotion in faces because current neural models (Vuilleumier et al., 2003; Johnson, 2005) make detailed predictions regarding the neural underpinnings of fast and slow responses. We then consider intermediate stimuli that share additional characteristics with spoken words (specifically, written words are also linguistic, and nonverbal sounds are also auditory). We also draw a comparison between the spatial cues of visual stimuli and the temporal cues of auditory stimuli, which convey dominantly emotional paralinguistic or linguistic information depending on their frequency.

We discuss the neural underpinnings of emotion perception within the framework of a “valence-general” hypothesis, according to which the perception of both positive and negative valences is realized by flexible neuronal assemblies in limbic and paralimbic brain regions (Barrett and Bliss-Moreau, 2009; Lindquist et al., 2016). In this framework, arousal (i.e., the degree of emotional salience) and not valence (i.e., degree of positive or negative emotional association) is the dominant variable according to which the level of activation in different neuronal assemblies varies. Another point of emphasis is that language provides the context for experiencing and understanding emotions (and the world in general) (Barrett et al., 2007). Thus, our neural model (depicted schematically in **Figure 1**) presumes that the perception of emotional speech is a product of neural interactions between limbic and paralimbic emotional, cortical auditory and semantic, as well as frontal cognitive control areas. The amygdala is thought to play a central role at the intersection of these networks, as a fast salience detector alerting limbic, paralimbic, endocrine, and autonomic nervous systems to highly arousing stimuli. But the amygdala is also involved in slower evaluation of stimulus valence and arousal, interactively with associative cortical networks. Primary evidence for the existence

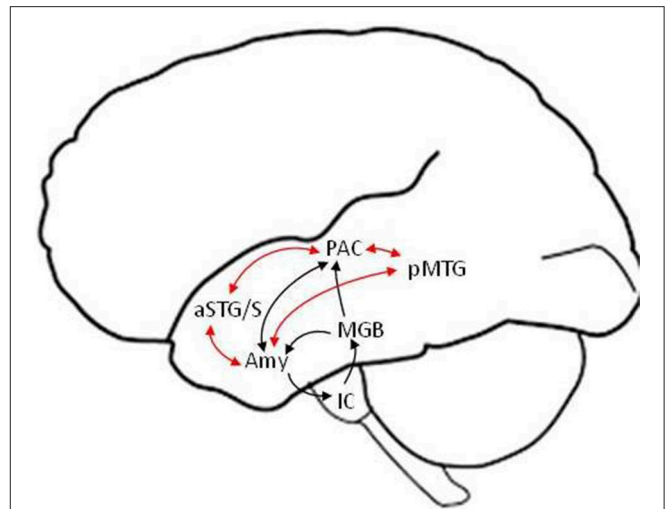


FIGURE 1 | Schematic model of putative fast (black arrows) and slow (red arrows) subcortical and temporal lobe pathways in the left hemisphere for the perception of emotional speech. The detection of basic emotional categories (e.g., joy, sadness) from brief and salient non-verbal utterances (e.g., laugh, cry) is suggested to be mediated by fast routes that bypass non-primary cortical areas and reach the amygdala within ~120 ms. The detailed evaluation of emotions based on the meaning and prosody of verbal utterances is suggested to involve slower efferent projections from non-primary auditory (e.g., aSTG/S voice) and language association (e.g., pMTG semantic) areas to the amygdala. Similar structural pathways are predicted in the right hemisphere (not shown), with differences in the strength of functional connections, including from higher-order association cortical areas, potentially accounting for differences in hemispheric lateralization between fast and slow pathways (see text for details). Efferent connections from cortical to subcortical areas (other than the amygdala) and output connections from the amygdala to paralimbic cortices and the hippocampus are not depicted, for clarity. IC, inferior colliculus; MGB, medial geniculate body; Amy, amygdala; PAC, primary auditory cortex; aSTG/S, anterior superior temporal gyrus and sulcus; pMTG, posterior middle temporal gyrus.

of fast direct, and slow indirect via non-primary cortical, routes for emotion perception comes from electrophysiological studies probing neural activity with high temporal resolution, as well as focal lesion studies of patients with focal subcortical or cortical lesions.

Face perception is thought to rely on both a fast neural pathway specialized for gross analysis of emotional expression, and a slower neural pathway for identity recognition and detailed evaluation of emotional expression. Evidence for different hemispheric lateralization of the amygdala is consistent with the possibility of a separation of neuroanatomical pathways for slow (left hemisphere dominance) and fast (right hemisphere dominance) processes underlying face perception (Morris et al., 1999; Wright et al., 2001). Electrophysiological studies of face perception suggest that early (<120 ms) differential responses to emotional expressions reflect the activity of fast direct routes, whereas later (>120 ms) differential responses to emotional expressions reflect the activity of indirect routes via non-primary visual cortex (Noesselt et al., 2002; Pourtois et al., 2004; West et al., 2011). This alleged division of labor for face emotion and

identity perception has been related to behavioral findings that low-spatial frequency global configurational cues are sufficient to convey coarse face emotional expressions, whereas high-spatial frequency fine-grained cues are needed to convey precise face identity features (Costen et al., 1996; Liu et al., 2000). Emotional aspects of stimuli are important for determining the level of significance and prioritizing time-sensitive salient input potentially critical for survival. Thus, an evolutionary advantage to faster processing of emotional input may have contributed to a differentiation of neural pathways for low and high spatial frequency cues. According to this theoretical framework, a direct pathway for extraction of low spatial frequencies evolved that can provide fast, subconscious appraisal of stimuli important for survival and for non-verbal communication (Vuilleumier et al., 2003; Johnson, 2005). Drawing on the findings in the visual modality, and recognizing that the timing of neural processing may depend on a variety of factors such as the sensory modality (e.g., basic auditory processing may be faster than basic visual processing), the stimulus complexity (e.g., linguistic stimuli may be processed more slowly than non-linguistic stimuli), and the stimulus category (e.g., some categories such as faces may confer special processing advantages), our working hypothesis is that neural processing within the first ~120 ms from stimulus presentation could be related to the activation of fast routes for prioritized processing.

Whether a processing advantage similar to that observed for emotional faces also extends to symbolic input such as written words, is controversial. While evidence exists for differential processing of emotional words early in the processing chain and subliminally (Gaillard et al., 2006; Kissler and Herbert, 2013), it remains unclear through what neural pathway/s. Compared to the perception of facial expressions which is acquired early in development and is perhaps even innate (Johnson, 2005), language comprehension is a learned skill that develops later. Thus, top-down modulation by semantic cortical networks and contextual learning have been suggested to play an important role in mediating prioritized emotional word perception (Barrett et al., 2007).

In the auditory system, the voice parallels the face in that it conveys a person's identity and current emotional status. Some aspects of voice emotions (in particular emotional category, e.g., anger, disgust, fear, sadness, joy) are thought to be perceived quickly based on coarse tone and intensity analysis of brief segments of familiar non-verbal vocalizations (e.g., shriek, cry, laugh etc...), and may be mediated by a fast direct route. However, other aspects of voice emotions (in particular emotional prosody), and identity recognition, may require sampling of longer voice segments and a more detailed spectral analysis thereof, and may involve slower routes via non-primary cortical areas.

The voice is also the natural carrier of speech. The voice paralinguistic and linguistic cues are separated such that the low-frequency band primarily carries prosodic cues important for communication of emotions, whereas the high-frequency band primarily carries phonemic cues critical for verbal communication (Remez et al., 1981; Scherer, 1986). Neural processing of the spectrally slow-varying emotional prosody cues

appears to involve more anterior auditory cortical areas in the superior temporal lobe than the processing of spectrally fast-varying phonemic cues (Belin et al., 2004; Liebenthal et al., 2005). Neural processing of emotional voice cues is also thought to involve auditory cortical areas predominantly in the right hemisphere, whereas that of phonemic cues predominantly auditory areas in the left hemisphere (Kotz et al., 2006; Scott and McGettigan, 2013). Which voice emotional cues confer a processing advantage, and through what neural routes, is an ongoing topic of investigation.

We conclude the paper with open questions that should be addressed in future research. In particular, the neural dynamics of direct and indirect routes for processing of different emotional cues require further study. For example, while both the amygdala and auditory cortices show sensitivity to various voice emotional cues, it remains unclear whether and under what circumstances, observed activation patterns are driven by the amygdala, are a result of cortical feedback connections to the amygdala, or both.

PERCEPTION OF FACE EMOTIONAL EXPRESSIONS

Various behavioral observations suggest that emotional stimuli are more likely to draw attention and be remembered than neutral stimuli, and that the emotional modulation of perception and memory is involuntary (Anderson, 2005; Phelps and LeDoux, 2005; Vuilleumier, 2005). For example, emotional faces are more readily detected than neutral faces in visual search (Eastwood et al., 2001; Fox, 2002) and spatial orienting (Pourtois et al., 2005) tasks. Face emotional expressions can be conveyed by coarse cues: low-spatial frequency cues (2–8 cycles/face) are important for processing visual input in the periphery, at a distance, or in motion (Livingstone and Hubel, 1988; Merigan and Maunsell, 1993), and may aid in the perception of threat. For example, the general outline of the eyes (e.g., degree of widening, coarse gaze direction) is visible at low spatial frequency, and can contribute to determining a person's emotional status (Whalen et al., 2004). Low frequency cues also carry crude facial information (face configuration, emotional expression), which can be perceived by newborn infants in the absence of a mature visual cortex (Johnson, 2005). This is in contrast to the high-spatial frequency cues (8–16 cycles/face) that are important for analysis of the visual shape and texture underlying accurate face identification (Fiorentini et al., 1983; Liu et al., 2000).

