ERC Starting Grant 2013
Research proposal [Part B2]
CREAM - Cracking the Emotional code of Music

"Music can lift us out of depression or move us to tears - it is a remedy, a tonic, orange juice for the ear. But for many of my neurological patients, music is even more - it can provide access, even when no medication can, to movement, to speech, to life. For them, music is not a luxury, but a necessity."
- Oliver Sacks, *Musicophilia: Tales of Music and the Brain*

Section a. State-of-the-art and objectives

State-of-the-art:

Music has incredible power on our emotions: it can make the lone listener cry with joy or sorrow, and a crowd of thousands jump and shout as one. The effects of music cover the complete range of reactions commonly associated with human emotions: subjective sensation (Zentner & Eerola, 2010), physiological changes (Bartlett, 1996), facial expression (Witvliet & Vrana, 2007) to name but a few. These effects are remarkable, because music as a stimulus is not directly associated with the furthering or hindrance of our physical needs and goals. Cognitive psychology has much to say about why one may cry, run and sweat upon encountering a bear during a walk in the woods (James, 1884). Why Beethoven or Lady Gaga should have the same effect on us, though, remains a biological mystery. This is a problem which science and general media alike have become fascinated with in recent years (see e.g. Levitin, 2007; Sacks, 2009; Mannes, 2011). Solving it holds promise for important advances in cognition and therapeutic applications (see e.g. Sarkamö et al., Brain, 2008)

With the increasing availability of brain imaging technology in the past ten years, cognitive neuroscience has gathered a wealth of insights into what patterns of activity musical emotions manifest in the human brain. Functional exploration has shown that musical emotions involve the sub-cortical regions of the limbic system (e.g. thalamus, hippocampus, amygdala - Blood & Zatorre, 2001) as well as cortical regions such as the pre-frontal and orbitofrontal cortices, the insula and nucleus accumbeus (Koelsch et al., 2006). Lesion studies have shown that emotional processing is at least partly distinct from more general musical skills: Peretz (2001) reports on a patient, I.R., with auditory cortex damage, who was unable to identify musical extracts with which she was yet well familiar (e.g. Albinoni's well-known Adaggio), but who could still identify the emotions they expressed. Finally, cognitive psychology has demonstrated that musical emotions can be processed at the pre-attentive, reflex level (Bigand et al, 2005), that they're at least partly universal (Fritz et al., 2009) and innate (Dalla Bella at al., 2001).

Yet, despite a growing body of literature on the subject (OUP's Handbook of Music and Emotion, 2010 Edition, has now reached a daunting 975 pages), the neural bases of our affective processing of music remain elusive. In fact, the more we know, the less we seem to understand: looking for a fixed neural architecture (Peretz & Zatorre, 2005), neuroscientists are instead finding vast and distributed networks of activity (Koelsch at al., 2006). Looking for hemispheric specialization (Peretz, 2001), we're finding left preference (Schmidt & Taylor, 2001), or right (Blood & Zatorre, 2001) or both (Peretz & Zatorre, 2005). Looking for dedicated circuits, we're finding large overlap with regions involved in the emotional interpretation of either expressive voice (Juslin & Laukka, 2003) or environmental sounds (Goydke et al. 2004).

Meanwhile, the field of informatics has also seen growing interest and expertise on how to model musical sounds, a discipline known as *Music Information Retrieval* (MIR). Identifying what acoustical or sensory features of musical sounds trigger emotional reactions has been an enduring concern of music emotion scientists, of course. In early studies, when the use of computers was still impractical, music was analysed "by ear" by musicologists and its features (pitch, tempo, harmony) correlated to what emotion the experimental participants may report (from Watson, 1942; up to e.g. Thompson & Balkwill, 1999). However, the recent progress of computer techniques to analyze sounds has lead to a rapid change of methodology: algorithms now exist to measure pitch (Klapuri, 2008) or tempo (Gouyon et al. 2006) automatically. Further, novel constructs have been introduced that had never been considered before:

physical properties of the musical signal (for instance, its root-mean-square, a measure of physical energy, or the ZCR, the rate of zero-crossings of the waveform), mathematical properties (e.g. the signal's entropy) and even models of human sensory processing (for instance, "Mel-Frequency Cepstrum Coefficients", a mathematical construct derived from the signal's Fourier transform, which approximates the response of the human cochlea - Peeters et al., 2010). A software library such as the MIRToolbox (Lartillot & Toiviainen, 2007) now offers more than 300 of such constructs, all algorithmic evaluations of the musical signal, making possible the automatic analysis of hundreds of music extracts, without needing subjective human evaluation.

This recent surge of analytical power has made new insights possible into musical emotions. Alluri and Toiviainen (2010) show that a multiple regression on 26 of such acoustical characteristics explain 60% of the variance of human evaluations of valence and arousal collected for 100 musical extracts. Characteristics that correlate most with their empirical data include e.g. for valence, the spectral centroid, the high frequency energy ratio, and the third statistical moment of the energy spectrum ("spectral skewness"). Similarly, Liu and Zhang (2006) simulate with more than 95% agreement human judgements of "depression" and "contentment" produced after listening to more than 250 songs, combining automatic characteristics describing their timbre (e.g. the spectral centroid), rhythm (e.g. the authors' own "average autocorrelation peak") and intensity (e.g. the root-mean-square).

Valence	
Audio characteristics	β
tonal_chromagram_peak_PeakMagPeriodEntropy	-0.75
tonal_mode_Mean	0.13
spectral_mfcc_PeriodAmp_8 (600 Hz +/- 66)	0.12
spectral_brightness_Slope	-0.12
spectral_ddmfcc_PeriodEntropy_1 (133.3)	-0.11
spectral_mfcc_PeriodEntropy_9 (666.6)	-0.11
spectral_dmfcc_Mean_3 (266.6)	-0.1
rhythm_attack_slope_PeriodEntropy	0.10
spectral_spread_PeriodAmp	-0.09
spectral_mfcc_PeriodAmp_11 (800)	-0.06
Arousal	
Audio characteristics	β
spectral_mfcc_Std_6 (466.6)	-0.34
spectral_mfcc_Mean_3 (266.6)	0.28
tonal_keyclarity_Std	-0.28
spectral_mfcc_Std_7 (533.3)	0.24
tonal_mode_Std	-0.19
spectral_irregularity_PeriodAmp	-0.18
spectral_mfcc_Mean_13 (933.3)	-0.15
spectral_dmfcc_PeriodAmp_4 (333.3)	0.14
tonal_chromagram_peak_PeakPosPeriodFreq	0.13
tonal_keyclarity_Slope	-0.11

Table 1: Acoustic characteristics that best correlate with the valence and arousal of a set of musical stimuli, as computed by the MIRToolbox. Despite near-perfect predictive power, these characteristics are too mathematical to yield any cognitive or physiological interpretation. Data reproduced from Aucouturier & Bigand, J. Intell. Inf. Syst., 2012.

However, despite reaching near-perfect predictive power (there is now an annual competition of algorithms for "music mood classification", see Kim et al. (2010) for a review), the MIR approach to musical emotions has failed to add much to what little understanding cognitive neurosciences had already brought to the problem. This is best illustrated with an example, taken from our own collaboration with the team of Bigand et al. (2005). Valence and arousal ratings were collected for a dataset of short musical extracts, and a multiple regression was conducted using the full predictive power of the MIRToolbox in an attempt to find what acoustical characteristics of sound *create* the emotions reported by the subjects. As shown in Table 1, we "found" that the valence of the music is very well "explained" by the *entropy of the period of the magnitude of the maximum peak detected every 50ms in the signal's chromagram* (the chromagram, yet-another MIR construct, gives, at every time step, the energy found in the signal's Fourier spectrum in the frequency bands corresponding to each note of the octave - c, c#, d, etc.). Arousal, on the other hand, was found to result from the variance of the 6th Mel-Frequency Cepstrum Coefficient and the average of the 3rd - but mind you, not reciprocally. Such explanations are obviously superficial. The factors found to explain emotions by such methods are too mathematical to yield any cognitive or physiological interpretation. Nothing seems to link them to the evidence gathered by cognitive neuroscience.

In summary, despite 10 years of progress of both fronts, we still do not know what type of musical signal and what type of listening condition is likely to activate this emotional mechanism or another, this brain area or another, this emotion or another. The "emotional code" of music remains to be cracked.

