The dynamics of music perception and emotional experience: a connectionist model

Eduardo Coutinho
Adaptive Behaviour & Cognition Group
University of Plymouth
Plymouth, UK
eduardo.coutinho@*plymouth.ac.uk*

Angelo Cangelosi
Adaptive Behaviour & Cognition Group
University of Plymouth
Plymouth, UK

ABSTRACT
In this paper we present a methodological framework for the study of musical emotions, incorporating psycho-physiological experiments and modelling techniques for data analysis. Our focus is restricted to the body implications as a possible source of information about the emotional experience, and responsible to certain levels of emotional engagement in music. We present and apply the use of spatiotemporal connectionist models, as a modelling technique. Simulation results using a simple recurrent network, demonstrate that our connectionist approach leads to a better fit of the simulated process, compared with previous models. We demonstrate that a spatiotemporal connectionist model trained on music and emotional rating data is capable of generalizing the level of arousal in response to novel music input. The model is also capable of identifying the main variables responsible for such an emotional rating behaviour.

Keywords
Music, emotion, brain, body, neural networks.

INTRODUCTION
As others (Panksepp & Bernatzky, 2002; Clynes, 1978) we believe that formal relationships between music psychoacoustic variables, body and brain dynamics exist. These embodiment factors are in part responsible for the emotional engagement in music. Without ignoring other processes involved in emotion induction through music (e.g. appraisal, memory), our focus is restricted to the body implications as a possible source of information about the emotional experience (e.g. proprioceptive feedback). Specifically we want to understand the relation between bodily arousal and emotional measurements, though requiring the use of both physiological and self-report frameworks.

In section 'Emotion' we discuss the physiological implications in the emotional experience, and we present our perspective. In the section 'Emotional Experience with Music', we refer to musical emotions and some of its prominent issues. Although the process that translates sounds into neural messages is nowadays quite well understood, several are the implications for the brain and the body during music listening that might be related with the emotion experience (Koelsch & Siebel, 2005). As in emotion experience, measurable effects are detectable on physiology (e.g. Iwanaga & Tsukamoto, 1997; Krumhansl, 1997), and brain dynamics (e.g. Patel & Balaban, 2000; Blood & Zatorre, 2001), during music listening. Clinical studies and music therapy explore the capacity of music to affect psychological and physiological states. They present evidence of helping patients with psycho-physiological symptoms, as in autism or Parkinson disease..

We intend to model the temporal dynamics of music psychoacoustics, emotional ratings, and physiological states, in order to study the possible relations between them. Through an analysis of previous models, we will propose the use of spatiotemporal neural networks (Kremer, 2001), as a modelling technique. Their properties are discussed in the section 'Connectionist Models'. Finally, in section 'Case Study', we present a set of simulations and results using an Elman network (Elman, 1990), and we use these to discuss the theoretical framework proposed and the future modelling work.

EMOTION
Charles Darwin, William James, Walter Cannon, Wilhelm Wundt, Susanne Langer, among others, introduced great developments in the study of emotions, and are still considered as reference points for many emotions theorists. Although much discussion exists around the mechanisms of emotion, current research generally accepts that emotions can be described as a multi-modal mechanism, with several processes involved including appraisal, basic emotions, physiological responses, and subjective feeling states (Panksepp, 2001; Damasio, 2000; Scherer, 1984; Dolan, 2002).

It is important to stress that emotions differ in important aspects from other psychological processes. For instance, emotion is an embodied experience with specific behavioural patterns (facial expressions, autonomic arousal, etc.); it is less susceptible to our control and also expressed at the unconscious level (Damasio, 2000; Ekman, 1973); has the capacity to affect other cognitive processes (e.g. decision making), though not confined to the old sub-cortical structures in our brains (Damasio, 2000). For a review on these issues refer to Dolan (2002) and Panksepp (2001). For the interests of this article we focus on the implications of the body as a source of information or elicitation of the emotional experience.

Emotion as Arousal

The autonomic nervous system (ANS) regulates the body and its readiness for action. The correspondent physiological variations are referred as changes in arousal. The resultant patterns, as seen in the previous section, can elicit emotional states or experiences by interacting with the brain. Regarding specifically the role of physiological responses on the emotional experience and elicitation (and the interaction with the brain), there have been several models proposed. As discussed in Philippot, Chapelle, & Blairy (2002) paper on current research on the role of the body in the emotional experience, three main models of emotion can be distinguished. They identify (i) the "*undifferentiated arousal model*", (ii) the "*cognitive appraisal model*", and (iii) the "*central network model*".