The degree of salience of emotional faces has been found to be positively related to level of activity in the amygdala and occipito-temporal visual cortex including the fusiform gyrus (Adolphs et al., 1998; Morris et al., 1998; Vuilleumier et al., 2001b; Pessoa et al., 2002). The amygdala is thought to play a key role in evaluating the significance and arousal associated with, and mediating automatic responses to, emotional stimuli (LeDoux, 2000; Sander et al., 2003a; Phelps and LeDoux, 2005), through its rich input and output connections to many subcortical and cortical regions (Amaral et al., 2003). Damage to the amygdala has been shown to eliminate the enhanced response in visual cortex for emotional faces (Vuilleumier et al., 2004), although

there are also contrary findings (Adolphs et al., 2005; Pessoa and Adolphs, 2010). Furthermore, emotional modulation of the neural response to faces arises very early, consistent with subcortical processing possibly at a subliminal level (Pourtois et al., 2005; Eimer and Holmes, 2007). To explain the automatic processing advantage of emotional faces, it has been proposed that coarse visual information relevant to emotional state (e.g., extent of eyes opening Whalen et al., 2004) is processed quickly via the amygdala and fed to visual cortex to enhance the processing of emotional information (Vuilleumier et al., 2003; Vuilleumier, 2005). Several lines of evidence, outlined below, support the existence of a direct subcortical pathway that mediates emotional influences on sensory processing (in parallel with the attentional modulation of sensory processing by frontal-parietal systems).

First, distinct patterns of spatial frequency sensitivity have been demonstrated in the fusiform cortex and the amygdala in several studies using functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) neuroimaging, consistent with the idea that distinct neural pathways operate on different subsets of cues available concurrently in face images (Vuilleumier et al., 2003; Winston et al., 2003). The fusiform cortex is more responsive to fine-grained high-spatial frequency information, whereas the amygdala is selectively modulated by coarse low-spatial frequency information. In extensive fusiform areas, neural adaptation was observed to repetition of low- after high- spatial frequency stimuli but not the reverse, suggesting that only the high frequency input established a long lasting representation in this area (Vuilleumier et al., 2003). The fusiform receives major inputs from parvocellular channels with fine resolution but slow processing (Livingstone and Hubel, 1988; Merigan and Maunsell, 1993), and this area is known to be important for fine visual shape and texture analysis and face recognition (Kanwisher et al., 1997; Vuilleumier et al., 2001a). In contrast, the amygdala receives major inputs from magnocellular channels with coarse resolution but fast processing, through a retinal—superior colliculus—pulvinar subcortical pathway (Schiller et al., 1979; Livingstone and Hubel, 1988; Merigan and Maunsell, 1993). This latter pathway is thought to bypass the slower cortical processing in the ventral visual pathway, and enable crude but fast processing of fear-related (LeDoux, 1996; Morris et al., 1999, de Gelder et al., 2003) and more generally emotion-related (Zald, 2003), aspects of visual input that determine stimulus salience.

A second line of evidence comes from electrophysiological (electro- and magneto- encephalography) studies that delineate the temporal course of neural processing of emotional input. An early differential event-related potential (ERP) response to emotional versus neutral faces (Eimer and Holmes, 2002; Eger et al., 2003; Streit et al., 2003; Pourtois et al., 2004, 2005; van Heijnsbergen et al., 2007; Rudrauf et al., 2008; Rotshtein et al., 2010) is observed around 120 ms latency, earlier than the ERP response associated with face recognition around 170–200 ms (Bentin et al., 1996). This differential ERP response to emotional faces is preserved even when the stimuli are filtered to include only low spatial frequencies (Pourtois et al., 2005). In the visual modality, the time around 120 ms from stimulus presentation

corresponds to the visual P1 response associated with pre-attentive perceptual processing (Di Russo et al., 2002, 2003; Liddell et al., 2004). The visual P1 is thought to be generated primarily in posterior occipito-temporal areas (Di Russo et al., 2002). However, amygdala responses to emotional stimuli demonstrated with intracranial recording within the same time range (Oya et al., 2002; Gothard et al., 2007), as well as findings of a diminished P1 response to emotional stimuli in patients with amygdala lesions (Rotshtein et al., 2010), are consistent with the possibility that neural generators in the amygdala also contribute to the P1 either directly or via modulation of cortical generators (Pourtois et al., 2013). In addition, earlier (<100 ms) responses to fearful versus happy faces have been recorded and localized to primary visual cortex (Noesselt et al., 2002; Pourtois et al., 2004; West et al., 2011), consistent with fMRI activation of this area by emotional faces (Vuilleumier et al., 2001b; Pessoa et al., 2002). The emotional enhancement of primary visual cortex at a time window preceding attentional enhancement is consistent with emotional modulation via a fast subcortical route. Overall, these findings are at the basis of models of visual perception placing the effects of prioritized processing of salient emotional material in the time range of about 120 ms from stimulus presentation (Vuilleumier et al., 2005).

Third, studies in patients with brain damage support the role of the amygdala in processing emotional cues in faces. Patients with extensive damage to the visual cortex resulting in hemispatial neglect, blindsight, or prosopagnosia, have been found to have residual ability for detection of faces and facial expressions (Morris et al., 2001; de Gelder et al., 2003; Pegna et al., 2005), suggesting that these functions can be accomplished subcortically. Patients with damage to the amygdala did not demonstrate the enhanced response in visual cortex for emotional faces (Vuilleumier et al., 2004). Furthermore, damage to the amygdala resulted in diminished early (100–150 ms) intracranial ERPs to fearful faces, consistent with a causal role for the amygdala in mediating the emotional enhancement of extrastriate visual cortex activity (Rotshtein et al., 2010). The amygdala response has also been shown to be modulated during subliminal processing of emotional faces (Whalen et al., 1998; Morris et al., 1999).

Finally, some evidence suggests that hemispheric lateralization may also differ between the slow cortical and fast subcortical face processing routes. In the amygdala, neural dynamics have been found to differ between the hemispheres such that the duration of response is shorter and the rate of adaptation is higher in the right compared to the left amygdala (Wright et al., 2001; Gläscher et al., 2004; Costafreda et al., 2008; Sergerie et al., 2008). Subliminal emotional stimuli activate predominantly the right amygdala (Morris et al., 1998; Gläscher and Adolphs, 2003; Pegna et al., 2005), whereas emotional information conveyed exclusively through language activates predominantly the left amygdala (Phelps et al., 2001; Olsson and Phelps, 2004). Taken together, these findings indicate the possibility that the right and left amygdala have somewhat different functions. Specifically, the right amygdala may play a primary role as an emotion detector, responding fast and at a subconscious level, possibly through the subcortical superior

colliculus-pulvinar-amygdala pathway (LeDoux et al., 1984; Morris et al., 1999). In contrast, the left amygdala may play a primary role in evaluating the significance of emotional stimuli, responding more slowly, possibly through an indirect, cortical route (Vuilleumier et al., 2003; Winston et al., 2003).

In summary, results from neuroimaging, electrophysiological, and lesion studies support the existence of a subcortical route for fast and subliminal processing of coarse emotional face cues. This route is thought to mediate the sensory processing, and attentional and memory enhancements observed for emotional faces. It is important to note however, that various findings in the literature are consistent with the existence of multiple parallel routes for visual emotional processing that may also contribute to rapid processing of salient information and may not involve the amygdala. Anatomically, “shortcut” connections within inferotemporal cortex and from the lateral geniculate nucleus to extrastriate visual cortex have been demonstrated in the monkey (Felleman and Van Essen, 1991) and could contribute to rapid visual processing. Pessoa and Adolphs (2010) use the finding of a patient with bilateral amygdala lesions who is able to normally process fearful faces (Adolphs et al., 2005; Tsuchiya et al., 2009) as evidence suggesting that the amygdala is not essential for exhibiting an emotional processing advantage. In a magnetoencephalography study, Rudrauf et al. (2008) demonstrate that the temporal course of processing arousing visual information is most accurately predicted by two-pathway models which include additional parallel shortcut pathways reaching the amygdala, temporal pole and orbitofrontal cortex more directly, either via cortical-cortical long-range fasciculi or via subcortical routes. While these studies do not negate the existence of a rapid amygdala route, they are consistent with the idea that there may be additional routes that mediate the prioritization of emotional stimuli.

