cortex 47 (2011) 1068–1081

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cortex

Special section on music in the brain: Research report

Emotion rendering in music: Range and characteristic values of seven musical variables

Roberto Bresin and Anders Friberg*

KTH School of Computer Science and Communication, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 7 October 2009
Reviewed 19 February 2010
Revised 24 May 2010
Accepted 16 February 2011
Published online 17 May 2011

Keywords:
Emotion
Musical variables
Music performance

ABSTRACT

Many studies on the synthesis of emotional expression in music performance have focused on the effect of individual performance variables on perceived emotional quality by making a systematical variation of variables. However, most of the studies have used a predetermined small number of levels for each variable, and the selection of these levels has often been done arbitrarily. The main aim of this research work is to improve upon existing methodologies by taking a synthesis approach. In a production experiment, 20 performers were asked to manipulate values of 7 musical variables simultaneously (tempo, sound level, articulation, phrasing, register, timbre, and attack speed) for communicating 5 different emotional expressions (neutral, happy, scary, peaceful, sad) for each of 4 scores. The scores were compositions communicating four different emotions (happiness, sadness, fear, calmness). Emotional expressions and music scores were presented in combination and in random order for each performer for a total of 5×4 stimuli. The experiment allowed for a systematic investigation of the interaction between emotion of each score and intended expressed emotions by performers.

A two-way analysis of variance (ANOVA), repeated measures, with factors emotion and score was conducted on the participants' values separately for each of the seven musical factors. There are two main results. The first one is that musical variables were manipulated in the same direction as reported in previous research on emotional expressive music performance. The second one is the identification for each of the five emotions the mean values and ranges of the five musical variables tempo, sound level, articulation, register, and instrument. These values resulted to be independent from the particular score and its emotion. The results presented in this study therefore allow for both the design and control of emotionally expressive computerized musical stimuli that are more ecologically valid than stimuli without performance variations.

* *Corresponding author*. KTH School of Computer Science and Communication, Department of Speech, Music and Hearing, Lindstedtsvägen 24, SE-114 28 Stockholm, Sweden.
E-mail address: roberto@kth.se (R. Bresin).
URL: http://www.speech.kth.se/~roberto
0010-9452/$ – see front matter © 2011 Elsevier Srl. All rights reserved.
doi:10.1016/j.cortex.2011.05.009

1. Introduction

In music, a central aspect for many listeners and performers is the induction and communication of emotion. Since the beginning of the twentieth century there have been a number of empirical studies focusing on the relation between music performance variables and emotional expression (Gabrielsson and Lindström, 2010). Recently, there has been an increase of research relating to music and emotion within neuropsychology (see e.g., Peretz, 2010). Many studies in neuropsychology and music have investigated reactions to specific musical parameters (Koelsch and Siebel, 2005), such as melody (Patel, 2003; Brattico, 2006), harmony including basic dissonance–consonance (Blood et al., 1999; Koelsch et al., 2006), modality in terms of major–minor (Mizuno and Sugishita, 2007; Pallesen et al., 2006), rhythm (Samson and Ehrlé, 2003), or musical timbre (Toiviainen et al., 1998; Caclin et al., 2006).

To construct appropriate test material in which musical parameters are studied in isolation is challenging. Given the unlimited compositional possibilities and a large covariability among the parameters, it is difficult to make independent variations of structural parameters such as rhythmic pattern without changing for example the melody. Additional challenges in neuropsychology are the technical limitations due to the measurement devices. For example, when using fMRI one has to cope with the excessive noise created by the magnets as well as avoid any magnetic equipment, such as normal headphones, in the scanner. Most methods also require rather short stimuli and numerous repetitions (Koelsch et al., 2010). In light of these and other methodological concerns it is understandable that many studies have used either rather simple computer-generated musical stimuli without natural performance variations or real music recordings. Real music recordings have high ecological validity but do not allow individual musical parameters to be isolated. We suggest that a better decomposition of musical parameters is needed in order to study their influence on the emotion perception/induction. This is also a prerequisite for isolating different brain regions active in for example music emotion induction. Isabelle Peretz in her recent review on neurobiology and musical emotions states that "… there is a need to fractionate the processing components involved in each basic musical emotion in order to understand the principles underlying hemispheric differences." (Peretz, 2010). A similar standpoint is given by Juslin and Laukka (2003) regarding behavioral studies on vocal and music expression: "… only by using synthesized sound sequences that are systematically manipulated in a factorial design can one establish that a given cue really has predictable effects on listeners' judgement of expression."

In order to improve stimuli quality the following three aspects can be identified:

1. *Ecological validity*. A lack of performance variations and a poor synthesizer quality make the stimuli less ecologically valid, less aesthetically appealing and may inhibit a positive emotional reaction. For example, in the pioneering experiment by Blood et al., (1999), the failure to induce a positive emotion reaction was attributed to the lack of musical expression in the computerized stimuli (Koelsch et al., 2010). In this case using real musical examples were more effective (Blood and Zatorre, 2001).
2. *Emotion impact*. A lack of performance variations will make the emotional communication less efficient. A number of studies have demonstrated that it is enough to vary the performance variables, such as tempo, dynamics, articulation, of one single musical score in order to efficiently communicate different emotions (Bresin and Friberg, 2000; Juslin and Timmers, 2010). Thus, a more effective communication is obtained by combining both structural aspects and performance variables. It is also important to find an optimal combination of performance variations for each score and each portrayed emotion.
3. *Optimal selection of stimuli*. Using one or a few experts to select the music examples and how they are performed may not result in an optimal set of music examples. There is an individual variation regarding the selection of features used for emotion decoding. For example, Juslin (2000) in modeling the emotion communication between performer and listener found that musicians are not using exactly the same features or features levels compared to listeners. An improved method is to use a larger database of music examples, rate them in a listening experiment with a group of listeners regarding the emotional communication and then select an optimal subset (Veillard et al., 2007; Eerola and Vuoskoski, 2010).

Recent developments in computer music software make it possible to generate musical performances similar to how musicians render a piece of music. The Director Musices (DM) software (Friberg et al., 2000) implements the KTH performance rules (overview in Friberg et al., 2006), which contains more than 20 different context-dependent rules for controlling the performance in relation to phrasing, melodic and harmonic accents, articulation, micro-timing, and intonation. This system has been extensively used to model music performance as well as making controlled stimuli for perceptual experiments (e.g., Juslin et al., 2002). It has also been used for automatic rendering of the seven different emotional expressions happiness, sadness, anger, solemnity, tenderness, fear, and neutral (Bresin and Friberg, 2000). Recently this system has been implemented in the real-time computer program pDM, where a user can control the amount of each rule while the music is playing (Friberg, 2006). Furthermore, different mappings to small semantic spaces are available, such as the activity–valence space. This opens up for a number of new applications such as expressive conducting in which the performance can be directly controlled by arm gestures using a video camera (Friberg, 2005), or by full-body expressive movements (Castellano et al., 2007).

By using the rule system, stimuli can be controlled also regarding musical expression, thus one may combine both structural and performance variations in a systematic way, thus improving the emotional impact. However, to produce convincing and optimal performances, there are a number of musical parameters that have to be determined. The present study is the first in a series of two in which the detailed relation between music variables and emotional expression is investigated. The main aim of the research work presented in this paper was to investigate the optimal combination of