Proceedings of the 3rd International Conference on Music & Emotion (ICME3), Jyväskylä, Finland, 11th - 15th June 2013. Geoff Luck & Olivier Brabant (Eds.)

WHAT EMOTIONS AND FREE ASSOCIATIONS CHARACTERIZE DIFFERENT MUSICAL STYLES?

Erica Bisesi, Florian Eckl, Richard Parncutt

Centre for Systematic Musicology, University of Graz, Austria
erica.bisesi@uni-graz.at

Abstract

Bisesi, Parncutt, & Friberg (2011) extended *Director Musices* (DM) - a software package for automatic rendering of expressive performance (Friberg et al., 2006) that relates the expressive features of a performance not only to global or intermediate structural properties, but also to local events (grouping, metrical, melodic and harmonic accents). Friberg & Bisesi (2013) applied the new rule system to model stylistic variations corresponding to different historical periods and styles in piano music. In this project, we extend pDM - a *pure data* system for real-time expressive control of music performance (Friberg, 2006) - by adjusting algorithmic formulations inspired by a theory of musical accents and by introducing new gestural interfaces based on qualitative data. First, different presets corresponding to different stylistic conventions and interpretations are implemented in DM, and corresponding sound files are generated and recorded. Second, we ask 20 participants (10 musicians and 10 amateurs) to rate the character of the performances on rating scales depending on a previous categorization involving emotions and free associations (Bisesi, Bodinger, & Parncutt, in progress). Terms involve basic and complex emotions, and free associations such as static/dynamic and living/non-living things. Results will be used to map music expression into multidimensional spaces of emotions and free associations, and to develop new gestural interfaces in pDM.

Keywords: performing styles, emotions, free associations

1. Introduction

The noteheads in a musical score may be equal in size, but when the music is performed, the notes do not sound equally important. Accents are local events that attract a listener's attention and are either evident from the score (immanent) or added by the performer (performed) (Parncutt, 2003). Immanent accents are associated with (temporal, serial) grouping (phrasing), metre (downbeats), melody (peaks, leaps), and harmony (or dissonance). Performed accents involve changes in timing, dynamics, articulation, and timbre; they vary in amplitude, form (amplitude as a function of time), and duration (the period of time during which the timing or dynamics are affected).

The usual relationship between immanent and performed accents is that performers tend to "bring out" immanent accents, i.e. to attract the listener's attention to them. For example, a performer may slow the tempo or add extra time in the vicinity of certain kinds of immanent accent, or change dynamics or articulation in consistent ways. This relationship is complex and depends on many factors such as musical and personal style, local and cultural context, intended emotion or meaning, and acoustical and technical constraints. In both cases, the perceptual *salience* of an accent is its perceptual or subjective importance, or the degree to which it attracts a listener's attention.

Following the analysis-by-synthesis approach of Sundberg (1988) and Friberg (1991), and their rule-based performance rendering system *Director Musices* (DM) (Friberg et al., 2006), we have developed a new algorithmic model of music expression that relates the expressive features of a performance not only to global or intermediate structural properties (i.e. different levels of phrasing), but also to local events (individual notes corresponding to accents) in a systematic way (Bisesi and Parncutt, 2011; Bisesi, Parncutt & Friberg, 2011; Friberg & Bisesi, 2013). DM is a software package for automatic rendering of expressive performance developed by Sundberg and Friberg that comprises performance rules (mathematically defined conventions of music performance), which change specific note properties, including timing, duration and intensity. By manipulating program parameters, meta-performers can change the degree and kind of expression by adjusting the extent to which each rule is (or all rules are) applied.

In its previous formulation, the main structural principle of DM is phrasing. The *Phrase Arc* rule assigns arch-like tempo and sound-level curves to phrases that are marked in the score. DM also models aspects of tonal tension. The *Melodic Charge* rule emphasizes tones that are far away from the current root of the chord on the circle of fifths, and the *Harmonic Charge* rule emphasizes chords that are far away from the current key on the circle of fifths. Several of the rules presented in DM have been interpreted in terms of Parncutt's (2003) taxonomy of accents. This suggests that a conflation of the two models may yield new insights into expressive performance and possibly lead to artistically superior computer-rendered performances.

On this basis, we have developed DM in a new direction, involving both the previous set of rules, and the new formulation based on accents. Our approach to performance rendering involves two separate stages. First, we automatically extract accents from the score. The output of the first stage comprises two independent hierarchical structures (for grouping and metrical accents) and a series of harmonic and melodic accents (considered as local events). For each hierarchical level or accent, our model estimates the musical importance or perceptual salience. In a second stage, we manipulate timing and dynamics in the vicinity of immanent accents (e. g., getting slower and/or louder near an accent). Let's now describe these two separate stages in detail.

