International Symposium on Performance Science
ISBN 978-94-90306-02-1

An accent-based approach to performance rendering: Music theory meets music psychology

Erica Bisesi[1], Richard Parncutt[1], and Anders Friberg[2]

[1] Centre for Systematic Musicology, University of Graz, Austria
[2] Department of Speech, Music, and Hearing, Royal Institute of Technology, Sweden

Accents are local events that attract a listener's attention and are either evident from the score (immanent) or added by the performer (performed). Immanent accents are associated with grouping, meter, melody, and harmony. In piano music, performed accents involve changes in timing, dynamics, articulation, and pedaling; they vary in amplitude, form, and duration. Performers tend to "bring out" immanent accents by means of performed accents, which attracts the listener's attention to them. We are mathematically modeling timing and dynamics near immanent accents in a selection of Chopin Preludes using an extended version of *Director Musices* (DM), a software package for automatic rendering of expressive performance. We are developing DM in a new direction, which allows us to relate expressive features of a performance not only to global or intermediate structural properties, but also accounting for local events.

Keywords: piano; expression; accents; timing; dynamics

Accents are local events that attract a listener's attention and are either evident from the score (immanent) or added by the performer (performed). Immanent accents are associated with grouping (phrasing), meter (downbeats), melody (peaks, leaps), and harmony (or dissonance; Parncutt 2003). In piano music, performed accents involve changes in timing, dynamics, articulation, and pedaling; 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). Performers tend to "bring out" immanent accents by means of performed accents, which attracts the listener's attention to them. For example, a performer may slow the tempo or add extra time in the vicinity of an immanent accent, or change dynamics or articulation in consistent ways. This relationship is complex and depends on musical and personal style, local and cultural context, intended emotion or meaning, and acoustical and technical constraints.

In a previous study, we asked ten music theorists to analyze a selection of Chopin Preludes by marking immanent accents on the score and evaluating their relative importance (salience). Agreement among participants was higher at phrase boundaries (grouping accents) than at melodic and harmonic accents. Phrase boundaries were determined by inter-onset interval (greater between than within phrases), contour (expected rise-fall arch shape), and meter (tendency for phrases to start on the beat).

In this study, we are mathematically modeling timing and dynamics near immanent accents in the central section of Chopin's *Prelude Op. 28 No. 13* using an extended version of *Director Musices* (DM), a software package for automatic rendering of expressive performance (Friberg *et al.* 2006). DM implements performance rules (mathematically defined conventions of music performance) that change the timing, duration, and intensity of individual tones. By manipulating program parameters, meta-performers can change

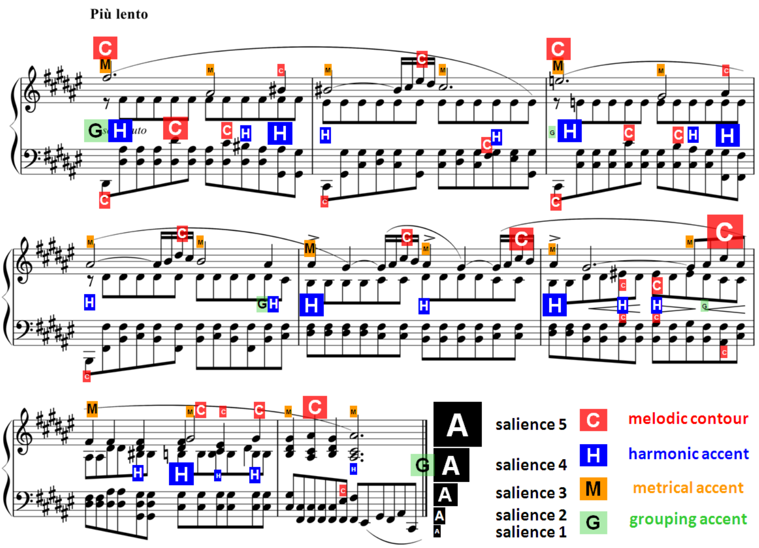

Figure 1. Subjective analysis of accents in the central section (bars from 21 to 28) of Chopin's *Prelude Op. 28 No. 13*. (See full color version at www.performancescience.org.)

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 (Sundberg *et al.* 2003). 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 can be interpreted in terms of Parncutt's (2003) taxonomy of accents, suggesting that a conflation of the two models may yield new insights into expressive performance and artistically superior computer-rendered performances. We are developing DM in this direction, which allows us to relate expressive features of a performance not only to global or intermediate structural properties (i.e. different levels of phrasing), but also accounting for local events (individual notes corresponding to accents) in a systematic way (Bisesi and Parncutt 2011).

MAIN CONTRIBUTION

Music analysis

The degree of accentuation of a note varies on a continuous scale which we call "salience." In Figure 1, immanent accents are divided into four types: melodic (or contour), harmonic, metrical, and grouping. The authors have subjectively assigned a salience rating between 1 and 5 to each accent, which is indicated by the size of the squares.

Melodic accents occur at the highest and lowest tones of the melody and at local peaks and valleys. For example, 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 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 low 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 lower salience.

The harmonic accent of a chord in a chord progression depends on its roughness, harmonic ambiguity, harmonic relationship to context, and familiarity or expectedness. The first chord in bar 1 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.