The main idea behind the first model (Reisenzein, 1983; Schachter, 1964) is that body responses increase with emotional intensity, but their pattern is not differentiated across the different emotional states. One practical prediction of this model is that the perception of the emotional intensity can be influenced by the arousal intensity. The main finding of this research has been the fact that after the exposure to an arousing stimulus, the following emotional feeling state is intensified. This phenomenon is called "activation transfer" (Zillmann, 1983). The second model focuses on the body changes as a function of cognitive appraisal processes (Scherer, 1984), or action readiness (Frijda, 1986). In this line of research, the patterns of body changes are the combined result of the several cognitive appraisal components. The fact that the body itself might generate emotional states is quite marginalized in this model. This line of research has already been the one that we followed in our previous work (Coutinho, Miranda, & Cangelosi, 2005). It comprises several mechanisms of emotion induction, by considering all processes as part of a whole body/mind interactive system. Finally, from the third model perspective, emotions share different neural and cognitive mechanisms and pathways, and their pattern of interaction defines the emotional nature. In short, the patterns of body changes are differentiable across emotions. The activation of the body with a pattern related with a specific emotion will, in certain conditions, elicit that emotion (the peripheral feedback). This last process is automatic at an implicit level (Damasio, 1994).

In summary, the underlying idea common to all these mechanisms is that a specific emotion can be elicited by creating specific body state patterns (by manipulating the body), even outside the awareness of the individual. An event (appraised via cortical or subcortical routes) elicits physiological changes that facilitate action and expressive behaviour. These changes are accompanied by, and contribute to, an affective feeling state. Motoric and visceral feedback can contribute to the intensity and valence of an emotional experience: consciously or subconsciously, individuals use their body state as a clue as to the valence and intensity of the emotion they feel (Damasio, 1994; Damasio, 2000).

Peripheral feedback

In some very interesting experiments, Philippot et al. (2002) tried to find empirical evidence of emotion experience through physiological induction. To achieve such a task "physiological state should be manipulated, not in intensity but in quality, in order to observe the impact of such manipulations on the nature rather than the intensity of feeling states. Further, it should be established whether this effect occurs without individuals explicitly using body state as a source of information to determine their emotional feeling state" (Phillippot et al., 2002, p. 6). Respiration is the variable used, since it is considered appropriate for the study aims, as breathing is both under voluntary and automatic control. There is also empirical evidence pointing to the fact that respiratory patterns are associated with positive and negative feelings, and it is well known that respiration affects many other physiological responses like skin conductance or cardio-vascular changes, and that it is related with anxiety states.

The first experiment in Philippot et al. (2002) indicates that people experience respiratory changes that are subjectively differentiated across different types of emotions. Experiment 2 showed that differentiated emotional feeling states were induced by respiration manipulations without participants' awareness of the process. Philippot et al. claim that this is the first demonstration that the alteration of respiration is sufficient to induce emotion. Concluding, peripheral feedback can, at least in part, modulate the quality of emotional feelings, and this can be achieved without the person awareness of the process.

EMOTIONAL EXPERIENCE WITH MUSIC

Current research supports the hypothesis that music perception is a distributed process within the brain and is involved in an interactive spatiotemporal system of different neural networks (Zatorre, 2005; Koelsch, Fritz, Cramon, Müller, & Friederici, 2006), including even networks specialized for other processes (e.g. emotion, memory, language).

Music and brain dynamics

Experimental evidence suggests the participation of both hemispheres in music perception (Wieser, 2003), although it is possible to observe some interesting specializations. For instance the left hemisphere seems to be related to perception of timing and rhythm, while the right hemisphere specializes in pitch and timbre perception. Along with the spatiotemporal patterns of neural activity presented to the primary auditory system, the brain engages in other processes, for instance in the motor system (Thaut, Kenyon, Schauer, & McIntosh, 1999), language (Patel, 2003), emotion and reward related areas (Blood et al., 2001; Blood et al., 1999).

Some researchers (Panksepp, 2001; Clynes, 1978; Janata & Grafton, 2003) suggest that music derives its affective power from dynamic aspects of the brain systems. These usually control emotional processes and are distinct, but interacting, with cognitive processes. This hypothesis follows Susanne Langer's ideas about the existence of shared properties in patterns of physical or mental states, emotion, and music. Support for these ideas comes from research with brain damage patients. This shows that the emotional appreciation of music can be maintained even in the presence of severe perceptual and memorization deficits, though reinforcing the idea that sub-cortical mediation is involved in "emotional judgments" (Blood et al., 1999; 2001).