According to our model, a *phrasing* (or *grouping*) accent occurs at the start of a phrase at any hierarchical level. The listener's attention is drawn to structural boundaries because they delineate the structure: the first note of a phrase or section is important because it announces something new, and the last note is important because it announces the end of a group. Repp (1998) demonstrated the psychological reality of phrases by showing that listeners are generally more sensitive to what happens near phrase boundaries than within phrases, and Bisesi, MacRitchie & Parncutt (2012) found that there is a general agreement about the position of phrase beginnings, endings and climaxes across performances and listeners. The salience of grouping accents corresponds to the number of levels at which a given event marks the start of a phrase. The phrasing structure can be determined by subdividing the entire piece into longer phrases, then dividing each phrase into subphrases and so on. The starts and ends of subphrases and sub-subphrases are marked according to repetitions of motives or introduction of new structural elements. Our model of phrasing-based timing is based on music analysis rather than physics (Todd, 1995). We begin by looking at accents within the phrase and apply timing/dynamic curves which are adjusted to start or end at phrase boundaries (just as phrase boundaries have been found in psychological studies to mark chunks of musical memory).

A *metrical* accent occurs at the start of important metrical units such as the start of a measure or group of measures (hypermeter). Again, salience is the number of different levels of pulsation (beats) to which an accent belongs.

A *melodic* accent occurs at local peaks and valleys of the melody and following leaps. To identify melodic accents, we first label the highest and lowest tones of the whole melody, and then label the local peaks and valleys, i.e. the highest and lowest pitches in a given phrase. Melodic accent salience depends on a number of factors. Peaks normally have more salience than valleys. Tones that are further away from the average pitch of the melody (or the local average of a phrase) are more salient.

Tones immediately preceding or following large intervals have more salience.

The *harmonic* accent of a chord in a progression has several components of dissonance: roughness, harmonic ambiguity, harmonic relationship to context, and familiarity or expectedness. Harmonic accents in major-minor tonal music are produced by tones foreign to the prevailing key. If the key of a passage is relatively clear, the salience of this kind of accent can be predicted using the key profiles of Krumhansl and Kessler (1982). These profiles may be considered as quantitative estimates of the harmonic stability of each tone in the chromatic scale in a given major or minor key. The lower the stability of a tone, the greater the harmonic accent at that tone. The harmonic accent of a chord may be estimated by combining accents for individual tones.

Figure 1 provides an example of subjective immanent accentuation. Here, the first accent in the upper voice in the first bar is a local peak relative to previous and following tones; because the peak is relatively prominent we have assigned a melodic accent (C) with intermediate salience (3). As peaks normally have more salience than valleys, and the melodic theme is played by the upper voice, the simultaneous melodic valley in the lower voice has lower salience (2). The second melodic peak in the inner voice of bar 1 is preceded by a smaller interval than the previous one, so the melodic accent has even lower salience (1). The first chord in the first bar feels new by comparison to the preceding context, so we marked a harmonic accent of salience 3. The harmonic accent at the end of bar 1 is a roughness accent.

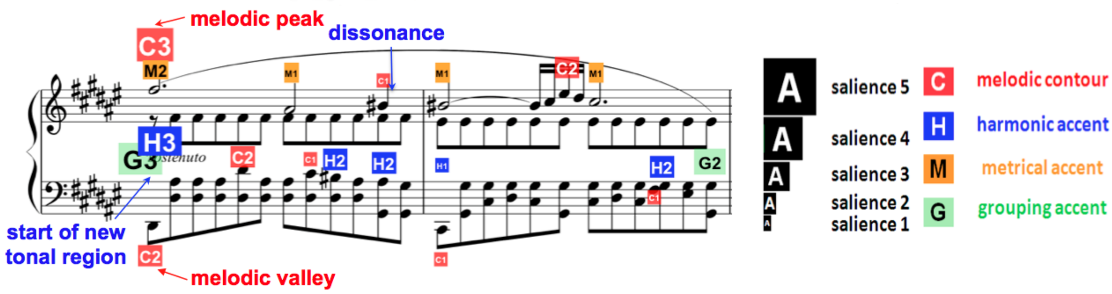

Figure 1. Subjective analysis of immanent accents and their salience in the first two measures of the central section of Prelude Op. 28 No. 13 by Frédérick Chopin.

On this basis, Parncutt, Bisesi & Friberg have developed a computational model of immanent accent salience in DM (Parncutt, Bisesi & Friberg, paper in progress). The new formulation of DM automatically extracts metrical, melodic and harmonic accents from a score and provides to each accent a salience from 1 to 5. Regards grouping accents analysis, as different theorists might offer different analyses for the same passage, the output is sometimes ambiguous and we have not attempted an automatic analysis here.