In our view, the current gap between the two approaches of cognitive neuroscience and music information retrieval has two causes. First, the emotional reactions examined by empirical research are not sufficiently well-controlled. There is growing agreement, in the cognitive neuroscience community, that music is in fact able to activate not only one, but more likely several, distinct mechanisms of emotional induction: it may prompt emotional images, remind emotional memories of past events, surprise by a particularly loud and sudden sound, etc. (see Juslin & Vastfjäll, 2008 for a tentative list of such mechanisms). Yet no methodological tool exists to selectively activate one mechanism or the other; most likely, we experimentalists trigger the whole range, indiscriminately. Hence the many, diverse, concurrent, contradictory patterns of activity that have been typically pondered upon by brain imaging, for which it is unrealistic to expect to find a common MIR invariant. Second, the sonic characteristics offered by MIR as an explanation are not targeted at any of such mechanisms in particular. Rather, they are in themselves a formidable mix of mechanisms, low and high-level (from the MFCC's model of the cochlea to the chromagram's theory of western music harmony), short or long-scale (from the Fourier spectrum's 50 milliseconds to the entropy thereof, every minute), some even cognitively implausible. It is unclear how any combination of such characteristics could be found explaining any mechanism of emotion induction in particular.

Objectives:

Our goal is to create a paradigm shift in how MIR signal processing is used in cognition research: not to *observe* stimuli of interest a posteriori (i.e. signal processing as a sophisticated probe), but to *generate* stimuli that selectively activate mechanisms (i.e. signal processing as a musical scalpel).

The originality of our approach is to combine a very high level of technicality in the computerized analysis of sound (based on our 10 years contributing to the emergence of the field of MIR) with in-depth understanding of the cognitive neurosciences of music and sound (made possible by our post-doctoral and early-career experience in the field). Using this rare inter-disciplinary expertise, we will introduce several new computational tools and make important and unprecedented progress towards effectively *cracking* the emotional code of music.

The project stands on two legs, one fundamental, the other applicative. Our fundamental research objectives are to study two important - yet, barely understood - mechanisms of emotional induction by music, and to characterize what exact type of musical signal is liable to activate them. Our applicative research objectives are, then, to use this operational knowledge to develop non-invasive cognitive technology aiming to improve cognitive functioning and emotional well-being for patients suffering from cognitive pathologies.

The two mechanisms we choose to study have come to the community's recent attention as the most likely candidates to explain why musical emotions activate neuronal circuits that are specialized for other types of stimuli (Juslin & Vastfjäll, 2008). According to mechanism 1, musical expression is, at least partly, an amplification of the emotional expression involved in spoken language: music's trembling notes, hesitating phrases, bright or dark timbres may well be "heard" and processed as if they were voice (Juslin & Laukka, 2003). According to mechanism 2, certain musical sounds, when loud, fast or dissonant, trigger the same reflex emotional reactions than would negative, survival-relevant environmental sounds (Goydke et al., 2004). Both mechanisms form a true bottleneck in our current understanding of musical emotions: they have been postulated, but never formally tested. It is our objective to do so.

<u>1) Mechanism 1: Recycling the circuits of emotional voice</u>

The hypothesis that music should invade circuits responsible for the emotional processing of speech originated in a meta-study by Juslin and Laukka (2003), which reviewed 104 psycho-acoustical studies of emotional expression in speech and 41 studies of music. The authors observe that many emotional cues are common to both modalities: for instance, both happy music and happy speech are often of a fast tempo/rate, of a medium to high intensity, of a high median pitch/F0, while sad music and sad speech are typically slower and of a lower pitch/F0. It was later proposed (Juslin & Västfjäll, 2008) that music could function as a super-stimulus which even surpassed the capacities of the human vocal apparel, to trigger even higher levels of speech-like emotions.

Despite receiving much attention in the community, the hypothesis still lacks empirical support. First, more recent studies have revealed important modality-specific cues which were not accounted for in Juslin and Laukka's meta-study (Ilie & Thompson, 2006), and it is unclear whether all or only specific types of music are processed as speech. Second, brain imaging studies have been inconclusive on whether speech and music emotional processing overlap or not: while emotional speech is known to activate a network of right frontal and temporal areas (Schirmer & Kotz, 2006), such areas have not always been observed in musical tasks (see Peretz, 2001 for a review). Proof is also lacking at the sub-cortical level: curiously, it is not yet established whether the amygdala, which is involved in fearful music processing (Blood & Zatorre, 2001), is implicated in emotional speech perception (Anderson & Phelphs, 1998).

Our first objective is to test the hypothesis that music invades circuits of emotional speech processing. To do so, we will introduce a new computational methodology to synthesize musical stimuli which *sound like* emotional speech, and compare the emotional reactions to both types of stimuli. In doing so, we will seek to, first, obtain a computational description of what exact type of music signals are liable to activate the mechanism (so refutable predictions can be made and therapeutic protocols can be developed); second, to determine the brain areas involved in this mechanism (our proposal uses a brain imaging technique called near-infrared spectroscopy - NIRS); third, to test the mechanism's boundaries in inter-cultural, inter-linguistic experiments.

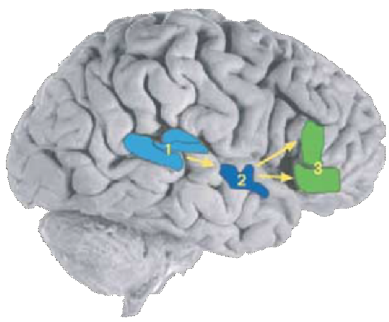


Figure 1: Processing of emotional speech is known to activate a network of right frontal and temporal areas, including (1) the Superior Temporal Gyrus, (2) the right Superior Temporal Sulcus where auditory representations for emotions are formed and (3) the right Inferior Front Gyrus, where emotions are evaluated and (possibly) made available to consciousness. It remains an open hypothesis whether music invades the same circuits when its emotions are processed. Figure adapted from Schirmer & Kotz, 2006

2) Mechanism 2: Recycling the circuits of environmental sounds

The hypothesis that music should invade circuits responsible for the emotional evaluation of environmental sounds has been proposed by Juslin and Vastfjall (2008) to account for several intriguing aspects of musical emotions: the fact that emotional induction happens even with very short musical extracts, in less than 250ms (Bigand et al. 2005); that it involves sub-cortical regions (brainstem, inferior colliculus) known to be attuned to environment evaluation (Blood & Zatorre, 2011); that it can be processed pre-attentively (Goydke, 2004); and it is partly universal (Fritz et al, 2009). According to this mechanism, emotion is triggered in a reflex-like manner when the central auditory system detects certain acoustic characteristics in the musical stimuli which qualifies it as a signal for urgent and important changes in the environment: for instance, a loud and sudden sound, a sharp transient, a dissonant tone.

As for the previous mechanism (recycling the circuits of emotional voice), this mechanism is often invoked as an explanation for empirical effects, but it has never been tested. A study like Blumstein, Bryant and Kaye (2012) shows for instance that noise and abrupt frequency transitions, when added to music, reduces its valence and arousal, and links these characteristics to the raucous sounds of animal distress calls, but it remains unclear whether it is indeed the same perceptive/cognitive mechanism at play in both types of stimuli. If anything, environmental sounds and music of similar emotional valence are often found to activate contradictory patterns of physiological responses: breathing rate increases with negative sounds, but positive music; skin conductance level increases with arousal ratings for music but not for noises (Gomez & Danuser, 2004). Careful stimulus control is therefore needed, which require a change of methodology.

Our second objective is to test the hypothesis that music invades circuits responsible for the emotional evaluation of environmental sounds. To do so, we will introduce a new computational methodology to select musical stimuli which *sound like* environmental sounds, and compare the emotional reactions to both types of stimuli. As before, we will seek to, first, to obtain a computational description of what exact type of music signals are liable to activate the mechanism; second, to determine the neuro-physiological reactions involved in this mechanism (using electroencephalography); third, to test the mechanism's boundaries in inter-cultural experiments and infant studies.

3) Therapeutic applications

Much of the recent interest in how music creates emotions is linked to the promise of new tools for cognitive stimulation and non-invasive therapies (Sacks, 2009). In clinical studies, music has been linked to effects on pain (Mitchell et al., 2007) and motor rehabilitation (Rojo et al., 2011). Because music interacts with the language centers of the brain (Patel, 2008), it is believed to facilitate strong transfer effects for linguistic rehabilitation (Sarkämö et al, 2008). Because it affects the limbic system, which is in part responsible for hormonal regulation, it is believed to stimulate neuronal plasticity (Fukui & Toyoshima, 2008). The non-invasive, non-pharmacological treatment protocols that could be based on such ideas hold much societal promise. However, until now, we still do not know what type of musical stimulus is likely to activate a given emotional mechanism or another. This is strongly limiting the practical implementation of such protocols.

