THE KINEMATIC RUBATO MODEL AS A MEANS OF STUDYING FINAL RITARDS ACROSS PIECES AND PIANISTS

Maarten Grachten
Dept. of Computational Perception
Johannes Kepler University, Linz, Austria
maarten.grachten@jku.at

Gerhard Widmer
Austrian Research Institute for
Artificial Intelligence, Vienna, Austria
Dept. of Computational Perception
Johannes Kepler University, Linz, Austria
gerhard.widmer@jku.at

ABSTRACT

This paper presents an empirical study of the performance of final ritards in classical piano music by a collection of famous pianists. The particular approach taken here uses Friberg and Sundberg's kinematic rubato model in order to characterize the variability of performed ritards across pieces and pianists. The variability is studied in terms of the model parameters controlling the depth and curvature of the ritard, after the model has been fitted to the data. Apart from finding a strong positive correlation of both parameters, we derive curvature values from the current data set that are substantially higher than curvature values deemed appropriate in previous studies. Although the model is too simple to capture all meaningful fluctuations in tempo, its parameters seem to be musically relevant, since performances of the same piece tend to be strongly concentrated in the parameter space. Unsurprisingly, the model parameters are generally not discriminative for pianist identity. Still, in some cases systematic differences between pianists are observed between pianists.

1 INTRODUCTION AND RELATED WORK

One of clearest manifestations of expressive timing in music is the final ritard, the slowing down toward the end of a musical performance to conclude the piece (or a part of it) gracefully. Several models have been proposed to account for the specific form of the ritard. These models typically come in the form of a mathematical function that describes how the tempo of the performance changes with score position. For example, Repp [9] found a quadratic function of score position to adequately describe IOI's measured in 28 performances. Honing [6] proposes a different kind of model, that consists in the combination of two computational models, one for tempo tracking, and one for rhythmic categorization. This model, rather than describing a single tempo curve, predicts the upper and lower boundary of the range of acceptable tempo curves for ritards. Since the two constituent models are intended to mimic perceptual processes involved in human listening, this can be called a *perceptual* model of expressive timing.

Another kind of models has arisen from the analogy of expressive timing with physical motion [11, 10, 3]. For example, Todd [11] describes a model for expressive timing where tempo is treated as the velocity of a particle that moves under constant acceleration or deceleration, depending on its position. The physical position of the particle is equated to score position with respect to phrase boundaries. Also lead by the analogy with physical motion, Friberg and Sundberg [4], derive a model for the velocity of human motion, when halting after running. An evaluation of the runner's stopping movement in terms of aesthetic quality yields that stopping motion with approximately constant deceleration power is rated highest. From the assumption of constant deceleration power, they derive a model of tempo as a function of score time.

As pointed out in [5], models that are dependent only on score position are incomplete in the sense that they ignore any characteristics of the musical material that is performed. Also, the physical motion metaphor ignores perceptual and production aspects of music performance that are relevant to the shaping of musical tempo [2, 5].

Nevertheless, the kinematic rubato models described above predict the evolution of tempo during the final ritard quite accurately, when matched to empirical data [4, 11]. An additional advantage of the models is their simplicity, both conceptually (they contain few parameters) and computationally (they are easy to implement).

In this paper we study the variability in the final ritards of Chopin's Nocturnes performed by multiple famous pianists, using Friberg and Sundberg's kinematic model. Rather than validating the model on empirical data, we use the model to learn about the data (as in [12]). More specifically, we investigate whether the identity of the piece or the pianist

SMC 2009, July 23-25, Porto, Portugal

is reflected in the parameters of the model. Given the simplicity of the two-parameter model, the existence of such an effect would be surprising, but would also shed some light on the interplay of personal interpretative freedom on the one hand, and performance practice and conventions on the other.

The data used for the study is described in section 2. Section 3 deals with the kinematic model and how it is applied to the measured data. Results are presented and discussed in section 4. Finally, section 5 states conclusions and remaining work.

2 DATA

The data used here consists in measurements of timing data of musical performances taken from commercial CD recordings of Chopin's Nocturnes. The contents of the data set are specified in table 1. We have chosen Chopin's Nocturnes since they exemplify classical piano music from the romantic period, a genre which is characterized by the prominent role of expressive interpretation in terms of tempo and dynamics. Furthermore, the music is part of a well-known repertoire, performed by many pianists, facilitating large scale studies.