Automatic notation of metrical accents starts by defining bars (the notated barline as the slowest pulse) and beats (pulses next to bars, as determined by key signature conventions). Then different pulses or metrical levels are defined including groups of bars (hypermetre) and each metrical level is weighted with a Gaussian curve with a peak near 600 msec. (Parncutt, 1994). Finally, the code evaluates the salience of any metrical level as determined by the weighting curve, and the metrical accent salience of each onset in the score by superposing two or three pulses and adding the saliences of pulses that coincide at a given point.

Melodic accent salience depends on two factors: the size of the leap preceding or following the accent, and the distance of the tone from the center of the melody's range. These are combined together so that for example repeated very high or low tones receive small melodic accent because the preceding interval is zero. Melodic peaks receive stronger accents than melodic valleys (arbitrary set at twice).

As since now there is no accepted general model for the dissonance of a sonority in western music, an algorithm for harmonic accents is difficult to formulate. At this stage, to predict the salience of harmonic accents in single tones or in chords, we use existing rules for Melodic and Harmonic Charge.

In the second stage, the code automatically manipulates timing and dynamics in the vicinity of accents. Each accent is modeled by means of two new functions, respectively for timing (duration) and dynamics (sound level) variations in the vicinity of the accent. These rules associate to each accent an arch-like tempo curve and a sound level curve according to a given parameterization. Each function admits 5 free parameters: the event peak, the width of the interval preceding the accent, the width of the interval following the accent, the shape of the curve before the peak, and the shape of the curve after the peak. Curves' shapes are linear, quadratic, cubic, exponential, Gaussian, cosine, and hand-gesture (a function using a mathematical model for approximating point-to-point hand gestures; Juslin et al., 2002). As most of the rules in DM, accent rules have a general quantity parameter k, which is used to control the overall degree to which each rule is applied. A value of k = 1 corresponds roughly to a neutral use of a rule, i.e. a noticeable but modest variation determined intuitively during the development phase. For timing and dynamics, an increment of k = 1 in a rule means an increase of 10% in the local tempo and an increase of 4 dB in the local sound level respectively. The quantity parameter is multiplicative in that k = 2 normally double the amount added (or multiplied) to the parameter in question. As the new accent rules are often triggered at the same position in the score (for example, a harmonic change often coincides with the start of a measure thus triggering both the harmonic and metrical accent), we also implemented a scheme for rule interaction (Friberg & Bisesi, 2013). Our system also gives the user the opportunity to manipulate emotional effect by changing the shapes of the curves.

A further aspect of our algorithmic approach to accent salience is the dependence on the stylistic context. Friberg & Bisesi (2013) addressed the question of how a performance style can be conceptualized as a collection of expressive gestures, and provided a set of examples belonging to different historical contexts (a baroque, a romantic and a modern piano piece respectively) and featuring different stylistic conventions as representative of the main performance techniques applied by musicians in their renditions. The aim there was to code specific performance styles as a set of DM performance rules, in order to model and automatically render these pieces by associating style-specific rule palettes. Based on that study, in the current project we asked musical participants to rate rule presets corresponding to different stylistic conventions and interpretations of pieces.

2. Aims

Our computational model of immanent accent salience in tonal music in DM (Bisesi, Parncutt & Friberg, forthcoming) addresses several questions, such as: (i) According to what general principles can immanent accents be identified? (ii) What general principles determine their salience? (iii) How are tempo, dynamics and other parameters typically varied in the vicinity of accents? (iv) Over what time period before and after an accent are they varied? (v) What is the shape of the timing or dynamic curve leading up to and away from the accent? (vi) How do musically acceptable performances fit with possible ranges of parameter values? (vii) Which parameter ranges correspond to particular qualities of performance such as emotions and free associations? (viii) How do all these findings depend on stylistic context?

In the remainder of this paper, we will discuss a preliminary experiment whose aim was to contribute to an answer to questions (vi), (vii) and (viii).

3. Main contribution and methods

To get insight into the relationship between performed accents and performed emotions, we related together different categories of accents, corresponding model parameters, and specific properties of a performance, such as musicality, expressivity, and evoked feelings. According to the analysis-by-synthesis procedure, musicians with different levels of expertise evaluated different renditions of selected passages and pieces, and parameters were adjusted accordingly.

Our methodology involves two different stages. In a first stage, we investigated the relationship between possible parameter values and musically acceptable performances. In the second stage, presets of parameters selected in the first stage will be linked to specific categories of emotions and free associations.

