



The role of the medial temporal limbic system in processing emotions in voice and music



Sascha Frühholz^{a,b,1,*}, Wiebke Trost^{a,b,1}, Didier Grandjean^{a,b}

^aNeuroscience of Emotion and Affective Dynamics Lab, Department of Psychology, University of Geneva, Geneva, Switzerland

^bSwiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland

ARTICLE INFO

Article history:

Received 5 June 2014

Received in revised form 16 September 2014

Accepted 29 September 2014

Available online 5 October 2014

Keywords:

Amygdala
Hippocampus
Emotion
Music
Voice
Neuroimaging

ABSTRACT

Subcortical brain structures of the limbic system, such as the amygdala, are thought to decode the emotional value of sensory information. Recent neuroimaging studies, as well as lesion studies in patients, have shown that the amygdala is sensitive to emotions in voice and music. Similarly, the hippocampus, another part of the temporal limbic system (TLS), is responsive to vocal and musical emotions, but its specific roles in emotional processing from music and especially from voices have been largely neglected. Here we review recent research on vocal and musical emotions, and outline commonalities and differences in the neural processing of emotions in the TLS in terms of emotional valence, emotional intensity and arousal, as well as in terms of acoustic and structural features of voices and music. We summarize the findings in a neural framework including several subcortical and cortical functional pathways between the auditory system and the TLS. This framework proposes that some vocal expressions might already receive a fast emotional evaluation via a subcortical pathway to the amygdala, whereas cortical pathways to the TLS are thought to be equally used for vocal and musical emotions. While the amygdala might be specifically involved in a coarse decoding of the emotional value of voices and music, the hippocampus might process more complex vocal and musical emotions, and might have an important role especially for the decoding of musical emotions by providing memory-based and contextual associations.

© 2014 Elsevier Ltd. All rights reserved.

Contents

1. Introduction	2
2. Which emotions are expressed and perceived in voices and music and how?	3
3. The auditory system and the temporal limbic system (TLS)	4
4. Vocal and musical emotions activate the TLS.	6
4.1. TLS activity in response to positive and negative vocal emotions	6
4.1.1. Amygdala	6
4.1.2. Hippocampus	8
4.2. TLS activity in response to musical emotions	8
5. Affective arousal, emotional intensity and multimodal processing.	9
5.1. TLS sensitivity to arousing and intense vocal emotions	10
5.2. TLS sensitivity to arousing musical emotions	10

Abbreviations: ABR, auditory brainstem response; AC/ac, auditory cortex; acb, accessory basal nucleus of the amygdala; AMY/amy, amygdala; b, basal nucleus of the amygdala; bg, basal ganglia; BOLD, blood oxygenation level dependent; c, central nucleus of the amygdala; ca, cornu ammonis; CA1, cornu ammonis, subregion 1; CA3, cornu ammonis, subregion 3; cbl, cerebellum; CN, cochlear nucleus; CN/cn, cochlear nucleus; DG/dg, dentate gyrus; ERC, entorhinal cortex; erc, entorhinal cortex; FO, fundamental frequency; fMRI, functional magnetic resonance imaging; h, healthy controls; HC/hc, hippocampus; HG, Heschl's gyrus; IC/ic, inferior colliculus; l, lateral nucleus of the amygdala; m, medial nucleus of the amygdala; MGB/mgb, medial geniculate body; MNI, Montreal Neurological Institute; NAc, nucleus accumbens; p, patients; pc, patient controls; PET, positron emission tomography; PHC/phc, parahippocampal cortex; PPo, planum polare; PRC/prc, perirhinal cortex; PTE, planum temporale; SOC/soc, superior olivary complex; STC/stc, superior temporal cortex; STG, superior temporal gyrus; STS, superior temporal sulcus; sub, subiculum; TLS, temporal limbic system.

* Corresponding author at: Swiss Center for Affective Sciences, University of Geneva, 9 Chemin des Mines, PO Box 60, 1211 Geneva 20, Switzerland. Tel.: +41 22 379 09 15.

E-mail address: sascha.fruehholz@unige.ch (S. Frühholz).

¹ These authors are contributed equally to this review paper.

6.	Acoustic and structural features of vocal and musical emotions.	10
6.1.	TLS sensitivity to acoustic features of vocal emotions	10
6.2.	TLS sensitivity to acoustic features of music	11
6.3.	TLS sensitivity to structural features of music	11
7.	A neural framework for processing vocal and musical emotions in the TLS.	11
7.1.	Functional pathways for the decoding of vocal emotions.	12
7.2.	Functional pathways for the decoding of musical emotions.	12
8.	Conclusions	13
	Acknowledgements	14
	References	14

1. Introduction

In order to signal and communicate our emotional states and feelings to other social individuals, humans utilize basic as well as elaborate means for social communication in different sensory modalities. Amongst other sensory modalities, auditory signals are important means for emotional communication. The voice, for example, is a very basic channel for expressing emotions that humans also share with nonhuman primates (Belin et al., 2004, 2008; Romanski and Averbeck, 2009) and with more distant ancestors, such as song birds (Williams, 2004). The human voice is a powerful way of expressing emotions, both as nonverbal short bursts of affective exclamations, such as laughter, cries, or screams (Sauter et al., 2010; Scherer, 1994; Schrober, 2003; Szameitat et al., 2010; Wattendorf et al., 2013), and as modulation of the intonation of speech utterances. The latter is often referred to as emotional prosody (Banse and Scherer, 1996; Grandjean et al., 2006; Patel et al., 2011; Sundberg et al., 2011) and describes the suprasegmental modulations of speech intonations predominantly in terms of pitch and intensity variations, but also along other features of the voice quality (Frühholz and Grandjean, 2013c; Frühholz et al., 2014; Wiethoff et al., 2008).

Besides vocal expressions humans also have developed more elaborated means to express emotional states in the auditory modality. One such example is the musical expression of emotions (Juslin and Vastfjäll, 2008; Koelsch, 2010, 2014; Trost et al., 2012; Trost and Vuilleumier, 2013). Compared to vocal expressions, music and musically expressed emotions are a rather late achievement in human phylogenesis, which largely depended on the cultural development in human societies. However, music still has some long developmental trajectory dating back to the time of the *Homo neanderthalensis* (Conard et al., 2009). Over several centuries, musical techniques became very sophisticated, and developed into a powerful mean to express and communicate emotions, especially of complex emotional states (Trost et al., 2012; Zentner et al., 2008). It seems that many of the ways to express emotions in music are shared with vocal expressions (Juslin and Laukka, 2003; Weninger et al., 2013). Moreover, at the cognitive and the neural level, human individuals might use similar mechanisms and the same brain systems to decode emotions from voices and from music (Escoffier et al., 2012; Schirmer et al., 2012).

The human brain has developed specific mechanisms for the perception and the decoding of emotional signals. One of the most important brain regions for emotional processing is the limbic system, comprising several cortical areas, but mainly structures located in the medial temporal lobe and in the brainstem (Nieuwenhuys, 1996). The amygdala is thought to be located at the center of the limbic system and the emotional brain (LeDoux, 2000, 2012). Recent studies have shown that the amygdala is sensitive to emotional signals across many sensory modalities, including auditory information. Specifically, recent studies have shown that auditory stimuli such as *vocal expressions* (Bach et al., 2008a; Beaucois et al., 2007; Ethofer et al., 2009; Fecteau et al.,

2007; Frühholz et al., 2012; Frühholz and Grandjean, 2013a; Grandjean et al., 2005; Leitman et al., 2010b; Morris et al., 1999; Mothes-Lasch et al., 2011; Phillips et al., 1998; Quadflieg et al., 2008; Sander et al., 2003b, 2005, 2007; Sander and Scheich, 2005; Schirmer et al., 2008; Wiethoff et al., 2009) and *musically expressed emotions* (Alluri et al., 2012; Ball et al., 2007; Baumgartner et al., 2006; Blood and Zatorre, 2001; Brown et al., 2004; Chapin et al., 2010; Dyck et al., 2011; Eldar et al., 2007; Engel and Keller, 2011; Khalfa et al., 2008; Kleber et al., 2007; Koelsch et al., 2006, 2008, 2013; Lehne et al., 2014; Lerner et al., 2009; Mitterschiffthaler et al., 2007; Mueller et al., 2011; Mutschler et al., 2010; Pallesen et al., 2005, 2009; Park et al., 2013; Salimpoor et al., 2013) can elicit activity in the amygdala, indicating that both vocal emotions and musical emotions share similar ways of expression, and involve similar brain mechanisms for the decoding of and adaptive response to emotional cues.

Caudally to the amygdala lies the hippocampus. The hippocampus is also part of the limbic system (Nieuwenhuys, 1996) and is mainly assumed to have an important role in different cognitive functions especially related to episodic memory processes (Maguire, 2001) and for the formation of memory associations (Henke et al., 1997). Accordingly, the hippocampus has also been suggested to be important for the formation of emotional memories, especially in connection with activity in the amygdala (Phelps and LeDoux, 2005; Richardson et al., 2004). In this functional interaction, the amygdala is supposed to provide the affective evaluation of the stimulus and to contribute to the emotional reaction, whereas the hippocampus is thought to be responsible for memory encoding. But the hippocampus seems also to specifically influence emotional processing in the amygdala, and is important for emotional and social processing in general (Immordino-Yang and Singh, 2013).

This possible role of the hippocampus for emotional processing has been partly neglected, but recent studies on processing visual (Fusar-Poli et al., 2009) and auditory emotional stimuli, such as emotional vocalizations (Alba-Ferrara et al., 2011; Beaucois et al., 2007; Kotz et al., 2012; Leitman et al., 2010b; Mitchell et al., 2003; Phillips et al., 1998; Rota et al., 2011; Sander et al., 2005; Szameitat et al., 2010; Wiethoff et al., 2008) and musical emotions (Alluri et al., 2012; Baumgartner et al., 2006; Blood and Zatorre, 2001; Chapin et al., 2010; Dyck et al., 2011; Eldar et al., 2007; Engel and Keller, 2011; Khalfa et al., 2008; Kleber et al., 2007; Koelsch et al., 2006; Lerner et al., 2009; Mitterschiffthaler et al., 2007; Mueller et al., 2011; Pallesen et al., 2009; Plailly et al., 2007; Salimpoor et al., 2013; Trost et al., 2012; Watanabe et al., 2008) have reported an involvement of the hippocampus in these processes as well. For the latter it has been suggested that hippocampal activity during emotional processing might add episodic memories and contextual associations to the perception of musically expressed emotions of complex social meaning (Trost et al., 2012). This however might also be the case for the perception of vocally expressed emotions, which have a certain level of social importance and complexity (Alba-Ferrara et al., 2011), and sometimes depend on

the social (Szameitat et al., 2010) and the temporal context for recognition (Bestelmeyer et al., 2010).

Here we review the recent literature and the empirical studies, which investigated the neural basis of decoding emotional cues from voices and from music. We will mainly describe the neural effects of processing emotional cues in nonverbal vocal expressions, and from suprasegmental features of intonations superimposed on speech (which we will refer to as “voice” studies), and mainly in instrumental music (which we will refer to as studies in “music”). They all have in common to primarily use acoustic features as their main medium of expression. We therefore will not discuss the effects of emotional cues encoded in the semantics of speech (Kissler, 2013), and we will only partly discuss the effects of lyrics in music (Brattico et al., 2011).

We will specifically focus on the amygdala and the hippocampus as the most important brain structures in the temporal limbic system (TLS). We report results from brain imaging studies as well as from lesions studies with patients during the processing of emotional voices and music. Furthermore, within an anatomical framework, we propose a model that might explain the common, but also the divergent empirical findings of amygdala and hippocampus activity in response to vocal and musical emotions. We specifically try to explain findings in the temporal limbic system (TLS) according to its relation and connectivity to the ascending and descending auditory pathways of the subcortical and cortical auditory system. We will argue that there might be four different functional pathways, or neural loops, connecting the auditory system and the TLS. These pathways might be differentially used for the decoding of vocal and musical emotions, but also within the different types of vocalizations and music.

Before we turn to review the recent literature on the neural basis of processing emotions from the voice and from music, we first give a short definition and description of the different emotions encoded in voices and in music. We also give a short introduction in the human neural auditory system and its relationship to the limbic system.

2. Which emotions are expressed and perceived in voices and music and how?

Compared to neutral voices and music, emotions *expressed* in both modalities show a diversity of different emotional qualities, generally ranging from positive (e.g. happiness, positive surprise) and negative emotions (e.g. anger, fear, disgust) expressed in voices, and from pleasant and unpleasant emotions expressed in music. These emotions expressed in voices and music can also elicit similar emotions in the listener, and therefore are also – to a variable degree – *felt* by the listener. Beyond this major negative-positive or pleasant-unpleasant distinction, there are also more fine-grained emotional categories, which largely correspond to the main categories of the “basic emotions” (Ekman, 1992), but also go beyond this constrained set of categories, such as in the case of sexual arousal and pleasure (Everitt, 1990; Rauch et al., 1999).

Vocal emotions that show effects in the limbic system, comprise simple emotions, such as *anger* (e.g. Bach et al., 2008a; Beaucousin et al., 2007; Ethofer et al., 2009; Frühholz et al., 2012; Leitman et al., 2010b; Sander et al., 2005; Wiethoff et al., 2009), *fear* (e.g. Bach et al., 2008a; Fecteau et al., 2007; Leitman et al., 2010b; Morris et al., 1999; Phillips et al., 1998; Wiethoff et al., 2009), *disgust* (e.g. Kreifelts et al., 2010; Phillips et al., 1998), *happiness* (e.g. Fecteau et al., 2007; Kotz et al., 2012; Leitman et al., 2010b; Mitchell et al., 2003; Sander et al., 2007; Szameitat et al., 2010; Wiethoff et al., 2009), *sadness* (e.g. Kotz et al., 2012; Kreifelts et al., 2010; Mitchell et al., 2003; Morris et al., 1999; Sander et al., 2007), *surprise* (Fecteau et al., 2007; Kotz et al., 2012), or *sexual pleasure* (Fecteau et al., 2007; Kreifelts et al., 2010; Wiethoff et al., 2009). For

example nonverbal screams, laughs, and crying most likely represent vocalizations of fear, happiness, and sadness, respectively (Sauter et al., 2010; Scherer, 1994; Schrober, 2003; Szameitat et al., 2010; Wattendorf et al., 2013). Socially more complex expressions of emotions are more commonly superimposed on speech and are encoded in speech intonations (i.e. emotional prosody), and comprise, for example, emotions of *pride*, *guiltiness*, or *boredom* (Alba-Ferrara et al., 2011) as well as emotions of *irony* and *doubt* (Beaucousin et al., 2007).