PERCEPTION OF EMOTIONAL WRITTEN-WORDS

Compared to non-linguistic stimuli such as faces, words can convey emotional states with greater accuracy and finer nuances. The emotional content of written words can systematically and continuously be deconstructed along several primary dimensions, and in particular valence (degree of positive or negative emotional association) and arousal (degree of emotional salience), that are separable but interact (Bradley and Lang, 1999; Warriner et al., 2013). The question of whether an expedited subcortical route exists for visual processing of symbolic, detailed emotional input such as written words is contentious (Naccache and Dehaene, 2001; Gaillard et al., 2006). Semantic processing of words is associated with activity across extensive cortical networks (Binder and Desai, 2011), but it is unclear whether some level of analysis related to emotional content is accomplished subcortically. Observations that compared to neutral words, emotional words are more likely to be attended (Williams et al., 1996; Mogg et al., 1997; Anderson and Phelps, 2001), are better remembered (Kensinger and Corkin, 2004; Krolak-Salmon et al., 2004; Strange and Dolan, 2004; Vuilleumier

et al., 2004; Kissler et al., 2006), and are also more quickly detected in a lexical decision task (Kanske and Kotz, 2007; Kousta et al., 2009; Scott et al., 2009; Vigliocco et al., 2014), have led to the suggestion that analysis of some emotional linguistic content (in particular, salience and emotional category) could be facilitated at a subcortical level. Connections from the amygdala to visual cortex (Amaral et al., 2003) and to the orbitofrontal cortex (Timbie and Barbas, 2015) could mediate the enhanced cortical processing of emotional words detected subliminally in the amygdala.

Greater activation of the amygdala for negative and positive valenced words relative to neutral words has been demonstrated with fMRI and PET in normal control subjects (Isenberg et al., 1999; Hamann and Mao, 2002; Kensinger and Schacter, 2006; Goldstein et al., 2007; Weisholtz et al., 2015). The level of activity in the amygdala was found to vary mostly with the level of word arousal, whereas activity in the orbitofrontal and subgenual cingulate cortex varied mostly with word valence (Lewis et al., 2007; Posner et al., 2009; Colibazzi et al., 2010), consistent with the hypothesized role of the amygdala as an emotional salience detector.

However, in general, language stimuli are less likely to activate the amygdala, particularly in the right hemisphere (Anderson and Phelps, 2001; Phelps et al., 2001; Olsson and Phelps, 2004; Goldstein et al., 2007; Costafreda et al., 2008). The weak response of the amygdala to language has been related to its reduced involvement in language processing, or even its inhibition by prefrontal cortex (Bechara et al., 1995; Rosenkranz et al., 2003; Pezawas et al., 2005; Blair et al., 2007). Another contributing factor could be that the amygdala response to language is highly dependent on the subjective relevance of words, and is therefore difficult to reliably detect across a group of individuals. Support for a high sensitivity of the amygdala response to individual variation in word processing comes from studies of patients with anxiety disorders. For example, elevated left amygdala activation and abnormal patterns of sensitization and habituation were observed in post-traumatic stress disorder (PTSD) relative to normal control subjects for trauma-related negative, but not panic-related negative, versus neutral written words (Protopopescu et al., 2005). Sensitivity of the amygdala response to individual variation is also demonstrated by a dependence on moment-by-moment subjective evaluation of emotional intensity and subsequent memory of stimuli (Canli et al., 2000; Protopopescu et al., 2005). A left amygdala preference for language could be due to the general dominance of the left hemisphere for language. Increased activation of the left amygdala for language could reflect increased functional connectivity with highly left-lateralized, higher-order semantic memory networks distributed across the temporal, parietal and frontal cortex (Binder and Desai, 2011). Effects of word frequency have been reported in the left amygdala (Nakic et al., 2006), also consistent with a linguistic basis for the lateralization pattern in this area for words. However, whether there exists a right amygdala advantage in a fast subcortical afferent route for subliminal processing of salient emotional words remains an entirely open question.

In terms of temporal course, emotionally arousing (positive and negative) relative to neutral words have most commonly been found to elicit a differential ERP response around 180–300 ms (Kissler et al., 2006; Thomas et al., 2007; Herbert et al., 2008; Schacht and Sommer, 2009; Scott et al., 2009; Hinojosa et al., 2010; Citron et al., 2011). The timing of the differential response to emotional written words is consistent with the timing of lexical access to written words (Schendan et al., 1998; Cohen et al., 2000b; Grossi and Coch, 2005) localized to the fusiform gyrus (Kissler et al., 2007; Schacht and Sommer, 2009). Lexical access occurs earlier for emotional (~220–250 ms) versus neutral (~320 ms) words (Kissler and Herbert, 2013), consistent with the behavioral enhancement of emotional words in lexical decision tasks. Earlier (80–180 ms) effects of arousal have been reported for highly familiar emotional words (Ortigue et al., 2004; Hofmann et al., 2009; Scott et al., 2009), and in individuals with elevated anxiety (Pauli et al., 2005; Li et al., 2007; Sass et al., 2010). These early effects are thought to reflect enhanced orthographic processing (Hauk et al., 2006), speeded lexical access (Hofmann et al., 2009), and even rudimentary semantic analysis (Skrandies, 1998), of high-frequency emotional words. Repeated association (i.e., contextual learning) of the visual orthographic form of the word with its emotional meaning may facilitate the processing of high-frequency emotional written words (Fritsch and Kuchinke, 2013).

Taken together, these findings suggest that the role of the amygdala in detecting and prioritizing time-sensitive salient input extends to written words. The bulk of current evidence, indicating predominant left lateralization of the amygdala response to words, and timing of emotional word effects attributable to speeded lexical access in extrastriate cortex, appears more consistent with an indirect cortical route to the amygdala than a direct route akin to that described for emotional faces. Nevertheless, faster afferent access to the amygdala may exist for specific words that are highly-familiar and highly emotionally-salient. Because the emotional relevance of words likely varies widely between individuals, this may lead to mixed or weak findings within and across studies.

PERCEPTION OF EMOTIONAL NON-VERBAL VOCALIZATIONS

The voice is a particularly important medium for conveying emotional state because it is relatively independent of the listener's distance from, and ability to view, the speaker (unlike face cues). The acoustic cues conveying voice emotion—consisting of pitch (fundamental frequency), loudness (intensity), rhythm (duration of segments and pauses), and timbre (distribution of spectral energy) (Banse and Scherer, 1996; Grandjean et al., 2006)—are modulated by physiological factors (e.g., heart rate, blood flow, muscle tension) that vary as a function of a person's emotional state. Two main aspects of the voice are thought to convey emotional state on different time scales. The prosody of speech (discussed in the next section), consisting of pitch, loudness contour, and rhythm of speech

articulation, evolves relatively slowly over suprasegmental speech intonations (>200 ms). The quality of non-speech vocalization (discussed in this section), consisting of timbre and abrupt, aperiodic spectral changes, emerges more rapidly (Pell et al., 2015), and has been shown to convey certain emotional categories (e.g., fear, disgust) potently (Banse and Scherer, 1996; Scott et al., 1997). Similar to emotional faces, emotional voices appear to confer perceptual advantages, as evidenced by improved memory for emotional over neutral nonspeech vocalizations (Armony et al., 2007) and priming effects across non-verbal vocalizations and faces or words conveying the same emotional category (Carroll and Young, 2005).

Similar to the increased activity observed in visual occipito-temporal cortex for emotional faces, emotional non-verbal vocalizations (e.g., scream, cry, laugh) produce increased activity in the auditory superior temporal cortex and the amygdala (Phillips et al., 1998; Morris et al., 1999; Sander and Scheich, 2001; Fecteau et al., 2007), albeit with a variable level and lateralization pattern in the amygdala. The mixed amygdala response to emotional vocalizations could be related to variations in the subjective level of arousal elicited by vocal stimuli (Schirmer et al., 2008; Leitman et al., 2010). The amygdala may be particularly responsive to short, nonverbal emotional vocalizations (Sander et al., 2003b; Fecteau et al., 2007; Frühholz et al., 2014) because they tend to carry higher emotional weight and be more emotionally salient than speech prosody which evolves over a longer suprasegmental time scale. The amygdala may also be activated particularly during implicit processing of vocal emotions (Sander et al., 2005; Bach et al., 2008; Frühholz et al., 2012). Rising sound intensity has been proposed as an elementary auditory warning cue (Neuhoff, 1998), and has been demonstrated to activate the right amygdala more than a comparable decline in sound intensity (Bach et al., 2008). This finding is compatible with findings in the visual modality associating the amygdala with emotional intensity detection (Bonnet et al., 2015), and more generally with emotional relevance detection (Sander et al., 2003a).

In terms of neural temporal course, ERP studies show that emotional non-verbal vocalizations are distinguished from neutral vocalizations as early as 150 ms after sound onset (Sauter and Eimer, 2010). In the auditory modality, this timing corresponds to obligatory processing of acoustic cues (e.g., pitch, intensity) in auditory cortex (Vaughan and Ritter, 1970; Näätänen and Picton, 1987), and has been linked to subliminal emotional salience detection based on integration of acoustic cues signaling the emotional significance of a sound (Paulmann and Kotz, 2008). The timing of these voice emotional effects is similar to the emotional effects seen in face perception (~120 ms), and this raises the possibility that attentional modulation of emotional voices and faces is mediated by common supramodal neural routes (Sauter and Eimer, 2010). A few studies have also reported earlier (in the 100 ms range) effects of emotions on vocalization perception. Interactions between sensory modality (auditory, visual, audiovisual) and valence (fear, anger, neutral) were seen on the amplitude of the N100 ERP response (Jessen and

Kotz, 2011). Another study showed that affective (positive and negative) auditory conditioning modulated the magnetic ERP response to brief tones in the time range <100 ms, reflecting the activity of auditory sensory, frontal, and parietal cortex regions suggested to be part of an auditory attention network (Bröckelmann et al., 2011). Overall, these findings are consistent with the possibility of emotional enhancement of vocalization perception via rapid auditory pathways. Experimentally, early enhancement may be limited to conditions in which the input is very familiar (e.g., due to a small stimulus set, or conditioning).