The fundamental breakthroughs brought by this project will unlatch the therapeutic potential of musical emotions. First, our study of mechanisms 1 and 2 will provide, for the first time in the history of the field, a precise characterization of the type of sounds that activate them, i.e. a cognitive technology able to select the most appropriate stimuli in order to mobilize one neuronal pathway or another. Second, the 2 mechanisms considered in our proposal in particular are especially well-adapted to emotional modulation in clinical contexts: they are fast, pre-attentive, and little prone to individual and cultural differences. Finally, the newly-gained operative knowledge over mechanism 1 in particular will give us a way to specifically activate brain areas related to voice processing (right frontal, temporal), yet without requiring access to the purely linguistic areas that may be impaired in stroke-related aphasia (left frontal). This targeted mobilization, by music, of language areas has potential to serve in a linguistic rehabilitation program for stroke patients.

Our third objective is to apply the knowledge gained on mechanisms 1 and 2 in the context of clinical studies. Using institutional partnerships with university hospital in Dijon, Lyon and Grenoble (France), we will implement and test 3 types of treatment protocols. First, use musical stimuli constructed to sound like emotional speech in order to stimulate the linguistic rehabilitation in patients of stroke-related aphasia. Second, use real-time emotional transformations of a patient's own speech to induce positive emotions in fibromyalgic patients, and third, to reduce the consolidation of traumatic memories in PTSD patients. The methodological details of these studies are to be found below.

On other mechanisms, and why we do not pursue them here

As a final note to this section, it is to be noted that the two phenomenon investigated in this proposal (music's interaction with emotional speech, and with environmental sounds) are only part of a potentially larger number of mechanisms by which music induces emotions. Leonard Meyer's theory of musical expectancy (Meyer, 1956) is one of such, according to which the syntactic sequence of musical events creates expectations of what follows next, and emotions are created when these expectations are met or violated (see Huron, 2006 for a recent review). Juslin & Vastfjall (2008) offered a list of at least 4 other candidates, including e.g. music's recalling happy or sad personal episodes of the listener's personal history or music's spontaneous invoking emotional mind-eye imagery. All of these mechanisms, and quite possibly more, deserve attention.

However, it is quite deliberate that this proposal should focus on only two. First, contrary to some more researched alternatives, these two mechanisms are in especially dire need of empirical validation: our work will fill a unique gap in the theory of emotions. Second, more than exhaustiveness, our goal with this proposal is to establish a general methodology which, once established, can then be extended to other mechanisms of musical emotions, and auditory cognition at large. Pragmatically, the two mechanisms of speech and environmental sounds are simply the best first steps in this research program: they involve sensory processes for which recent signal processing models offer particularly creative solutions. Finally, we focused this proposal on two mechanisms in an effort to balance the project's impact on both fundamental and applicative research, the latter of which has too often been postponed to wishful "future work". Rather than pure fundamental knowledge, our project purposes to create cognitive technology (techniques to "engineer" musical stimuli that selectively activate one neural pathway or another), and we are therefore affecting time and resources in the explicit intent to use it in clinical studies.

Impact:

If successful, the project will impact the field of music cognition to the point of changing both *what* we know of musical emotions, but also *how* we acquire this knowledge. The project will:

- stop the current trend of using MIR signal processing techniques to correlate musical stimuli to their indistinct emotional reactions, a strategy that after 10 years of development has remained mostly fruitless, and
- reposition MIR as a powerful means of experimental control, able to generate stimuli that can selectively activate specific mechanisms of musical emotions.

This new methodological framework is a profound change of paradigm, which has potential to change the field of music cognition and establish new avenues for interdisciplinary research between the life sciences and information sciences and technologies.

In the life sciences, this new paradigm will generate, for the first time in the history of the field, a computational characterization of what exact type of musical stimuli is able to activate one emotional pathway or another, one emotion or another. In other words, we will be able to create musical stimuli tailored to induce a specific emotion in a specific individual. Such "cognitive technology" will provide disruptive new tools for empirical research (e.g. functional exploration with brain imaging) and for non-pharmacological therapeutic applications (e.g. emotion modulation for patients with affective disorders).

In the information sciences, the new formulation of MIR as a tool for experimental control in cognitive neurosciences will trigger programs of research aiming to establish novel algorithms with better biological and cognitive validity - initiating in audio research the same movement that changed the field of image processing in the past 5-10 years (Serre et al., 2007). MIR algorithms will evolve to become more robust, more tailored to individual listeners, and more efficient, all challenges currently faced by the community without much of a program to address them so far (Aucouturier & Bigand, J. Intell. Inf. Syst., 2012).

Section b. Methodology

Three new tools to crack the code

The originality of our methodology is to combine the experimental logics of cognitive psychology and neuroscience with the algorithms of Music Information Retrieval. This is made possible by our unique career trajectory, spanning the 2 fields. This interdisciplinary will allow a double breakthrough: 1) to empirically isolate 2 specific mechanisms of emotion induction and 2) to provide a computational characterization of the type of stimuli likely to activate them. To this aim, we introduce several new experimental and computational tools.

1) Very short sounds

First, we take a radical stance and propose to start our study with the special case of very short (<1 sec.) musical sounds. While the vast majority of music emotion research utilizes stimuli long enough for a melodic or rhythmic structure to emerge (tens of seconds, or even minutes), a small number of recent studies suggest that emotions can be induced by much shorter extracts. Watt & Ash (1998) found consistent emotional ratings with extracts of Wagner Opera of duration 3-5 seconds. Peretz, Gagnon & Bouchard (1998) extended this study to even shorter sounds, and showed that only 250ms of sounds were sufficient to differentiate sad and happy music. Bigand et al. (2005) later confirmed this ability with a wider range of emotions, while Gjerdingen & Perrot (2008) replicated these results with durations down to an impressive 100ms. Our own study in JASA (Aucouturier & Defreville, 2009) extended these results to environmental sounds.

Our first experimental tool, and the first originality of our approach, is therefore to focus on very short, infra-second musical extracts. This is a bold-but-powerful means of experimental control, related to stimulus impoverishment: within one second, no melody can be recognized, no expectation can be built up, no mental image can be conjured; many mechanisms that would otherwise confound our study of mechanism 1 and 2 simply do not have time to be activated (Table 2). In addition, short signals lend themselves to easier computational characterization than longer textures (as our own work first demonstrated - Pattern Recognition Letter, 2007). Starting with short sounds is therefore a powerful wedge to crack into the complexity of the general problem: first, understand the code at the scale of 1 second, then study its temporal integration to tens of seconds and minutes.

2) Real-time acoustic manipulations for spoken voice

The second methodological innovation of our project is to use a new signal processing technique that we recently developed (for a different purpose) in collaboration with colleagues in UCL (London, UK) and Lund (Sweden). In order to study emotional influences on cognition (see e.g. Aucouturier et al., Emotion, under review), we created algorithms able to manipulate the emotional content of spoken speech in real-time, so as it make it happier, sadder, more afraid than how it was originally uttered. Each manipulation combines a series of signal processing transformations of pitch (e.g. vibrato to make voice tremble, communicating fear) and spectrum (e.g. low-pass filtering to make voice appear darker, duller).

These manipulations have psychological and physiological validity: manipulated voices are recognized by external judges as emotional, both using numerical scales and lexical description (Figure 2), and they create action tendencies that are coherent with the emotion being portrayed (Aucouturier et al, Applied Cognitive Psychology, under review). In addition, when speakers listen to themselves talking with manipulated voices, they self-attribute the manipulated emotion, and show coherent physiological responses in e.g. their galvanic skin response (Aucouturier et al., in prep., 2012). But such techniques have never been applied to music, and never been conceptualized as a means of experimental control.