Tempo in music is usually estimated from the interonset intervals of successive events. A problematic aspect of this is that when a musical passage contains few events, the obtained tempo information is sparse, and possibly unreliable, thus not very suitable for studying tempo. Therefore, through inspection of the score, we selected those Nocturnes whose final passages have a relatively high note density, and are more or less homogeneous in terms of rhythm. In two cases (Op. 9 nr. 3 and Op. 48 nr. 1), the final passage consists of two clearly separated parts, both of which are performed individually with a ritard. These ritards are treated separately (see table 1). In one case (Op. 27 nr. 1), the best-suited passage is at the end of the first part, rather than at the end (so strictly speaking, it is not a *final* ritard).

The data were obtained in a semi-automated manner, using a software tool [8] for automatic transcription of the audio recordings. From the transcriptions generated in this way, the segments corresponding to the final ritards were extracted and corrected manually by the authors, using *Sonic Visualizer*, a software tool for audio annotation and analysis [1].

3 METHOD

As mentioned in section 1, we wish to establish whether the specific form of the final ritard in a musical performance is dependent on the identity of the piece being played, or the performing pianist. We address this question by fitting a model to the data, and investigating the relation between the piece/pianist identity and the parameter values of the fitted model. We employ the kinematic model by Friberg & Sundberg [4], mainly for it's simplicity.

3.1 Friberg & Sundberg's kinematic model

The model is based on the hypothesized analogy of musical tempo and physical motion, and is derived from a study of the motion of runners when slowing down. From a variety of decelerations by various runners, the decelerations judged by a jury to be most aesthetically pleasing turned out to be those where the deceleration force is held roughly constant. This implies that velocity is proportional to square root function of time, and to a cubic root function of position. Equating physical position to score position, Friberg and Sundberg use this velocity function as a model for tempo in musical ritards. Thus, the model describes the tempo $v(x)$ of a ritard as a function of score position x:

$$v(x) = (1 + (w^q - 1)x)^{1/q} \quad (1)$$

The parameter q is added to account for variation in curvature (that is, the function is not necessarily a cubic root of position). The parameter w represents the final tempo, and was added since the tempo in music cannot reach zero. The model is designed to work with normalized score position and tempo. More specifically, the ritard is assumed to span the score positions in the range $[0, 1]$, and the initial tempo is defined to be 1.

The effect of the parameters w and q is illustrated in figure 1, which shows plots of tempo curves defined by the model for different values of w and q. Note that values of $q > 1$ lead to convex tempo curves, whereas values of $q < 1$ lead to concave curves. The latter is not expected to occur under normal circumstances, since tempo curves of ritards are typically convex. Note also that w determines the vertical end position of the curve.

3.2 Fitting the model to the data

The parameters of the model allow it to be fitted to ritards performed by particular pianists. As explained above, for this it is necessary to normalize the data. When normalizing the score position, it is important to make normalized position 0 coincide with the actual start of the ritard. Although in most cases there is a ritard instruction written in the score, the ritard may start slightly before or after this instruction. A manual inspection of the data showed that the starting position of the ritards strongly tended to coincide among pianists. For each piece, the predominant starting position was determined and the normalization of score positions was done accordingly.

When normalizing tempo, it is important to notice that normalizing should be done globally for the data set, rather

Pianist	Year	Op.9 nr.3 rit1	Op.9 nr.3 rit2	Op.15 nr.1	Op.15 nr.2	Op.27 nr.1	Op.27 nr.2	Op.48 nr.1 rit1	Op.48 nr.1 rit2
Argerich	1965			X					
Arrau	1978	X	X	X	X	X	X	X	X
Ashkenazy	1985	X	X	X	X	X	X	X	X
Barenboim	1981	X	X	X	X	X	X	X	X
Biret	1991	X	X	X	X	X	X	X	X
Engerer	1993	X	X	X	X	X	X	X	X
Falvai	1997	X	X	X	X	X	X	X	X
Harasiewicz	1961	X	X	X	X	X	X	X	X
Hewitt	2003	X	X	X	X	X	X	X	X
Horowitz	1957			X		X			
Kissin	1993					X	X		
Kollar	2007	X	X	X	X		X	X	X
Leonskaja	1992	X	X	X	X	X	X	X	X
Maisenberg	1995			X					
Mertanen	2001	X	X	X	X	X	X		
Mertanen	2002							X	X
Mertanen	2003							X	X
Ohlsson	1979	X	X	X	X	X	X	X	X
Perahia	1994			X					
Pires	1996	X	X	X	X	X	X	X	X
Pollini	2005	X	X	X	X	X	X	X	X
Richter	1968			X					
Rubinstein	1937	X	X	X	X	X	X	X	X
Rubinstein	1965	X	X	X	X	X	X	X	X
Tsong	1978	X	X	X	X	X	X	X	X
Vasary	1966	X	X	X	X	X		X	X
Woodward	2006	X	X	X	X	X	X	X	X
d´Ascoli	2005	X	X	X	X	X	X	X	X