Similar to vocal emotions, musical emotions could be classified as pleasant and unpleasant on a general level, and several studies have investigated the neural basis for this distinction (Blood and Zatorre, 2001; Blood et al., 1999; Dellacherie et al., 2009; Koelsch et al., 2006; Mueller et al., 2011; Pallesen et al., 2005; Wieser and Mazzola, 1986). Especially listening to highly pleasant music is supposed to lead to intensive feeling of “chills” (Blood and Zatorre, 2001; Panksepp, 1995; Panksepp and Bernatzky, 2002), which have the strongest effect with self-selected highly pleasurable music (Blood and Zatorre, 2001; Brattico et al., 2011; Salimpoor et al., 2011, 2013). Additional to the dimension of pleasantness, studies also investigated neural responses to music portraying categorical emotions of *happiness* or *joy* (Brattico et al., 2011; Eldar et al., 2007; Khalifa et al., 2008; Koelsch et al., 2013; Mitterschiffthaler et al., 2007), *fear* (Eldar et al., 2007; Gosselin et al., 2007, 2011; Koelsch et al., 2013; Park et al., 2013), *sadness* (Brattico et al., 2011; Khalifa et al., 2008; Mitterschiffthaler et al., 2007), or *tension* (Lehne et al., 2014). Musical tension is often induced by expectancy violations due to the mismatches of the temporal progression of music (James et al., 2008; Koelsch, 2014; Lehne et al., 2014).

Besides investigations on basic musical emotions, there have also been investigations on domain specific models, such as the Geneva Emotional Music Scale (GEMS) (Zentner et al., 2008). The GEMS model includes both more basic emotion categories such as *joy* or *sadness*, but also more complex musical emotions, such as *wonder*, *peacefulness*, *transcendence* or *nostalgia* (Trost et al., 2012; Zentner et al., 2008), with the intention to cover the full range of basic and complex emotions encoded in, decoded from music or felt during musical listening. Especially, complex musical emotions are discussed in relation to the conceptual distinction of *utilitarian* and *esthetic* emotions (Scherer, 2004). While the former are strongly related to adaptive behavior in terms of immediate emotional and bodily reactions to an environmental or social concern, esthetic emotions are appreciations or feelings related to the intrinsic qualities of stimuli that are unrelated to any immediate concerns (Scherer, 2004; Trost et al., 2012). Esthetic emotions usually have less well defined triggers, rarely elicit immediate goal-oriented responses, and the source of esthetic emotions is often encountered for the own sake of the consumer (Zentner, 2010). These characteristics are supposed to be the dominant mode while listening to music. However, esthetic appreciations of music, although unrelated to an object of concern, can still consequently lead to basic emotional reactions in the listener, which can be measured by different physiological and cognitive means (Juslin and Västfjäll, 2008), and which are accompanied by feelings of “liking–disliking” or “pleasantness–unpleasantness” as well as by neural and bodily responses (Gomez and Danuser, 2007; Grewe et al., 2007; Trost et al., 2012). Furthermore, stimulus appraisal processes that are supposed to be essential for affective states related to utilitarian emotions, such as novelty or expectancy violations (Scherer, 1984), seem also important in esthetic appreciations of music, such as in musical improvisations (Engel and Keller, 2011; Liu et al., 2012).

Unlike categorical models of emotions, dimensional models propose that all emotions can be represented by a small number of emotional dimensions. The most prominent dimensional model

represents emotions in a two dimensional space, which is defined by the dimension *valence* and *arousal*, ranging from negative to positive and from low to high arousal, respectively (Russell, 1980, 2003). This two-dimensional model has been extended by the *potency* dimension (Fontaine et al., 2007). These three dimensions, that are *valence*, *arousal*, and *potency*, represent the emotional quality, the level of excitement elicited by the stimulus, and the level of potential control over the stimulus, respectively. The potency dimension is specifically related to subjective feelings of power vs. weakness or dominance vs. submission related to the other individual from which we perceive the emotional signals (Fontaine et al., 2007). This dimensional emotion model has been successfully applied to vocal expressions of emotions (Goudbeek and Scherer, 2010) and to emotions encoded in music (Eerola and Vuoskoski, 2011), but musical emotions might even be of a higher-dimensional space (Trost et al., 2012; Zentner et al., 2008). Especially the arousal dimension of emotions is of high interest here, and it has been suggested that there might be two different arousal dimensions represented by “energy arousal” and “tension arousal” (Thayer, 1989). Instead of the emotional valence of stimuli, it has been suggested that especially the amygdala is more generally sensitive to the intensity or arousal level of emotional expressions irrespectively of the emotional valence (see Zald, 2003). The intensity or arousal could signify the general relevance of emotional stimuli to the listener (Sander et al., 2003a,b). This will be discussed in more detail in Section 5.

Besides the important question about *what* emotions are encoded in voices and music and probably felt by the listener, a second question concerns of *how* these emotions are encoded in the acoustic features of these emotional sounds. Voices and music consist of patterns of acoustic features, which actually help to decode their emotional meaning, especially in terms of their affective valence (Banse and Scherer, 1996; Hammerschmidt and Jürgens, 2007; Juslin and Laukka, 2003; Patel et al., 2011; Sauter et al., 2010; Weninger et al., 2013). There seem to be some common features for voices and music from which the emotional meaning can be decoded, respectively (Juslin and Laukka, 2003; Weninger et al., 2013). The spectral (along the frequency axis) and the temporal dimension (along the time axis; see Fig. 1) are the most common acoustic domains including many different features that

help to decode the emotional meaning. The most important perceived acoustical features defined by the spectral and temporal properties of sounds are the pitch, the loudness, the tempo and duration, the timbre as well as the temporal development of these features. For listeners, these perceived acoustic features provide a valid cue to discriminate and classify different emotions expressed in voices and in music (Banse and Scherer, 1996; Juslin and Laukka, 2003; Sauter et al., 2010), that is also reflected in corresponding brain activity (Frühholz et al., 2012; Leitman et al., 2010a,b, 2011; Wiethoff et al., 2008). How the human brain and especially how the auditory and the limbic system deal with these features will be discussed in detail in Section 6. We first turn to the general principles of acoustic decoding of external sounds in the human ascending auditory system.

3. The auditory system and the temporal limbic system (TLS)

The auditory signal of voices and music is represented by a complex sound wave (Fig. 1). This sound wave has some central features, such as the spectral pattern and sound intensity, which are perceived as pitch and loudness, respectively. Voices and music enter the human ear and trigger a cascade of signal transpositions along the ascending auditory pathway (Fig. 2). Incoming sound waves are transmitted by the outer and middle ear to the cochlea of the inner ear, which transposes them into neural signals. Downstream from an initial auditory processing in the cochlear nucleus (CN) and the superior olivary complex (SOC), the inferior colliculi (IC) are the first brain structures in the ascending pathway serving more complex acoustic roles, such as spectral and temporal decoding of acoustic signals (Baumann et al., 2011). Specifically, the IC is the first processing node, which extracts complex sound features from the incoming sensory information and serves multisensory integration (Aitkin et al., 1978). The IC might also be the source (Warrier et al., 2011) of a scalp recorded brainstem potential referred to as the auditory brainstem response (ABR) (Chandrasekaran and Kraus, 2010), which is a brain potential that locks in with essential temporal patterns of auditory stimuli. The next processing node is the medial geniculate body (MGB), located postero-medial to the thalamus, which responds to tones and to complex sounds. The

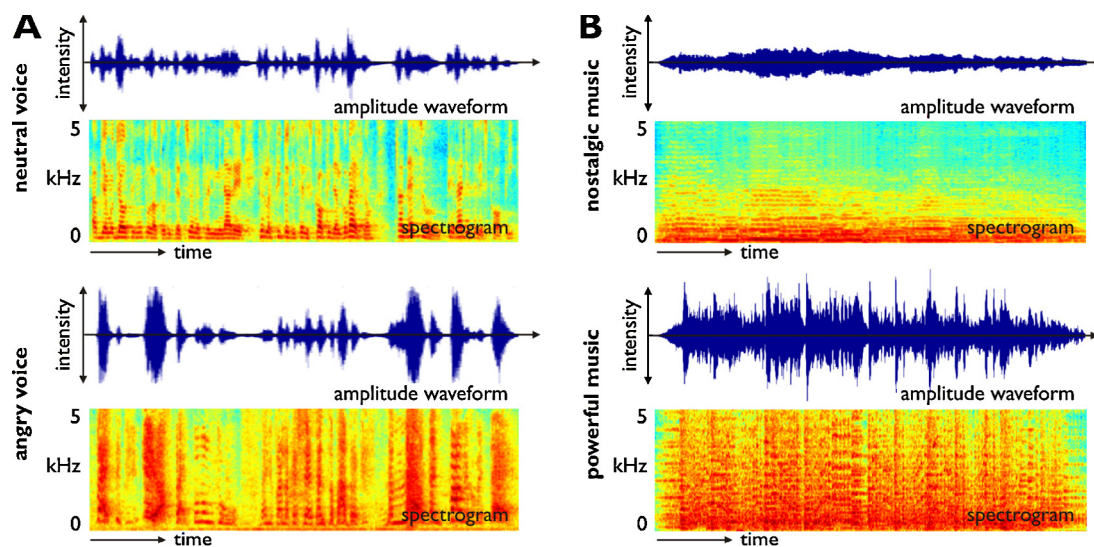


Fig. 1. Example stimuli for the expression of vocal and musical emotions. (A) Amplitude waveform and spectrogram of an angry voice (lower panel) and neutral voice (upper panel). Typical features of angry compared with neutral voices are increased intensity (i.e. amplitude) as well as increased mean and standard deviation of the pitch, which is mainly determined by the fundamental frequency (f_0). The pitch results from the specific spectral pattern at each time point. (B) Example amplitude waveform and spectrogram of a musical piece rated as expressing strong feelings of “nostalgia” (upper panel), which usually involve low level of musical intensity (i.e. amplitude). The lower panel shows the amplitude waveform and the spectrogram of a “powerful” music piece (lower panel) showing increased intensity, increased power in higher frequency bands, and increased temporal variability.

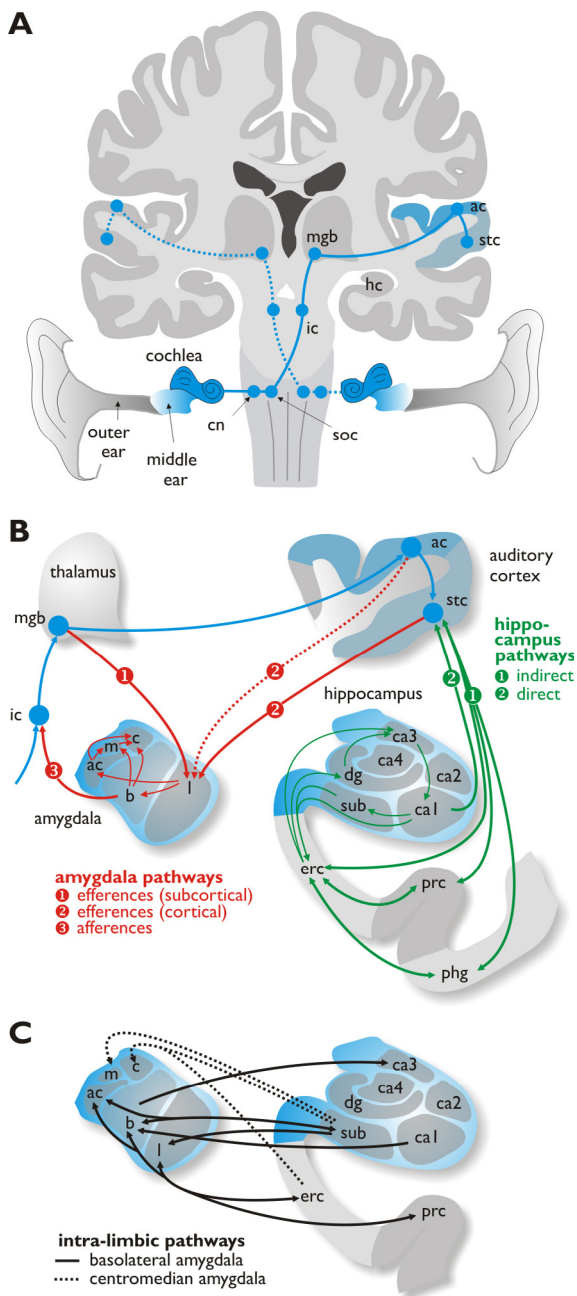


Fig. 2. (A) The neural auditory pathway consists of an interconnecting cascade of processing nodes from the cochlear nucleus (CN) up to primary auditory cortex (AC) and higher-level auditory regions in superior temporal cortex (STC). All pathways are shown here only for the right hemisphere for illustration purposes, but we assume an identical architecture in the left hemisphere. (B) Amygdala connections (thin lines indicate intra-amygdalar connections) with the auditory system are shown in red. The amygdala receives direct input from the MGB of the thalamus and from higher-level auditory cortex in STC, which both project to the lateral nucleus (l) of the basolateral complex of the amygdala. Tracing studies in animals also report a connection between AC and the amygdala (dashed line), but there is no evidence yet for this connection in humans. The basal nucleus (b) of the basolateral complex has efferent connection to the IC. The accessory nucleus (ac), the medial (m) and the central nucleus (c) are not directly connected to the auditory system. Hippocampal connections (thin lines indicate intra-hippocampal connections) to the auditory system are shown in green. The hippocampus shows direct and indirect connections to the auditory cortex. A direct connection exists from the CA1 region to higher-level auditory cortex. Indirect connections mainly provide input to the hippocampal formation by connections from the STC to the parahippocampal gyrus (phg), to the perirhinal cortex (prc) and the entorhinal cortex (erc), which figure as input relays to the hippocampus. (C) Intra-limbic pathway between the amygdala and the hippocampal formation. The basolateral complex of the amygdala (l, b, ac) shows bidirectional connections to the hippocampal formation, except for unidirectional connections from the subiculum (sub) to

MGB already shows less tonotopic organization (i.e. sensitivity to specific spectral frequencies) compared to previous processing nodes, at least in some of its subnuclei (Jones, 2007). The MGB is sensitive to relatively fast modulation rates in sounds up to 128 Hz (Preuss and Muller-Preuss, 1990). Activity in the MGB can be top-down modulated during speech processing indicating that it already processes complex sound information (Von Kriegstein et al., 2008). The MGB is the last subcortical node in the ascending pathway, projecting directly to the auditory cortex (AC).

The AC in humans is located in the posterior part of the superior temporal lobe covering Heschl's gyrus (HG), parts of the planum temporal (PTe) and parts of the posterior superior temporal gyrus (STG) (Hackett, 2011). The primary AC is largely located on the medial two-thirds of the HG, but the lateral boundaries are less well defined (Rivier and Clarke, 1997). The primary AC usually shows tonotopic mapping (Da Costa et al., 2011), such that different subregions of AC are sensitive to different frequency bands in complex sound waves. There is also a region in antero-lateral primary AC, which seems specifically sensitive to sound pitch (Bendor and Wang, 2005). Besides the sensitivity to spectral sound information, the sensitivity to temporal modulation rates of sounds is far less in the AC (<20 Hz) compared to subcortical structures (Malone et al., 2007). Nonprimary AC also seems to show lateralized sensitivity to different temporal modulation rates (Liem et al., 2014) following the suggestion of asymmetric temporal sampling in both hemispheres (Poeppel, 2003), which a left hemispheric sensitivity to faster modulations and a right hemispheric sensitivity to slower temporal modulations. After this decoding in primary AC, auditory information is fed forward to secondary AC (i.e. the planum temporale (PTe) and the planum polare (PPo)) and to higher-level associative cortex in the superior temporal cortex (STC) (Frühholz and Grandjean, 2013b) consisting of the STG and superior temporal sulcus (STS), which serve more complex sound processing including auditory object such as a voice as well as music and musical structure processing. For example, several regions along the STG are more sensitive to human voices compared to other environmental sounds (Belin and Zatorre, 2000). Similarly, regions in STG are also more sensitive when listening to music as compared to other sounds (Trost et al., 2012), and these regions for voice and music processing seem to overlap in STG (Schirmer et al., 2012).