The second originality of our approach is to use this technique as a tool to create musical stimuli that sound like emotional voice, in order to study mechanism 1. We will do a psycho-acoustical study of the properties of such manipulations for spoken voice, and then use the same effects to manipulate musical stimuli: making e.g. a violin tremble, stutter and crackle as if it was frightened speech. We will then study the emotional reactions to these manipulated musical stimuli, and test whether the same perceptive processes are at play than for emotional speech. This will provide, for the first time, an empirical tool to activate selectively, with music, the neural pathways of emotional speech

3) Neuro-mimetic sub-cortical representation for environmental sounds

The third methodological originality of our proposal is to use recent innovations in signal processing to model sensory processing in the human peripheral auditory system. While little is known still about the functional connectivity of the areas of the auditory pathway from the primary auditory cortex (A1) upwards (Read, Winer & Schreiner, 2002 for a review), the sensory properties of the cells involved up to A1 are rather well understood (Ghazanfar & Nicolelis, 2001), and can be simulated with computers. Cochlear processing can be approximated by a bank of band-pass filters followed by compression and temporal integration simulating the spiking behavior of the auditory nerve (Lyons & Mead, 1988). Cortical processing, in which individual cells function as modulation detectors, each tuned to a given timescale and frequency (*spectro-temporal receptive fields*), can be approximated by a wavelet transform in the time-frequency plane (Chi, Ru & Shamma, 2005) or simulated by sparse coding (Klein, Konig & Kording, 2003).

Such "neuro-mimetic" computer modeling isn't new: it is state-of-art for image processing (Serre et al, 2007), for which it permits higher recognition speed than the standard statistical approaches (Thorpe, Delorme & VanRullen, 2001). In the audio modality, neuro-mimetic modeling is used in coding and compression for speech (Chi, Ru & Shamma, 2005) and music (Plumbley et al, 2010). However, these models have never been used in empirical research as what they really are: models of sub-cortical auditory processing. This is an important opportunity: while this representation would be insufficient to model, e.g. some of the more sophisticated (and cortical) mechanisms of emotion induction, it is entirely appropriate for the study of mechanism 2, which is sub-cortical.

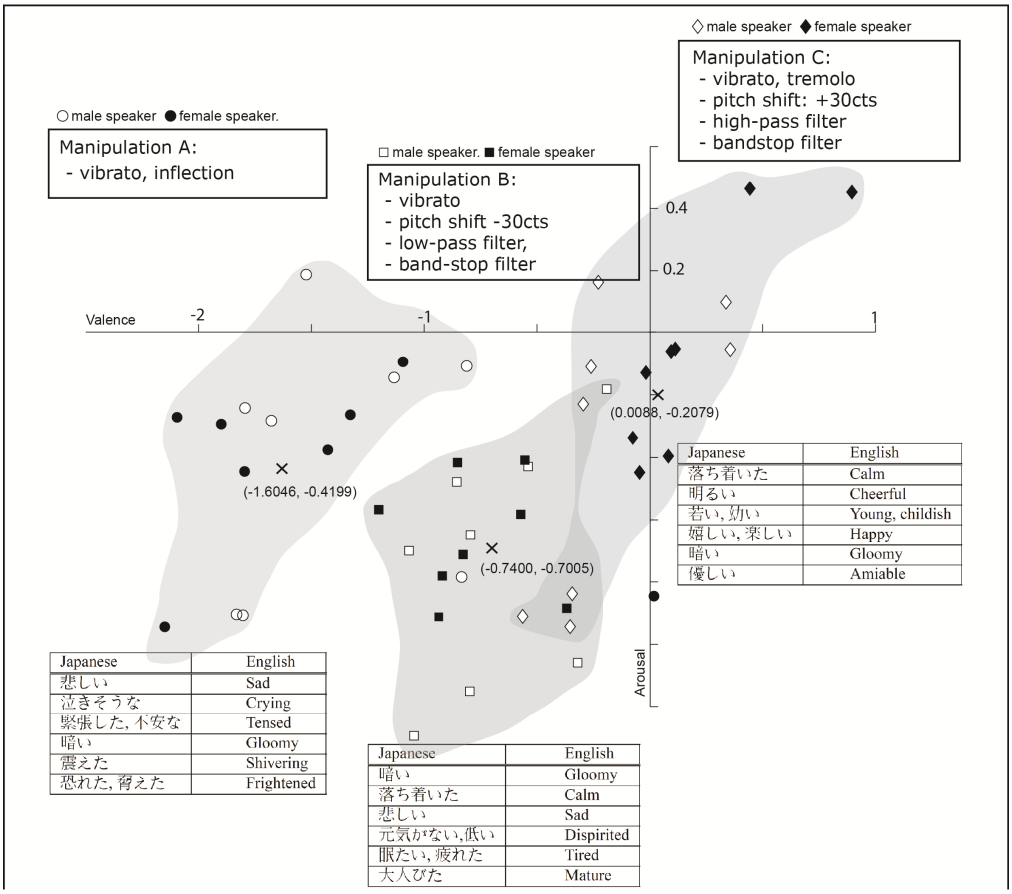


Figure 2: Three examples of our recent work on voice manipulation algorithms able to impart the emotions of fear, sadness and happiness onto a neutral spoken voice. Female and male neutral voices modified by manipulation A (which uses the acoustical transformations of vibrato and inflection) are perceived by both Japanese and American external judges as the most negative (left of the valence axis) and are associated in free-choice task to adjectives such as "frightened" or "tensed". Similarly, voices manipulated with B (vibrato, pitch decrease, low-pass filter) are linked to adjectives such as "sad" and "tired" and manipulation C (pitch increase, high-pass filter) to e.g. "cheerful" and "childish". Although this would fill a gaping void in the current knowledge of musical emotions, such techniques have never been applied to musical stimuli, and never been considered as a means of experimental control for cognitive neuroscience. Our proposal builds on this unique opportunity. Figure reproduced from Aucouturier et al, in prep., 2012

The third originality of our approach is, therefore, to use such neuro-mimetic representation as a common model for both environmental sounds and music. We will study the emotional reactions to a dataset of (short) environmental sounds and then select musical sounds from a large dataset which present auditory characteristics that are maximally similar to these environmental sounds: musical sounds that look like e.g. impact noises or roars from the point of view of the computational model. This will provide, for the first time, an empirical tool to activate selectively, with music, the neural pathways of emotional evaluation of environmental sounds.

As summary, Table 2 provides a schematic view of the experimental logics of our proposal. By focusing on very short sounds, we inhibit all emotional induction mechanisms (list in Table 2 adapted from Juslin & Vastfjäll, 2008) but the fastest ones. Further, by using computational techniques such as vocal transformations and neuro-mimetic modeling, we selectively activate either mechanism 1 or mechanism 2. Note that neither computational technique actually inhibits other mechanisms. Even when combined with short sounds, other fast mechanisms may exist that could confound the emotional reactions (e.g. in Table 2, fear conditioning). For further control, we will then turn to state-of-art tools of cognitive psychology (intercultural and infant studies) and cognitive neuroscience (fNIRS brain imaging, and electroencephalography). See research plan for details.

	WP1	WP2					
Mechanism:	**recycling voice**	**recycling environment. sounds**	**fear conditioning**	**rhythmic entrainment**	**musical expectancy**	**episodic memory**	**mental imagery**
Activated by short sounds (<1 sec.)	yes	yes	yes	no	no	no	no
Activated by voice-like musical sounds	yes	(maybe)	(maybe)	(maybe)	(maybe)	(maybe)	(maybe)
Activated by noise-like musical sounds	(maybe)	yes	(maybe)	(maybe)	(maybe)	(maybe)	(maybe)
Intercultural variation	yes	no	yes				
Ontologic development	first year	prior to birth	prior to birth				
Cortical activity	yes	no	no				

Table 2: our experimental logics to isolate mechanisms 1 and 2, from other confounding mechanisms such as fear conditioning or musical expectancy. By using short sounds and vocal transformations, we selectively activate mechanism1 while inhibiting a large number of longer-term, confounding mechanisms. If the emotional reactions to such sounds are subjected to intercultural variation and developmental variation within first year of life, we will exclude a brainstem mechanism such as mechanism2. If NIRS imaging shows cortical activation, we will ascertain mechanism1. Similarly, by using short sounds and neuromimetic modeling, we will selectively activate mechanism2. If the reactions are identical across cultures, appear even with neonates, and brain imaging is consistent with sub-cortical activation, we will ascertain mechanism2.

All the empirical methodologies to be used in the project are some in which the PI has gathered hands-on, practical experience during his postdoctoral years in cognitive neuroscience research units, albeit in projects which did not address musical emotions. We have for instance studied the psychoacoustics of short environmental sounds to study categorization (JASA, 2009); used EEG measures and ERP paradigms to study interactions between music and osteopathic treatment (Psychophysiology, under review, 2012); used Near-Infrared Spectroscopy (NIRS) brain imaging to study the context effect of music and environmental sounds on word encoding in memory (in prep., 2012); neonates behavioral paradigms to study the acoustic development of baby cries (JASA, 2011); and physiological measures (galvanic skin response) to study emotional influences on cognition (Emotion, under review).