Table 1. Performances used in this study. The symbol “X” denotes the presence of the corresponding combination of pianist/piece in the data set. The additions “rit1” and “rit2” refer to two distinct ritards within the same piece

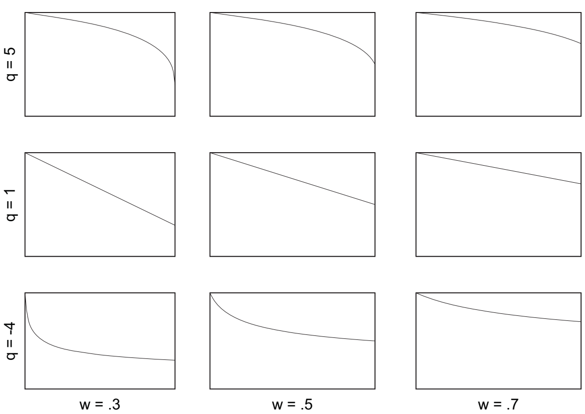

Figure 1. Ritards produced by the model using different values for the parameters w and q; In each plot, the x and y axis represent score position and tempo respectively, both in arbitrary units

than individually, since the latter would render the w parameter useless (the final tempo of every ritard would be 0). The result of global normalization is that the tempo value 1 corresponds to the highest tempo occurring in the data set, and the tempo value 0 to the lowest. Although this procedure maintains the relative scaling of the ritards, the majority of the ritards will not start with a tempo value of 1, whereas a constraint of the model is that it starts at tempo 1. An additional problem is that in some cases the first tempo value is not always the maximal value. This implies that shifting the data to make either the first value or the maximal value equal to 1, will result in a poor fit. To illustrate this, a problematic case is presented in the left plot of figure 2, where the maximal tempo value is not equal to 1. The fitted model is clearly a very poor approximation of the data. To alleviate these problems, an additional offset parameter is included added to the model while fitting. The right plot of figure 2 shows the same data with the fitted model using the offset parameter. Within its capabilities, the model now fits the data relatively well. Note that the offset parameter is only used for calibration purposes and is not regarded as a meaningful part of the model.

The model is fitted to the data by non-linear least-squares fitting through the Marquardt-Levenberg algorithm, using the *gnuplot* implementation [1]. The model fitting is applied to each performance individually, so for each combination of pianist and piece a value is obtained for w, q, and the root mean square of the error after fitting (this serves as a goodness-of-fit measure).

4 RESULTS AND DISCUSSION

The values obtained from fitting are displayed as a scatterplot on the two-dimensional parameter space q versus w, in

[1] The fitting must be done by numerical approximation since the model is non-linear in the parameters w and q.

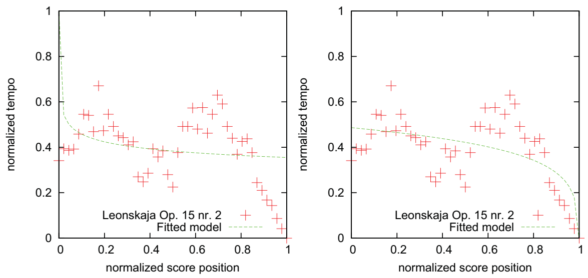

Figure 2. The necessity of an offset parameter for fitting the model (dashed line) to the data ('+' symbols). left: fitting without offset compensation; right: fitting with offset compensation (see text)