Due to these interactions certain basic mechanisms related to motivation/emotion in the brain can be elicited by music. This gives rise to the changes in the body and brain dynamics, and to the interference with ongoing mental and bodily processes (Panksepp et al., 2002; Patel et al., 2000). This multi-modal integration of musical and non-musical information might take place in the brain (Koelsh, 2005), opening a window for associations between the role of the body and the emotional experience during music listening. This indicates the existence of a possible relationship in the dynamics of musical emotion and the cognition of musical structure.

Music and body dynamics

Another quantifiable aspect of emotional responses to music is its effect on hormone levels in the body (Brownley, McMurray, & Hackney, 1995). There is evidence that music can lower levels of cortisol (associated with arousal and stress), and raise levels of melatonin (which can induce sleep). It can also cause the release of endorphins (Van der Ark & Ely, 1993), and can therefore help relieve pain. Krumhansl (1997), reports that sad excerpts are related to large changes in heart rate, blood pressure, skin conductance, and temperature, while fear excerpts are related with changes in the rate and amplitude of the blood flow. Happy ratings were associated with changes in respiration measurements. Although the correlation is fairly low and variable across individuals, some physiological changes related to musically induced emotions were found (Krumhansl, 2002).

In physiology, arousal is the term to define the body's readiness for action. Increased arousal is associated with increased heart rate, increased body temperature, increased respiration rate (increased oxygen consumption), and many other physiological changes. From a neurological perspective, increased arousal is associated with the release of adrenaline and noradrenaline, affecting for instance the amygdala. As a reference, the amygdala, deep in the limbic system, can influence cortical areas via feedback from proprioceptive, visceral or hormonal signals, via projections to various networks. In summary, hormones secreted in the body affect bodily processes (e.g. cardiovascular, muscular and immune systems) and the brain as well. Music can interact with both.

Peripheral feedback and Music

Nicola Dibben (2004) conducted a study to analyze the role of peripheral feedback in emotional experience with music. In a first experiment she tries to discover whether arousal causes intensification of emotional feeling when listening to music, comparing with the emotion thought to be express by the excerpt. Then she analyzes how the valence and arousal character of the music might mediate that process. Pulse rate was used as measure of arousal. To differentiate further variables, a second experiment was carried out to study the effect of the physiological arousal on emotion. It also looked at the origin of this effect, whether it was due to peripheral feedback or due to mood changes associated with the experiment. From the analysis of both experiments Dibben (2004) concludes that increased arousal influences listeners' experiences of emotion. Experiment 1 showed that increased physiological arousal intensified the dominant valence of emotions felt when listening to music, but not the arousal dimension of their emotional ratings. Experiment 2 indicated that if music has a positive valence, positive emotions feelings and thought are intensified more clearly than with negative dimension. Overall these results show that arousal intensifies the dominant valence response (in this case to music).

Measuring musical emotions

The findings presented above are particularly interesting as they demonstrate that the perception of the body state can be the source of experienced emotional feeling in music. From another perspective, the fact that the body state was manipulated shows that the source of the emotional arousal can be misattributed (in the above experiments they were attributed to music). This is consistent with the fact that listeners explore the environment to find clues for the type of emotion expressed by music, especially when it is not evident or complex (Juslin & Sloboda, 2001). Peripheral feedback may then be one of the sources used to classify that emotional experience.

Scherer (2004) posits some important questions on the study of musical emotions, mainly issue related to experimental frameworks. He suggests a combined methodology for the study of feelings induced by music integrating their cognitive and physiological effects. Frameworks to measure emotional experiences include discrete models (e.g. lists basic emotions: Ekman, 1999; Plutchik, 1991), eclectic approaches (e.g. Scherer, 2004), and dimensional models (e.g. Russell, 1989). In this article we focus on the last model, due to the temporal dynamics approach to emotion, as outlined earlier. For a discussion about the different methodologies, refer to Scherer (2004) and Schubert (1999).

The dimensional approach: arousal and valence

Wundt's (1897) initial dimensional model of emotion suggested the division of the emotion into three dimensions of feeling: pleasantness-unpleasantness, rest-activation, and tension-relaxation. This classification can be simplified using the two-dimensional space of arousal and valence. Both can be defined as subjective experiences (Russell, 1989). Arousal corresponds to a subjective state of feeling activated or deactivated (bodily activation); valence stands for a subjective feeling of pleasantness or unpleasantness (hedonic value). The conscious affective experience may be associated with a tendency to attend to the internal sensations associated with an affective experience, both of activation and of hedonic impact (Feldman, 1995).

Among the advantages of this model we can consider the simplicity (also for the experiment participants), and good reliability (Scherer, 2004). More recently Schubert (1999) has applied this concept to music creating the two dimensional EmotionSpace experimental software. While listening to music, participants were asked to continuously rate the emotion thought to be expressed by music. Each rating would correspond to a point in on the arousal/valence dimensional space.