Animal studies show that many neurons in the amygdala respond to broad-band sounds, with some neurons tuned to specific frequency bands, albeit not as narrowly as, and at a higher response threshold than, neurons in the tonotopically organized lemniscal pathway from medial geniculate body to auditory cortex (Bordi and LeDoux, 1992). A large proportion of amygdala neurons responding to sounds also exhibit high habituation rates (Bordi and LeDoux, 1992). The amygdala receives fast, direct auditory thalamic input from extralemniscal areas that are one synapse away from the amygdala and weakly encode sound spectral properties. The amygdala also receives slow, indirect auditory cortical input from association areas that are several synapses removed from the amygdala and encode more detailed acoustic patterns of sounds (Bordi and LeDoux, 1992; LeDoux, 2000). The direct thalamic pathway to the amygdala could be important for fast, subliminal detection and evaluation of emotional cues in short vocalizations based on coarse spectral properties (LeDoux, 2000; Frühholz et al., 2014). Indeed, a recent neuroimaging study in humans found that amygdala activation is sensitive to voice fundamental frequency and intensity variations relevant to emotional state in short nonword utterances (Frühholz et al., 2012). On the other hand, emotional prosody in longer speech segments may be evaluated on longer time scales (Pell and Kotz, 2011) via an indirect cortical route to the amygdala (Frühholz et al., 2014).

In summary, the comparatively small body of work investigating the neural basis of voice perception indicates the possibility of a fast route for prioritized perception of emotional non-verbal vocalizations. This route appears to be responsive particularly to brief vocalizations in which emotions are conveyed categorically by voice tone and intensity. However, the precise physical and perceptual attributes of vocalizations potentially processed via a direct route to the amygdala, and the degree of overlap with neural processing described for emotional faces, require further study.

PERCEPTION OF EMOTIONAL SPOKEN WORDS

Compared to written words, spoken words contain additional non-verbal emotional information (i.e., emotional prosody) that is physically and perceptually intertwined with the verbal information (Kotz and Paulmann, 2007; Pell and Kotz, 2011). The verbal and emotional cues in speech differ in their spectrotemporal properties. The phonemic cues consist primarily

of relatively fast spectral changes occurring within 50 ms speech segments, whereas the prosodic cues consist of slower spectral changes occurring over more than 200 ms speech segments (syllabic and suprasegmental range). Emotional speech confers processing advantages such as improved intelligibility in noise background as well as faster repetition time for words spoken with congruent emotional prosody (Nygaard and Queen, 2008; Gordon and Hibberts, 2011; Dupuis and Pichora-Fuller, 2014).

Similar to brief emotional non-verbal vocalizations, emotional prosody in speech and speech-like sounds produces increased activity in the auditory superior temporal cortex (Grandjean et al., 2005; Sander et al., 2005; Beaucousin et al., 2007; Ethofer et al., 2009) and less consistently, in the amygdala (Wildgruber et al., 2005; Wiethoff et al., 2008). The amygdala is more likely to be activated by concurrent and congruent face and voice emotional cues than by emotional voices alone (Ethofer et al., 2006; Kreifelts et al., 2010). Damage to the amygdala has also only inconsistently been associated with impaired perception of emotion in voices (Scott et al., 1997; Anderson and Phelps, 1998; Sprengelmeyer et al., 1999; Adolphs et al., 2005). A recent fMRI study showed that damage to the left, but not the right, amygdala resulted in reduced cortical processing of speech emotional prosody, suggesting that only the left amygdala plays a causal role in auditory cortex activation for this type of input (Frühholz et al., 2015). Given the association of the left amygdala with controlled, detailed evaluation of emotional stimuli including language (Phelps et al., 2001; Olsson and Phelps, 2004; Costafreda et al., 2008; Sergerie et al., 2008), this latter result is consistent with slower cortical processing of speech emotional prosody.

In terms of neural temporal course, the processing of emotional speech has been shown to diverge from that of neutral speech around 200 ms after word onset (Schirmer and Kotz, 2006; Paulmann and Kotz, 2008; Paulmann and Pell, 2010). This time range is similar to that described for emotional written words (Kissler et al., 2006; Schacht and Sommer, 2009; Scott et al., 2009; Hinojosa et al., 2010; Citron et al., 2011) and considered to reflect lexical processing in non-primary cortex (Schendan et al., 1998; Cohen et al., 2000a; Grossi and Coch, 2005). A differentiation between emotional categories (e.g., anger, disgust, fear, etc...) based on emotional prosody occurs later, around 300–400 ms (Paulmann and Pell, 2010), and with a different latency for different categories (Pell and Kotz, 2011).

Neurons across auditory cortical fields have differential spectrotemporal response properties that are consistent with the existence of separate processing streams for low- and high-spectral bands in complex sounds. In the core region of primate auditory cortex, neurons in anterior area R integrate over longer time windows than neurons in area A1 (Bendor and Wang, 2008; Scott et al., 2011), and neurons in the lateral belt have preferential tuning to sounds with wide spectral bandwidths compared to the more narrowly-tuned neurons in the core (Rauschecker et al., 1995; Rauschecker and Tian, 2004; Recanzone, 2008). Thus, a posterior-anterior auditory ventral stream from the core is thought to process sounds at increasing longer time scales, and a medial-lateral auditory ventral stream from the core processes sounds at increasing larger spectral bandwidth

(Rauschecker et al., 1995; Bendor and Wang, 2008; Rauschecker and Scott, 2009). Indeed, more anterior areas in the superior temporal cortex show sensitivity to increasingly longer chunks of speech (DeWitt and Rauschecker, 2012). Anterior and middle areas of the superior temporal gyrus and sulcus (STG/S) show sensitivity to voice prosody (Kotz et al., 2003; Belin et al., 2004; Humphries et al., 2014) and voice emotional cues (Grandjean et al., 2005; Schirmer and Kotz, 2006), which tend to be slow-varying. In contrast, the middle STG/S is thought to be specifically tuned to the faster spectral transitions relevant to phonemic perception (Liebenthal et al., 2005, 2010, 2014; Obleser et al., 2007; DeWitt and Rauschecker, 2012; Humphries et al., 2014), and more posterior areas in STG/S are important for phonological processing (Wise et al., 2001; Buchsbaum et al., 2005; Hickok and Poeppel, 2007; Chang et al., 2010; Liebenthal et al., 2010, 2013).

In addition to differences in spectrotemporal response properties within auditory cortex in each hemisphere there are differences between the two hemispheres. The right hemisphere has been suggested to be more sensitive to fine spectral details over relatively long time scales and the left hemisphere more sensitive to brief spectral changes (Zatorre and Belin, 2001; Boemio et al., 2005; Poeppel et al., 2008). A related theory is that resting state oscillatory properties of neurons predispose the left auditory cortex for processing at short time scales relevant to the rate of phonemes (gamma band) and the right auditory cortex for processing at longer time scales relevant to the rate of syllables (theta band) (Giraud et al., 2007; Giraud and Poeppel, 2012). Such differences in auditory cortex spectrotemporal sensitivity have been suggested as the basis for the common fMRI finding of right hemisphere dominance for emotional prosody perception, and left hemisphere dominance for speech comprehension (Mitchell et al., 2003; Grandjean et al., 2005). However, whether lateralization differences originate in auditory cortex or result from lateralized feedback connections to auditory cortex cannot be determined without examining the neural dynamics of the involved functional networks. A simultaneous fMRI/ERP study (Liebenthal et al., 2013) found that left lateralization during phonological processing of ambiguous speech syllables occurred early (120 ms) in inferior parietal and ventral central sulcus areas, and only later (380 ms) in the superior temporal gyrus, consistent with left lateralized feedback projections from articulatory somatomotor areas to auditory cortex. Attention to the spectral (non-linguistic) properties of the same sounds elicited early right lateralized activity in homologous parietal regions and bilateral activity in superior temporal gyrus. Drawing from these findings, we suggest that differences in the lateralization of auditory cortex responses to emotional prosody cues in speech could result from hemispheric differences in higher order areas feeding back to auditory cortex rather than from inherent hemispheric differences in auditory cortex spectrotemporal resolution.

In summary, while it appears that short, familiar non-verbal vocalizations can be processed quickly via a subcortical route similar to that described for emotional faces, there is no evidence to date that emotional prosody in speech confers the same advantage. Emotional prosody evolves on longer (suprasegmental) time scales and is conveyed by fine-grained

spectral cues. Emotional prosody input may therefore reach the amygdala primarily via slower, indirect routes from auditory and association cortices. There is also no evidence that the verbal information in speech can be processed via a fast route to the amygdala (except, as for written words, possibly for highly-familiar, frequent and salient spoken-words).