The auditory system is directly and indirectly connected to the TLS with different connection patterns to the amygdala and the hippocampus. The amygdala is efferently and afferently connected to this ascending auditory system (see Fig. 2B). On the most general level the amygdala can be subdivided into a basolateral complex, a centromedial complex, and a superficial complex (Amunts et al., 2005; Frühholz and Grandjean, 2013a). The MGB has direct efferent connections (i.e. the direct route) to the lateral part of the basolateral complex, which receives also afferences from regions in higher-level STC (i.e. the indirect route) (LeDoux, 2007), probably also from the AC as found in rodents (Reser et al., 2009). The basal part of the basolateral complex, in turn, projects to the IC, thus possibly influencing early levels of auditory processing.

The amygdala is also heavily connected with the hippocampus (HC) and the hippocampal formation as found in rats (Pitkanen

basal nucleus, from CA1 to the lateral nucleus as well as from the basal nucleus to the CA3 region. Unidirectional connections exist from the subiculum and the CA1 regions to the centromedial amygdala. Abbreviations: ac auditory cortex; acb accessory basal nucleus of the amygdala; amy amygdala; b basal nucleus of the amygdala; bg basal ganglia; c central nucleus of the amygdala; ca cornu armonis; cbl cerebellum; cn cochlear nucleus; dg dentate gyrus; erc entorhinal cortex; hc hippocampus; ic inferior colliculus; l lateral nucleus of the amygdala; m medial nucleus of the amygdala; mgb medial geniculate body; phg parahippocampal gyrus; prc perirhinal cortex; soc superior olivary complex; stc superior temporal cortex; sub subiculum.

et al., 2000). The hippocampus can be divided into a ventral part located caudally to the amygdala, and a dorsal part located more posteriorly (Fanselow and Dong, 2010). These subregions seem to serve emotional and more cognitive processing in the ventral and the dorsal part, respectively. The hippocampus itself is surrounded by the parahippocampal formation, which mainly provides sensory input to the TLS as found in monkeys (Mohedano-Moriano et al., 2007). The hippocampus has also connections with the auditory system, but this connectivity differs considerably from the connectivity of the amygdala with the auditory system (Fig. 2B). First, the hippocampus shows mainly indirect connections to the auditory system (i.e. indirect route), which are mediated by the parahippocampal gyrus (PHG), the perirhinal cortex (PRC) and the entorhinal cortex (ERC). The PHG, the PRC, and the ERC are the main input relays to the hippocampus. A direct efferent connection might exist between the CA1 regions of the hippocampus and the auditory cortex as found in rats (Cenquizca and Swanson, 2007). Second, the hippocampus is only connected to the cortical regions of the auditory system, and shows no connectivity to subcortical auditory brain regions (Mohedano-Moriano et al., 2007).

The amygdala and the hippocampus show manifold interconnections (Pitkanen et al., 2000) (Fig. 2C). The main processing cascade in the hippocampus comprises neuronal input provided by the ERC and the PRC to the dentate gyrus (DG) and the CA3 region, which feed the signal forward to the CA1 region and to the subiculum. The subiculum is the major output node of the hippocampus. Afferent connections from the amygdala terminate at early levels of this cascade, probably influencing how signals are fed forward to the hippocampus by the ERC and the PRC, as well as affecting early processing in the CA3 region. After signal processing in the hippocampus, this information might be transferred to the basolateral and the centromedial complex of the amygdala via the CA1 region and the subiculum (Fig. 2C).

Thus the amygdala is closely connected with subcortical and cortical regions of the auditory system. The hippocampus is closely connected with the cortical auditory system and with the amygdala. This interconnection puts the amygdala and the hippocampus in a suitable position within the auditory network to extract the emotional value from incoming sensory information and to prepare adaptive responses.

4. Vocal and musical emotions activate the TLS

As mentioned above, different kinds of emotions can be expressed in voices and in music. These emotions range from basic emotions, such as emotions of fear, happiness, anger, disgust, surprise, sadness, up to more complex emotions. Complex emotions are, for example, vocal expressions of pride or guilt (Alba-Ferrara et al., 2011), or musically expressed emotions of transcendence or nostalgia (Troost et al., 2012; Zentner et al., 2008). The amygdala and the hippocampus are thought to receive coarse as well as more complex acoustic and emotional information along a direct and an indirect route (Jarrell et al., 1987; Phelps and LeDoux, 2005), respectively. From this information they might decode the emotional value of these different kinds of emotional expression.

4.1. TLS activity in response to positive and negative vocal emotions

4.1.1. Amygdala

Most of the recent neuroimaging studies reported amygdala activity in response to vocal emotions (Fig. 3; see also Table 1). However, results from early lesion studies provided quite inconsistent evidence (Adolphs and Tranel, 1999; Anderson and Phelps, 1998; Bach et al., 2013; Scott et al., 1997; Sprengelmeyer et al., 1999). Scott et al. (1997) were the first to show that amygdala damage can lead to impairments for the processing of vocal emotions. In a patient with bilateral amygdala damage they showed specific impairments in decoding fear and anger cues from voices. This finding was replicated for fearful voices in another patient with bilateral amygdala damage (Sprengelmeyer et al., 1999), but other patients with bilateral amygdala damage did not exhibit such an impairment (Adolphs and Tranel, 1999; Anderson and Phelps, 1998; Bach et al., 2013). Thus, most of the studies including patients with lesions in unilateral or bilateral amygdala seem to indicate that the amygdala is not crucial for the decoding of the emotional tone from vocal expressions. This was mainly evident for the vocal expression of fear, while for vocal expressions of anger the amygdala seems to have some kind of importance.

As for lesions studies, data from functional imaging studies similarly revealed inconsistent findings. Recent studies have shown that the amygdala responds to emotional vocal intonations

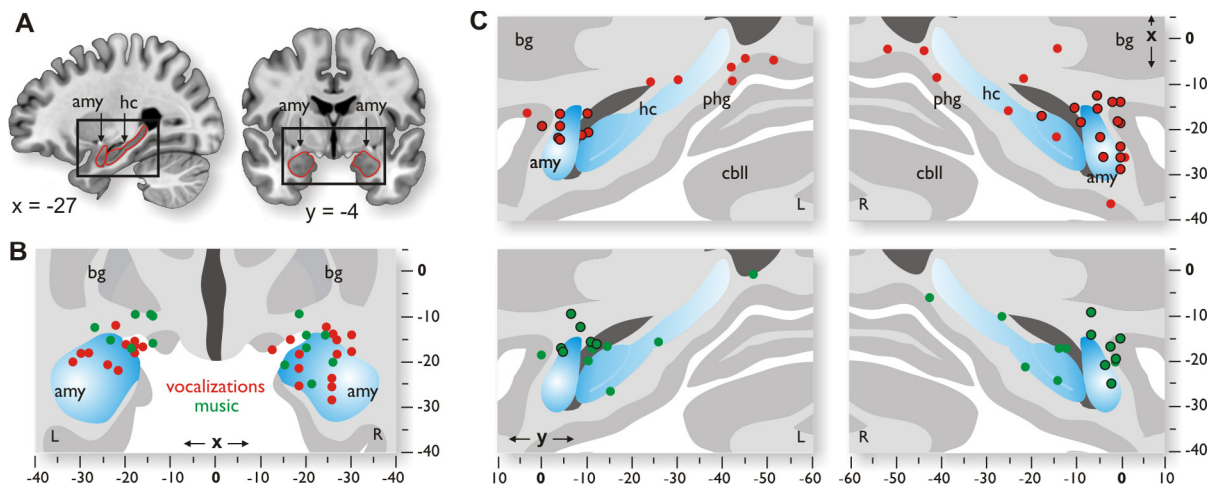


Fig. 3. Functional peak activations of studies on processing emotions from vocal expressions and from music. (A) Sagittal and coronal view to show the position of the amygdala and the hippocampus (red outlines) as part of the temporal limbic system. (B) Functional peak activations reported for the bilateral amygdala in response to emotional vocalizations (red) and in response to music (green). The image shows a coronal view ($y = -4$) on the amygdala as indicated by the black box in A, and the axis are in MNI standard space. (C) Functional peak activations for the amygdala and for the hippocampus, shown in two sagittal views for the left (left panel, $x = -27$) and the right brain (right panel, $x = 27$). Peak activations encircled in black indicate peak activations, which have been reported to be located in the amygdala in the original studies. For abbreviations see Fig. 2; bg basal ganglia.

Table 1

Functional imaging and lesions studies, which found an involvement of the amygdala (Amy) and the hippocampus (HC) for the decoding of vocal emotions. The table lists the number of subjects (*N*), stimulus material, imaging method used (functional resonance imaging (fMRI), (PET), or lesion study), and contrasts from which foci were derived.

Authors	<i>N</i>	Amy	HC	Stimulus	Method	Contrasts
Alba-Ferrara et al. (2011)	19		•	Neutral semantic speech	fMRI	Complex vs. simple emotional voices
Bach et al. (2008a,b)	16	•		Pseudo-language-based speech	fMRI	Implicit (gender ^d) vs. explicit (valence ^e)
Beaucousin et al. (2007)	23		•	Emotional semantic speech	fMRI	Human actors vs. Synthetic speech
Ethofer et al. (2009)	24	•		Neutral semantic speech	fMRI	Emotional vs. neutral voices ^f
Fecteau et al. (2007)	14	•		Nonverbal expressions	fMRI	Implicit (gender ^d): anger vs. neutral
Frühholz et al. (2012)	17	•		Pseudo-language-based speech	fMRI	Anger vs. neutral voices; explicit (valence ^e): anger vs. neutral
Frühholz and Grandjean (2013a,b,c)	17	•		Pseudo-language-based speech	fMRI	Anger vs. neutral voices; explicit (valence ^e): anger vs. neutral
Kotz et al., 2012			•	Neutral semantic speech	fMRI	Surprise
Kreifelts et al. (2010)	24	•		Neutral semantic speech; facial expressions	fMRI	Audio-visual > (auditory+ visual)
Leitman et al. (2010b)	20	•		Neutral semantic speech	fMRI	Happy + fear voices and cue interaction
Mitchell et al. (2003)	13		•	Emotional semantic speech	fMRI	Implicit (linguistic ^f) vs. explicit (valence ^e)
Morris et al. (1999)	6	•		Nonverbal expressions	PET	Emotional vs. neutral voices ^f
Mothes-Lasch et al. (2011)	24	•		Neutral semantic speech	fMRI	Implicit (gender ^d): anger vs. neutral; auditory vs. visual task
Phillips et al. (1998)	6	•	•	Nonverbal expressions	fMRI	Amygdala: Fearful vs. mild happy voices ^f ; Hippocampus: Fearful vs. neutral voices; Fearful vs. disgust voices
Quadflieg et al. (2008)	12 p; ^a 12 h ^b	•		Neutral semantic speech	fMRI	Emotional vs. neutral voices
Rota et al. (2011)	12		•	Emotional semantic speech	fMRI	Functional connectivity target of the rIFG as seed region
Sander et al. (2003a,b)	10	•		Nonverbal expressions	fMRI	Passive listening (detect oddball pitch events)
Sander and Scheich (2005)	20	•		Nonverbal expressions	fMRI	Passive listening (detect oddball spectral events)
Sander et al. (2005)	15	•	•	Pseudo-language-based speech	fMRI	Amygdala: Anger vs. neutral voices; anger (spatial ^g) vs. neutral voices; anger (no spatial) vs. neutral voices; anger (spatial ^g) vs. anger (no spatial); Hippocampus: anger (no spatial) vs. anger (spatial ^g)
Sander et al. (2007)	17	•		Nonverbal expressions	fMRI	Passive listening (detect oddball sine tone events)
Schirmer et al. (2008)	14	•		Pseudo-language-based speech	fMRI	Emotional vs. neutral deviant events
Scott et al. (1997)	1 p ^a	•		Neutral semantic speech; nonverbal expressions	Lesion	Impaired recognition of vocal fear and anger cues
Sprengelmeyer et al. (1999)	1 p ^a	•		Nonverbal expressions	Lesion	Impaired recognition of vocal fear cues
Szameitat et al. (2010)	18		•	Nonverbal expressions	fMRI	Joy vs. tickle; taunt vs. tickle
Wiethoff et al. (2009)	24	•	•	Neutral semantic speech	fMRI	Emotional vs. neutral voices

^a p = patients.

^b h = healthy controls.

^c The contrast “emotional vs. neutral voices” includes the comparison of emotional expression of different valence with neutral expressions. Studies differ in the amount and the kind of emotional expression, which were entered in this contrast.

^d “Gender” indicates an implicit gender discrimination task.

^e “Valence” indicates an explicit valence discrimination task.

^f “Linguistic” indicates an implicit linguistic discrimination task.

^g “Spatial” indicates an explicit processing of emotional voices by directing attention to the spatial location of these voices (operationalized as attentional focus to the left or right ear). “No spatial” means direction of attention away from the spatial location of emotional voices.

expressed by nonverbal utterances (Fecteau et al., 2007; Morris et al., 1999; Phillips et al., 1998; Sander et al., 2003b, 2007; Sander and Scheich, 2005) or superimposed on speech or speech-like verbal utterances (Bach et al., 2008a; Beaucousin et al., 2007; Ethofer et al., 2009; Frühholz et al., 2012; Frühholz and Grandjean, 2013a; Grandjean et al., 2005; Leitman et al., 2010b; Mothes-Lasch et al., 2011; Quadflieg et al., 2008; Sander et al., 2005; Schirmer et al., 2008; Wiethoff et al., 2009). Although most of these studies found an increase of amygdala activity in response to vocal emotions, some also found a decrease of activity, especially for fearful expressions (Morris et al., 1999).

In terms of sensitivity to vocal emotions of different emotional valence, the amygdala is often thought to be especially sensitive to negative emotions (Vuilleumier, 2005) and in particular to fear (Ohman and Mineka, 2001). This notion is usually embedded within an evolutionary framework assuming that the fast and prioritized detecting of imminent threat was of evolutionary importance. Not surprisingly, many studies found that the amygdala is generally sensitive to vocal cues of fear (Leitman et al., 2010b; Phillips et al., 1998; Scott et al., 1997) or anger (Ethofer et al., 2009; Frühholz and Grandjean, 2013a; Grandjean

et al., 2005; Mothes-Lasch et al., 2011; Quadflieg et al., 2008; Sander et al., 2005; Schirmer et al., 2008; Scott et al., 1997). The amygdala is also sensitive to vocal cues of sadness (Sander and Scheich, 2005), but not to disgust (Phillips et al., 1998). Given these recent studies, the most consistent evidence of amygdala sensitivity to specific vocal emotions yet has been found for angry expressions (Ethofer et al., 2009; Frühholz and Grandjean, 2013a; Grandjean et al., 2005; Mothes-Lasch et al., 2011; Quadflieg et al., 2008; Sander et al., 2005; Schirmer et al., 2008; Scott et al., 1997). Thus, unlike its proposed role to mainly decode *fear*, the amygdala seems more sensitive to vocal cues of *anger* (Bach et al., 2008a). Accordingly, results from lesions studies showed that amygdala damage does not necessarily impair the detection and processing of vocal cues for fear (Adolphs and Tranel, 1999; Anderson and Phelps, 1998; Sprengelmeyer et al., 1999). But amygdala activity seems not to be exclusive for negative vocal expressions, because amygdala activity has also been found in response to high arousing expressions of happiness (Fecteau et al., 2007; Leitman et al., 2010b; Sander et al., 2003b; Sander and Scheich, 2005), indicating that the amygdala might respond to high arousing emotions in general.