Despite our first-hand experience with each of these methodologies, the PI and his team will also surround themselves with a suitable network of expert practitioners from which we will seek one-off, punctual expertise in well-defined perimeters at various stages of the project: for instance, for brain imaging paradigms (WP 1.4 and 2.4) , intercultural (WP1.3) and infant studies (WP2.3). These experts, with most of whom we already have collaboration history, are listed in the detailed research plan below for each punctual action with which they will be associated.

Research plan:

WP1. Voice mechanism (mechanism 1)

WP1 is the fundamental research unit devoted to the empirical study of mechanism1, according to which musical emotions invade circuits responsible for the evaluation of emotional speech. The research in WP1 focuses on the special case of very short sounds, to be extended to the general case in WP3.

WP 1.1 Emotional psychoacoustics of voice transformation effects

Software development and psychoacoustical study of participants' emotional reactions to short stimuli of emotional voice, obtained by the acoustic manipulation of neutral voice recordings. The space of emotional voice stimuli will be explored using the parameters of the manipulations, e.g. the modulation speed of vibrato or the cut frequency of a filter. Emotional judgements will be collected with explicit self-report methodologies (Self-Assessment Manikin - Bradley & Lang, 1994) and physiological measures (heart rate, galvanic skin response - Bartlett, 1996).

Punctual expertise: Dr G. Peeters and Dr A. Röbel (Senior Researchers, IRCAM), whose team is a world-leader in voice synthesis technology.

WP 1.2. Comparison with voice-transformed musical stimuli

Behavioral study of emotional reactions obtained to musical stimuli that are manipulated with the acoustic manipulations developed and normed in WP1.1. We will use the corpuses of short musical sounds utilized by Bigand et al. (2005) and Peretz, Gagnon and Bouchard (1998), for which emotional norms are available. Emotional judgements will be measured using the same apparatus as WP1.1, to allow the comparison of results.

WP 1.3. Intercultural validation

Intercultural comparison of emotional reactions obtained in WP1.2. We will compare the reactions of both western and indian listeners to musical sounds extracted from both Western and Indian (hindustani) music, manipulated to reproduce the acoustic cues for emotions expressed in both Western (French - the result of WP1.2) and Indian (Kannad/Tamil) languages. Measure apparatus similar to 1.1 and 1.2**.** We expect intercultural validity with intra-group advantage (Thompson & Balkwill, 2006) favoring the recognition of musical emotions manipulated with vocal cues that were identified from the participants' native linguistic culture. Control stimuli not expected to show intercultural variation will be also tested using the stimuli of WP2.2.

Punctual expertise: WP1.3. will be conducted in collaboration with Dr. Shantala Hegde, a clinical psychologist at National Institute for Mental Health and Neuro Sciences (NIMHANS), Bangalore, India (ICPC country).

WP 1.4. fNIRS brain imaging validation

Mapping of the activation in the right temporo-frontal areas of the vocal prosody networks, compared for emotional stimuli of type: voice, voice-transformed music and non voice-transformed (control) music (the stimuli of WP2.2), using functional Near-infrared spectroscopy (fNIRS). Oddball paradigm inspired by fNIRS study on phoneme categorization (Minagawa-Kawai et al., 2002). We expect significant activation in the ROI for both vocal and voice-transformed music, but not for non voice-transformed musical stimuli

Punctual expertise: Dr Aurelia Bugaiska (Assistant professor of cognitive psychology, LEAD, Dijon), an expert in the functional exploration with NIRS of frontal activity in healthy adults.

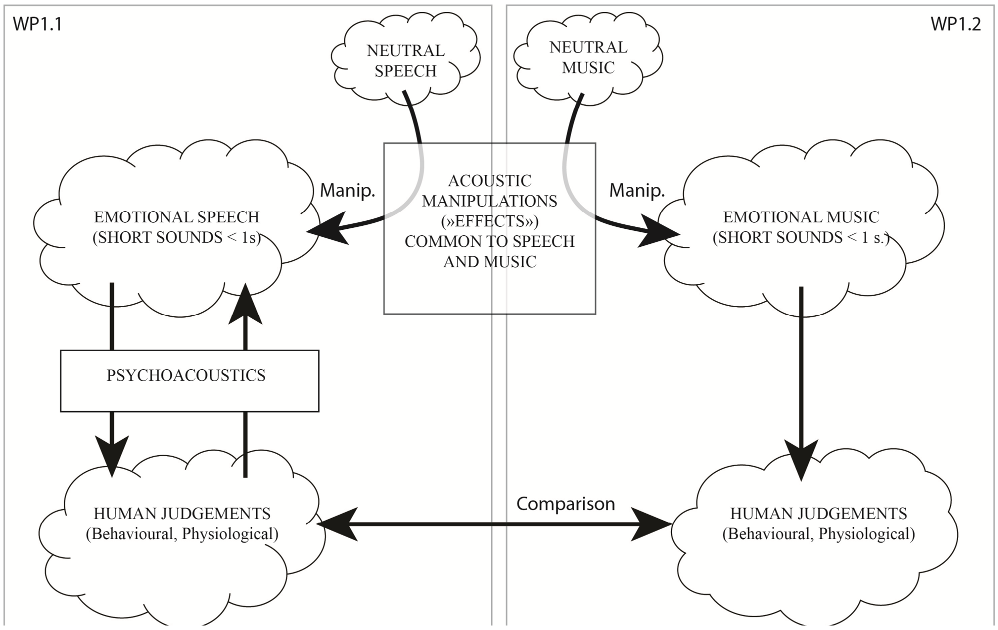


Figure 3: The experimental procedure of units WP1.1 and WP1.2, and how they interact

WP 2. Environmental sound mechanism (mechanism 2)

WP2 is the fundamental research unit devoted to the study of mechanism2, according to which musical sounds are interpreted sub-cortically as emotion-relevant sounds from the environment. As before, research in WP2 focuses on the special case of very short sounds, to be extended to the general case in WP3.

WP 2.1. Emotional psychoacoustics of environmental sounds

Software development and psychoacoustical study of participants' emotional reactions to short environmental sounds, selected for their biological relevance: impact sounds, cries, sounds indicative or rapid changes or fast movements, etc. As in WP1.1 and 1.2, emotional judgements will be collected with explicit and physiological measures. Psychoacoustical mapping will be seeked using a neuromimetic computational representation if the sounds which simulate the STRF encoding of the peripheral auditory system (consistent with the sub-cortical implementation of the mechanism).

Punctual expertise: at Ecole Normale Supérieure in Paris, Prof. Shihab Shamma and Dr. Daniel Pressnitzer (ERC Advanced grant ADAM, 2011), world-leaders in neuromimetic models of audition (Chi, Ru & Shamma, 2005)

WP 2.2. Comparison with similar musical stimuli

Behavioral study of emotional reactions obtained to musical stimuli selected for their sensory similarity to the environmental sounds identified in WP2.1. As before, we will use the corpuses of Bigand et al (2005) and Peretz, Gagnon and Bouchard (1998). From these corpuses, we will select musical sounds that have the most similar neuromimetic representation to the target environmental sounds, using the similarity techniques that we previously developed for MIR (Aucouturier & Pachet, JNRSAS 2004 - 270 citations). Emotional judgements will be measured with the same apparatus as WP2.1, WP1.1 and WP1.2, to allow for comparison.