The first stage comprised three tasks. In each task, sound stimuli corresponding to different combinations of accents and model parameters were presented to expert listeners and amateurs for evaluation. Data taking is still in progress and will involve 20 participants (10 musicians and 10 amateurs). The aim of the first task was to quantify the musical acceptability of the pieces. Participants were asked to rate on a scale from 1 to 10 how much they liked the way the performers were playing, where 1 stood for "not at all" and 10 stood for "a lot". We asked participants to imagine they were hearing the sound stimuli in a concert or on a CD and to rate only the interpretation and not the music itself. The aim of the second task was to quantify the degree of expressivity of different renditions as compared with deadpan performances. We asked participants to rate on a scale from 1 to 10 the difference between two sound stimuli (a deadpan and an expressive performance), where 1 stood for "small" and 10 stood for "big". Deadpan performances correspond exactly to the score with the same MIDI velocity for each tone. The aim of the third task was to describe automatic performance renditions from the viewpoint of the listeners. Participants were asked to describe the sound stimuli subjectively from their viewpoint and in their native language with the first adjectives (feelings, images, or other descriptions) they think of. Between the three tasks, participants had the possibility to take short breaks. Overall, the experiment was about 40 minutes long.

To design the experiment, we used the open source software Psychopy v1.76.00 (Peirce, 2008). Before hearing to the stimuli, participants read instructions on the computer screen. During the experiment, they listened to different stimuli and evaluated each stimulus by clicking on a rating scale, which appeared on the screen at the end of the sound.

We used 27 different sound stimuli. There were 9 different interpretations for each of 3 short excerpts belonging to different musical styles - from the baroque, the classical and the romantic periods respectively: the Bach Bourrée from Cello Suite No. 3 BWV 1009, the first movement of Haydn Quartet Op. 74 No. 2, and the first movement of Mendelssohn Violin Concert Op. 64. Scores for the three excerpts are shown in **Figure 2**.

Figure 2. Musical scores for the three pieces (1: Bach, 2: Haydn, 3: Mendelssohn; details in the text) used in this study (adapted with the open source music composition and notation software Musescore 1.1).

Interpretations were produced by the new version of DM (based on automatic mark of accents). For each piece, there were 8 different expressive performances and a deadpan performance. Expressive performances differed from one another by emphasis on different kinds of accents (metrical and melodic/harmonic respectively), by the presence of timing or dynamics variations in the vicinity of accents, and by the mathematical curves employed to model timing and dynamics variations on accented notes or chords (steep versus smooth profiles). Stimuli were produced by systematically adjusting model parameters in the new version of DM, and corresponding midi files were listened to on a Clavinova CLP 370 previously calibrated with our system. After having been recorded through the open source software Audacity 2.0.2, stimuli were stored in wav format and presented to participants. Each stimulus lasted approximately 15 sec.

In task 1 and 3, each of the 27 performances was provided to participants separately. In task 2, for each piece, each stimulus consisted of a combination of a deadpan performance and one of the 8 expressive performances, separated by a gap of 3 sec. In all the three tasks, stimuli were presented in a random order.

Table 1 lists stimuli provided to participants in each task. Columns 1-2 and 3-4 refer to task 1 and 3, and task 2, respectively. Numbers in columns 1 and 3 indicate the stimuli as referred to in the data analysis. Columns 2 and 4 illustrate the expressive content of the stimuli by mean of acronyms. Acronyms "1", "2", and "3" refer to different pieces of music (Bach, Haydn, and Mendelssohn, respectively); "N" stands for nominal (or deadpan); "M" and "CH" mean that expressive variations are applied to metrical or melodic/harmonic accents respectively; "T" and "D" stand for local deviations in the tempo or in the dynamics; "W1" and "W2" indicate that the profiles of the curves employed to model expressive variations are steep or smooth respectively. In order to make values comparable, all performances were produced with the same default values (as currently set in DM) and the same quantity (k = 1).

Table 1. Stimuli provided to participants in tasks 1, 2, and 3. Acronyms' explanation: "1": Bach; "2": Haydn; "3": Mendelssohn; "N": nominal (or deadpan) performance; "M": expressive performance based on metrical accents; "CH": expressive performance based on melodic and harmonic accents; "T": local deviations in the tempo; "D": local deviations in the dynamics; "W1": steep curve profile; "W2": smooth curve profile.

TASK 1, TASK 3		TASK 2	
#	**performance**	**#**	**performance**
1	1N		
2	1MTW1	1	1N – 1MTW1
3	1MTW2	2	1N – 1MTW2
4	1MDW1	3	1N – 1MDW1
5	1MDW2	4	1N – 1MDW2
6	1CHTW1	5	1N – 1CHTW1
7	1CHTW2	6	1N – 1CHTW2
8	1CHDW1	7	1N – 1CHDW1
9	1CHDW2	8	1N – 1CHDW2
10	2N		
11	2MTW1	9	2N – 2MTW1
12	2MTW2	10	2N – 2MTW2
13	2MDW1	11	2N – 2MDW1
14	2MDW2	12	2N – 2MDW2
15	2CHTW1	13	2N – 2CHTW1
16	2CHTW2	14	2N – 2CHTW2
17	2CHDW1	15	2N – 2CHDW1
18	2CHDW2	16	2N – 2CHDW2
19	3N		
20	3MTW1	17	3N – 3MTW1
21	3MTW2	18	3N – 3MTW2
22	3MDW1	19	3N – 3MDW1
23	3MDW2	20	3N – 3MDW2
24	3CHTW1	21	3N – 3CHTW1
25	3CHTW2	22	3N – 3CHTW2
26	3CHDW1	23	3N – 3CHDW1
27	3CHDW2	24	3N – 3CHDW2