Despite this abundant evidence, the topic of amygdala sensitivity to vocal emotions remains debated, as several studies did not obtain activity in the amygdala in response to vocal emotions (Buchanan et al., 2000; Dietrich et al., 2007; Escoffier et al., 2012; Ethofer et al., 2006b; Kotz et al., 2003, 2012; Mitchell, 2006; Mitchell et al., 2003; Schlaug et al., 2010; Szameitat et al., 2010; Wiethoff et al., 2008; Wildgruber et al., 2002, 2005; Wittfoth et al., 2010). One critical factor to elicit amygdala activity in response to vocal emotions might be a task condition where the emotional cue of a voice is presented outside the focus of attention, and is thus processed on an “implicit” level (Wildgruber et al., 2009). For example, when focusing on the gender of the voice, the emotional value of the voice is still processed on an implicit level and elicits amygdala activity (Bach et al., 2008a; Fecteau et al., 2007; Hurt et al., 1998; Mothes-Lasch et al., 2011; Phillips et al., 1998). Similar amygdala activity is found during incidental processing of vocal emotions when focusing on the occurrence of other auditory events (Sander et al., 2003b, 2007; Sander and Scheich, 2005), or when the emotional cues was not in the spatial focus of attention (Sander et al., 2005). However, other studies report amygdala activity in response to vocal emotions only when the emotional cue is directly in the focus of attention (Beaucousin et al., 2007; Frühholz and Grandjean, 2013a; Leitman et al., 2010b).

4.1.2. Hippocampus

The hippocampus and the PHG are usually widely neglected in studies on processing vocal emotions. This is surprising, because a non-negligible number of studies report activity in the hippocampal formation in response to vocal emotions (Alba-Ferrara et al., 2011; Beaucousin et al., 2007; Kotz et al., 2012; Leitman et al., 2010b; Mitchell et al., 2003; Phillips et al., 1998; Rota et al., 2011; Sander et al., 2005; Szameitat et al., 2010; Wiethoff et al., 2008) (Fig. 3). Some of this activity is located in the ventral hippocampus in close vicinity to the amygdala (Beaucousin et al., 2007; Leitman et al., 2010b; Rota et al., 2011). The hippocampus seems also more responsive when vocal emotions are presented outside the focus of attention (Mitchell et al., 2003; Sander et al., 2005). When vocal emotions are presented inside the focus of attention the hippocampus however seems also to be part of a network including the right inferior frontal cortex (see Rota et al., 2011).

Similar to the amygdala, the hippocampus is sensitive to negative and positive vocal expressions, such as fear (Phillips et al., 1998), happiness (Leitman et al., 2010b), surprise (Kotz et al., 2012), or anger (Sander et al., 2005). The hippocampus is even able to discriminate between different negative (Phillips et al., 1998) and positive vocal emotions (Szameitat et al., 2010). For the latter it has been shown that the hippocampus, but not the amygdala, is sensitive to different types of laughter, such as positively (i.e. emotional laughter) and negatively connoted laughter, which is socially more complex (i.e. taunt laughter) (Szameitat et al., 2010). A valid categorization of both types of laughter might strongly depend on the context, on memory associations or on contextual conditions. Thus the hippocampus might provide the appropriate contextual associations during the perception of different types of laughter. This might be especially the case for more complex types of laughter, such as taunt laughter. This is in accordance with the observation that the hippocampus is more responsive to complex (i.e. pride or guilt) compared with simple vocal emotions (i.e. anger or happiness) (Alba-Ferrara et al., 2011). The former more strongly depend on cultural and contextual connotations and associations, similar to the processing emotions from musical pieces.

4.2. TLS activity in response to musical emotions

Several neuroimaging studies have demonstrated that the TLS is involved in the processing of emotions induced by music

(Alluri et al., 2012; Ball et al., 2007; Baumgartner et al., 2006; Blood and Zatorre, 2001; Brown et al., 2004; Chapin et al., 2010; Dyck et al., 2011; Eldar et al., 2007; Engel and Keller, 2011; Khalfa et al., 2008; Kleber et al., 2007; Koelsch et al., 2006, 2008, 2013; Lehne et al., 2014; Lerner et al., 2009; Mitterschiffthaler et al., 2007; Mueller et al., 2011; Mutschler et al., 2010; Pallesen et al., 2005, 2009; Park et al., 2013; Salimpoor et al., 2013) (see also Table 2). This suggests that the TLS is sensitive to music, and particularly to music that has the power to induce emotions. The amygdala and hippocampus respond to unpleasant or negative musical emotions (Koelsch et al., 2006; Lehne et al., 2014; Mitterschiffthaler et al., 2007), but also to pleasant music (Koelsch et al., 2013; Mueller et al., 2011; Salimpoor et al., 2013; Trost et al., 2012). These results are, similarly to the results obtained with vocal emotions, not in accordance with the general view that the TLS and the amygdala are mainly responsive to negative emotions (Ohman and Mineka, 2001; Vuilleumier, 2005). This discrepancy could be partly explained by different signal behavior of the amygdala in response to pleasant and unpleasant music (see below), but also by regional activity especially within subregions of the amygdalar complex (Ball et al., 2007; Koelsch et al., 2013). Some amygdala nuclei seem to respond to negative musical emotions, while others respond to pleasant music (Koelsch et al., 2013).

In terms of emotional valence, activity in the amygdala and the hippocampus seems to be generally sensitive to the level of subjective pleasantness of music, such as during chills. Chills are musically evoked moments of extreme pleasure (Panksepp, 1995). Amygdala and hippocampus activity seems to correlate negatively with the subjectively experienced intensity of chills (Blood and Zatorre, 2001). This suggests that the amygdala and the hippocampus might decrease their activity in response to positive and pleasant music, and to increase it in response to unpleasant music, such as dissonant music (Koelsch et al., 2006). Dissonance has often been used as a model to test the perception of unpleasantness and the negative valence induced by music (Blood et al., 1999; Dellacherie et al., 2009; Koelsch et al., 2006; Pallesen et al., 2005), and the amygdala indeed responds more strongly to dissonant and minor chords in comparison to major chords (Pallesen et al., 2005).

Unlike the simple distinction of pleasant and unpleasant music, several studies investigated more specifically the brain processes of certain emotions (Brattico et al., 2011; Lerner et al., 2009; Mitterschiffthaler et al., 2007; Park et al., 2013). The amygdala and the hippocampus are sensitive to sad compared to neutral music (Brattico et al., 2011; Mitterschiffthaler et al., 2007), and the amygdala responds to scary compared to relaxing music (Park et al., 2013). Similarly, lesions in the TLS can impair the recognition of sad and scary music (Gosselin et al., 2005, 2007). These results again seem to suggest that the amygdala and the hippocampus are rather sensitive to unpleasant music. However, there is increasing evidence that this association should be revisited (Khalifa et al., 2008; Koelsch et al., 2013; Mueller et al., 2011; Trost et al., 2012; Zatorre and Salimpoor, 2013). A recent study found that activity in the right hippocampus increased when listening to pleasant music (Trost et al., 2012). Interestingly, the hippocampus differentiated also between the levels of arousal, responding preferentially to complex musical emotions evaluated as pleasant and not very arousing (i.e. tenderness, peacefulness, nostalgia, or wonder). This indicates that the TLS might be sensitive not only to the arousal and the intensity level of musical emotions, but also to music with a certain level of emotional complexity. This would point to an emotional processing in the TLS that goes beyond simple effects of arousal and intensity.

Patient studies suggest that the TLS plays a decisive role in musical emotion processing (Dellacherie et al., 2008, 2011;

Table 2

Functional imaging and lesions studies, which found an involvement of the amygdala (Amy) and the hippocampus (HC) for the decoding of musical emotions. The table lists the number of subjects (*N*), stimulus material, imaging method used (functional resonance imaging (fMRI), positron emission tomography (PET), voxel-based morphometric (VBM), local field potential (LFP), or lesion study), and contrasts from which foci were derived.

Study	<i>N</i>	Amy	HC	Stimuli	Method	Contrast
Alluri et al. (2012)	11	•	•	Instrumental music	fMRI	Decreasing pulse clarity
Ball et al. (2007)	14	•	•	Instrumental music	fMRI	ROI analysis
Baumgartner et al. (2006)	9	•	•	Instrumental music; IAPS ^a	fMRI	IPAS+ music vs. IAPS; arousal ratings (Amy)
Blood and Zatorre (2001)	10	•	•	Music ^b (w/o lyrics)	PET	Decrease of chill intensity
Brattico et al. (2011)	15	•	•	Music ^b (w/o lyrics)	fMRI	Sad music+ lyrics vs. sad music
Chapin et al. (2010)	21	•	•	Instrumental music	fMRI	Expressive vs. mechanic music
Dellacherie et al. (2008)	43 p ^c , 19 h ^d	•	•	Instrumental music	Lesion	Impaired recognition of neutral music
Dellacherie et al. (2009)	1 p	•	•	Instrumental music	LFP	Dissonant music
Dellacherie et al. (2011)	19	•	•	Instrumental music	Lesion	No impairment
Dyck et al. (2011)	30	•	•	Instrumental music, facial expressions	fMRI	Music+ face vs. face; intensity of negative stimuli (Amy)
Engel and Keller (2011)	22	•	•	Instrumental music	fMRI	Improvisations vs. Imitations (Amy); Imitations vs. Improvisations (HC)
Eldar et al. (2007)	14	•	•	Instrumental music, film excerpts	fMRI	Music+ film vs. film; Music+ film vs. music
Fletcher et al. (2013)	37 p,		•	-	VBM	HC gray matter in musicophilia
Gosselin et al. (2005)	16 p, 16 h	•	•	Instrumental music	Lesion	Scary music
Gosselin et al. (2006)	17 p, 19 h	•	•	Instrumental music	Lesion	Dissonant music
Gosselin et al. (2007)	1 p	•	•	Instrumental music	Lesion	Scary and sad music
Gosselin et al. (2011)	16 p, 16 h	•	•	Instrumental music, facial expressions	Lesion	Fearful faces+ scary music
Hsieh et al. (2012)	23 p, 20 h	•	•	Instrumental music, facial expression	VBM	Impaired emotion recognition
James et al. (2008)	26	•	•	Instrumental music	EEG	Experts vs. laymen
Kleber et al. (2007)	16	•	•	Vocal music	fMRI	Imagined singing
Khalifa et al. (2008)	26 p, 60 h	•	•	Instrumental music	Lesion	Dissonant, sad and happy (left resections) music
Koelsch et al. (2006)	11	•	•	Instrumental music	fMRI	Unpleasant vs. pleasant music
Koelsch et al. (2008)	24	•	•	Instrumental music	fMRI	Syntactical violations vs. regular endings
Koelsch et al. (2013)	18	•	•	Instrumental music	fMRI	Happy vs. fear
Lehne et al. (2014)	25	•	•	Instrumental music	fMRI	Increasing vs. decreasing tension
Lerner et al. (2009)	12	•	•	Instrumental music	fMRI	Eyes closed vs. open (with negative music)
Levitin et al. (2003)	5 p, 5 h	•	•	Instrumental music, environmental sounds	fMRI	Music vs. noise
Mitterschiffthaler et al. (2007)	16	•	•	Instrumental music	fMRI	Sad vs. neutral music
Mueller et al. (2011)	20	•	•	Instrumental music	fMRI	Consonant vs. dissonant music
Mutschler et al. (2010)	20	•	•	Instrumental music	fMRI	Long-term habituation
Omar et al. (2011)	26 p, 21 h	•	•	Instrumental music, facial expressions, vocal expressions	VBM	Gray matter loss
Pallesen et al. (2005)	21	•	•	Instrumental music	fMRI	Minor vs. major chords; Dissonant vs. major chords
Pallesen et al. (2009)	10	•	•	Instrumental music	fMRI	Passive listening vs. working memory task
Park et al. (2013)	12	•	•	Instrumental music	fMRI	Scary vs. relaxing music
Pereira et al. (2011)	14	•	•	Music (w/o lyrics)	fMRI	Familiar vs. unfamiliar music
Plailly et al. (2007)	13	•	•	Music (w/o lyrics), odors	fMRI	Familiar vs. unfamiliar music and odors
Polston et al. (2011)	Rats	•	•	Instrumental music	Dopamine release	Music as conditioned stimulus
Trost et al. (2012)	15	•	•	Instrumental music	fMRI	Low arousal+ high valence
Salimpoor et al. (2013)	19	•	•	Music (w/o lyrics)	fMRI	Increasing connectivity to NAc during highly rewarding music processing
Watanabe et al. (2008)	18	•	•	Instrumental music	fMRI	Musical memory retrieval
Wieser and Mazzola (1986)	1	•	•	Instrumental music	LFP	Dissonances, higher-order tonality

Abbreviation: NAc nucleus accumbens.

^a Selection of images from the International Affective Picture System (IAPS; <http://csea.php.ufl.edu/Media.html>) picture set.

^b Musical pieces were selected by the participants.

^c p = patients.

^d h = healthy controls.

Gosselin et al., 2005, 2006, 2007, 2011; Khalifa et al., 2008). Lesions in the bilateral TLS can lead to an impaired recognition of scary and fearful music (Gosselin et al., 2005, 2007), but this impairment is not found with unilateral amygdala lesions (Dellacherie et al., 2011). This suggests that only bilateral damage in the TLS leads to an impairment of emotion recognition in music. While some patients show impaired recognition of musical emotions, others have shown increased sensitivity to music. Patients with Williams Syndrome, for example, show increased emotional reactions and increased amygdala activity when listening to music, an effect that was not found in healthy controls (Levitin et al., 2003). Furthermore, in patients with fronto-temporal degeneration it has been shown that the gray matter loss in the TLS is associated

with impaired recognition of musical emotions (Omar et al., 2011). Moreover, patients with fronto-temporal degeneration, who present musicophilia (i.e. a special interest in music), show increased gray matter volume in the left posterior hippocampus (Fletcher et al., 2013), indicating that the emotional attachment to music as a trait factor might be related to the hippocampus volume.

5. Affective arousal, emotional intensity and multimodal processing

Rather than decoding and discriminating different emotional valences, it has been proposed that the TLS and the amygdala are

more generally sensitive to the intensity level of emotional expressions irrespectively of the emotional valence (see Zald, 2003). This is closely related to the arousal effect of emotional signals. There is some evidence that the TLS might partly decode some kind of emotional intensity expressed in voices and in music.