Continuous measurements

Another important issue in music and emotion research regards the experimental methodology used to define the music stimulus. As pointed out by Schubert (2004), two main perspectives have been chosen for music stimulus in experiments: (i) the atomistic approach, based on the use of short auditory stimulus specifically produced for certain experiments, and (ii) the ecologically valid approach based on the use of "real" music. This idea of continuous measurements is consistent with the fact that music unfolds on time, since both body and brain engage in temporal processes by interacting with the musical cues. In line with Scherer (2004), this supports a framework to complement the study of the temporal dynamics of both physiological and brain processes, during music listening.. For example, rhythm and beat affect body rhythms and motor dynamics (e.g. Byers, 1976). A change in respiration rate due to musical rhythm, through the cardiovascular function, could affect several neurophysiologic systems (Boiten, Frijda, & Wientjes, 1994). This process is very similar to an emotion-induced arousal changes, in this case elicited by music.

We support the second approach based on ecological stimuli and continuous measurements. This is because continuous measurements of both music and emotion temporal dynamics better represent non linear relationships between body, music and emotions (Schubert, 1999). In the following part we will focus on a modelling techniques proposed to analyze the results from the framework suggested here.

SPATIOTEMPORAL MODELS

Several researchers (e.g. Gabrielsson & Lindström, 2001) agree on the point that the way musical elements are organized in time by the composer can evoke emotional responses in the listener. Parameters such as mood, physiological state, cultural background, and preferences of the listener influence this process. Other variables include performance style (Juslin, 2001) and musical style. This literature implies the existence of a causal, underlying relationship between musical features and emotional response. Hevner (1936) produced the first comprehensive work that attempted a systematic explanation of the relationship between musical features and perceived emotion (For a detailed review, see Gabrielsson & Lindström, 2001). This line of research found interesting relations between musical variables and elicited emotional states. However, the hypotheses proposed are not conclusive, especially due to the complexity of the psychoacoustic properties of music involved, and their linear and non-linear relations. We will now focus on the contribution of mathematical and computational modelling techniques for understanding the relationship between music elements and evoked emotions. In particular, we support the use of spatiotemporal models, i.e. approaches where the model at the same time includes a temporal dimension (e.g. musical sequences and continuous emotional ratings) and a spatial component (e.g. the parallel contribution of various music and psychoacoustic factors).

Previous models

Schubert (1999) and Korhonen (2004) were the first to study the interaction between music psychoacoustics and emotion ratings. They focus on the relationship between musical/psychoacoustic variables and emotional ratings using the continuous response methodology. Schubert proposes a methodology based on the use of combinations of time series analysis techniques (e.g. linear regression) to analyze the data and to model such process. Korhonen (2004) proposes the use of the System Identification technique (Ljung, 1999). Both authors focus on the issues of the complexity of the data and their possible nonlinear relationships. Schubert monitors melodic pitch, tempo, loudness, timbral sharpness, and texture. Then he applies first order autoregressive adjustments for serial correlation, building a regression model of emotional ratings and selected musical features. The extension of the model to a non-linear space with other possible sources of analysis is also suggested. As it will be discussed below, neural networks provide such an extension.

Korhonen (2004) extends the music features space and the musical repertoire. The experimental setup is similar to Schubert's one, and the modelling techniques consider the increased complexity of the data and generalization properties. The System Identification technique was chosen as it can model time-varying patterns, and it offers generalization properties. Again, for future work, the use of non-linear models is suggested, with more generalization properties. Schubert highlights also the problem correlation between psychoacoustic variables, suggesting the identification of spatial patterns, along with temporal.

Spatiotemporal connectionist models

The study of the interaction between musical features and emotion ratings requires models that are capable of identifying spatiotemporal relationships among the data both in individual temporal patterns and in musical spatial (e.g. psychoacoustic) features correlations. This constitutes a complex dynamical system to which we can add generalization capabilities. Spatiotemporal connectionist models (Kremer, 2001) are an ideal methodology to investigate the dynamic spatiotemporal relationship in music and emotion..