CONCLUSION

In conclusion, a review of current literature on emotion perception suggests that a fast route to the amygdala akin to that described for facial expressions may also function for other classes of non-linguistic stimuli such as brief emotional non-verbal vocalizations. Although, whether afferent access to the amygdala is specialized for visual biological input and in particular faces, and “borrowed” for auditory (and other sensory) input, is unclear. For language, current evidence points to the importance of lexico-semantic (non-sensory), and perhaps contextual top-down, processing for prioritized perception of emotions. However, fast afferent processing may apply under specific circumstances to highly-familiar and emotionally-salient words. One of the primary challenges in future work will be to determine the neural dynamics of amygdala activation for various types of inputs and under different conditions. Another challenge will be to determine the importance of different parallel routes for emotion processing. These issues should be addressed with methods that can provide both high temporal and high spatial resolution of neural activity in order to identify the time course of activation of the amygdala. Future work will also benefit from taking into account individual differences in the perception of stimulus valence, arousal, and familiarity. In particular, negative or inconclusive findings with respect to amygdala involvement in word and voice perception may in some cases be related to weaker emotional charge and personal relevance of the stimulus material. Future work on the neural dynamics of emotion perception will contribute to our understanding of anxiety and other emotional disorders, as well as help identify the neural circuits that should be targeted for effective training and rehabilitation of disorders which affect emotional function. An important issue will be to identify potent classes of stimuli for direct and indirect activation of the amygdala. In particular, indirect routes may be controlled by fronto-parietal executive circuits that are more amenable to training (for example, Kreifelts et al., 2013; Cohen et al., 2016), whereas direct routes may be less amenable to executive control and training. Another potential implication of research on subcortical routes for perception will be to understand the neural basis of hallucinations. In patients with schizophrenia and auditory or verbal hallucinations, subcortical structures including the thalamus and hippocampus have been proposed to play an important role in generating salient and emotionally charged sensations, whereas the cortical structures with which they are interconnected have been suggested to supply the detailed sensory content of hallucinations (Silbersweig et al., 1995; Silbersweig and Stern, 1998). Resolving the neural dynamics of emotion perception will contribute to our understanding of hallucinations and how they can be treated.

AUTHOR CONTRIBUTIONS

EL wrote the review, with substantial intellectual input from DS and ES.

REFERENCES

- Adolphs, R., Tranel, D., and Buchanan, T. W. (2005). Amygdala damage impairs emotional memory for gist but not details of complex stimuli. *Nat. Neurosci.* 8, 512–518. doi: 10.1038/nn1413
- Adolphs, R., Tranel, D., and Damasio, A. R. (1998). The human amygdala in social judgment. *Nature* 393, 470–474. doi: 10.1038/30982
- Amaral, D. G., Behniea, H., and Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience* 118, 1099–1120. doi: 10.1016/S0306-4522(02)01001-1
- Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *J. Exp. Psychol. Gen.* 134, 258–281. doi: 10.1037/0096-3445.134.2.258
- Anderson, A. K., and Phelps, E. A. (1998). Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport* 9, 3607–3613.
- Anderson, A. K., and Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411, 305–309. doi: 10.1038/35077083
- Armony, J. L., Chochol, C., Fecteau, S., and Belin, P. (2007). Laugh (or cry) and you will be remembered: influence of emotional expression on memory for vocalizations. *Psychol. Sci.* 18, 1027–1029. doi: 10.1111/j.1467-9280.2007.02019.x
- Bach, D. R., Schächinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., et al. (2008). Rising sound intensity: an intrinsic warning cue activating the amygdala. *Cereb. Cortex* 18, 145–150. doi: 10.1093/cercor/bhm040
- Banase, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636. doi: 10.1037/0022-3514.70.3.614
- Barrett, L. F., and Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Adv. Exp. Soc. Psychol.* 41, 167–218. doi: 10.1016/S0065-2601(08)00404-8
- Barrett, L. F., Lindquist, K. A., and Gendron, M. (2007). Language as context for the perception of emotion. *Trends Cogn. Sci.* 11, 327–332. doi: 10.1016/j.tics.2007.06.003
- Beaucousin, V., Lacheret, A., Turbelin, M. R., Morel, M., Mazoyer, B., and Tzourio-Mazoyer, N. (2007). fMRI study of emotional speech comprehension. *Cereb. Cortex* 17, 339–352. doi: 10.1093/cercor/bhj151
- Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., and Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 269, 1115–1118. doi: 10.1126/science.7652558
- Belin, P., Fecteau, S., and Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.* 8, 129–135. doi: 10.1016/j.tics.2004.01.008
- Bendor, D., and Wang, X. (2008). Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J. Neurophysiol.* 100, 888–906. doi: 10.1152/jn.00884.2007
- Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565. doi: 10.1162/jocn.1996.8.6.551
- Binder, J. R., and Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn. Sci.* 15, 527–536. doi: 10.1016/j.tics.2011.10.001
- Blair, K. S., Smith, B. W., Mitchell, D. G., Morton, J., Vythilingam, M., Pessoa, L., et al. (2007). Modulation of emotion by cognition and cognition by emotion. *Neuroimage* 35, 430–440. doi: 10.1016/j.neuroimage.2006.11.048
- Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8, 389–395. doi: 10.1038/nn1409
- Bonnet, L., Comte, A., Tatu, L., Millot, J. L., Moulin, T., and Medeiros de Bustos, E. (2015). The role of the amygdala in the perception of positive emotions: an “intensity detector”. *Front. Behav. Neurosci.* 9:178. doi: 10.3389/fnbeh.2015.00178

ACKNOWLEDGMENTS

The work was supported by Brain & Behavior Foundation grant 22249 and NIH R01 DC 006287.

- Bordi, F., and LeDoux, J. (1992). Sensory tuning beyond the sensory system: an initial analysis of auditory response properties of neurons in the lateral amygdaloid nucleus and overlying areas of the striatum. *J. Neurosci.* 12, 493–2503.
- Bradley, M. M., and Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*. Gainesville, FL: NIMH Center for Emotion and Attention, University of Florida.
- Bröckelmann, A. K., Steinberg, C., Elling, L., Zwanzger, P., Pantev, C., and Junghöfer, M. (2011). Emotion-associated tones attract enhanced attention at early auditory processing: magnetoencephalographic correlates. *J. Neurosci.* 31, 7801–7810. doi: 10.1523/JNEUROSCI.6236-10.2011
- Buchsbaum, B. R., Olsen, R. K., Koch, P., and Berman, K. F. (2005). Human dorsal and ventral auditory streams subservise rehearsal-based and echoic processes during verbal working memory. *Neuron* 48, 687–697. doi: 10.1016/j.neuron.2005.09.029
- Canli, T., Zhao, Z., Brewer, J., Gabrieli, J. D., and Cahill, L. (2000). Event-related activation in the human amygdala associated with later memory for individual emotional experience. *J. Neurosci.* 20:RC99.
- Carroll, N. C., and Young, A. W. (2005). Priming of emotion recognition. *Q. J. Exp. Psychol. A* 58, 1173–1197. doi: 10.1080/02724980443000539
- Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nat. Neurosci.* 13, 1428–1432. doi: 10.1038/nn.2641
- Citron, F. M., Oberecker, R., Friederici, A. D., and Mueller, J. L. (2011). Mass counts: ERP correlates of non-adjacent dependency learning under different exposure conditions. *Neurosci. Lett.* 487, 282–286. doi: 10.1016/j.neulet.2010.10.038
- Cohen, H., Benjamin, J., Geva, A. B., Matar, M. A., Kaplan, Z., and Kotler, M. (2000a). Autonomic dysregulation in panic disorder and in post-traumatic stress disorder: application of power spectrum analysis of heart rate variability at rest and in response to recollection of trauma or panic attacks. *Psychiatry Res.* 96, 1–13.
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000b). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123(Pt 2), 291–307.
- Cohen, N., Margulies, D. S., Ashkenazi, S., Schaefer, A., Taubert, M., Henik, A., et al. (2016). Using executive control training to suppress amygdala reactivity to aversive information. *Neuroimage* 125, 1022–1031. doi: 10.1016/j.neuroimage.2015.10.069
- Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., et al. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion* 10, 377–389. doi: 10.1037/a0018484
- Costafreda, S. G., Brammer, M. J., David, A. S., and Fu, C. H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Res. Rev.* 58, 57–70. doi: 10.1016/j.brainresrev.2007.10.012
- Costen, N. P., Parker, D. M., and Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Percept. Psychophys.* 58, 602–612. doi: 10.3758/BF03213093
- de Gelder, B., Frissen, I., Barton, J., and Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13105–13110. doi: 10.1073/pnas.1735530100
- DeWitt, I., and Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proc. Natl. Acad. Sci. U.S.A.* 109, E505–E514. doi: 10.1073/pnas.1113427109
- Di Russo, F., Martínez, A., and Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb. Cortex* 13, 486–499. doi: 10.1093/cercor/13.5.486

- Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., and Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Hum. Brain Mapp.* 15, 95–111. doi: 10.1002/hbm.10010
- Dupuis, K., and Pichora-Fuller, M. K. (2014). Intelligibility of emotional speech in younger and older adults. *Ear Hear.* 35, 695–707. doi: 10.1097/AUD.0000000000000082
- Eastwood, J. D., Smilek, D., and Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Percept. Psychophys.* 63, 1004–1013. doi: 10.3758/BF03194519
- Eger, E., Jedynak, A., Iwaki, T., and Skrandies, W. (2003). Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. *Neuropsychologia* 41, 808–817. doi: 10.1016/S0028-3932(02)00287-7
- Eimer, M., and Holmes, A. (2002). An ERP study on the time course of emotional face processing. *Neuroreport* 13, 427–431. doi: 10.1097/00001756-200203250-00013
- Eimer, M., and Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia* 45, 15–31. doi: 10.1016/j.neuropsychologia.2006.04.022
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum. Brain Mapp.* 27, 707–714. doi: 10.1002/hbm.20212
- Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., et al. (2009). Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *J. Cogn. Neurosci.* 21, 1255–1268. doi: 10.1162/jocn.2009.21099
- Fecteau, S., Belin, P., Joannette, Y., and Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *Neuroimage* 36, 480–487. doi: 10.1016/j.neuroimage.2007.02.043
- Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47. doi: 10.1093/cercor/1.1.1
- Fiorentini, A., Maffei, L., and Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception* 12, 195–201. doi: 10.1068/p120195
- Fox, E. (2002). Processing emotional facial expressions: the role of anxiety and awareness. *Cogn. Affect. Behav. Neurosci.* 2, 52–63. doi: 10.3758/CABN.2.1.52
- Fritsch, N., and Kuchinke, L. (2013). Acquired affective associations induce emotion effects in word recognition: an ERP study. *Brain Lang.* 124, 75–83. doi: 10.1016/j.bandl.2012.12.001
- Frühholz, S., Ceravolo, L., and Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cereb. Cortex* 22, 1107–1117. doi: 10.1093/cercor/bhr184
- Frühholz, S., Hofstetter, C., Cristinzio, C., Saj, A., Seeck, M., Vuilleumier, P., et al. (2015). Asymmetrical effects of unilateral right or left amygdala damage on auditory cortical processing of vocal emotions. *Proc. Natl. Acad. Sci. U.S.A.* 112, 1583–1588. doi: 10.1073/pnas.1411315112
- Frühholz, S., Trost, W., and Grandjean, D. (2014). The role of the medial temporal limbic system in processing emotions in voice and music. *Prog. Neurobiol.* 123, 1–17. doi: 10.1016/j.pneurobio.2014.09.003
- Gaillard, R., Del Cul, A., Naccache, L., Vinckier, F., Cohen, L., and Dehaene, S. (2006). Nonconscious semantic processing of emotional words modulates conscious access. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7524–7529. doi: 10.1073/pnas.0600584103
- Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127–1134. doi: 10.1016/j.neuron.2007.09.038
- Giraud, A. L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063
- Gläscher, J., and Adolphs, R. (2003). Processing of the arousal of subliminal and supraliminal emotional stimuli by the human amygdala. *J. Neurosci.* 23, 10274–10282.
- Gläscher, J., Tuschler, O., Weiller, C., and Buchel, C. (2004). Elevated responses to constant facial emotions in different faces in the human amygdala: an fMRI study of facial identity and expression. *BMC Neurosci.* 5:45. doi: 10.1186/1471-2202-5-45
- Goldstein, M., Brendel, G., Tuescher, O., Pan, H., Epstein, J., Beutel, M., et al. (2007). Neural substrates of the interaction of emotional stimulus processing and motor inhibitory control: an emotional linguistic go/no-go fMRI study. *Neuroimage* 36, 1026–1040. doi: 10.1016/j.neuroimage.2007.01.056
- Gordon, M. S., and Hibberts, M. (2011). Audiovisual speech from emotionally expressive and lateralized faces. *Q. J. Exp. Psychol.* 64, 730–750. doi: 10.1080/17470218.2010.516835
- Gothard, K. M., Battaglia, F. P., Erickson, C. A., Spitler, K. M., and Amaral, D. G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *J. Neurophysiol.* 97, 1671–1683. doi: 10.1152/jn.00714.2006
- Grandjean, D., Bänziger, T., and Scherer, K. R. (2006). Intonation as an interface between language and affect. *Prog. Brain Res.* 156, 235–247. doi: 10.1016/S0079-6123(06)56012-1
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* 8, 145–146. doi: 10.1038/nn1392
- Grossi, G., and Coch, D. (2005). Automatic word form processing in masked priming: an ERP study. *Psychophysiology* 42, 343–355. doi: 10.1111/j.1469-8986.2005.00286.x
- Hamann, S., and Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *Neuroreport* 13, 15–19. doi: 10.1097/00001756-200201210-00008
- Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., and Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage* 30, 1383–1400. doi: 10.1016/j.neuroimage.2005.11.048
- Herbert, C., Junghofer, M., and Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology* 45, 487–498. doi: 10.1111/j.1469-8986.2007.00638.x
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113
- Hinojosa, J. A., Méndez-Bértolo, C., Carretie, L., and Pozo, M. A. (2010). Emotion modulates language production during covert picture naming. *Neuropsychologia* 48, 1725–1734. doi: 10.1016/j.neuropsychologia.2010.02.020
- Hofmann, M. J., Kuchinke, L., Tamm, S., Vo, M. L., and Jacobs, A. M. (2009). Affective processing within 1/10th of a second: high arousal is necessary for early facilitative processing of negative but not positive words. *Cogn. Affect. Behav. Neurosci.* 9, 389–397. doi: 10.3758/9.4.389
- Humphries, C., Sabri, M., Lewis, K., and Liebenthal, E. (2014). Hierarchical organization of speech perception in human auditory cortex. *Front. Neurosci.* 8:406. doi: 10.3389/fnins.2014.00406
- Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malavade, K., Beattie, B., et al. (1999). Linguistic threat activates the human amygdala. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10456–10459. doi: 10.1073/pnas.96.18.10456
- Jessen, S., and Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage* 58, 665–674. doi: 10.1016/j.neuroimage.2011.06.035
- Johnson, M. H. (2005). Subcortical face processing. *Nat. Rev. Neurosci.* 6, 766–774. doi: 10.1038/nrn1766
- Kanske, P., and Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* 1148, 138–148. doi: 10.1016/j.brainres.2007.02.044
- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Kensinger, E. A., and Corkin, S. (2004). Two routes to emotional memory: distinct neural processes for valence and arousal. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3310–3315. doi: 10.1073/pnas.0306408101
- Kensinger, E. A., and Schacter, D. L. (2006). Processing emotional pictures and words: effects of valence and arousal. *Cogn. Affect. Behav. Neurosci.* 6, 110–126. doi: 10.3758/CABN.6.2.110
- Kissler, J., Assadollahi, R., and Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Prog. Brain Res.* 156, 147–183. doi: 10.1016/S0079-6123(06)56008-X
- Kissler, J., and Herbert, C. (2013). Emotion, Etmnooi, or Emittoon?—Faster lexical access to emotional than to neutral words during reading. *Biol. Psychol.* 92, 464–479. doi: 10.1016/j.biopsycho.2012.09.004
- Kissler, J., Herbert, C., Peyk, P., and Junghofer, M. (2007). Buzzwords: early cortical responses to emotional words during reading. *Psychol. Sci.* 18, 475–480. doi: 10.1111/j.1467-9280.2007.01924.x