5.1. TLS sensitivity to arousing and intense vocal emotions

Only few studies yet have investigated the effect of arousal level on brain responses when listening to vocal emotions. Warren et al. (2006) presented positive and negative nonverbal vocal expressions and analyzed the brain responses in relationship to the arousal level of these vocal expressions. While they found a strong relationship between the arousal level and premotor areas, they did not find any relationship between the arousal level and activity in the amygdala or the hippocampus. Wiethoff et al. (2008) presented neutral words that were spoken in different emotional tones. They found a linear relationship between emotional arousal and signal activity in bilateral STG (see also Ethofer et al., 2006c) and the hypothalamus, but reported no such effect in the amygdala or the hippocampus. Interestingly, in their study these arousal effects in STG were strongly associated with specific acoustic features of vocal expressions, such as the fundamental frequency (F0) and the intensity (i.e. loudness) of vocal emotions.

However, the intensity level of vocal emotions might be influenced not only by the intensity level of the vocal expressions alone, but can also be influenced by additional affective cues during the processing of emotional signals. In everyday life, emotional signals usually do not occur in one single modality, but rather in multiple sensory channels. Vocal expressions most often are accompanied by facial emotional cues (Kreifelts et al., 2010) and by body posture cues (Stienen et al., 2011). Ethofer et al. (2006a) have shown that the perceived emotional intensity of fearful, but not of happy facial expressions is increased when accompanied by congruent vocal expressions. And this increased intensity of face-voice cues was correlated with activity in the left amygdala. Similarly, Kreifelts et al. (2010) found an increase in bilateral amygdala activity during the processing of vocal emotions accompanied by facial expressions, compared to vocal emotions alone. No such effects were found in the hippocampus. Thus multimodal studies of vocal emotions presented together with emotional cues from other modalities provide some evidence that the amygdala, but not the hippocampus, is sensitive to the arousal or intensity level of vocal emotions (Milesi et al., 2014).

5.2. TLS sensitivity to arousing musical emotions

Direct evaluations for the arousal dimensions in music studies have rarely been used. One study (Trost et al., 2012) found an association between low arousing musical emotions (i.e. peacefulness, tenderness, nostalgia) and hippocampal activity, but no direct correlations were found between amygdala activity and arousal evaluations. A recent study (Lehne et al., 2014) investigated musical tension in terms of expectations and predictions induced by the temporal musical structure. Musical tension could also be regarded as a form of arousal, as these expectations might cause the suspense and curiosity of how the music will develop over time. This study found an involvement of the amygdala in moments of increasing tension compared to decreasing tension. Changes in musical tension might also indicate a change in the relevance of the music, and therefore the amygdala might interpret these changes as significant events (Sander et al., 2003a).

More indirect evidence for brain processes in response to emotional intensity of music comes from multimodal studies (Baumgartner et al., 2006; Brattico et al., 2011; Eldar et al., 2007). While the amygdala and the hippocampus seem to be inversely

related the level of chill intensity and thus the pleasantness of music, multimodal studies found a positive relationship when listening to music accompanied by visual stimulation (Dyck et al., 2011). Amygdala activity, for example, was found in response to sad music only when lyrics were added (Brattico et al., 2011). Both the amygdala and the hippocampus are active when viewing emotional pictures combined with music (Baumgartner et al., 2006), when viewing neutral films combined with positive or negative music (Eldar et al., 2007), or when listening to negative emotional music with closed eyes, but not with open eyes (Lerner et al., 2009). The latter finding might also point to the strength of imagination, which might be increased when individuals have their eyes closed. Similarly the amygdala and hippocampus are involved when professional opera singers imagine themselves singing (Kleber et al., 2007), which suggests that the TLS might to be involved in the imagination of emotionally rich contexts, which need the retrieval of memory associations.

The most classical multimodal condition however is given in the case of instrumental music, which is accompanied by singing voices. Although most of popular music contains voices and lyrics this effect has only been investigated in a few studies. It could be imagined that combining both, instrumental music and voices, could yield an additive emotional effect and thus an increased involvement of the TLS. Furthermore, studies on musical memories have shown that the integration of lyrics and melodies involves the TLS, notably the hippocampus (Alonso et al., 2014; Samson and Zatorre, 1991). Comparing brain responses to instrumental music and music with lyrics, has shown also a stronger recruitment of the amygdala and the hippocampus when lyrics are present (Alluri et al., 2013). This effect might be due to the increased emotional intensity in music with lyrics. More explicitly, when presenting the same songs with and without lyrics, the presence of lyrics seems to involve the TLS more strongly in sad music but not in happy music. This might be explained such that in happy music both music with and without lyrics are equally arousing, whereas in sad music the singing voice might provide additional emotional intensity and thus increase the multimodal arousal effect. Taken together, the amygdala and the hippocampus are sensitive to the intensity of musical emotions, but this effect was mostly shown in multimodal studies.

6. Acoustic and structural features of vocal and musical emotions

Emotions in voices and in music are encoded in specific acoustic features, such as pitch, intensity, tempo, and in rhythm. Cells in the amygdala have been shown to respond to basic auditory stimulation, such as simple tones (Bordi and LeDoux, 1992). As outlined above, the sensitivity of the amygdala, maybe of the TLS in general, to emotional cues in voices and in music might result from a specific sensitivity of the TLS to the specific acoustic features of vocal and musical emotions.

6.1. TLS sensitivity to acoustic features of vocal emotions

Emotional compared with neutral vocal expressions comprise a pattern of acoustic features, which make them clearly distinguishable from neutral expressions and help to distinguish between different emotions (Banse and Scherer, 1996; Juslin and Laukka, 2003; Patel et al., 2011; Sauter et al., 2010). While expressions of anger, fear, and happiness, for example, are accompanied by increased vocal intensity including an enhanced intensity variation, expressions of sadness are usually accompanied by low intensity. Similarly, angry, happy, and fearful voices show higher levels of pitch as mainly determined by the F0 as well as pitch variability compared to sad voices. Beside pitch and intensity many

more features of the voice quality, such as pitch envelope (i.e. voice melody), speech rate (i.e. tempo), harmonics-to-noise ratio (HNR), jitter and shimmer are considerably increased in expressions of emotions (Banse and Scherer, 1996; Juslin and Laukka, 2003; Patel et al., 2011).

Only a few studies yet have tested the brain sensitivity to the acoustic features of vocal emotions. An early study did not find any sensitivity to the variation of the F0 as well as to the vowel length of the stressed syllable on vocal emotions across the whole brain (Wildgruber et al., 2002). A subsequent study reported that a complex pattern of acoustic features, such as the mean and the variation of both the F0 and the intensity, can influence signal activity in the STG, but not in the amygdala and the hippocampus (Wiethoff et al., 2008). However, a recent study found that the amygdala is indeed sensitive to specific acoustic features especially to the level and the variations of the F0 and the acoustical intensity and mainly for angry voices (Frühholz et al., 2012; Frühholz and Grandjean, 2013a). Similarly, activity in posterior hippocampus correlates with the level and the variations in acoustical intensity (Wiethoff et al., 2008). Thus, the amygdala and the hippocampus are sensitive to acoustic features that are most relevant to decode and distinguish vocal emotions.

6.2. TLS sensitivity to acoustic features of music

Similar to human voices, emotions in music are encoded in specific acoustical features, which help the listener to decode the emotional value of music (Coutinho and Dibben, 2013; Juslin and Laukka, 2003). Compared to voices however, the relation between acoustic features and emotions in music is more complex, because cultural connotations play an important role in music perception. Also powerful emotional reactions to music are usually experienced at time scale of several seconds (Guhn et al., 2007), which is beyond the temporal resolution of single acoustic features. Nevertheless, some expressions of specific musical emotions can be attributed to certain acoustical features. For example, the tempo of the music is clearly associated with the arousal dimension, whereas mode and dissonance are linked to the valence dimension (Gabrielson and Juslin, 2003).

In music perception, dissonance is an acoustic feature that has been repeatedly reported to have an impact on amygdala activity. The degree of dissonance of the music is often used to modulate the level of unpleasantness (Blood et al., 1999). Dissonance can be described as the roughness of sound, which is generated by the inharmoniousness of close frequencies (McDermott et al., 2010). Several neuroimaging studies showed that both the amygdala and the hippocampus are sensitive to the dissonance of music (Koelsch et al., 2006; Pallesen et al., 2005; Wieser and Mazzola, 1986), and that the amygdala is specifically relevant for the perception of dissonant musical chords (Dellacherie et al., 2009). The PHG too seems to be sensitive to dissonance (Blood et al., 1999; Gosselein et al., 2006; Green et al., 2008).

The TLS has not been reported to play a dominant role in the processing of musical sound intensity or timbre (Alluri et al., 2012). Regarding the amygdala, this relative insensitivity to intensity is an interesting finding, because for non-musical acoustic stimuli it is well known that the amygdala, for example, responds to rising intensities in alerting sounds (Bach et al., 2008b). Hence, musical sound intensity seems not to have the same impact on the amygdala compared to simple sounds. This might be explained by the fact that music listening represents a safe context, and therefore the alerting function of the amygdala might be reduced. Another reason for the insensitivity of amygdala and the hippocampus to basic acoustic features of music might be the nature of music, which is a stimulus that develops in time. Emotional responses to music might therefore depend on larger

time scale structural elements rather than on instantaneous acoustical features. Meyer (1956) suggests that implicit expectations are induced during the time course of music listening and an affective response is elicited when these expectations are either met or violated based on the structural features of music.

6.3. TLS sensitivity to structural features of music

The structural features of music that seem to elicit the TLS are syntactical violations or temporal variability, whose decoding largely depends on the temporal context. The amygdala is sensitive to harmonic violations (Koelsch et al., 2008), and both the amygdala and the hippocampus are sensitive to harmonically unexpected endings of cadences (James et al., 2008). Violations of the temporal structure, such as expressively played music with more temporal variability and slight deviations from the metronomic metrical structure (Chapin et al., 2010) or the clarity of musical pulses (Alluri et al., 2012), seem to elicit activity in the TLS. Also jazz improvisations with temporal variability and unexpected events elicit activity in the amygdala, whereas an opposite effect is found in the hippocampus (Engel and Keller, 2011). This is in accordance with a general sensitivity of the amygdala to temporal unpredictability (Herry et al., 2007), which largely depends on the temporal and situational context of sound processing. Together, these findings suggest that both the amygdala and the hippocampus are sensitive to structural elements of music, which partly define its emotional character.

7. A neural framework for processing vocal and musical emotions in the TLS

As an interim summary, the evidence discussed so far suggests that both the amygdala and the hippocampus respond to vocal and musical emotions. However there are considerable differences. For vocal emotions both the amygdala and the hippocampus are sensitive to negative and positive expressions. The amygdala, but not the hippocampus, seems to respond to some acoustic features of vocal emotions and might be sensitive to some types of emotional intensity in vocal emotions.

Unlike for vocal emotions, the amygdala and the hippocampus mainly decode negative and unpleasant emotions in music. Both regions are predominantly sensitive to structural musical features, and partly also to dissonance as an acoustic feature of unpleasant music. Although the amygdala and the hippocampus are more active when music listening is accompanied by additional sensory information, direct evidence only suggests that they are sensitive to rather low levels of emotional intensity, but of higher social complexity.

Given the commonalities and differences in the neural processing of emotions in voices and music in the TLS, we suggest a model based on the differential use of four major processing loops and pathways during the decoding of vocal and musical emotions (see Fig. 4). A first loop between the auditory system and the TLS (see ① in Fig. 4) is represented by the subcortical route (i.e. the direct route) that connects the MGB, the amygdala, and the IC. This route is likely to provide a fast and coarse evaluation of simple sounds, or of short but meaningful sound segments. This evaluation is immediately fed back to influence low-level sound processing in the IC. A cortical route (i.e. indirect route; see ② in Fig. 4) feeds auditory signal forward to the amygdala from higher-level auditory cortex (STC), and probably also from AC. The indirect route sends simple and especially more complex information of preprocessed sound information to the amygdala. The amygdala is bidirectionally connected to the hippocampus (see ③ in Fig. 4). This connectivity might subserve the retrieval of emotional associations resulting from previous experiences with an emotional voice or music, might

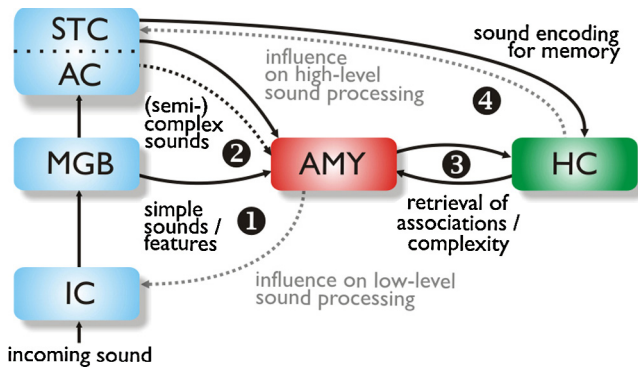


Fig. 4. Functional model and loops for decoding emotional cues from voices and from music. Any sound enters the human brain by the ascending auditory pathway (blue, right panel). A first auditory system-TLS loop (1) is represented by the subcortical or direct route and which connects the MGB, the amygdala and the IC. A cortical or indirect route (2) feeds auditory signal to the amygdala forward from higher-level auditory cortex (STC), and probably also from AC. The amygdala is bidirectionally connected to the hippocampus (3). Finally, the hippocampus is bidirectionally connected with the STC (4). For abbreviations see Fig. 2.

provide contextual information, and might be especially engaged during for the evaluation of complex vocal and musical emotions. Finally, the hippocampus is bidirectionally connected with the STC (see 4 in Fig. 4), and might influence high-level auditory processing of sound information.

7.1. Functional pathways for the decoding of vocal emotions

We propose that each of these functional loops play a distinct role in the neural decoding of vocal emotions. We suggest that the first loop performs a coarse stimulus evaluation, and is important for the fast evaluation of emotional cues in voices. Several lines of evidence are consistent with this model. First, both voices of negative and positive valence can activate the amygdala, and the fast evaluation of vocal emotions of any valence is important for an adaptive social behavior. Second, the amygdala is sensitive to acoustic features of voices (Frühholz et al., 2012). This information might be provided by the MGB, as suggested by studies reporting thalamic activation in response to vocal emotions (Kreifelts et al., 2010; Wildgruber et al., 2004), and some of these thalamic regions in the posterior thalamus (Kreifelts et al., 2010) were located close to the MGB. This findings indicate that decoding of emotional cues from vocal emotions can already occur at the level of MGB (Wildgruber et al., 2009). Third, vocal emotions drive amygdala activity preferentially during implicit processing, which points to the fact that emotional information is decoded by the amygdala already at early stages of processing prior to cortical processing. This allows the detection and processing of important emotional information outside the current focus of attention (Bach et al., 2008a; Grandjean et al., 2005; Wildgruber et al., 2009), or even during sleep (Ising and Kruppa, 2004). Fourth, especially short affective bursts of nonverbal vocal expressions seem to predominantly activate the amygdala (Fecteau et al., 2007; Morris et al., 1999; Phillips et al., 1998; Sander et al., 2003b; Sander and Scheich, 2005), but less the hippocampus (Phillips et al., 1998; Szameitat et al., 2010), and these short affect burst might already receive fast decoding in the first subcortical loop.