The design for the second stage is based on results from the first stage. In the following, we will discuss results from a preliminary investigation concerning the first stage, while design and results from the second stage will be presented at the conference.

4. Preliminary results

Results for an explorative study conducted on a small number of listeners (5) are displayed in **Figures 3** and **4.**

Figure 3 shows ratings provided to the 24 expressive (i.e., non deadpan) performances presented in task 1, as a function of the piece (upper-left), the typology of accent (upper-right), the parameter deviating from the nominal value (bottom-left), and the slope of the curve used to model expressive deviations (bottom-right). For these selected pieces, we found that participants tended to prefer baroque- and classical-style performances (p=0.0056), whose expressive variations from the nominal score are provided by changes in the dynamics (p=0.0005).

As performances were produced with the same default values in DM, we assume that they are comparable and interpret differences in the ratings as due to the different way different parameters affect perceptual listening. Nevertheless, the case that default values determined intuitively during the development phase in DM are not correct is possible, and should be further tested.

A possible reason why system worked better in the cases listed above is that, in current formulation, our model is not yet appropriate for romantic music (of which piece 3 is a typical example). A possible explanation of higher ratings provided to excerpts based on dynamical variations is that these pieces may imply a motor quality that is better achieved if the tempo is regular.

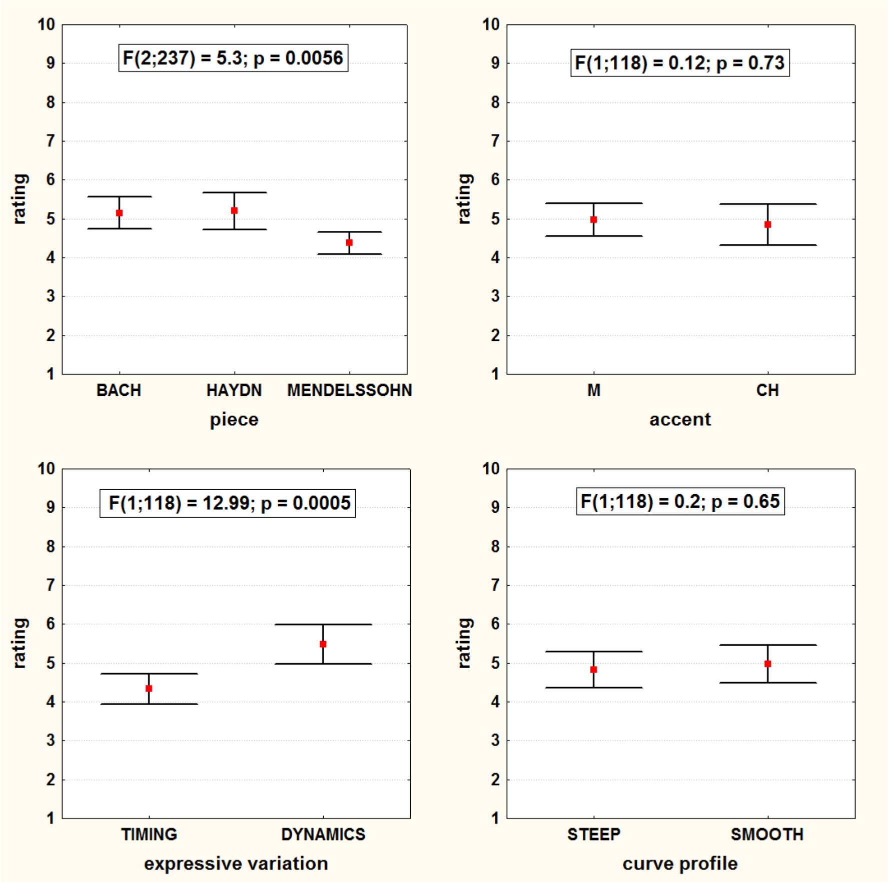

Figure 3. Ratings provided to the 24 expressive (i.e., non deadpan) performances of task 1, as a function of the piece (upper-left), the typology of accent (upper-right), the parameter deviating from the nominal value (bottom-left), and the slope of the curve used to model expressive deviations (bottom-right). Vertical bars denote 0.95 confidence interval.