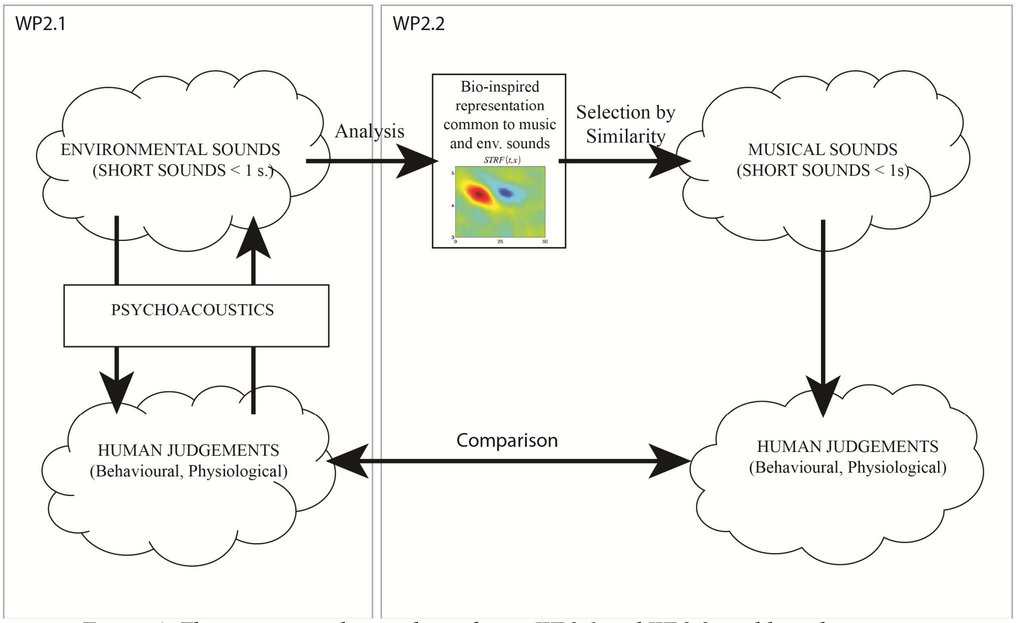


Figure 4: The experimental procedure of units WP2.1 and WP2.2, and how they interact

WP 2.3. Infant developmental study

The ontological development of mechanism 2 is expected to happen prior to birth, while mechanism 1 is expected to develop during the first year of life (Juslin & Vastfjäll, 2008). We will compare emotional reactions to the stimuli created in WP2.2 and 1.2. with neonates and infants of 0-12 months of age. The experimental paradigm of WP2.1&2 will be adapted to infants, using e.g. non-nutritive sucking (NNS) rhythm (Melher et al., 1978) and ERP (Kudo et al., 2011). We expect evidence of early development for WP2 stimuli, and gradual improvement of recognition for WP1 stimuli.

Punctual expertise: WP2.3 will be conducted in collaboration with Prof. Kazuo Okanoya, a developmental neuroscientist at University of Tokyo and Riken brain Science Institute, Saitama, Japan.

WP 2.4. ERP validation

Study of Event-Related Potential (ERP) evoked during passive listening to a series of musical and control sounds, using the mismatch negativity (MMN) paradigm (see Goydke et al. (2004) for a similar study on musical timbre). Stimuli will be selected to control for various degrees of similarity to environmental sound targets identified in WP2.1. A sub-cortical (or early auditory) mechanism should be associated with a pre-attentive mismatch response (N100). On the contrary, higher-level cortical mechanisms for e.g. attentive evaluation and categorization, if present, would involve a larger P300 response. The amplitude of N100 could correspond to the sensory distance between stimuli, computed using the neuromimetic representation of WP2.1 (see Toiviainen et al, 1998 for a similar effect on musical timbre)

Punctual expertise: Dr Bénedicte Poulain-Charronat (CNRS Senior researcher, LEAD, Dijon), an expert in music ERP studies.

WP 3. Extension to normal-length sounds

WP3 is the fundamental research unit devoted to the generalization of research results obtained in WP1 and WP2, from the special case of very short sounds (<1s.) to the general case of tens-of-seconds, or minute-long stimuli. While the studies in WP3 are difficult to specify for now before actual results can be considered from WP1 and 2, we foresee several types of experimental paradigms:

WP 3.1. Temporal integration with neutral distractors

Emotionally valenced stimuli (for which an emotional response was calibrated in WP1&2) are presented in the midst of a texture otherwise composed of neutral sounds (i.e. which normative response in WP1&2 was calibrated as neutral). The emotional reactions in this context are then compared to the isolated reactions obtained in WP1&2.

WP 3.2. Temporal integration with valenced distractors

Stimuli normed with emotion X (for instance, positive) are presented in the midst of a texture composed of other sounds normed with emotion Y (for instance, negative). We then test for the emergence of emotions linked to X or Y. Target sounds and distractors are organized in time in order to test various hypotheses of temporal integration, e.g. moving average, leaky integrator, etc. (see e.g. models by McAngus Todd, 1996).

WP 3.3. Interaction of 2 mechanisms

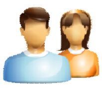
We will use the computational characterizations developed in WP1&2 to investigate the possibility that a same stimuli activates multiple-mechanisms at once: are there (potentially rare) musical sounds which sound both like environmental sounds and emotional voices? How are emotional effects integrated in such cases? Are some reactions faster than others, as predicted by theory?

WP 4. Clinical applications

WP4 is the applied research unit of the project. It is devoted to the utilization of the fundamental research results of WP1,2&3 as cognitive technology to develop therapeutic applications, to be validated with clinical studies.

WP 4.1. Effect of musical stimulation on stroke-induced aphasia

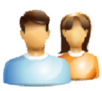
Daily stimulation with music provides larger degrees of linguistic and cognitive rehabilitation than speech for aphasic stroke-victims (Särkämo et al, 2008). We expect that voice-transformed music will increase this effect. We will establish a treatment protocol consisting of daily listening sessions to 4 experimental types of stimuli: (a) neutral speech, (b) emotional speech, (c) neutral music and (d) emotional music. Speech and music stimuli presented in groups (b) and (d) will be produced by manipulation of the stimuli in (a) and (c), using the research results of WP1. We will evaluate the effect of the 4 types of treatment after 2 weeks, 3 months and 6 months on the strength of the aphasia symptoms using standard clinical questionnaires, as well as on the emotional state of the patients.



Punctual expertise: Prof. Emmanuel Bigand (Professor of Cognitive Psychology, LEAD, Dijon), expert of linguistic rehabilitation with music. Collaboration with Dijon Univ. Hospital (cerebrovascular accident unit)

WP 4.2. Emotional induction in fibromyalgia patients using emotional vocal feedback

The signal-processing manipulations of voice used in WP1 appear to be detected but not recognized (Aucouturier et al., in prep. 2012). We will use this technique to develop a paradigm of manipulated vocal feedback, in which patients participate in a low-demand task (dialog, reading out loud) while listening to their own voice artificially manipulated to convey an emotion they had not intended, such as joy or calmness. This procedure will provide unconscious emotion modulation, which can circumvent the emotion avoidance biases associated with many affective disorders, such as in fibromyalgia (Van Middendorp et al. 2008). The short-term and medium-term impact of such sessions on fibromyalgic patients will be measured on standard indices of clinical evaluations, such as verbal fluency or profile of mood states.



Punctual expertise: Dr Gérard Mick, M.D. Rheumatology (Associate Researcher, LEAD, Dijon), expert of pain and affective disorders in fibromyalgia. Collaboration with Rheumatology units at Dijon & Lyon Univ. Hospitals.

WP 4.3. Emotional vocal feedback and the consolidation of traumatic memories in PTSD

Post-traumatic stress disorder (PTSD) patients suffer from hypermnesia linked to a traumatic episode, leading to emotional disorders (Tapia et al., 2011). Positive music was found able to modulate the consolidation of PTSD memories if listened to during or immediately after the traumatic episode (Rickard, Wong, & Velik 2011). We will establish a treatment protocol in which PTSD patients will be tasked with recalling by speech the traumatic episode under modified vocal feedback: unconsciously, they will hear themselves recall their memories in a vocal tone that is manipulated artificially to convey more serenity or happiness than they intended. The short and medium-term impact of the sessions will be measured with standard indices of clinical evaluation, such as subjective scales of anxiety, heart and salivary IgA secretion rates.

Punctual expertise: Dr Aurelia Bugaiska (Assistant professor of cognitive psychology, LEAD, Dijon), expert of memory and post-traumatic stress disorders.

Section c. Resources (incl. project costs)

Two research sites: The project will be implemented on two research sites, distant 250km apart in France: the primary site, Institute of Research/Coordination in Musical Acoustics (IRCAM), is based in central Paris (http://www.ircam.fr). The secondary site, the Learning and Developmental Psychology Laboratory (LEAD), is based in University of Burgundy in Dijon (http://leadserv.u-bourgogne.fr)

This 2-site montage is tailored to the unique interdisciplinary requirements of our proposal: on one front, IRCAM is the world-leading research centre for musical informatics, with notable expertise in the analysis and synthesis of musical signals and Music Information Retrieval (it was the organizer of the 2002 international ISMIR conference); on the other front, LEAD is the only French psychology laboratory devoted to music cognition, and the European coordinator of several high-profile European projects on clinical applications of music (incl. FP7 EBRAMUS). The two institutions have a history of joint collaboration, and the PI has institutional ties with both.