Connectionist models (based on artificial neural networks) are computational paradigms inspired by the highly complex, nonlinear, and parallel information processing that occurs in the brain. A spatiotemporal connectionist networks can be defined as "a parallel distributed information processing structure that is capable of dealing with input data presented across time as well as space" (Kremer, 2001, pp. 2). There are several approaches to develop algorithms that achieve such a task, focusing on the dynamical processing of temporal patterns across a distributed system. Neural networks have been extensively used in computer music research (e.g. Todd & Loy, 1991; Martins & Miranda, 2006)

Neural Networks and Time: Elman networks

Neural networks that have to process time-based tasks, such as time series forecasting, often make use of recurrent connections to endow the network with a kind of dynamic memory. This way, the network can detect not only static patterns (not changing in time), but also add another dimension, containing temporal related information. Various proposals and architectures can be found in literature for time-based neural networks (see Kremer 2001 for a review). In this work we have selected Elman network (Elman, 1990), also called Simple Recurrent Network (SRN).

An Elman Neural Network (ENN) is based on the basic feed-forward architecture (multi-layer perceptron) with an additional layer called "context" or "memory" layer. The units in this layer receive a copy of the previous internal state of the hidden layer. They are connected back to the same hidden layer, through adjustable (learning) weights. These units endow the network with a dynamic memory, achieved through recursive access to past information. The basic functional assumption is that the next element in a time-series sequence can be predicted by accessing the previous hidden state of the system. This network has been extensively applied in areas such as language (e.g. Elman, 1990) and financial forecasting systems (e.g. Giles, Lawrence, & Tsoi, 2001), among others.

We have chosen an ENN because it allows the simulation of a task in which the input time series consists of a music piece, and the output a rating of emotional state (e.g. arousal). Moreover, Elman networks have very good generalization capabilities.

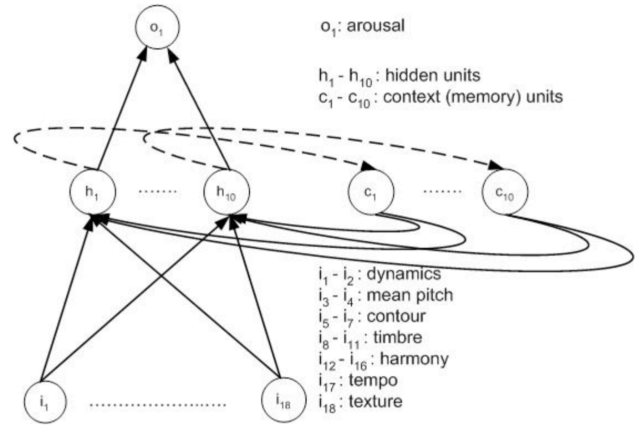

Figure 1. Neural network architecture (ENN).

CASE STUDY

Our work consists on the simulation modelling of Korhonen's (2004) experimental data. He used 6 pieces of classical music[1] for his experiments. Volunteers used a continuous arousal and valence scales (with a sample rate of 1s), to rate the emotion thought to be expressed by the music. We use the same 18 psychoacoustic variables chosen by Korhonen to encode the musical input to the system, and, for the purpose of this preliminary analysis of the model, as output, we selected only the arousal ratings (the focus of this preliminary analysis of the model). Music 1 to 5 are used to train the system, while Music 6 is used to test the generalization response to novel stimuli.

The music input stimuli were encoded into 18 psychoacoustic variables, clustered into 7 major groups: dynamics (loudness), mean pitch (centroid), pitch variation (contour), timbre, harmony (western consonance/dissonance), tempo, and texture (no. instruments playing). These correspond to the inputs of the ENN, while the single output refers to arousal. This way we are trying to find the relations between music and emotion appraisal, assuming the last one as a non-linear function of the former. The network architecture and its specifications are shown in Fig. 1.

Methodology and Simulations

We carried out two sets of simulations to evaluate the connectionist framework proposed here, and to be able to compare it with the other referenced models. Experiment 1 focuses on a quantitative analysis of the model performance in learning (training data) and generalizing (validation data) the musical temporal sequences, and relating them with emotional appraisals. Experiment 2 consists in evaluating

---

[1] Music 1 - Concierto de Aranjuez – Adagio (Rodrigo)
Music 2 - Fanfare for the Common Man (Copland)
Music 3 - Moonlight Sonata – Adagio Sostenuto (Beethoven)
Music 4 - Peer Gynt – Morning (Grieg)
Music 5 - Pizzicato Polka (Strauss)
Music 6 - Piano Concerto No. 1 – Allegro Maestoso (Liszt)

the relative contribution of each individual group of psychoacoustic music variables to arousal dynamics, in order to find relationships between them. For all experiments we present the results of 5 representative replications, each corresponding to the training of an ENN with different randomly initialized weights. In Experiment 1 we train we trained the ENN with Music 1 to 5, and we calculate the errors for the train data and for the novel data of Music 6 (see Table 1 and Fig. 2). Then, in Experiment 2, we removed the input values of each individual psychoacoustic group from the network, in a set of 7 simulations (see Table 1 and Fig. 2). This allowed us to analyze the influence of each individual psychoacoustic lag on the output prediction, therefore identifying their contribution to the arousal rating. All simulations were run for 2000 epochs (803 training sweeps per epoch, i.e. the combined total length of the 5 training music pieces), with a learning rate of 0.01 and momentum of 0. The context layer was reset every 20 training sweeps (corresponding to 20s of music).