The four panels in **Figure 4** display ratings provided to stimuli presented in task 2. Consistently with results for task 1, baroque- and classical-style performances were rated as more different from deadpan performances than romantic performances (p=0.03), and ratings were higher when differences were due to changes in the dynamics than in the timing (p=0.007). Furthermore, renditions based on emphasis on melodic and harmonic accents modelled by mean of steep curve profiles were perceived as more different from deadpan performances than renditions based on metrical accents and/or smooth curve profiles. A possible explanation is that melodic and harmonic accents occur less regularly than metrical accents, and therefore they are perceived as more unexpected events. The more a performance sounds unexpected, the higher it is rated as different from a deadpan performance. The degree of unexpectation is higher when expressive features are achieved suddenly (by mean of a steep curve) than gradually (by mean of a smooth curve).

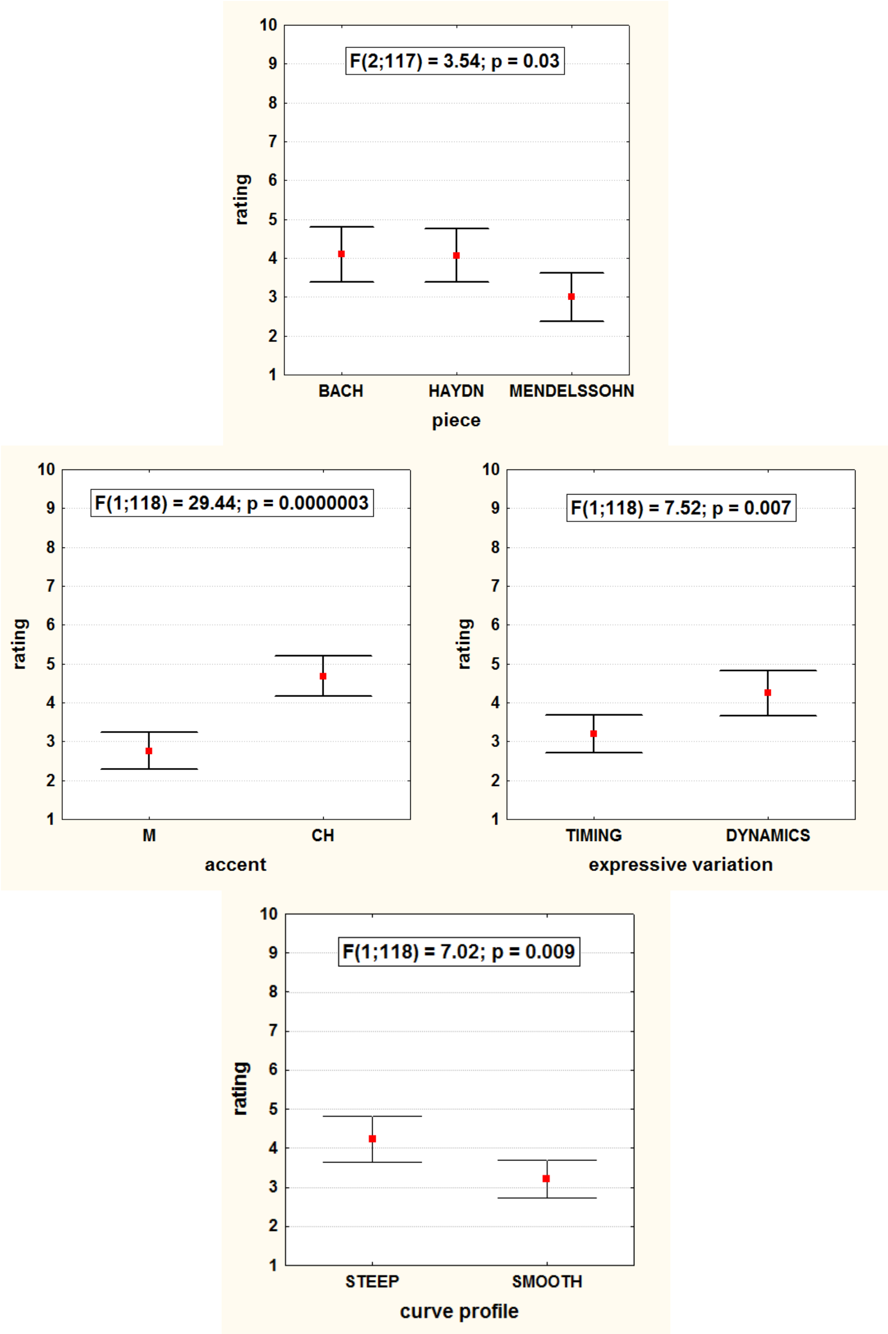

Figure 4. Ratings provided to the 24 performances of task 2, as a function of the piece (upper-left), the typology of accent (upper-right), the parameter deviating from the nominal value (bottom-left), and the slope of the curve used to model expressive deviations (bottom-right). Vertical bars denote 0.95 confidence interval.