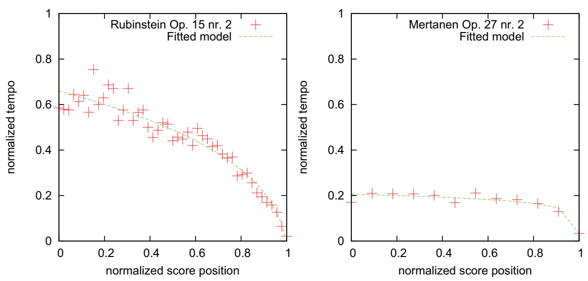

Figure 3. Two extremes of the effective parameter space; left: Rubinstein's ritard in Op. 15 nr. 2 (low w, low q); right: Mertanen's ritard in Op. 27 nr. 2 (high w, high q)

figures 4 and 5. [2] To facilitate the interpretation of points in specific locations of the plane, the reader is referred to figure 1, in which the relative location of the plots corresponds to the topology of the w-q plane.

In figure 4, the symbols group the data points by piece, such that the performances of the same piece by different pianists have identical symbols. The sizes of the symbols are proportional to the goodness-of-fit. That is, bigger symbols are represent a better fit of the model to the data, and are therefore to be considered more reliable than smaller symbols.

The scatter plot reveals a strong positive correlation between the w and the q parameter. In musical terms, this implies that the tempo decrease in deep ritards (low w) is more gradual (low q), whereas in shallow ritards (high w), the tempo decrease is more sudden, and postponed to the last notes of the ritard (high q). These two situations are illustrated by the ritards shown in figure 3.

Notable is also that virtually all performances correspond to q values above 3. This value (marked in the figure as a black horizontal line), corresponds to the model setting that mimics the motion of a physical body under constant braking power. This setting, together with $q = 2$ (constant braking force, assumed in [7] and [11]), is claimed by Friberg and Sundberg [4] to yield ritards that are aesthetically preferred by listeners. They suggest that this preference is due the fact that we are familiar with these conditions from our perception of physical motion. In contrast, the higher q values that are measured in the current study suggest model settings where braking power increases with time. Interestingly, the scatter plot shows a strong ridge close to $q = 4$, where the range above the boundary is highly populated, whereas the range below it is virtually empty. This means that, independent of the depth of the ritard, curvatures below $q = 4$ (approximately the curvature displayed in the left plot of figure 3) are very uncommon.

The distribution of the symbols indicate that, even if the data points of some pieces overlap, they are clearly clustered according to piece. For example, the performances of Op. 15 nr. 2, are all located in the lower ranges of w and q (deep and gradual ritards), whereas those of Op. 27 nr. 2 are all in the higher ranges (shallow and sudden ritards).

Another notable aspect of the results is that the ritards of Op. 48 nr. 1 rit. 2 (except for one, by Leonskaja) are played with various depths (w), but always with low curvature (q).

Figure 5 shows the same data, but labeled according to pianist. In this case clustering is less apparent from the plot. In part this may be due to the amount of different pianists (and thus symbols) displayed in the figure. However, since the data shows a considerable clustering along piece, a strong clustering along pianists is not to be expected. Still, upon more detailed inspection, the w-q plane conveys some differences between pianists.

Firstly, some pianists tend to concentrate in distinct areas of the w-q plane. This is the case for Leonskaja and Vasary. Their performances are displayed jointly in figure 6. Note that the pianists are almost separable based on their w and q coordinates.

As a second example of differences between pianists, consider the performances of Rubinstein and Pollini (figure 7). Rubinstein's w-q coordinates span a much larger part of the plane, suggesting that he plays ritards in a more diverse ways, whereas Pollini's coordinates are concentrated in a smaller area, suggesting a more uniform way of playing ritards. It is interesting to note that the relative locations of the pieces are roughly the same for Rubinstein and Pollini.