Results

The network learning errors are calculated using the Root Mean Square (RMS) error. This measure gives us an indication on how well the network is capable of representing the desired output. We then can calculate the squared multiple correlation coefficient R^2 (or fit), as in Eq. 1, where N corresponds to the length of the time series, and $y(t)$ to the ideal arousal value at the output (the train target at time t). This measure allows us to discuss the amount of output variation that it is explained by the model, and to compare it with the performance of Schubert's and Korhonen's models. The average training and generalization errors of experiment 1 are presented in the first column (*All variables*) of Fig. 2. The individual errors for the simulations of Experiment 2 correspond to the other columns in the histogram of Fig. 2.

From Table 1, we can see the significantly improved performance of the ENN model (R^2 values of 97% and 98%) compared with the other two approaches (R^2 values below 90%). This supports the validity of the spatiotemporal connectionist model. Another important result is the capability of the network to generalize to new novel music input, as it is the case with Music 6. Similar high performance to that of the training was obtained for the validation tests. This indicates that the neural network was able to infer the underlying dynamics for rating the arousal level produced by the music input.

(Equation 1)
$$R^2 = \left(1 - \frac{RMS^2}{\frac{1}{N}\sum_{t=1}^{N}|y(t)|^2}\right) * 100\%$$

Table 1. Experiment 1 (all variables in input): Training error and correlation coefficient in the ENN and Korhonen's (2004) and Schubert's (1999) models.

Music	Average RMS	R^2 Exp. 1	R^2 Korhonen	R^2 Schubert
Train data				
1	0.069	98	90	57
2	0.075	97	-170	n. a.
3	0.058	98	25	n. a.
4	0.097	96	66	67
5	0.060	98	65	36
6	n. a.	n. a.	86.5	n. a.
Novel data				
6	0.092	97	n. a.	n. a.

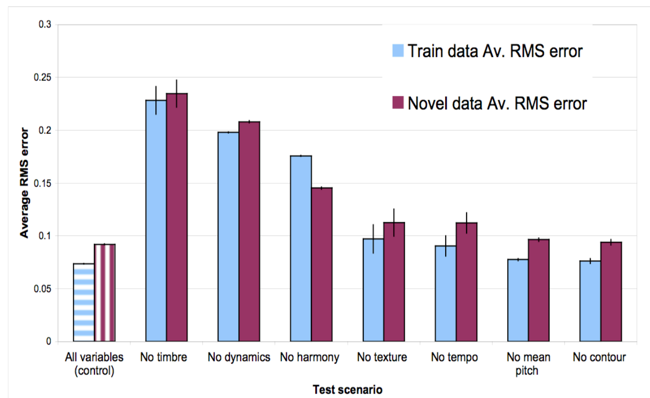

Figure 2. Experiment 2: train and test average RMS and standard errors.

Another analysis is reported in Fig. 2. As presented before we wanted to understand to which extent specific musical variables might be closely related with emotional judgments. For this we conducted individual analysis for each musical variable group. The histogram shows 3 cases of relevant error increase when removing these specific psychoacoustic groups: dynamics (loudness), timbre, and harmony. The significant drop in performance (i.e. error increase) indicates the stronger correlation between these variables and arousal ratings, suggesting their close relatedness for the repertoire and population used. This result agrees with research in music and emotion, as extensively reviewed by Schubert (1999). In his work he reports the strong relationship between arousal and loudness, as well with tempo. Here we report that timbre as well as loudness and harmony, can explain great part of the variations in arousal ratings. Texture and tempo didn't have a relevant impact on the process; also mean pitch and melody contour do not seem to affect the model performance. Nevertheless our results support and extend Schubert thesis. We will apply more extensive methodologies to analyze their temporal dynamics, variables correlations, and their relation with musical context.

SUMMARY AND DEVELOPMENTS

In this paper we presented a methodological framework for the study of musical emotions, incorporating psycho-physiological experiments and modelling techniques for data analysis. Then our focus is restricted to the body implications as a possible source of information about the emotional experience, and responsible to certain levels of emotional engagement in music. Along with this, we presented a framework to model the temporal dynamics of music psychoacoustics, emotional ratings, and physiological states, in order to study the possible relations between them, through an analysis of previous models. We present and apply the use of spatiotemporal connectionist models, as a modelling technique.