Unlike short affect bursts, vocal cues superimposed on speech have a broader time scale, similar to the modulation rate sensitivity of the auditory cortical regions. They also need a broader temporal integration for a proper recognition (Pell and Kotz, 2011). Emotional prosody might therefore mainly be decoded using the indirect route (i.e. the second loop), which is involved in the decoding of more complex sounds. Furthermore, multimodal

integration of emotional voices accompanied by sensory information in other modalities is accomplished in the STC (Kreifelts et al., 2009) as well as the decoding of the arousal level of vocal emotions (Ethofer et al., 2006c; Wiethoff et al., 2008). After preprocessing and integration in the STC, vocal emotional cues might be fed forward to the amygdala for further emotional decoding. Interestingly, compared to nonverbal expressions emotional cues superimposed on speech seem to more consistently activate the hippocampus (Alba-Ferrara et al., 2011; Beaucousin et al., 2007; Kotz et al., 2012; Mitchell et al., 2003; Rota et al., 2011; Sander et al., 2005; Wiethoff et al., 2008). Also, more complex emotional prosodies, such as expression of pride and guilt, activate the hippocampus (Alba-Ferrara et al., 2011). Thus, in addition to the predominant use of the second loop, acoustically and perceptually more complex emotional prosody might also use the third loop between the amygdala and the hippocampus. Hippocampal activity might subserve the temporal integration of temporally more complex prosodic information, and might provide important contextual and memory associations. This might be the case also for some types of short nonverbal vocal expression, such as laughter. Taunt laughter, for example, is a highly socially and contextually dependent type of laughter, and strongly activates the hippocampus instead of the amygdala (Szameitat et al., 2010).

The importance of the fourth loop between the STC and the hippocampus might be lesser for the immediate decoding of vocal emotions. The forward connections between the STC and the hippocampus might however be relevant for the memory related encoding of episodic events, for example during the perception of relevant vocal emotions expressed by certain or socially familiar individuals (Nakamura et al., 2001) or in certain contexts (Hasselmo and Eichenbaum, 2005). These memory associations might influence the cortical processing of vocal utterances by feedback connections from hippocampus to STC. Thus, prior experience might influence future encounters with similar vocal expressions or with the same individuals.

7.2. Functional pathways for the decoding of musical emotions

Like for vocal emotions, the TLS seems sensitive for negative, but also for positive emotions expressed in music. Instead of single musical features, the TLS is especially sensitive to structural features of music (James et al., 2008; Koelsch et al., 2008) and to the temporal variability as a mean of musical expressivity (Chapin et al., 2010; Engel and Keller, 2011). The former usually show a relative late modulatory effect in the TLS (James et al., 2008) beyond time intervals of cortical processing (Koelsch and Friederici, 2003). Compared to vocal expressions of emotions, both musical structure and musical variability usually have a wider time range of temporal integration and depend on expectations, which are usually created at a cortical level (for example in the IFG, Lehne et al., 2014) based on the context provided by prior musical events. Given these findings and the fact that the amygdala seems to react preferably to short and relevant events, we think that cortical loops involving the STC and the hippocampus play a more important role in music processing than the subcortical pathways. It has also been reported that signal in the amygdala in response to music habituates more rapidly than in the hippocampus (Mutschler et al., 2010). Moreover, in most of the studies on musical emotions, both amygdala and hippocampus are reported to be involved, which suggest a close connection and supposedly an implication of an internal feedback loop between the two structures.

However there might be some cases of music listening where the first subcortical loop might be relevant. For example, when music is used in conditioning paradigms, the amygdala is involved, supposedly via the first loop (Polston et al., 2011). The amygdala might especially respond to short and meaningful musical events

with sudden onsets, extreme loudness, perceptual discomfort, or unexpected musical changes. Conditioning paradigms partly use sudden and uncomfortable sounds as an unconditioned stimulus with considerable power to induce affective reactions (Buchel et al., 1999). Music sometimes mimics these emotional properties of simple sounds as a means of emotional expressions (Scherer, 1995), and through the first loop the amygdala might perform a fast evaluation of these short, potentially meaningful, and especially negative musical events.

Although the first loop might have some relevance for the decoding of musical emotions, we think that the other loops mainly involving the STC, the amygdala, and the hippocampus are much more relevant during music listening. Specifically, the fast evaluation of music is often not important, because music listening usually represents a “safe” listening context, in which no immediate threat is expected. Furthermore, music listening is mostly a pleasant and often social leisure activity. Thus cortical processing of music in the STC and functional loops with the amygdala (the second loop) and the hippocampus (the fourth loop) might have a predominant role in processing musical emotions. Cortical processing of music in the STC is relevant, because musical emotions are often described as esthetic emotions, which require cognitive appraisal processes, and given that music appreciation is a highly individual process (Zentner et al., 2008). Also, affective reactions induced by music require temporal integration into a percept represented at the cortical level, because the musical structure represents an important ingredient of musical emotions. The amygdala has also shown to be sensitive to temporal unpredictability in music (Engel and Keller, 2011). This processing involving unpredictability as well as the detection of novel events might need longer temporal integration windows and thus might be a functional role of the second loop between the STC and the amygdala. Finally, amygdala and hippocampus activity often depends on the multimodal processing of music, such as when music is accompanied by additional sensory information in another modality (Baumgartner et al., 2006; Brattico et al., 2011; Dyck et al., 2011; Eldar et al., 2007; Gosselin et al., 2011).

Besides the relevance of cortical processing of music, we think that the hippocampus and loops involving the hippocampus (i.e. the third and fourth loop) are important for processing musical emotions. The hippocampus is involved in musical memory evaluations (Pereira et al., 2011; Plailly et al., 2007; Trost et al., 2012; Watanabe et al., 2008), and responds to musical emotions characterized by low arousal and high complexity, such as peacefulness or nostalgia (Trost et al., 2012), but also sadness (Brattico et al., 2011; Mitterschiffthaler et al., 2007). Thus, the hippocampus might be associated with musical memory retrieval, which seems relevant for emotional reactions to music, especially for more complex emotions, which involve episodic memory, contextual, and cultural associations. Furthermore, the involvement of the hippocampus might also be linked to the social function of music (Immordino-Yang and Singh, 2013). Music is regarded as a social activity that may foster communication, cooperation, and coordination between people through its power to synchronize behavior and minds (Hove and Risen, 2009; Koelsch, 2010). Moreover, studies have shown that the feeling of being in synchrony with a virtual tapping partner in a sensorimotor synchronization task correlated with the BOLD activity in the hippocampus (Fairhurst et al., 2012). Emotional processing of musical emotions in the amygdala might therefore be strongly linked, via the third loop, to memory associations provided by hippocampal activity. This is consistent with the fact that many of the (unimodal) studies on musical emotions find both amygdala and hippocampus activity (Alluri et al., 2012; Blood and Zatorre, 2001; Chapin et al., 2010; Dellacherie et al., 2011; Dyck et al., 2011;

Engel and Keller, 2011; Gosselin et al., 2006; James et al., 2008; Khalfa et al., 2008; Kleber et al., 2007; Koelsch et al., 2006; Lerner et al., 2009; Mitterschiffthaler et al., 2007; Mutschler et al., 2010; Pallesen et al., 2009; Pereira et al., 2011). Furthermore, the third loop might be initiated either in the amygdala or in the hippocampus such that not only the emotional evaluation may trigger the retrieval of memory associations, but also memories associated with a specific music may trigger specific emotional reactions. It has been proposed that memory encoding for negative memories might be enhanced and associated with an increased recruitment of the TLS (Kensinger, 2007). This could corroborate the notion that the TLS seems preferentially responsive to unpleasant music and that emotion related memories might be especially triggered during listening of music of negative and unpleasant emotional valence. Negative musical emotions are often induced by dissonant music, and both, the amygdala and the hippocampus, as well as the PHG, as an important input relay to the hippocampus are decoding dissonance in music (Blood et al., 1999; Gosselin et al., 2006; Green et al., 2008). On the other hand, as discussed above, there is accumulating evidence for an involvement of the TLS also for pleasant music as well as for complex emotions (Koelsch et al., 2013; Mueller et al., 2011; Salimpoor et al., 2013; Trost et al., 2012). These findings suggest a more sophisticated association between the TLS and musical emotions, but more research is needed to disentangle the role of valence for the TLS in the processing of musical emotions.

8. Conclusions

Vocal and musical expressions of emotions are powerful means of signaling and expressing emotions. The TLS shows strong reactivity to emotions expressed in voices and in music. The amygdala and the hippocampus are central brain structures in the limbic system and specifically in the TLS for the decoding of the emotional value from incoming sensory information. Emotional cues in voices and in music could be transferred to, exchanged with and decoded by the TLS using several pathways and loops. We reviewed the recent literature about the TLS sensitivity to vocal and musical emotions, which indicates a common, but also differential sensitivity of the amygdala and the hippocampus to the valence, the intensity and acoustic and structural features of vocal and musical emotions.

Given this differential encoding of vocal and musical emotions, we proposed a differential functional use of four different pathways connecting the auditory system and the TLS. Whereas a subcortical and direct route to the amygdala might be predominantly used for the decoding of short vocal affect burst, but partly also for the decoding of rapid, salient, and meaningful musical events, the cortical pathways connecting higher-level auditory cortex to the amygdala and the hippocampus might be equally used for the decoding of vocal and music emotions. The hippocampus might support the decoding of complex vocal emotions and might provide associations for the decoding of contextually dependent vocal expressions. The decoding of musical emotions might even more strongly depend on the hippocampus, as well as on hippocampal connections to the amygdala and to the auditory system. Processing esthetic emotions expressed in music considerably depend on the temporal context of music itself, but also on the social and cultural context. The hippocampus might be strongly sensitive to these social and contextual settings.

Taken together, our proposed network summarizes and points out that musical emotions and vocal expressions have common processing pathways for auditory processing, but that they are partly distinct in terms of emotional processing in the TLS given the differential nature of their communicational value.

Acknowledgments

S.F. was supported by a grant from the Swiss National Science Foundation (SNSF 105314.146559/1). DG was supported by the NCCR in Affective Sciences (51NF40-104897). The authors thank Alessia Pannese for helpful comments on the manuscript.