5 CONCLUSIONS AND FUTURE WORK

In this study we have used a kinematic rubato model [4] to investigate the performance of final ritards in Chopin's Nocturnes, played by 25 pianists. To our knowledge this is the first application of the model to data gathered from famous pianists. Studying the value range of model parameters that represent the measured ritards, we found that there is a strong positive correlation between the depth of ritards

[2] The figures are best viewed in color

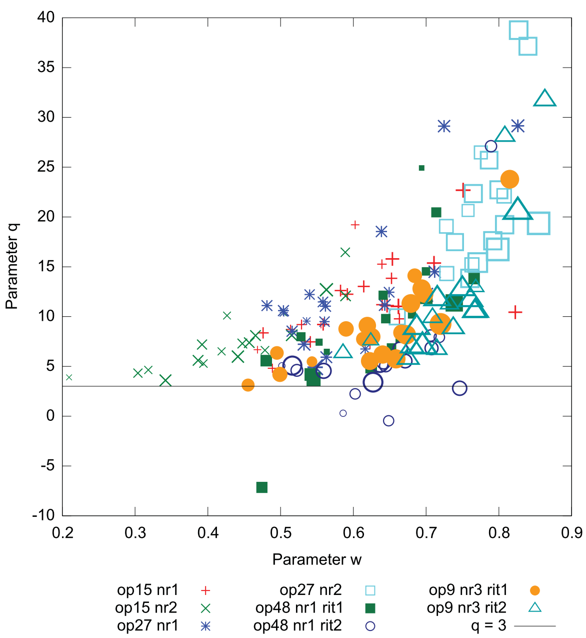

Figure 4. Distribution of rubato model parameter values over pieces; the size of the symbols is proportional to goodness of fit of the parameters to the data

(w) and their curvature (q). In addition, fitted q values are only sporadically below 4. This contrasts with earlier studies stating $q = 2$ and $q = 3$ as plausible settings [4, 7, 11].

Furthermore, ritards of the same piece by different pianists tend to be concentrated in the w-q plane, suggesting that the musical material being played is an important factor in the determination of the depth and curvature of the ritard. Although in general the model parameters are not discriminative for pianists, in some cases the differences in the parameter ranges for individual pianists are considerable. In order to make more decisive claims about pianist-specific differences however, more performances per pianist are needed.

An important issue that we have not addressed in this paper is that in many cases the structure of the ritards are more complex than the model can accommodate. More specifically, the measured tempo data in addition to a simple tempo decrease often shows internal structure that seems to be related to rhythmical patterns or motivic grouping in the music. This affects the goodness-of-fit of the model, and shows the need for a more elaborate modeling approach, either by using more sophisticated models (such as the one proposed by [6]), or by an analysis of the residual information after the model has been fitted.

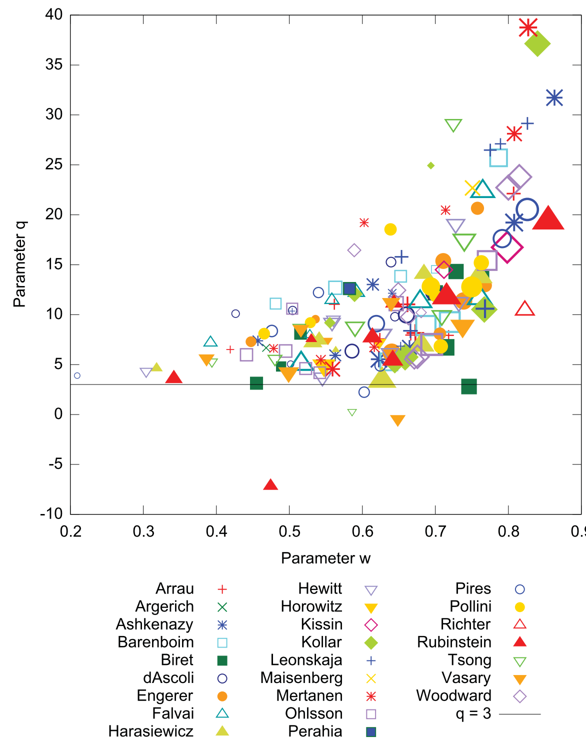

Figure 5. Distribution of rubato model parameter values over pianists; the size of the symbols is proportional to goodness of fit of the parameters to the data

Acknowledgments

We wish to thank Werner Goebl and Bernhard Niedermayer for their help in the acquisition of the timing data from the audio recordings. This work is funded by the Austrian National Research Fund (FWF) under project number P19349-N15.

6 REFERENCES

[1] CANNAM, C., SANDLER, M., AND BELLO, J. The sonic visualiser: A visualisation platform for semantic descriptors from musical signals. In *Proceedings of the 7th International Conference on Music Information Retrieval* (2006), ISMIR.