- 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. doi: 10.1016/S0093-934X(02)00532-1
- Kotz, S. A., Meyer, M., and Paulmann, S. (2006). Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design. *Prog. Brain Res.* 156, 285–294. doi: 10.1016/S0079-6123(06)56015-7
- Kotz, S. A., and Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Res.* 1151, 107–118. doi: 10.1016/j.brainres.2007.03.015
- Kousta, S. T., Vinson, D. P., and Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 112, 473–481. doi: 10.1016/j.cognition.2009.06.007
- 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. doi: 10.1002/hbm.20913
- Kreifelts, B., Jacob, H., Brück, C., Erb, M., Ethofer, T., Wildgruber, D. (2013). Non-verbal emotion communication training induces specific changes in brain function and structure. *Front. Hum. Neurosci.* 7:648. doi: 10.3389/fnhum.2013.00648
- Krolak-Salmon, P., Hénaff, M. A., Vighetto, A., Bertrand, O., and Mauguière, F. (2004). Early amygdala reaction to fear spreading in occipital, temporal, and frontal cortex: a depth electrode ERP study in human. *Neuron* 42, 665–676. doi: 10.1016/S0896-6273(04)00264-8
- LeDoux, J. (1996). Emotional networks and motor control: a fearful view. *Prog. Brain Res.* 107, 437–446. doi: 10.1016/S0079-6123(08)61880-4
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184. doi: 10.1146/annurev.neuro.23.1.155
- LeDoux, J. E., Sakaguchi, A., and Reis, D. J. (1984). Subcortical efferent projections of the medial geniculate nucleus mediate emotional responses conditioned to acoustic stimuli. *J. Neurosci.* 4, 683–698.
- Leitman, D. I., Wolf, D. H., Ragland, J. D., Laukka, P., Loughhead, J., Valdez, J. N., et al. (2010). “It’s Not What You Say, But How You Say it”: A Reciprocal Temporo-frontal Network for Affective Prosody. *Front. Hum. Neurosci.* 4:19. doi: 10.3389/fnhum.2010.00019
- Lewis, P. A., Critchley, H. D., Rotshtein, P., and Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cereb. Cortex* 17, 742–748. doi: 10.1093/cercor/bhk024
- Li, W., Zinbarg, R. E., and Paller, K. A. (2007). Trait anxiety modulates supraliminal and subliminal threat: brain potential evidence for early and late processing influences. *Cogn. Affect. Behav. Neurosci.* 7, 25–36. doi: 10.3758/CABN.7.1.25
- Liddell, B. J., Williams, L. M., Rathjen, J., Shevrin, H., and Gordon, E. (2004). A temporal dissociation of subliminal versus supraliminal fear perception: an event-related potential study. *J. Cogn. Neurosci.* 16, 479–486. doi: 10.1162/089892904322926809
- Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., and Medler, D. A. (2005). Neural substrates of phonemic perception. *Cereb. Cortex* 15, 1621–1631. doi: 10.1093/cercor/bhi040
- Liebenthal, E., Desai, R., Ellingson, M. M., Ramachandran, B., Desai, A., and Binder, J. R. (2010). Specialization along the left superior temporal sulcus for auditory categorization. *Cereb. Cortex* 20, 2958–2970. doi: 10.1093/cercor/bhq045
- Liebenthal, E., Desai, R. H., Humphries, C., Sabri, M., and Desai, A. (2014). The functional organization of the left STS: a large scale meta-analysis of PET and fMRI studies of healthy adults. *Front. Neurosci.* 8:289. doi: 10.3389/fnins.2014.00289
- Liebenthal, E., Sabri, M., Beardsley, S. A., Mangalathu-Arumana, J., and Desai, A. (2013). Neural dynamics of phonological processing in the dorsal auditory stream. *J. Neurosci.* 33, 15414–15424. doi: 10.1523/JNEUROSCI.1511-13.2013
- Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., and Barrett, L. F. (2016). The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. *Cereb. Cortex* 26, 1910–1922. doi: 10.1093/cercor/bhv001
- Liu, C. H., Collin, C. A., Rainville, S. J., and Chaudhuri, A. (2000). The effects of spatial frequency overlap on face recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 956–979. doi: 10.1037/0096-1523.26.3.956
- Livingstone, M., and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240, 740–749. doi: 10.1126/science.3283936
- Merigan, W. H., and Maunsell, J. H. (1993). How parallel are the primate visual pathways? *Annu. Rev. Neurosci.* 16, 369–402.
- Mitchell, R. L., Elliott, R., Barry, M., Cruttenden, A., and Woodruff, P. W. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia* 41, 1410–1421. doi: 10.1016/S0028-3932(03)00017-4
- Mogg, K., Bradley, B. P., de Bono, J., and Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behav. Res. Ther.* 35, 297–303. doi: 10.1016/S0005-7967(96)00109-X
- Morris, J. S., DeGelder, B., Weiskrantz, L., and Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain* 124, 1241–1252. doi: 10.1093/brain/124.6.1241
- Morris, J. S., Ohman, A., and Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature* 393, 467–470. doi: 10.1038/30976
- Morris, J. S., Scott, S. K., and Dolan, R. J. (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia* 37, 1155–1163. doi: 10.1016/S0028-3932(99)00015-9
- Näätänen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x
- Naccache, L., and Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition* 80, 215–229. doi: 10.1016/S0010-0277(00)00139-6
- Nakic, M., Smith, B. W., Busis, S., Vythilingam, M., and Blair, R. J. (2006). The impact of affect and frequency on lexical decision: the role of the amygdala and inferior frontal cortex. *Neuroimage* 31, 1752–1761. doi: 10.1016/j.neuroimage.2006.02.022
- Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature* 395, 123–124. doi: 10.1038/25862
- Noesselt, T., Hillyard, S. A., Woldorff, M. G., Schoenfeld, A., Hagner, T., Jäncke, L., et al. (2002). Delayed striate cortical activation during spatial attention. *Neuron* 35, 575–587. doi: 10.1016/S0896-6273(02)00781-X
- Nygaard, L. C., and Queen, J. S. (2008). Communicating emotion: linking affective prosody and word meaning. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1017–1030. doi: 10.1037/0096-1523.34.4.1017
- Obleser, J., Zimmermann, J., Van Meter, J., and Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb. Cortex* 17, 2251–2257. doi: 10.1093/cercor/bhl133
- Olsson, A., and Phelps, E. A. (2004). Learned fear of “unseen” faces after Pavlovian, observational, and instructed fear. *Psychol. Sci.* 15, 822–828. doi: 10.1111/j.0956-7976.2004.00762.x
- Ortigue, S., Michel, C. M., Murray, M. M., Mohr, C., Carbonnel, S., and Landis, T. (2004). Electrical neuroimaging reveals early generator modulation to emotional words. *Neuroimage* 21, 1242–1251. doi: 10.1016/j.neuroimage.2003.11.007
- Oya, H., Kawasaki, H., Howard, M. A. III, and Adolphs, R. (2002). Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *J. Neurosci.* 22, 9502–9512.
- Pauli, P., Amrhein, C., Mühlberger, A., Dengler, W., and Wiedemann, G. (2005). Electrocorical evidence for an early abnormal processing of panic-related words in panic disorder patients. *Int. J. Psychophysiol.* 57, 33–41. doi: 10.1016/j.jpsycho.2005.01.009
- Paulmann, S., and Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport* 19, 209–213. doi: 10.1097/WNR.0b013e3282f454db
- 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. doi: 10.3758/CABN.10.2.230
- Pegna, A. J., Khateb, A., Lazeyras, F., and Seghier, M. L. (2005). Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nat. Neurosci.* 8, 24–25. doi: 10.1038/nn1364