Preliminary results using a simple recurrent network, improve previous models results. Moreover they open an interesting window for temporal dynamical analysis of a non-linear system, as it seems to be the case of our investigation. In summary we demonstrated that a spatiotemporal connectionist model trained on music and emotional rating data is capable of generalizing the level of arousal in response to novel music input. The model is also capable of identifying the main variables responsible for such an emotional rating behaviour.

The next steps include a complete analysis of the data and improvement of the model. Along with that we are working on the psycho-physiological framework to establish a set of experiments, exploring also new techniques. We intend to explore the framework proposed in order to obtain better experimental setups. It is interesting to analyze to what extent this can be also applied to a more diverse musical scenario. One of the difficulties relies on colinearity of musical stimuli, affecting the model capability to represent the desired process. Namely the case of tempo is one of the evident facts in this article, since apparently no relevant quantitative correlations were found with emotional ratings. We believe that this can be related with both with representation or variables co-linearity issues. Nevertheless an extensive analysis is required to find if effectively this and other variables are correlated with the emotional reports in this experiment. Another important issue is the music representation, since the transformation for the neural network inputs may also introduce some problems. Current developments include strategies to solve both the problem of music stimuli and their diversity of contents and contexts (allowing individual and mutual correlation analysis), and the music representation. These issues are extremely relevant to endow our framework with a reliable basis to conclude about the influence of music on emotional ratings, in a more "universal" basis. Finally, valence will also be modelled, although some preliminary analysis has shown this to be a more difficult task to achieve. We believe that physiological measurements will highlight and facilitate this process, helping to detect and ground important effects. Another possibility after the complete analysis is to use different dimensions for the emotional ratings.

AKNOWLEDGEMENTS

The authors would like to acknowledge the financial support of the Portuguese Foundation for Science and Technology (FCT, Portugal). The acknowledgements are extended to Mark Korhonen for making available his experimental data, and also to Guido Bugmann, Andrew Hennell and Roman Borisyuk for their relevant comments and helpful suggestions.

REFERENCES

Bateson & P. Kopfer (Eds.). *Perspectives in Ethology* (pp. 135-164). New York: Plenum.

Blood, A. J., Zatorre, R. J., Bermudez, P. & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activities in paralimbic brain regions. *Nature Neuroscience, 2(4)*, 382-387.

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 the National Academy of Sciences*, *98(20)*, 11818-11823.

Boiten, F. A., Frijda, N. H. & Wientjes, C. J. (1994). Emotions and Respiratory Patterns: review and critical analysis. *Int. J. Psychophysiology, 17*, 103-128.

Brownley K.A., McMurray ,R.G. & Hackney, A.C. (1995). Effects of music on physiological and affective responses to graded treadmill exercise in trained and untrained runners. *Int. Journal of Psychophysiology, 19(3)*,193-201

Byers, P. (1976). Biological rhythms as information channels in interpersonal communication behavior. In P.

Clynes, M (1978). *Sentics: the Touch of Emotions*. New York: Doubleday.

Coutinho, E., Miranda, E. & Cangelosi, A. (2005). Towards a Model for Embodied Emotions. *Proceedings of the Workshop on Affective Computing: Towards Affective Intelligent Systems (AC 2005)*, Covilhã (Portugal).

Damasio, A. (1994). *Descarte's error: emotion , reason and the human brain*. New York: Grosset/Putnam Books.

Damasio, A. (2000). *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. London: Vintage.

Dibben, N. (2004). The role of peripheral feedback in emotional experience with music. *Music Perception, 22(1)*, 79-115.

Dolan, R. J. (2002). Emotion, Cognition, and Behavior. *Science 298(5596)*, 1191-1194.

Ekman, P. (1973). *Darwin and facial expression: A century of research in review*. New York: Academic Press.

Ekman, P. (1999). Basic Emotions. In T. Dalgleish & T. Power (Eds.). *The Handbook of Cognition and Emotion*. Sussex: John Wiley and Sons, Ltd.

Elman, J. L (1990). Finding structure in time. *Cognitive Science, 14*,179-211.

Feldman, L. A. (1995). Valence-focus and arousal-focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology*, *69*, 153-166.

Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and Research* (pp. 223-248). New York: Oxford University Press.

Giles, C. L., Lawrence, S. & Tsoi, A. C. (2001). Noisy Time Series Prediction Using a Recurrent Neural Network and

Grammatical Inference. *Machine Learning, 44(1/2)*, 161-183.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48*, 246-268.