Both sites are managed by the same host institution, the French CNRS (Centre National de la Recherche Scientifique, http://www.cnrs.fr). CNRS is a government-funded research organization under the administrative authority of France's Ministry of Research. It is the main fundamental research organization in Europe and is largely involved in national, European, and international projects covering all fields of knowledge. CNRS is organized in different types of laboratories, often in partnership with universities, other research organizations or industry. The project's 2 sites, IRCAM and LEAD, are two of such laboratories. In practice, the project will be managed by the "Délégation Paris A" Regional Office of CNRS which is fully experienced in this field with dedicated staff in legal, financial and administrative issues.

Team members: The PI's team members will be supported on both sites: at IRCAM, where the project will be coordinated, the PI will be supported at 80% full-time, with the addition of one software engineer (18 months) and one research assistant (for 2 periods of 18months). At LEAD, we will recruit a post-doctoral researcher (acting as the project's local scientific lead, 60months) and 2 PhD students (for a consecutive 36months each). The PI will manage coordination between the 2 teams.

Division of work: The team in IRCAM will be responsible for the signal processing developments involved in WP1.1 and WP2.1, as well as the psycho-acoustical studies of WP1.1, WP1.2, WP2.1 and WP2.2. The team will also be responsible for the later, similar developments in WP3. The team in LEAD will be responsible for the cognitive psychological studies of WP1.3 and WP2.3, the brain imaging experiments of WP1.4 and WP2.4 as well as the clinical studies of WP4.

Prior infrastructure: Without direct cost, IRCAM provides the technical environment of a computer science laboratory (software licenses, servers, etc.) and facilities for psychoacoustics (2 soundproof rooms fitted with high-end audio systems). LEAD provides a technical environment for cognitive psychology (EEG room, one 8-channel NIRS room, one eye-tracking system), as well as institutional agreements for clinical research with the university hospitals of Dijon, Lyon and Grenoble (France).

Equipment direct costs: beyond basic computer equipment for team members, the project will add several key pieces of hardware to the existing facilities: (1) brain imaging (extension of LEAD's 8-channel NIRS system to a 24-channel system, required to do the functional studies in WP1.4, WP2.4 and WP4. 100k€ in period 1, quotation obtained from Artinis Medical Systems B.V), (2) system for physiological measures (BioPaC MP150 with GSR, ECG, respiration transducers. 15k€ in period 1, quotation: CEROM S.A.R.L),

(3) professional audio (2 TC Electronics VoiceWorks Plus audio processors or equivalent, Sennheiser microphones and noise cancellation headphones. 10k€ in period 1, quotation: AudioSolutions, S.A.R.L.)

Travel direct costs: beyond regular travel for conferences for team members, the project will support frequent PI travel between both sites (weekly, 7k€ per period), as well as two 3-week research stays (for two team members) aboard, one in National Institute for Mental Health and Neuro Sciences, Bangalore, India for the purpose of the intercultural studies of WP1.3, the other in RIKEN Brain Science Institute in Saitama, Japan for the purpose of the developmental study of WP2.3 (4k€ per trip and per person, based on market price for flights and the host institution's daily standard allowance of $150).

Other costs of note: the project will endorse EU's and CNRS's positions on Open Access dissemination of research results, and provisions for author-supported manuscript charges (4,5k€ per period). Psycho-acoustical studies in IRCAM require participant fees, which will be handled by subcontracting (5k€ for period 1&2 on the basis of 50€x100 participants); participants for studies in LEAD do not require such fees.

	Cost Category	**Month 1-18**	**Month 19-36**	**Month 37-54**	**Month 55-60**	**Total (M1-60)**
Direct Costs:	*Personnel:*					
	PI	83 000	83 000	83 000	28 000	277 000
	Senior Staff	0	0	0	0	0
	Post docs	72 000	72 000	72 000	24 000	240 000
	Students	0	108 000	108 000	0	216 000
	Other (engineer)	180 000	0	60 000	60 000	300 000
	Total Personnel:	335 000	263 000	323 000	112 000	1 033 000
	Other Direct Costs:					
	Equipment	39 000	39 000	39 000	13 000	130 000
	Consumables	3 000	2 000	2 000		7 000
	Travel	15 000	20 000	20 000	2 400	57 400
	Publications, etc	4 000	4 500	4 500	1 260	14 260
	Other	0	0	0	0	0
	Total Other Direct Costs:	61 000	65 500	65 500	16 660	208 660
	Total Direct Costs:	396 000	328 500	388 500	128 660	1 241 660
Indirect Costs (overheads):	Max 20% of Direct Costs	79 200	65 700	77 700	25 732	248 332
Subcontracting Costs:	(No overheads)	5 000	5 000	0	0	10 000
Total Costs of project:	(by reporting period and total)	480 200	399 200	466 200	154 392	1 499 992
Requested Grant:	(by reporting period and total)	480 200	399 200	466 200	154 392	1 499 992

% of working time the PI dedicates to the project over the period of the grant:	**80%**

Section d. Ethical and security-sensitive issues

ETHICS ISSUES TABLE

	Research on Human Embryo/ Foetus	YES	Page
	Does the proposed research involve human Embryos?		
	Does the proposed research involve human Foetal Tissues/ Cells?		
	Does the proposed research involve human Embryonic Stem Cells (hESCs)?		
	Does the proposed research on human Embryonic Stem Cells involve cells in culture?		
	Does the proposed research on Human Embryonic Stem Cells involve the derivation of cells from Embryos?		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	X	

	Research on Humans	YES	Page
	Does the proposed research involve children?	X	12
	Does the proposed research involve patients?	X	13-14
	Does the proposed research involve persons not able to give consent?		
	Does the proposed research involve adult healthy volunteers?	X	10-14
	Does the proposed research involve Human genetic material?		
	Does the proposed research involve Human biological samples?		
	Does the proposed research involve Human data collection?	X	10-14
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL		

	Privacy	YES	Page
	Does the proposed research involve processing of genetic information or personal data (e.g. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)?		
	Does the proposed research involve tracking the location or observation of people?		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	X	

	Research on Animals	YES	Page
	Does the proposed research involve research on animals?		
	Are those animals transgenic small laboratory animals?		
	Are those animals transgenic farm animals?		
	Are those animals non-human primates?		
	Are those animals cloned farm animals?		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	X	

	Research Involving non-EU Countries (ICPC Countries)	YES	Page
	Is the proposed research (or parts of it) going to take place in one or more of the ICPC Countries?	X	10
	Is any material used in the research (e.g. personal data, animal and/or human tissue samples, genetic material, live animals, etc) : a) Collected and processed in any of the ICPC countries?		
	b) Exported to any other country (including ICPC and EU Member States)?		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL		

	Dual Use	YES	Page
	Research having direct military use		
	Research having the potential for terrorist abuse		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	X	

References

Alluri, V & Toiviainen, P. (2010) Exploring perceptual and acoustic correlates of polyphonic timbre. Music Perception 27(3), 223-41

Anderson, A. K. & Phelps, E. A. (1998) Intact recognition of vocal expression of fear following bilateral lesions of the human amygdala, Neuroreport, 9, 3607-13

Balkwill, L.L. & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. Music Perception, 17, 43-64

Bartlett, D. (1996) Physiological responses to music and sound stimuli. In D. Hodges (ed.), Handbook of music psychology (2nd edn, pp. 343-385). San Antonio, TX: IMR Press

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J. & Dacquet, A. (2005) Multidimensional scaling of emotions responses to music: effect of musical expertise and duration. Cognition & Emotion, 19, 1113-39

Blood, A. J. & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proceedings of National Academy of Sciences, 98, 11818-23

Blumstein, D., Bryant, G. & Kaye, P. (2012) The sound of arousal in music is context-dependent. Biology Letters 8(5), pp.744-7.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavioral Therapy and Experimental Psychiatry, 25, 49–59.

Chi, T., Ru, P. & Shamma, S. (2005) Multiresolution spectrotemporal analysis of complex sounds Journal of Acoustical Society of America, Volume 118, Issue 2, pp.887-906.