As amounts of different rules are still under testing, a detailed statistical analysis may seem inappropriate. Nevertheless, we carried out the following preliminary analysis, in order to get insight into general tendencies.

According to the Kolmogorov-Smirnov test, in both task 1 and 2 performance ratings are not normally distributed (Z=0.19, p=0.000); therefore, we run non-parametric tests. In task 1, the Kruskall-Wallis test showed no significant main effect on the performance factor (X2=28.175, df=26, p=0.35) (probably due to the small number of participants). Mann-Whitney-U test returned significant differences between performances 6 and 9 (U=2, p=0.03), 7 and 9 (U=2, p=0.03), 9 and 16 (U=1, p=0.016), 9 and 21 (U=1, p=0.016), and 9 and 24 (U=2, p=0.03). On average, performance 9 received higher ratings (M=6.8) than performance 6 (M=4.4), 7 (M=3.8), 16 (M=4), 21 (M=3.8), and 24 (3.6). In task 2, the Kruskall-Wallis test showed a significant effect on the performance factor (X2=58.66, df=23, p=0.0). According with the Mann-Whitney-U test, there are significant differences between performances 1 and 5 (U=0, p=0.008), 1 and 6 (U=0, p=0.008), 1 and 7 (U=0, p=0.008), 1 and 8 (U=0, p=0.008), 1 and 9 (U=1.5, p=0.016), 1 and 11 (U=1.5, p=0.016), 1 and 13 (U=0, p=0.008), 1 and 14 (U=0, p=0.008), 1 and 15 (U=1.5, p=0.016), 1 and 23 (U=1.5, p=0.016), 1 and 24 (U=1.5, p=0.016), 2 and 5 (U=1.5, p=0.016), 4 and 5 (U=0.5, p=0.008), 5 and 10 (U=1, p=0.016), 5 and 12 (U=2, p=0.03), 5 and 17 (U=1.5, p=0.016), 5 and 18 (U=0, p=0.008), 5 and 20 (U=1, p=0.016), and 5 and 22 (U=1, p=0.016). On average, participants' rating for performance 1 (M=1.6) are closer to ratings for performance 5 (M=6), 6 (M=5.4), 7 (M=5.6), 8 (M=4.8), 9 (M=3.2), 11 (M=5.4), 13 (M=4.2), 14 (M=3.8), 15 (M=6), 23 (M=5), and 24 (M=3.8). Participants' ratings for performance 5 (M=6) are more distant to ratings for performance 2 (M=2.6), 4 (M=2.2), 10 (M=2), 12 (M=2.6), 17 (M=1.8), 18 (M=1.6), 20 (M=2.2), and 22 (M=2.2).

Specific qualities of performances as provided by participants' descriptions in task 3 will be used in a further stage of the study. In general, stimuli were described with words concerning musical aspects related to the structure, feelings and free associations. For instance, they described performances with words like "slow", "experiencing something new", "light-footed", "waking up" (performance 1), "harmonious", "relaxing", "developing ideas" (performance 14), "very accented", "not fluent", "being sick" (performance 20), "too slow", "not fluent", "more accented", "nervous", "confusion" (performance 24), "exiting", "vital", "stressful", "danger" (performance 25), "harmonious", "cozy", "meditative", "nursery rhyme", "children birthday" (performance 27).

Final results, carried on the whole set of participants, will be presented at the conference.

5. Implications

Stimuli associated with best ratings in task 1 and 2 will be used in the second stage of the experiment, whose aim is to relate parameters' presets to specific words describing emotions and free associations. These words will be selected in two ways. First, we will use descriptions provided by participants in task 3. Second, we will include new categories of emotions and free associations obtained in separate studies.

Bisesi & Parncutt (2010) explored listeners' informal vocabularies for describing expression in piano music. Most words felt into three categories: musical structure, emotion, and free association. Subcategories included immanent/performed attributes of the sound (tempo, dynamics, pitch and timbre), accentuation, character and meaning, basic/complex emotions and static/dynamic or living/non-living associations. Bisesi, Bodinger, & Parncutt (2013, in progress) are now investigating the relationship between emotions and free associations in more detail. 21 participants (7 musicians, 7 amateurs and 7 non-musicians) are listening to 8 piano pieces in different "classical" styles. They are then asked to describe their experience of the music (open question). After that we ask them to focus on emotions and free associations (or abstract images). Results are currently being processed. We are exploring qualitative relationships and quantitative correlations between subcategories of emotions and images by grounded theory, multidimensional scaling, and correspondence analysis.