[2] DESAIN, P., AND HONING, H. Physical motion as a metaphor for timing in music: the final ritard. In *Proceedings of the 1996 International Computer Music Conference* (San Francisco, 1996), ICMA, pp. 458–460.

[3] FELDMAN, J., EPSTEIN, D., AND RICHARDS, W. Force dynamics of tempo change in music. *Music Perception 10*, 2 (1992), 185–204.

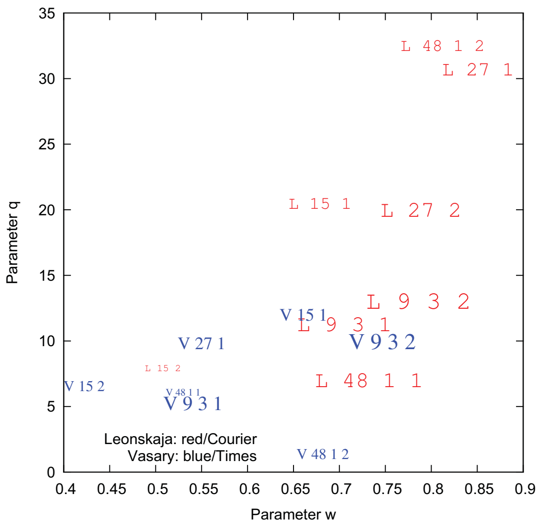

Figure 6. Two pianists occupying different areas of the parameter space; the size of the labels is proportional to goodness of fit of the parameters to the data; “L 48 1 1” = “Leonskaja, Op. 48, nr. 1, rit 1”, etc.

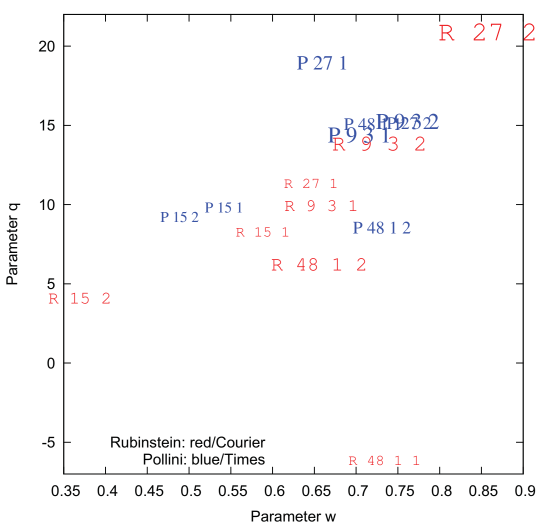

Figure 7. Two pianists spanning overlapping, but differently sized areas of the parameter space; the size of the symbols is proportional to goodness of fit of the parameters to the data

[4] FRIBERG, A., AND SUNDBERG, J. Does music performance allude to locomotion? a model of final ritardandi derived from measurements of stopping runners. *Journal of the Acoustical Society of America 105*, 3 (1999), 1469–1484.

[5] HONING, H. When a good fit is not good enough: a case study on the final ritard. In *Proceedings of the ICMPC* (2004), pp. 510–513.

[6] HONING, H. Is there a perception-based alternative to kinematic models of tempo rubato? *Music Perception 23*, 1 (2005), 79–85.

[7] KRONMAN, U., AND SUNDBERG, J. Is the musical ritard an allusion to physical motion? In *Action and Perception in Rhythm and Music*, A. Gabrielsson, Ed., no. 55. Royal Swedisch Academy of Music, 1987, pp. 57–68.

[8] NIEDERMAYER, B. Non-negative matrix division for the automatic transcription of polyphonic music. In *Proceedings of the 9th International Conference on Music Information Retrieval* (2008), ISMIR.

[9] REPP, B. H. Diversity and commonality in music performance - An analysis of timing microstructure in Schumann’s “Träumerei”. *Journal of the Acoustical Society of America 92*, 5 (1992), 2546–2568.

[10] SUNDBERG, J., AND VERRILLO, V. On the anatomy of the retard: A study of timing in music. *Journal of the Acoustical Society of America 68*, 3 (1980), 772–779.

[11] TODD, N. The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America 91* (1992), 3540–3550.

[12] WINDSOR, W. L., AND CLARKE, E. F. Expressive timing and dynamics in real and artificial performances: using an algorithm as an analytical tool. *Music Perception 15*, 2 (1997), 127–152.