- Pell, M. D., and Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS ONE* 6:e27256. doi: 10.1371/journal.pone.0027256
- Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., and Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biol. Psychol.* 111, 14–25. doi: 10.1016/j.biopsycho.2015.08.008
- Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nat. Rev. Neurosci.* 11, 773–783. doi: 10.1038/nrn2920
- Pessoa, L., McKenna, M., Gutierrez, E., and Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11458–11463. doi: 10.1073/pnas.172403899
- Pezawas, L., Meyer-Lindenberg, A., Drabant, E. M., Verchinski, B. A., Munoz, K. E., Kolachana, B. S., et al. (2005). 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nat. Neurosci.* 8, 828–834. doi: 10.1038/nn1463
- Phelps, E. A., and LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187. doi: 10.1016/j.neuron.2005.09.025
- Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., and Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nat. Neurosci.* 4, 437–441. doi: 10.1038/86110
- Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., et al. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proc. Biol. Sci.* 265, 1809–1817. doi: 10.1098/rspb.1998.0506
- Poeppl, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071–1086. doi: 10.1098/rstb.2007.2160
- Posner, J., Russell, J. A., Gerber, A., Gorman, D., Colibazzi, T., Yu, S., et al. (2009). The neurophysiological bases of emotion: an fMRI study of the affective circumplex using emotion-denoting words. *Hum. Brain Mapp.* 30, 883–895. doi: 10.1002/hbm.20553
- Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., and Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. *Hum. Brain Mapp.* 26, 65–79. doi: 10.1002/hbm.20130
- Pourtois, G., Grandjean, D., Sander, D., and Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cereb. Cortex* 14, 619–633. doi: 10.1093/cercor/bhh023
- Pourtois, G., Schettino, A., and Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biol. Psychol.* 92, 492–512. doi: 10.1016/j.biopsycho.2012.02.007
- Protopopescu, X., Pan, H., Tuescher, O., Cloitre, M., Goldstein, M., Engelien, W., et al. (2005). Differential time courses and specificity of amygdala activity in posttraumatic stress disorder subjects and normal control subjects. *Biol. Psychiatry* 57, 464–473. doi: 10.1016/j.biopsych.2004.12.026
- Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. doi: 10.1038/nn.2331
- Rauschecker, J. P., and Tian, B. (2004). Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *J. Neurophysiol.* 91, 2578–2589. doi: 10.1152/jn.00834.2003
- Rauschecker, J. P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114. doi: 10.1126/science.7701330
- Recanzone, G. H. (2008). Representation of con-specific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey. *J. Neurosci.* 28, 13184–13193. doi: 10.1523/JNEUROSCI.3619-08.2008
- Remez, R. E., Rubin, P. E., Pisoni, D. B., and Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science* 212, 947–949. doi: 10.1126/science.7233191
- Rosenkranz, J. A., Moore, H., and Grace, A. A. (2003). The prefrontal cortex regulates lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. *J. Neurosci.* 23, 11054–11064.
- Rotshtein, P., Richardson, M. P., Winston, J. S., Kiebel, S. J., Vuilleumier, P., Eimer, M., et al. (2010). Amygdala damage affects event-related potentials for fearful faces at specific time windows. *Hum. Brain Mapp.* 31, 1089–1105. doi: 10.1002/hbm.20921
- Rudrauf, D., David, O., Lachaux, J. P., Kovach, C. K., Martinerie, J., Renault, B., et al. (2008). Rapid interactions between the ventral visual stream and emotion-related structures rely on a two-pathway architecture. *J. Neurosci.* 28, 2793–2803. doi: 10.1523/JNEUROSCI.3476-07.2008
- Sander, D., Grafman, J., and Zalla, T. (2003a). The human amygdala: an evolved system for relevance detection. *Rev. Neurosci.* 14, 303–316.
- Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage* 28, 848–858. doi: 10.1016/j.neuroimage.2005.06.023
- Sander, K., Brechmann, A., and Scheich, H. (2003b). Audition of laughing and crying leads to right amygdala activation in a low-noise fMRI setting. *Brain Res. Brain Res. Protoc.* 11, 81–91.
- Sander, K., and Scheich, H. (2001). Auditory perception of laughing and crying activates human amygdala regardless of attentional state. *Brain Res. Cogn. Brain Res.* 12, 81–198. doi: 10.1016/S0926-6410(01)00045-3
- Sass, S. M., Heller, W., Stewart, J. L., Siltan, R. L., Edgar, J. C., Fisher, J. E., et al. (2010). Time course of attentional bias in anxiety: emotion and gender specificity. *Psychophysiology* 47, 247–259. doi: 10.1111/j.1469-8986.2009.00926.x
- Sauter, D. A., and Eimer, M. (2010). Rapid detection of emotion from human vocalizations. *J. Cogn. Neurosci.* 22, 474–481. doi: 10.1162/jocn.2009.21215
- Schacht, A., and Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cogn. Affect. Behav. Neurosci.* 9, 28–43. doi: 10.3758/CABN.9.1.28
- Schendan, H. E., Ganis, G., and Kutas, M. (1998). Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology* 35, 240–251. doi: 10.1111/1469-8986.3530240
- Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychol. Bull.* 99, 143–165. doi: 10.1037/0033-2909.99.2.143
- Schiller, P. H., Malpeli, J. G., and Schein, S. J. (1979). Composition of geniculostriate input of superior colliculus of the rhesus monkey. *J. Neurophysiol.* 42, 1124–1133.
- Schirmer, A., Escoffier, N., Zysset, S., Koester, D., Striano, T., and Friederici, A. D. (2008). When vocal processing gets emotional: on the role of social orientation in relevance detection by the human amygdala. *Neuroimage* 40, 1402–1410. doi: 10.1016/j.neuroimage.2008.01.018
- Schirmer, A., and Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10, 24–30. doi: 10.1016/j.tics.2005.11.009
- Scott, B. H., Malone, B. J., and Semple, M. N. (2011). Transformation of temporal processing across auditory cortex of awake macaques. *J. Neurophysiol.* 105, 712–730. doi: 10.1152/jn.01120.2009
- Scott, G. G., O'Donnell, P. J., Leuthold, H., and Sereno, S. C. (2009). Early emotion word processing: evidence from event-related potentials. *Biol. Psychol.* 80, 95–104. doi: 10.1016/j.biopsycho.2008.03.010
- Scott, S. K., and McGettigan, C. (2013). Do temporal processes underlie left hemisphere dominance in speech perception? *Brain Lang.* 127, 36–45. doi: 10.1016/j.bandl.2013.07.006
- Scott, S. K., Young, A. W., Calder, A. J., Hellawell, D. J., Aggleton, J. P., and Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature* 385, 254–257. doi: 10.1038/385254a0
- Sergerie, K., Chochol, C., and Armony, J. L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neurosci. Biobehav. Rev.* 32, 811–830. doi: 10.1016/j.neubiorev.2007.12.002
- Silbersweig, D. A., and Stern, E. (1998). Towards a functional neuroanatomy of conscious perception and its modulation by volition: implications of human auditory neuroimaging studies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1883–1888. doi: 10.1098/rstb.1998.0340
- Silbersweig, D. A., Stern, E., Frith, C., Cahill, C., Holmes, A., Grootoonk, S., et al. (1995). A functional neuroanatomy of hallucinations in schizophrenia. *Nature* 378, 176–179. doi: 10.1038/378176a0
- Skrandies, W. (1998). Evoked potential correlates of semantic meaning—A brain mapping study. *Brain Res. Cogn. Brain Res.* 6, 713–183. doi: 10.1016/S0926-6410(97)00033-5

- Sprengelmeyer, R., Young, A. W., Schroeder, U., Grossenbacher, P. G., Federlein, J., Büttner, T., et al. (1999). Knowing no fear. *Proc. Biol. Sci.* 266, 2451–2456. doi: 10.1098/rspb.1999.0945
- Strange, B. A., and Dolan, R. J. (2004). Beta-adrenergic modulation of emotional memory-evoked human amygdala and hippocampal responses. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11454–11458. doi: 10.1073/pnas.0404282101
- Streit, M., Dammers, J., Simsek-Kraues, S., Brinkmeyer, J., Wölwer, W., and Ioannides, A. (2003). Time course of regional brain activations during facial emotion recognition in humans. *Neurosci. Lett.* 342, 101–104. doi: 10.1016/S0304-3940(03)00274-X
- Thomas, S. J., Johnstone, S. J., and Gonsalvez, C. J. (2007). Event-related potentials during an emotional Stroop task. *Int. J. Psychophysiol.* 63, 221–231. doi: 10.1016/j.ijpsycho.2006.10.002
- Timbie, C., and Barbas, H. (2015). Pathways for emotions: specializations in the amygdalar, mediodorsal thalamic, and posterior orbitofrontal network. *J. Neurosci.* 35, 11976–11987. doi: 10.1523/JNEUROSCI.2157-15.2015
- Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., and Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nat. Neurosci.* 12, 1224–1225. doi: 10.1038/nn.2380
- van Heijnsbergen, C. C., Meeren, H. K., Grèzes, J., and de Gelder, B. (2007). Rapid detection of fear in body expressions, an ERP study. *Brain Res.* 1186, 233–241. doi: 10.1016/j.brainres.2007.09.093
- Vaughan, H. G. Jr., and Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalogr. Clin. Neurophysiol.* 28, 360–367. doi: 10.1016/0013-4694(70)90228-2
- Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777. doi: 10.1093/cercor/bht025
- Vuilleumier, P. (2005). Cognitive science: staring fear in the face. *Nature* 433, 22–23. doi: 10.1038/433022a
- Vuilleumier, P., Armony, J. L., Driver, J., and Dolan, R. J. (2001a). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 30, 829–841.
- Vuilleumier, P., Armony, J. L., Driver, J., and Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nat. Neurosci.* 6, 624–631. doi: 10.1038/nn1057
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., and Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat. Neurosci.* 7, 1271–1278. doi: 10.1038/nn1341
- Vuilleumier, P., Sagiv, N., Hazeltine, E., Poldrack, R. A., Swick, D., Rafal, R. D., et al. (2001b). Neural fate of seen and unseen faces in visuospatial neglect: a combined event-related functional MRI and event-related potential study. *Proc. Natl. Acad. Sci. U.S.A.* 98, 3495–3500. doi: 10.1073/pnas.051436898
- Vuilleumier, P., Schwartz, S., Duhoux, S., Dolan, R. J., and Driver, J. (2005). Selective attention modulates neural substrates of repetition priming and “implicit” visual memory: suppressions and enhancements revealed by fMRI. *J. Cogn. Neurosci.* 17, 1245–1260. doi: 10.1162/0898929055002409
- Warriner, A. B., Kuperman, V., and Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* 45, 1191–1207. doi: 10.3758/s13428-012-0314-x
- Weisholtz, D. S., Root, J. C., Butler, T., Tüscher, O., Epstein, J., Pan, H., et al. (2015). Beyond the amygdala: linguistic threat modulates peri-sylvian semantic access cortices. *Brain Lang.* 151, 12–22. doi: 10.1016/j.bandl.2015.10.004
- West, G. L., Anderson, A. A., Ferber, S., and Pratt, J. (2011). Electrophysiological evidence for biased competition in V1 for fear expressions. *J. Cogn. Neurosci.* 23, 3410–3418. doi: 10.1162/jocn.2011.21605
- Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., et al. (2004). Human amygdala responsivity to masked fearful eye whites. *Science* 306, 2061. doi: 10.1126/science.1103617
- Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., and Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J. Neurosci.* 18, 411–418.
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., et al. (2008). Cerebral processing of emotional prosody—influence of acoustic parameters and arousal. *Neuroimage* 39, 885–893. doi: 10.1016/j.neuroimage.2007.09.028
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., et al. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage* 24, 1233–1241. doi: 10.1016/j.neuroimage.2004.10.034
- Williams, J. M., Mathews, A., and MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychol. Bull.* 120, 3–24. doi: 10.1037/0033-2909.120.1.3
- Winston, J. S., O’Doherty, J., and Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *Neuroimage* 20, 84–97. doi: 10.1016/S1053-8119(03)00303-3
- Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., and Warburton, E. A. (2001). Separate neural subsystems within ‘Wernicke’s area’. *Brain* 124, 83–95. doi: 10.1093/brain/124.1.83
- Wright, C. I., Fischer, H., Whalen, P. J., McInerney, S. C., Shin, L. M., and Rauch, S. L. (2001). Differential prefrontal cortex and amygdala habituation to repeatedly presented emotional stimuli. *Neuroreport* 12, 379–383. doi: 10.1097/00001756-200102120-00039
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res. Brain Res. Rev.* 41, 88–123. doi: 10.1016/S0165-0173(02)00248-5
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953. doi: 10.1093/cercor/11.10.946

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Liebenthal, Silbersweig and Stern. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.