We are mapping out possible ranges of parameter values in a multidimensional parameter space that correspond to musically acceptable performances, and specifying parameter ranges that correspond to particular qualities of performance as expressed by words as bright and dark, joyful and sad, static and dynamic, expected and surprising. Our target is to associate a given accent's salience (as estimated by automatic musical analysis in DM) to specific "effective" combinations or presets of model parameters, then describe and classify presets according with emotions and free associations by mean of a semantic differential (a collection of rating scales to evaluate different aspects of an object, event or concept).

Results will be used to develop new gestural interfaces in pDM - a *pure data* system for real-time expressive control of music performance (Friberg, 2006). The current version of pDM is based on the old formulation of DM and includes several examples of emotional expression mappers. They are the arousal-valence two-dimensional space commonly suggested for describing emotional expression in the dimensional approach (happy, tender, sad, and angry; Sloboda & Juslin, 2001), a kinematics-energy space (hard, light, heavy, soft; Canazza et al., 2003), and a gesture-energy space (gentle, expressive, light, and strong; Friberg, 2006). We plan to implement a new version of pDM based on our model of accents by adjusting algorithmic formulations inspired by accent theory, by introducing new gestural interfaces based on our qualitative data, and by relating different regions in multidimensional spaces with specific music and performing styles.

This research is supported by the FWF Stand-Alone Project P 24336-G21 "Expression, emotion and imagery in music performance".

References

Bisesi, E., & Parncutt, R. (2010). The informal vocabulary of professional musicians for describing expression and interpretation. Paper presented at ICMPC11 - 11th International Conference on Music Perception and Cognition, Seattle, WA, USA, 23-27 August 2010.

Bisesi, E., & Parncutt, R. (2011). An accent-based approach to automatic rendering of piano performance: Preliminary auditory evaluation. *Archives of Acoustics,* 36(2), 1-14.

Bisesi, E., Parncutt, R., & Friberg, A. (2011). An accent-based approach to performance rendering: Music theory meets music psychology. In A. Williamon, D. Edwards and L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science, Toronto, Canada, 24-27 August 2011* (pp. 27-32). Utrecht, The Netherlands: European Association of Conservatoires (AEC). ISBN: 978-9-490306-02-1.

Bisesi, E., Bodinger, M., & Parncutt, R. (2013). Listeners' informal vocabulary for emotions and free associations in piano music. *Abstract accepted for spoken presentation at ICME3 - 3rd International Conference on Music & Emotion,* Jyväskylä, Finland, 11-15 June 2013.

Bisesi, E., MacRitchie, J. & Parncutt, R. (2012). Recorded interpretations of Chopin Preludes: Performer's choice of score events for emphasis and emotional communication. In E. Cambouropoulos, C. Tsougras, P. Mavromatis and K. Pastiadis (Eds.), *Proceedings of ICMPC-ESCOM, Thessaloniki, Greece, 23-28 July 2012* (pp. 106-107). ISBN: 978-960-99845-1-5.

Canazza, S., De Poli, G., Rodà, A., & Vidolin, A. (2003). An abstract control space for communication of sensory expressive intentions in music performance. *Journal of New Music Research*, 32(3), 281-294.

Friberg, A. (1991). Generative rules for music performance. *Computer Music Journal,* 15(2), 56-71.

Friberg, A. (2006). pDM: An expressive sequencer with real-time control of the KTH music performance rules. *Computer Music Journal*, 30 (1), 37-48.

Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology,* 2(2-3), 145-161.

Friberg, A., & Bisesi, E. (2013, in press). Using computational models of music performance to model stylistic variations. In D. Fabian, E. Schubert and R. Timmers (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures.* Oxford: Oxford University Press.

Juslin, P. N., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in performance: The GERM model. *Musicae Scientiae,* Special issue 2001-2002, 63-122.

Parncutt, R. (1994). Perceptual model of pulse salience and metrical accents. *Music Perception*, 11, 409–464.

Parncutt, R. (2003). Accents and expression in piano performance, In K. W. Niemöller (Ed.), *Perspektiven und Methoden einer Systemischen Musikwissenshaft* (Festschrift Fricke) (pp. 163-185). Frankfurt/Main, Germany: Peter Lang.

Peirce, J. W. (2008). Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in neuroinformatics*, 2, 10.

Repp, B. H. (1998). Variations on a theme by Chopin: Relations between perception and production of timing in music. *Journal of Experimental Psychology, Human Perception and Performance*, 24, 791-811.

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

Sundberg, J. (1988). Computer synthesis of music performance. In J. A. Sloboda (Ed.), *Generative processes in music*. Oxford: Clarendon.

Todd, N. P. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*, 97, 1940–1949.