References

- Adolphs, R., Tranel, D., 1999. Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia* 37, 1285–1292.
- Aitkin, L.M., Dickhaus, H., Schult, W., Zimmermann, M., 1978. External nucleus of inferior colliculus: auditory and spinal somatosensory afferents and their interactions. *J. Neurophysiol.* 41, 837–847.
- Alba-Ferrara, L., Hausmann, M., Mitchell, R.L., Weis, S., 2011. The neural correlates of emotional prosody comprehension: disentangling simple from complex emotion. *PLoS ONE* 6, e28701.
- Alluri, V., Toiviainen, P., Jaaskelainen, I.P., Glerean, E., Sams, M., Brattico, E., 2012. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage* 59, 3677–3689.
- Alluri, V., Toiviainen, P., Lund, T.E., Wallentin, M., Vuust, P., Nandi, A.K., Ristaniemi, T., Brattico, E., 2013. From Vivaldi to Beatles and back: predicting lateralized brain responses to music. *Neuroimage* 83, 627–636.
- Alonso, I., Sammler, D., Valabregue, R., Dinkleracker, V., Dupont, S., Belin, P., Samson, S., 2014. Hippocampal sclerosis affects fMR-adaptation of lyrics and melodies in songs. *Front. Hum. Neurosci.* 8, 111.
- Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N.J., Habel, U., Schneider, F., Zilles, K., 2005. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anat. Embryol. (Berl.)* 210, 343–352.
- Anderson, A.K., Phelps, E.A., 1998. Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport* 9, 3607–3613.
- Bach, D.R., Grandjean, D., Sander, D., Herdener, M., Strik, W.K., Seifritz, E., 2008a. The effect of appraisal level on processing of emotional prosody in meaningless speech. *Neuroimage* 42, 919–927.
- Bach, D.R., Hurlmann, R., Dolan, R.J., 2013. Unimpaired discrimination of fearful prosody after amygdala lesion. *Neuropsychologia* 51, 2070–2074.
- Bach, D.R., Schachinger, H., Neuhoff, J.G., Esposito, F., Di Salle, F., Lehmann, C., Herdener, M., Scheffler, K., Seifritz, E., 2008b. Rising sound intensity: an intrinsic warning cue activating the amygdala. *Cereb. Cortex* 18, 145–150.
- Ball, T., Rahm, B., Eickhoff, S.B., Schulze-Bonhage, A., Speck, O., Mutschler, I., 2007. Response properties of human amygdala subregions: evidence based on functional MRI combined with probabilistic anatomical maps. *PLoS ONE* 2, e307.
- Banse, R., Scherer, K.R., 1996. Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636.
- Baumann, S., Griffiths, T.D., Sun, L., Petkov, C.I., Thiele, A., Rees, A., 2011. Orthogonal representation of sound dimensions in the primate midbrain. *Nat. Neurosci.* 14, 423–425.
- Baumgartner, T., Lutz, K., Schmidt, C.F., Jancke, L., 2006. The emotional power of music: how music enhances the feeling of affective pictures. *Brain Res.* 1075, 151–164.
- Beaucousin, V., Lacheret, A., Turbelin, M.R., Morel, M., Mazoyer, B., Tzourio-Mazoyer, N., 2007. FMRI study of emotional speech comprehension. *Cereb. Cortex* 17, 339–352.
- Belin, P., Fecteau, S., Bedard, C., 2004. Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.* 8, 129–135.
- Belin, P., Fecteau, S., Charest, I., Nicastro, N., Hauser, M.D., Armony, J.L., 2008. Human cerebral response to animal affective vocalizations. *Proc. R. Soc. Lond. B Biol. Sci.* 275, 473–481.
- Belin, P., Zatorre, R.J., 2000. Voice-selective areas in human auditory cortex. *Nature* 403, 309.
- Bendor, D., Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. *Nature* 436, 1161–1165.
- Bestelmeyer, P.E.G., Rouger, J., DeBriune, L.M., Belin, P., 2010. Auditory adaption in vocal affect perception. *Cognition* 117, 217–223.
- Blood, A.J., Zatorre, R.J., 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11818–11823.
- Blood, A.J., Zatorre, R.J., Bermudez, P., Evans, A.C., 1999. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2, 382–387.
- Bordi, F., 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, 2493–2503.
- Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., Tervaniemi, M., 2011. A functional MRI study of happy and sad emotions in music with and without lyrics. *Front. Psychol.* 2, 308.
- Brown, S., Martinez, M.J., Parsons, L.M., 2004. Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport* 15, 2033–2037.
- Buchanan, T.W., Lutz, K., Mirzazade, S., Specht, K., Shah, N.J., Zilles, K., Jancke, L., 2000. Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Res. Cogn. Brain Res.* 9, 227–238.
- Buchel, C., Dolan, R.J., Armony, J.L., Friston, K.J., 1999. Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *J. Neurosci.* 19, 10869–10876.
- Cenquizca, L.A., Swanson, L.W., 2007. Spatial organization of direct hippocampal field CA1 axonal projections to the rest of the cerebral cortex. *Brain Res. Rev.* 56, 1–26.
- Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47, 236–246.
- Chapin, H., Jantzen, K., Kelso, J.A., Steinberg, F., Large, E., 2010. Dynamic emotional and neural responses to music depend on performance expression and listener experience. *PLoS ONE* 5, e13812.
- Conard, N.J., Malina, M., Munzel, S.C., 2009. New flutes document the earliest musical tradition in southwestern Germany. *Nature* 460, 737–740.
- Coutinho, E., Dibben, N., 2013. Psychoacoustic cues to emotion in speech prosody and music. *Cognit. Emot.* 27, 658–684.
- Da Costa, S., van der Zwaag, W., Marques, J.P., Frackowiak, R.S., Clarke, S., Saenz, M., 2011. Human primary auditory cortex follows the shape of Heschl's gyrus. *J. Neurosci.* 31, 14067–14075.
- Dellacherie, D., Bigand, E., Molin, P., Baulac, M., Samson, S., 2011. Multidimensional scaling of emotional responses to music in patients with temporal lobe resection. *Cortex* 47, 1107–1115.
- Dellacherie, D., Ehrle, N., Samson, S., 2008. Is the neutral condition relevant to study musical emotion in patients? *Music Percept.* 25, 285–294.
- Dellacherie, D., Pfeuty, M., Hasboun, D., Lefevre, J., Hugueville, L., Schwartz, D.P., Baulac, M., Adam, C., Samson, S., 2009. The birth of musical emotion: a depth electrode case study in a human subject with epilepsy. *Ann. N. Y. Acad. Sci.* 1169, 336–341.
- Dietrich, S., Hertrich, I., Alter, K., Ischebeck, A., Ackermann, H., 2007. Semiotic aspects of human nonverbal vocalizations: a functional imaging study. *Neuroreport* 18, 1891–1894.
- Dyck, M., Loughhead, J., Kellermann, T., Boers, F., Gur, R.C., Mathiak, K., 2011. Cognitive versus automatic mechanisms of mood induction differentially activate left and right amygdala. *Neuroimage* 54, 2503–2513.
- Eerola, T., Vuoskoski, J.K., 2011. A comparison of the discrete and dimensional models of emotion in music. *Psychol. Music* 39, 18–49.
- Ekman, P., 1992. An argument for basic emotions. *Cognit. Emot.* 6, 169–200.
- Eldar, E., Ganor, O., Admon, R., Bleich, A., Hendler, T., 2007. Feeling the real world: limbic response to music depends on related content. *Cereb. Cortex* 17, 2828–2840.
- Engel, A., Keller, P.E., 2011. The perception of musical spontaneity in improvised and imitated jazz performances. *Front. Psychol.* 2, 83.
- Escoffier, N., Zhong, J., Schirmer, A., Qiu, A., 2012. Emotional expressions in voice and music: same code, same effect? *Hum. Brain Mapp.* 34, 1796–1810.
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., Reiterer, S., Grodd, W., Wildgruber, D., 2006a. Impact of voice on emotional judgment of faces: an event-related fMRI study. *Hum. Brain Mapp.* 27, 707–714.
- Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., Wildgruber, D., 2006b. Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. *Neuroimage* 30, 580–587.
- Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., Grodd, W., Wildgruber, D., 2006c. Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport* 17, 249–253.
- Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., Wildgruber, D., 2009. Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *J. Cogn. Neurosci.* 21, 1255–1268.
- Everitt, B.J., 1990. Sexual motivation: a neural and behavioural analysis of the mechanisms underlying appetitive and copulatory responses of male rats. *Neurosci. Biobehav. Rev.* 14, 217–232.
- Fairhurst, M.T., Janata, P., Keller, P.E., 2012. Being and feeling in sync with an adaptive virtual partner: brain mechanisms underlying dynamic cooperativity. *Cereb. Cortex* 25, 2592–2600.
- Fanselow, M.S., Dong, H.-W., 2010. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* 65, 7–19.
- Fecteau, S., Belin, P., Joanette, Y., Armony, J.L., 2007. Amygdala responses to nonlinguistic emotional vocalizations. *Neuroimage* 36, 480–487.
- Fletcher, P.D., Downey, L.E., Witoonpanich, P., Warren, J.D., 2013. The brain basis of musicophilia: evidence from frontotemporal lobar degeneration. *Front. Psychol.* 4, 347.
- Fontaine, J.R., Scherer, K.R., Roesch, E.B., Ellsworth, P.C., 2007. The world of emotions is not two-dimensional. *Psychol. Sci.* 18, 1050–1057.
- Frühholz, S., Ceravolo, L., Grandjean, D., 2012. Specific brain networks during explicit and implicit decoding of emotional prosody. *Cereb. Cortex* 22, 1107–1117.
- Frühholz, S., Grandjean, D., 2013a. Amygdala subregions differentially respond and rapidly adapt to threatening voices. *Cortex* 49, 1394–1403.
- Frühholz, S., Grandjean, D., 2013b. Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: a quantitative meta-analysis. *Neurosci. Biobehav. Rev.* 37, 24–35.
- Frühholz, S., Grandjean, D., 2013c. Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neurosci. Biobehav. Rev.* 2847–2855.
- Frühholz, S., Klaas, H., Patel, S., Grandjean, D., 2014. Talking in fury: the cortico-subcortical network underlying angry vocalizations. *Cereb. Cortex*.
- Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., Benedetti, F., Abbamonte, M., Gasparotti, R., Barale, F., Perez, J., McGuire, P., Politi, P., 2009. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatry Neurosci.: JPN* 34, 418–432.

- Gabrielson, A., Juslin, P.N., 2003. Emotional expression in music. In: Davidson, R.J., Scherer, K., Goldsmith, H.H. (Eds.), *Handbook of Affective Sciences*. Oxford University Press, New York, pp. 503–534.
- Gomez, P., Danuser, B., 2007. Relationships between musical structure and psychophysiological measures of emotion. *Emotion* 7, 377–387.
- Gosselin, N., Peretz, I., Hasboun, D., Baulac, M., Samson, S., 2011. Impaired recognition of musical emotions and facial expressions following anteromedial temporal lobe excision. *Cortex* 47, 1116–1125.
- Gosselin, N., Peretz, I., Johnsen, E., Adolphs, R., 2007. Amygdala damage impairs emotion recognition from music. *Neuropsychologia* 45, 236–244.
- Gosselin, N., Peretz, I., Noulhiane, M., Hasboun, D., Beckett, C., Baulac, M., Samson, S., 2005. Impaired recognition of scary music following unilateral temporal lobe excision. *Brain* 128, 628–640.
- Gosselin, N., Samson, S., Adolphs, R., Noulhiane, M., Roy, M., Hasboun, D., Baulac, M., Peretz, I., 2006. Emotional responses to unpleasant music correlates with damage to the parahippocampal cortex. *Brain* 129, 2585–2592.
- Goudbeek, M., Scherer, K., 2010. Beyond arousal: valence and potency/control cues in the vocal expression of emotion. *J. Acoust. Soc. Am.* 128, 1322–1336.
- Grandjean, D., Banziger, T., Scherer, K.R., 2006. Intonation as an interface between language and affect. *Prog. Brain Res.* 156, 235–247.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.R., Vuilleumier, P., 2005. The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* 8, 145–146.
- Green, A.C., Baerentsen, K.B., Stodkilde-Jørgensen, H., Wallentin, M., Roepstorff, A., Vuust, P., 2008. Music in minor activates limbic structures: a relationship with dissonance? *Neuroreport* 19, 711–715.
- Grewe, O., Nagel, F., Kopiez, R., Altenmüller, E., 2007. Emotions over time: synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion* 7, 774–788.
- Guhn, M., Hamm, A., Zentner, M., 2007. Physiological and musico-acoustic correlates of the chill response. *Music Percept.* 24, 473–483.
- Hackett, T.A., 2011. Information flow in the auditory cortical network. *Hear. Res.* 271, 133–146.
- Hammerschmidt, K., Jürgens, U., 2007. Acoustical correlates of affective prosody. *J. Voice: Off. J. Voice Found.* 21, 531–540.
- Hasselmo, M.E., Eichenbaum, H., 2005. Hippocampal mechanisms for the context-dependent retrieval of episodes. *Neural Netw.* 18, 1172–1190.
- Henke, K., Buck, A., Weber, B., Wieser, H.G., 1997. Human hippocampus establishes associations in memory. *Hippocampus* 7, 249–256.
- Herry, C., Bach, D.R., Esposito, F., Di Salle, F., Perrig, W.J., Scheffler, K., Luthi, A., Seifritz, E., 2007. Processing of temporal unpredictability in human and animal amygdala. *J. Neurosci.* 27, 5958–5966.
- Hove, M.J., Risen, J.L., 2009. It's all in the timing: interpersonal synchrony increases affiliation. *Soc. Cognit.* 27, 949–960.
- Hsieh, S., Hornberger, M., Piguet, O., Hodges, J.R., 2012. Brain correlates of musical and facial emotion recognition: evidence from the dementias. *Neuropsychologia* 50, 1814–1822.
- Hurt, C.P., Rice, R.R., McIntosh, G.C., Thaut, M.H., 1998. Rhythmic auditory stimulation in gait training for patients with traumatic brain injury. *J. Music Ther.* 35, 228–241.
- Immordino-Yang, M.H., Singh, V., 2013. Hippocampal contributions to the processing of social emotions. *Hum. Brain Mapp.* 34, 945–955.
- Ising, H., Kruppa, B., 2004. Health effects caused by noise: evidence in the literature from the past 25 years. *Noise Health* 6, 5–13.
- James, C.E., Britz, J., Vuilleumier, P., Hauert, C.A., Michel, C.M., 2008. Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *Neuroimage* 42, 1597–1608.
- Jarrell, T.W., Gentile, C.G., Romanski, L.M., McCabe, P.M., Schneiderman, N., 1987. Involvement of cortical and thalamic auditory regions in retention of differential bradycardiac conditioning to acoustic conditioned stimuli in rabbits. *Brain Res.* 412, 285–294.
- Jones, E.G., 2007. *The Thalamus*, vol. 2. Cambridge University Press, Cambridge.
- Juslin, P.N., Laukka, P., 2003. Communication of emotions in vocal expressions and music performance: different channels, same code? *Psychol. Bull.* 129, 770–814.
- Juslin, P.N., Vastfjäll, D., 2008. Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* 31, 559–575 (discussion 575–621).
- Kensinger, E.A., 2007. Negative emotion enhances memory accuracy: behavioral and neuroimaging evidence. *Curr. Dir. Psychol. Sci.* 16, 213–218.
- Khalifa, S., Guye, M., Peretz, I., Chapon, F., Girard, N., Chauvel, P., Liegeois-Chauvel, C., 2008. Evidence of lateralized anteromedial temporal structures involvement in musical emotion processing. *Neuropsychologia* 46, 2485–2493.
- Kissler, J., 2013. Love letters and hate mail: cerebral processing of emotional language content. In: Armony, J., Vuilleumier, P. (Eds.), *Human Affective Neuroscience*. Cambridge University Press, Cambridge, pp. 304–328.
- Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., Lotze, M., 2007. Overt and imagined singing of an Italian aria. *Neuroimage* 36, 889–900.
- Koelsch, S., 2010. Towards a neural basis of music-evoked emotions. *Trends Cogn. Sci.* 14, 131–137.
- Koelsch, S., 2014. Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* 15, 170–180.
- Koelsch, S., Friederici, A.D., 2003. Toward the neural basis of processing structure in music. Comparative results of different neurophysiological investigation methods. *Ann. N. Y. Acad. Sci.* 999, 15–28.
- Koelsch, S., Fritz, T., Schlaug, G., 2008. Amygdala activity can be modulated by unexpected chord functions during music listening. *Neuroreport* 19, 1815–1819.
- Koelsch, S., Fritz, T., Cramon, V., Müller, K., Friederici, A.D., 2006. Investigating emotion with music: an fMRI study. *Hum. Brain Mapp.* 27, 239–250.
- Koelsch, S., Skouras, S., Fritz, T., Herrera, P., Bonhage, C., Kussner, M.B., Jacobs, A.M., 2013. The roles of superficial amygdala and auditory cortex in music-evoked fear and joy. *Neuroimage* 81, 49–60.
- Kotz, S.A., Kalberlah, C., Bahlmann, J., Friederici, A.D., Haynes, J.D., 2012. Predicting vocal emotion expressions from the human brain. *Hum. Brain Mapp.* 34, 1971–1981.
- Kotz, S.A., Meyer, M., Alter, K., Besson, M., von Cramon, D.Y., Friederici, A.D., 2003. On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Lang.* 86, 366–376.
- Kreifelts, B., Ethofer, T., Huberle, E., Grodd, W., 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.
- Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., Wildgruber, D., 2009. Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia* 47, 3059–3066.
- LeDoux, J., 2007. The amygdala. *Curr. Biol.* 17, R868–R874.
- LeDoux, J., 2012. Rethinking the emotional brain. *Neuron* 73, 653–676.
- LeDoux, J.E., 2000. Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184.
- Lehne, M., Rohrmeier, M., Koelsch, S., 2014. Tension-related activity in the orbitofrontal cortex and amygdala: an fMRI study with music. *Soc. Cogn. Affect. Neurosci.* 9, 1515–1523.
- Leitman, D.I., Laukka, P., Juslin, P.N., Saccente, E., Butler, P., Javitt, D.C., 2010a. Getting the cue: sensory contributions to auditory emotion recognition impairments in schizophrenia. *Schizophr. Bull.* 36, 545–556.
- Leitman, D.I., Wolf, D.H., Laukka, P., Ragland, J.D., Valdez, J.N., Turetsky, B.I., Gur, R.E., Gur, R.C., 2011. Not pitch perfect: sensory contributions to affective communication impairment in schizophrenia. *Biol. Psychiatry* 70, 611–618.
- Leitman, D.I., Wolf, D.H., Ragland, J.D., Laukka, P., Loughead, J., Valdez, J.N., Javitt, D.C., Turetsky, B.I., Gur, R.C., 2010b. It's not what you say, but how you say it: a reciprocal temporo-frontal network for affective prosody. *Front. Hum. Neurosci.* 4, 1–13.
- Lerner, Y., Papo, D., Zhdanov, A., Belozersky, L., Hendler, T., 2009. Eyes wide shut: amygdala mediates eyes-closed effect on emotional experience with music. *PLoS ONE* 4, e6230.
- Levitin, D.J., Menon, V., Schmitt, J.E., Eliez, S., White, C.D., Glover, G.H., Kadis, J., Korenberg, J.R., Bellugi, U., Reiss, A.L., 2003. Neural correlates of auditory perception in Williams syndrome: an fMRI study. *Neuroimage* 18, 74–82.
- Liem, F., Hürschler, M.A., Jancke, L., Meyer, M., 2014. On the planum temporale lateralization in suprasegmental speech perception: evidence from a study investigating behavior, structure, and function. *Hum. Brain Mapp.* 35, 1779–1789.
- Liu, S., Chow, H.M., Xu, Y., Erkkänen, M.G., Swett, K.E., Eagle, M.W., Rizik-Baer, D.A., Braun, A.R., 2012. Neural correlates of lyrical improvisation: an fMRI study of freestyle rap. *Sci. Rep.* 2, 834.
- Maguire, E.A., 2001. Neuroimaging, memory and the human hippocampus. *Rev. Neurol. (Paris)* 157, 791–794.
- Malone, B.J., Scott, B.H., Semple, M.N., 2007. Dynamic amplitude coding in the auditory cortex of awake rhesus macaques. *J. Neurophysiol.* 98, 1451–1474.
- McDermott, J.H., Lehr, A.J., Oxenham, A.J., 2010. Individual differences reveal the basis of consonance. *Curr. Biol.* 20, 1035–1041.
- Meyer, L.B., 1956. *Emotion and Meaning in Music*. The University of Chicago Press, Chicago.
- Milesi, V., Cecic, S., Peron, J., Frühholz, S., Cristinzio, C., Seck, M., Grandjean, D., 2014. Multimodal emotion perception after anterior temporal lobectomy (ATL). *Front. Hum. Neurosci.* 8, 275.
- Mitchell, R.L., 2006. How does the brain mediate interpretation of incongruent auditory emotions? The neural response to prosody in the presence of conflicting lexico-semantic cues. *Eur. J. Neurosci.* 24, 3611–3618.
- Mitchell, R.L., Elliott, R., Barry, M., Cruttenden, A., Woodruff, P.W., 2003. The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia* 41, 1410–1421.
- Mitterschiffthaler, M.T., Fu, C.H., Dalton, J.A., Andrew, C.M., Williams, S.C., 2007. A functional MRI study of happy and sad affective states induced by classical music. *Hum. Brain Mapp.* 28, 1150–1162.
- Mohedano-Moriano, A., Pro-Sistiaga, P., Arroyo-Jimenez, M.M., Artacho-Perula, E., Insausti, A.M., Marcos, P., Cebada-Sanchez, S., Martinez-Ruiz, J., Munoz, M., Blaizot, X., Martinez-Marcos, A., Amaral, D.G., Insausti, R., 2007. Topographical and laminar distribution of cortical input to the monkey entorhinal cortex. *J. Anat.* 211, 250–260.
- Morris, J.S., Scott, S.K., Dolan, R.J., 1999. Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia* 37, 1155–1163.
- Mothes-Lasch, M., Mentzel, H.J., Miltner, W.H., Straube, T., 2011. Visual attention modulates brain activation to angry voices. *J. Neurosci.* 31, 9594–9598.
- Mueller, K., Mildner, T., Fritz, T., Lepsien, J., Schwarzbauer, C., Schroeter, M.L., Moller, H.E., 2011. Investigating brain response to music: a comparison of different fMRI acquisition schemes. *Neuroimage* 54, 337–343.
- Mutschler, I., Wieckhorst, B., Speck, O., Schulze-Bonhage, A., Hennig, J., Seifritz, E., Ball, T., 2010. Time scales of auditory habituation in the amygdala and cerebral cortex. *Cereb. Cortex* 20, 2531–2539.

- Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Ito, K., Kojima, S., 2001. Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia* 39, 1047–1054.
- Nieuwenhuys, R., 1996. The greater limbic system, the emotional motor system and the brain. *Prog. Brain Res.* 107, 551–580.
- Ohman, A., Mineka, S., 2001. Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522.
- Omar, R., Henley, S.M., Bartlett, J.W., Hailstone, J.C., Gordon, E., Sauter, D.A., Frost, C., Scott, S.K., Warren, J.D., 2011. The structural neuroanatomy of music emotion recognition: evidence from frontotemporal lobar degeneration. *Neuroimage* 56, 1814–1821.
- Pallesen, K.J., Brattico, E., Bailey, C., Korvenoja, A., Koivisto, J., Gjedde, A., Carlson, S., 2005. Emotion processing of major, minor, and dissonant chords: a functional magnetic resonance imaging study. *Ann. N. Y. Acad. Sci.* 1060, 450–453.
- Pallesen, K.J., Brattico, E., Bailey, C.J., Korvenoja, A., Gjedde, A., 2009. Cognitive and emotional modulation of brain default operation. *J. Cogn. Neurosci.* 21, 1065–1080.
- Panksepp, J., 1995. The emotional sources of chills induced by music. *Music Percept.* 13, 171–207.
- Panksepp, J., Bernatzky, G., 2002. Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behav. Processes* 60, 133–155.
- Park, M., Hennig-Fast, K., Bao, Y., Carl, P., Poppel, E., Welker, L., Reiser, M., Meinl, T., Gutyrchik, E., 2013. Personality traits modulate neural responses to emotions expressed in music. *Brain Res.* 1523, 68–76.
- Patel, S., Scherer, K.R., Bjorkner, E., Sundberg, J., 2011. Mapping emotions into acoustic space: the role of voice production. *Biol. Psychol.* 87, 93–98.
- Pell, M.D., Kotz, S.A., 2011. On the time course of vocal emotion recognition. *PLoS ONE* 6, e27256.
- Pereira, C.S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S.L., Brattico, E., 2011. Music and emotions in the brain: familiarity matters. *PLoS ONE* 6, e27241.
- Phelps, E.A., LeDoux, J.E., 2005. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187.
- Phillips, M.L., Young, A.W., Scott, S.K., Calder, A.J., Andrew, C., Giampietro, V., Williams, S.C., Bullmore, E.T., Brammer, M., Gray, J.A., 1998. Neural responses to facial and vocal expressions of fear and disgust. *Proc. Biol. Sci./R. Soc.* 265, 1809–1817.
- Pitkanen, A., Pikkariainen, M., Nurminen, N., Ylinen, A., 2000. Reciprocal connections between the amygdala and the hippocampal formation, perirhinal cortex, and postrhinal cortex in rat – a review. *Parahippocampal Reg.* 911, 369–391.
- Plailly, J., Tillmann, B., Royet, J.P., 2007. The feeling of familiarity of music and odors: the same neural signature? *Cereb. Cortex* 17, 2650–2658.
- Poeppl, D., 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun.* 41, 245–255.
- Polston, J.E., Rubbinaccio, H.Y., Morra, J.T., Sell, E.M., Glick, S.D., 2011. Music and methamphetamine: conditioned cue-induced increases in locomotor activity and dopamine release in rats. *Pharmacol. Biochem. Behav.* 98, 54–61.
- Preuss, A., Muller-Preuss, P., 1990. Processing of amplitude modulated sounds in the medial geniculate body of squirrel monkeys. *Exp. Brain Res.* 79, 207–211.
- Quadflieg, S., Mohr, A., Mentzel, H.J., Miltner, W.H., Straube, T., 2008. Modulation of the neural network involved in the processing of anger prosody: the role of task-relevance and social phobia. *Biol. Psychol.* 78, 129–137.
- Rauch, S.L., Shin, L.M., Dougherty, D.D., Alpert, N.M., Orr, S.P., Lasko, M., Macklin, M.L., Fischman, A.J., Pitman, R.K., 1999. Neural activation during sexual and competitive arousal in healthy men. *Psychiatry Res.* 91, 1–10.
- Reser, D.H., Burman, K.J., Richardson, K.E., Spitzer, M.W., Rosa, M.G., 2009. Connections of the marmoset rostrotemporal auditory area: express pathways for analysis of affective content in hearing. *Eur. J. Neurosci.* 30, 578–592.
- Richardson, M.P., Strange, B.A., Dolan, R.J., 2004. Encoding of emotional memories depends on amygdala and hippocampus and their interactions. *Nat. Neurosci.* 7, 278–285.
- Rivier, F., Clarke, S., 1997. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage* 6, 288–304.
- Romanski, L.M., Averbeck, B.B., 2009. The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu. Rev. Neurosci.* 32, 315–346.
- Rota, G., Handjaras, G., Sitaram, R., Birbaumer, N., Dogil, G., 2011. Reorganization of functional and effective connectivity during real-time fMRI-BCI modulation of prosody processing. *Brain Lang.* 117, 123–132.
- Russell, J.A., 1980. A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178.
- Russell, J.A., 2003. Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172.
- Salimpoor, V.N., Benovoy, M., Larcher, K., Dagher, A., Zatorre, R.J., 2011. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* 14, 257–262.
- Salimpoor, V.N., van den Bosch, I., Kovacevic, N., McIntosh, A.R., Dagher, A., Zatorre, R.J., 2013. Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science* 340, 216–219.
- Samson, S., Zatorre, R.J., 1991. Recognition memory for text and melody of songs after unilateral temporal lobe lesion: evidence for dual encoding. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 793–804.
- Sander, D., Grafman, J., 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., Vuilleumier, P., 2005. Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage* 28, 848–858.
- Sander, K., Brechmann, A., 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., Frome, Y., Scheich, H., 2007. FMRI activations of amygdala, cingulate cortex, and auditory cortex by infant laughing and crying. *Hum. Brain Mapp.* 28, 1007–1022.
- Sander, K., Scheich, H., 2005. Left auditory cortex and amygdala, but right insula dominance for human laughing and crying. *J. Cogn. Neurosci.* 17, 1519–1531.
- Sauter, D.A., Eisner, F., Calder, A.J., Scott, S.K., 2010. Perceptual cues in nonverbal vocal expressions of emotion. *Q. J. Exp. Psychol.* 63, 2251–2272.
- Scherer, K., 2004. Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them?. *J. New Music Res.* 33, 239–251.
- Scherer, K.R., 1984. On the nature and function of emotion: a component process approach. In: Scherer, K.R., Ekman, P. (Eds.), *Approaches to Emotion*. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, pp. 293–317.
- Scherer, K.R., 1994. Affect bursts. In: van Goozen, S., van de Poll, N.E., Sergeant, J.A. (Eds.), *Emotions: Essays on Emotion Theory*. Erlbaum, Hillsdale, NJ, pp. 161–196.
- Scherer, K.R., 1995. Expression of emotion in voice and music. *J. Voice* 9, 235–248.
- Schirmer, A., Escoffier, N., Zysset, S., Koester, D., Striano, T., 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.
- Schirmer, A., Fox, P.M., Grandjean, D., 2012. On the spatial organization of sound processing in the human temporal lobe: a meta-analysis. *Neuroimage* 63, 137–147.
- Schlag, G., Altenmüller, E., Thaut, M., 2010. Music listening and music making in the treatment of neurological disorders and impairments. *Music Percept.* 27, 249–250.
- Schrober, M., 2003. Experimental study of affect bursts. *Speech Commun.* 40, 99–116.
- Scott, S.K., Young, A.W., Calder, A.J., Hellowell, D.J., Aggleton, J.P., Johnson, M., 1997. Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature* 385, 254–257.
- Sprengelmeyer, R., Young, A.W., Schroeder, U., Grossenbacher, P.G., Federlein, J., Buttner, T., Przuntek, H., 1999. Knowing no fear. *Proc. Biol. Sci.* 266, 2451–2456.
- Stienen, B.M., Tanaka, A., de Gelder, B., 2011. Emotional voice and emotional body postures influence each other independently of visual awareness. *PLoS ONE* 6, e25517.
- Sundberg, J., Patel, S., Bjorkner, E., Scherer, K.R., 2011. Interdependencies among voice source parameters in emotional speech. *IEEE Trans. Affect. Comput.* 2, 162–174.
- Szameitat, D.P., Kreifelts, B., Alter, K., Szameitat, A.J., Sterr, A., Grodd, W., Wildgruber, D., 2010. It is not always tickling: distinct cerebral responses during perception of different laughter types. *Neuroimage* 53, 1264–1271.
- Thayer, R.E., 1989. *The Biopsychology of Mood and Arousal*. Oxford University Press, New York.
- Trost, W., Ethofer, T., Zentner, M., Vuilleumier, P., 2012. Mapping aesthetic musical emotions in the brain. *Cereb. Cortex* 22, 2769–2783.
- Trost, W., Vuilleumier, P., 2013. Rhythmic entrainment as a mechanism for emotion induction by music: a neurophysiological perspective. In: Cochrane, T., Fantini, B., Scherer, K.R. (Eds.), *The Emotional Power of Music*. Oxford University Press, Oxford.
- Von Kriegstein, K., Patterson, R.D., Griffiths, T.D., 2008. Task-dependent modulation of medial geniculate body is behaviourally relevant for speech recognition. *Curr. Biol.* 18, 1855–1859.
- Vuilleumier, P., 2005. How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594.
- Warren, J.E., Sauter, D.A., Eisner, F., Wiland, J., Dresner, M.A., Wise, R.J., Rosen, S., Scott, S.K., 2006. Positive emotions preferentially engage an auditory-motor mirror system. *J. Neurosci.* 26, 13067–13075.
- Warrier, C.M., Abrams, D.A., Nicol, T.G., Kraus, N., 2011. Inferior colliculus contributions to phase encoding of stop consonants in an animal model. *Hear. Res.* 282, 108–118.
- Watanabe, T., Yagishita, S., Kikyo, H., 2008. Memory of music: roles of right hippocampus and left inferior frontal gyrus. *Neuroimage* 39, 483–491.
- Wattendorf, E., Westermann, B., Fiedler, K., Kaza, E., Lotze, M., Celio, M.R., 2013. Exploration of the neural correlates of ticklish laughter by functional magnetic resonance imaging. *Cereb. Cortex* 23, 1280–1289.
- Weninger, F., Eyben, F., Schuller, B.W., Mortillaro, M., Scherer, K., 2013. On the acoustics of emotion in audio: what speech, music, and sound have in common. *Front. Psychol.* 4, 1–12.
- Wieser, H.G., Mazzola, G., 1986. Musical consonances and dissonances: are they distinguished independently by the right and left hippocampi? *Neuropsychologia* 24, 805–812.
- Wiethoff, S., Wildgruber, D., Grodd, W., Ethofer, T., 2009. Response and habituation of the amygdala during processing of emotional prosody. *Neuroreport* 20, 1356–1360.
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., Ethofer, T., 2008. Cerebral processing of emotional prosody – influence of acoustic parameters and arousal. *Neuroimage* 39, 885–893.

- Wildgruber, D., Ethofer, T., Grandjean, D., Kreifelts, B., 2009. A cerebral network model of speech prosody comprehension. *Int. J. Speech Lang. Pathol.* 11, 277–281.
- Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., Ackermann, H., 2004. Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. *Cereb. Cortex* 14, 1384–1389.
- Wildgruber, D., Pihan, H., Ackermann, H., Erb, M., Grodd, W., 2002. Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters, emotional valence, and sex. *Neuroimage* 15, 856–869.
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., Ackermann, H., 2005. Identification of emotional intonation evaluated by fMRI. *Neuroimage* 24, 1233–1241.
- Williams, H., 2004. Birdsong and singing behavior. *Ann. N. Y. Acad. Sci.* 1016, 1–30.
- Wittfoth, M., Schroder, C., Schardt, D.M., Dengler, R., Heinze, H.J., Kotz, S.A., 2010. On emotional conflict: interference resolution of happy and angry prosody reveals valence-specific effects. *Cereb. Cortex* 20, 383–392.
- Zald, D.H., 2003. The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res.: Brain Res. Rev.* 41, 88–123.
- Zatorre, R.J., Salimpoor, V.N., 2013. From perception to pleasure: music and its neural substrates. *Proc. Natl. Acad. Sci. U. S. A.* (110 Suppl), 10430–10437.
- Zentner, M., 2010. Homer's prophecy: an essay on music's primary emotions. *Music Anal.* 29, 102–125.
- Zentner, M., Grandjean, D., Scherer, K.R., 2008. Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion* 8, 494–521.