Dalla Bella, S., Peretz, I., Rousseau, L. & Gosselin, N. (2001) A developmental study of the affective value of tempo and mode in music. Cognition, 80, B1-B10

Elhilali M, Shamma SA, Thorpe SJ, & Pressnitzer D (2007) "Models of timbre using spectrotemporal receptive fields: investigation of coding strategies", Proceedings of the 19th International Congress on Acoustics, Madrid, Spain.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Friederici, A.D. & Koelsch, S. (2009) Universal recognition of three basic emotions in Music. Current Biology, 19,573-76

Fukui H. & Toyoshima K. (2008) Music facilitate the neurogenesis, regeneration and repair of neurons. Med Hypotheses. 71(5):765-9.

Ghazanfar & Nicolelis (2001), The Structure & Function of Dynamic Cortical & Thalamic Receptive Fields, Cerebral Cortex, 11(3):183-93.

Gjerdingen, R. & Perrott, D. (2008) Scanning the Dial: The Rapid Recognition of Music Genres, J. New Music Research, vol. 37, No. 2, pp. 93–100

Gomez, P. & Danuser, B. (2004) Affective and physiological responses to environmental noises and music. Int. J. Psychophysiol. 53(2), pp.91-103.

Goydke, KN, Altenmüller, E, Möller, J & Münte, TF (2004) Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. Brain Res Cogn Brain Res. 21(3):351-9.

Gouyon, F., Klapuri, A., Dixon, S., Alonso, M., Tzanetakis, G., Uhle, C. & Cano, P. (2006) An experimental comparison of audio tempo induction algorithms. IEEE Transactions on Audio, Speech and Language Processing 14(5), pp.1832-1844.

Huron, D. (2006) Sweet anticipation: Music and the psychology of expectation.Cambridge, MA: MIT Press

Ilie, G. & Thompson, W.F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. Music Perception, 23, 319-329.

James, William (1884). What is an emotion? Mind, 9, 188-205

Juslin, P. & Laukka, P. (2003) Communication of emotions in vocal expression and music performance: Different channels, same code? Psychological Bulletin, 129, 770-814

Juslin, P. & Västfjäll, D. (2008) Emotional responses to music: the need to consider underlying mechanisms. Behav. & Brain Sciences, 31, 559-621.

Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, B. G., Richardson, P., Scott, J., Speck, J. A. & Turnbull, D. (2010). Music emotion recognition: a state of the art review. Proc. Int. Soc. Music Information Retrieval, Utrecht, Netherlands

Klapuri, A. (2008), Multipitch analysis of polyphonic music and speech signals using an auditory model, IEEE Trans. Audio, Speech and Language Processing, 16(2), pp. 255-266.
Klein, König & Körding (2003), Sparse Spectrotemporal Coding of Sounds, EURASIP Journal of Applied Signal processing, 2003
Koelsch, S., Fritz, T., von Cramon, D; Y., Müller, K. & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. Human Brain Mapping, 27, 239-50
Kudo N., Nonaka Y., Mizuno N., Mizuno K., and Okanoya K.(2011) On-line statistical segmentation of a non-speech aud auditory stream in neonates as demonstrated by event-related brain potentials. Developmental Science, 14(5), 1100-1106
Lartillot, O. & Toiviainen, P. (2007) MIR in Matlab (II): A Toolbox for Musical Feature Extraction From Audio, in proc. International Conference on Music Information Retrieval, Vienna.
Levitin, D. J. (2007) This is your brain on music. Plume/Penguin.
Liu, D. & Zhang, H.-J. (2006) Automatic mood detection and tracking of music audio signal, IEEE Tr. Speech and Audio Proc. 14(1)
Lyon, R. F. & Mead, C. (1988) , “An analog electronic cochlea,” IEEE Transactions on Acoustics Speech and Signal Processing, vol. 36, no. 7, pp.1119–1134, 1988.
Mannes, E. (2011). The Power of Music: Pioneering Discoveries in the new science of song. Walker & Co.
McAngus Todd, N. (1994) A computational model of temporal integration J. Acoust. Soc. Am. Volume 96, Issue 5, pp. 3258-3258 (1994);
Mehler, J., Bertoncini, J., Barriere, M., Jassik-Gerschenfeld, D. (1978). Infant Recognition of Mother's Voice, Perception, 7, pp.491-497
Meyer, L. B. (1956) Emotion and Meaning in music. Chicago: CUP
Minagawa-Kawai, Y., Mori, K., Furuya, I., Hayashi, R. & Sato, Y. (2002). Assessing cerebral representations of short and long vowel categories by NIRS. Neuroreport, Vol 13 No 5 16, pp. 581-4
Mitchell, L.A., MacDonald, R.A.R, Serpell, M.G. & Knussen, C. (2007). A survey investigation of the effects of music listening on chronic pain. Psychology of Music. 35(1), 39-59.
Peeters, G., Giordano, B., Susini, P., Misdariis, N. & McAdams, S. The Timbre Toolbox: Audio descriptors of musical signals, J. Acoustical Society of America, 130 (5) 2011
Peretz, I. (2001) Listen to the brain: The biological perspective on musical emotions. In P.N. Juslin & J.A. Sloboda (eds) Music and Emotion: Theory and Research (pp. 105--34).Oxford: Oxford University Press
Peretz, I., Gagnon, L. & Bouchard (1998) Music emotion: Perceptual determinants, immediacy & isolation after brain damage. Cognition, 68, 111-41
Peretz, I & Zatorre, R. (2005) Brain organization for music processing. Ann. Rev. of Psychology, 57, 27-53
Plumbley, M., Blumensath, T., Daudet, L., Gribonval, R. & Davies, M. E. (2010) . Sparse representations in audio and music: from coding to source separation. Proceedings of the IEEE 98, 995-1005
Read, H., Winer, J. & Schreiner, C. (2002) Functional architecture of auditory cortex, Current Opinion in Neurobiology, Volume 12, Issue 4, 1.
Rickard, N. A., Wong, W., & Velik, L. (2011) Relaxing music counters heightened consolidation of emotional memory. Neurobiol Learn Mem. 2011
Rojo, N., Amengual, J., Juncadella, M., Rubio, F., Camara, E. & Marco-Pallares, J (2011). Music-supported therapy induces plasticity in the sensorimotor cortex in chronic stroke. Brain injury, 25(78), 787-93.
Sacks, O. (2008) Musicophilia: Tales of Music and the Brain. Vintage.
Särkämö, T., Tervaniemi, M., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., Autti, T., Silvennoinen, H., Erkkilä, J., Laine, M., Peretz, I. & Hietanen, M. (2008) Music listening enhances cognitive recovery and mood after middle cerebral artery stroke. Brain vol. 131, pp. 866-876
Schirmer, A. & Kotz, S.A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. Trends in Cognitive Sciences, 10, 24-30.
Schmidt, L.A. & Trainor, L.J. (2001) Frontal brain electrical activity distinguishes valence and intensity of musical emotions. Cognition & Emotion, 15, 487-500
Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M. & Poggio, T.(2007). Object recognition with cortex-like mechanisms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29 (3), pp. 411-426
Tapia, G., Clarys, D., Bugaiska, A. & El-Hage, W. (2011) Recollection of negative information in PTSD. Journal of Traumatic Stress.
Thompson, W & Balkwill, L.(2006) Decoding speech prosody in five languages. Semiotica, 158, 407-24
Thorpe S, Delorme A & Van Rullen R. (2001). Spike-based strategies for rapid processing. Neural Networks 14: 715-25
Toiviainen, P., Tervaniemi, M., Louhivuori, J., Saher, M., Huotilainen, M. & Nâätänen, R. (1998) Timbre similarity: convergence of neural, behavioral and computational approaches, Music Perception 16, p.223-241

Van Middendorp H, Lumley MA, Jacobs JW, van Doornen LJ, Bijlsma JW & Geenen R (2008), Emotions and emotional approach and avoidance strategies in fibromyalgia. J Psychosom Res. 2008 Feb;64(2):159-67.
Watson, K.(1942). The nature and measurement of musical meanings. Psychological Monographs, 54, 1-43
Watt RJ & Ash RL (1998). A Psychological Investigation of Meaning in Music. Musicae Scientiae, 1, 33-53
Witvliet, C. & Vrana, S. (2007) Play it again Sam: Repeated exposure to emotionally evocative music polarizes liking and smiling responses, and influences other affects reports, facial EMG, and heart rate. Cognition & Emotion, 21, 3-25
Zentner, M. & Eerola, T. (2010) Self-report measures and models, in P.N. Juslin & J.A. Sloboda (eds) Music and Emotion: Theory and Research (chap. 8). Oxford: Oxford University Press