Ljung, L. (1999). *System Identification: theory for the user*. New Jersey: Prentice-Hall.

Iwanaga, M. & Tsukamoto, M. (1997) Effects of excitative and sedative music on subjective and physiological relaxation. Percept Mot Skills; 85: 287-296.

Janata, P. & Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience, 6(7)*, 682-687.

Juslin, P. N. & Sloboda, J. A. (2001). Psychological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.). *Music and emotion: theory and research* (pp. 71-104). New York: Oxford University Press.

Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 305-333). New York: Oxford University Press.

Koelsch, S. & Siebel, W.A (2005).Towards a neural basis of music perception. *Trends in Cognitive Sciences, 9(12)*, 578-584.

Koelsch, S., Fritz, T., Cramon, D. Y., Müller, K. & Friederici, A. D. (2006). Investigating emotion with music: an fMRI study. *Human Brain Mapping, 27*, 239-250.

Kremer S. C. (2001). Spatiotemporal Connectionist Networks: A Taxonomy and Review. *Neural Computation, 13,* 249-306.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51(4)*, 336-353.

Krumhansl, C. L. (2002). Music: a link between cognition and emotion. *Current Directions in Psychological Science, 11*, 45-50.

Korhonen, M. (2004). *Modeling Continuous Emotional Appraisals of Music Using System Identification*. Phd thesis, University of Waterloo.

Korhonen, M. (2004). *Modeling Continuous Emotional Appraisals of Music Using System Identification*. Retrieved August 3, 2005, from http://www.sauna.org/kiulu/emotion.html.

Martins, J. M. and Miranda, E. R. (2006). A Connectionist Architecture for the Evolution of Rhythms. *Proceedings of EvoWorkshops 2006, Lecture Notes in Computer Sciente*. Berlin: Springer-Verlag.

Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago: Chicago University Press.

Panksepp, J. (2001). The neuro-evolutionary cusp between emotions and cognitions. *Evolution and cognition, 7(2)*, 141-163.

Panksepp, J. & Bernatzky, G. (2002). Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioural Processes*, *60(2)*, 133-155.

Patel, A.D. & Balaban, E. (2000). Temporal patterns of human cortical activity reflect tone sequence structure. *Nature, 404*, 80-84.

Patel, A. (2003). Language, Music, syntax and the brain. *Nature Neuroscience, 6(7)*, 674-681.

Philippot, P., Chapelle, G. & Blairy, S. (2002). Respiratory feedback in the generation of emotion. *Cognition and Emotion 16(5)*, 605-627.

Plutchik, R. (1991). *The Emotions*. Lanham, MD: University Press of America.

Reisenzein, R. (1983). The Schachter theory of emotion: two decades later. *Psychological bulletin, 94*, 239-264.

Russell, J. A. (1989). Measures of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion:Theory, research, and experience* (4, pp. 83-111). Toronto: Academic.

Schachter, S. (1964). The interaction of cognitive and psysiological determinants of emotional state. In L. Berkiwitz (Ed.), *Advances in experimental Social Psychology* (1, 44-79). New York: Academic Press.

Scherer, K. R. (1984). On the nature and function of emotion: a component process approach. In K. Scherer and P. Ekman (ed.). *Approaches to emotion* (pp. 293-317). Hillsdale: Erlbaum.

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research, 33(3)*, 239-251.

Schubert, E. (1999). *Measurement and time series analysis of emotion in music*. PhD thesis, University of New Southy Wales.

Thaut, M. H., Kenyon, G. P., Schauer, M. L. & McIntosh G. C. (1999). The connection between rhythmicity and brain function. *IEEE Engineering in Medicine and Biology, 18(2)*, 101-108.

Todd, P. M. & Loy, D. G. (Eds.) (1991). *Music and connectionism*. Cambridge, MA: MIT Press.

Van der Ark, S. D. & Ely, D. (1993). Cortisol, biochemical, and galvanic skin responses to music stimuli of different preference values by college students in biology and music. *Perceptual and Motor Skills, 77(1)*, 227-234.

Wieser, H. G. (2003). Music and the Brain: Lessons from Brain Diseases and Some Reflections on the "Emotional" Brain. *Annals. N.Y. Academy of Science,* 99**,** 76–94.

Wundt, W. (1897). *Outlines of Psychology*. Leipzig: Wilhelm Engelmann.

Zatorre, R. J. (2005). Music, the food of neuroscience?. *Nature, 434*, 312-315.

Zillmann, D. (1983). Transfer of Excitation in emotion behavior. In J.T. Cacioppo & R.E. Petty (Ed.), *Social psychophysiology* (215-240). New York: Guilford.