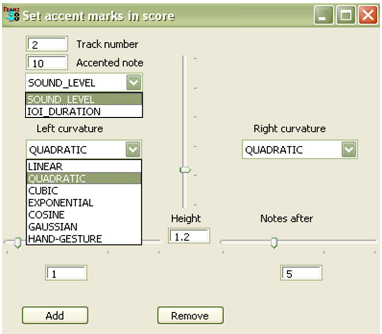

Figure 2. Graphical interface in the accent-based formulation of *Director Musices*.

Metrical and grouping accents depend on hierarchical metrical and phrasing structure. At the highest level, Figure 1 is one long phrase. It can be divided into two 4-bar sub-phrases of nominally equal importance, which in turn can be divided into sub-sub-phrases of two bars. In this case, the hypermetrical structure is indistinguishable from the phrasing structure.

Mathematical and computational model

We are modeling the timing and dynamics in the vicinity of an accent by two separate mathematical functions. Once the accents are marked in the score, the *Accent-Sl* and *Accent-Dr* rules give to them arch-like tempo and sound-level curves (here, the suffixes *Sl* and *Dr* respectively stand for sound level and duration). Each function admits five free parameters: the height of the peak, the duration before and after the peak, and the shape before and after the peak. Shapes may be linear, quadratic, cubic, exponential, Gaussian, cosine, or hand-gesture (Juslin *et al.* 2002). A graphical interface enables the performer to choose any combination of parameters (see Figure 2).

We are systematically evaluating different combinations of these parameters in given musical contexts, based on our artistic and professional experience as pianists. Different combinations of height, width, and curvature of both timing and dynamics can account for different performance qualities. The perceptual salience of the performed accent function depends on the area under a graph of beat duration or loudness against time. The greater the accent salience, the greater the height and/or width of the function. The curvature is not only connected with the perceptual salience, but also with the motion and emotional content (Juslin *et al.* 2002). We model the relationship among width, height, musical function, and expressive content in the follow-

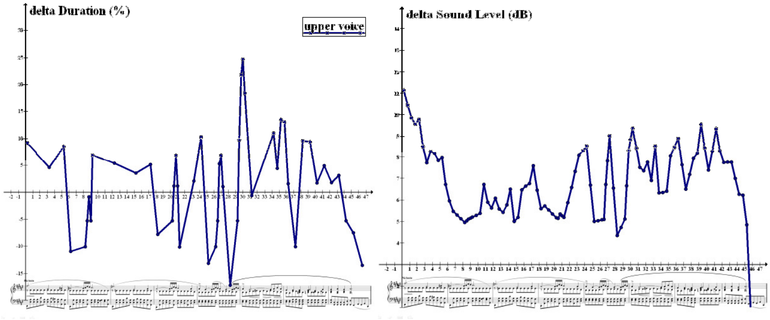

Figure 3. Example of mathematical modeling of timing (left panel) and dynamics (right panel) in the central section of Chopin's *Prelude Op. 28 No. 13*, according to the analysis of Figure 1. See text for description.

ing way: for linear function, we associate a combination of peak and width to a given salience according with the algorithm P+(W1+W2)/2=S+1, where P is the peak amplitude, W1 is the width interval preceding the accent, W2 is the width interval following the accent, and S is the salience. Units for P, W1 and W2 are defined so that a value of 1 corresponds to an increment of 4 dB in the sound level and 20% timing deviations, respectively. According to the algorithm above, any value of salience can correspond to many combinations of peak and width. For non-linear functions, salience is modeled by adapting any combination of peak and width to provide the same area below the graph as in the linear case (for each combination of P, W1, and W2). When a tone or a chord has more than an accent, profiles in timing and dynamics account for the root mean square of all the accents.

Figure 3 provides an example of rendition of the central section of Chopin's *Prelude Op. 28 No. 13*, according with the analysis of Figure 1: left panel shows the duration relative to the nominal duration of each note of the upper voice as a function of its position in the score, and right panel corresponds to the difference in sound level from the default value as a function of the note position (here, values of different voices are superimposed).

IMPLICATIONS

In a future study, different renditions of selected passages and pieces will be evaluated by pianists, theorists, and musicologists, and model parameters will be adjusted accordingly. We will map out possible ranges of parameter values or fields in multidimensional parameter space that correspond to musically acceptable performances. We will also specify small parameter ranges that correspond to particular qualities of performance as expressed by words obtained from a separate qualitative study, such as bright and dark, joyful and sad, static and dynamic, expected and surprising.

The theory can be applied in expressive music performance pedagogy. Students can learn the theory by working with a computer interface to create renderings of pieces that they are currently studying. In the process they will select immanent accents for accentuation and adjust the corresponding model parameters to achieve a desired result. They will then be in a position to apply the ideas behind the model in their performance and teaching.

Acknowledgments

This research is supported by Lise Meitner Project M 1186-N23 "Measuring and modelling expression in piano performance" of the Austrian Research Fund (FWF, Fonds zur Förderung der wissenschaftlichen Forschung).

Address for correspondence

Erica Bisesi, Centre for Systematic Musicology, University of Graz, Merangasse 70, Graz 8010, Austria; *Email:* erica.bisesi@uni-graz.at

References

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

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

Juslin P., Friberg A., and Bresin R. (2002). Toward a computational model of expression in performance: The GERM model. *Musicae Scientiae, Special issue 2001-02*, pp. 63-122.

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

Sundberg J., Friberg A., and Bresin, R. (2003). Attempts to reproduce a pianist's expressive timing with Director Musices performance rules. *Journal of New Music Research, 32*, pp. 317-325.