Proceedings of the 8th International Conference on Music Perception & Cognition, Evanston, IL, 2004
ICMPC8
S.D. Lipscomb, R. Ashley, R.O. Gjerdingen, & P. Webster (Eds.)
Adelaide, Australia: Causal Productions

WHEN A GOOD FIT IS NOT GOOD ENOUGH: A CASE STUDY ON THE *FINAL RITARD**

Henkjan Honing

Institute for Logic, Language, and Computation (ILLC)
Department of Musicology
University of Amsterdam
www.hum.uva.nl/mmm/hh/

ABSTRACT

The relation between music and motion has been a topic of much theoretical and empirical research. An important contribution is made by a family of computational theories, so-called *kinematic models,* that make an explicit relation between the laws of physical motion in the real world and expressive timing in music performance (see Friberg & Sundberg, 1999). These models were shown to have a good fit with a variety of empirical data, most notably that of the *final ritard* in music performance: the typical slowing down at the end of a music performance. However, the predictions of these kinematic models are independent of (1) the number of events, (2) the rhythmic structure, and (3) the overall tempo of the performance; These factors have no effect on the predicted shape of the *ritardando*. Computer simulations of a number of rhythm perception models show, however, a large effect of these structural and temporal factors. They are therefore proposed as a perception-based alternative to the kinematic approach. While a final ritard might coarsely resemble a square root function (according to a kinematic model), the predictions made by perception-based models are also influenced by the temporal structure of the musical material that constraints possible shapes of the ritard, and it can therefore be considered a potentially stronger theory than one that simply has a good fit (Roberts & Pashler, 2000).

1. INTRODUCTION

The relation between music and motion has been studied in a large body of theoretical and empirical work (see Shove & Repp, 1995 for a partial overview). However, it is very difficult to specify – let alone validate –, what is the nature of this long-assumed relationship (Honing, 2003). An important contribution to this topic is made by a family of computational theories, so-called *kinematic models,* that make an explicit relation between the laws of physical motion in the real world and expressive timing in music performance (Sundberg & Verillo, 1980; Kronman & Sundberg, 1987; Longuet-Higgins & Lisle, 1989; Feldman, Epstein & Richards 1992; Todd, 1992; Epstein, 1994; Todd, 1995; Friberg & Sundberg, 1999). A large number of these theories focus on modeling the *final ritard*, the typical slowing down at the end of a music performance, especially in music from the Western Baroque and Romantic periods. But this characteristic slowing down can also be observed in, for instance, Javanese gamelan music or some pop and jazz genres. These models were shown to produce a good fit with a variety of empirical performance data (Friberg & Sundberg, 1999; but see also alternatives proposed by Repp, 1992, and Feldman, Epstein, & Richards, 1992, *unrelated* to the laws of physical motion, see Figure 1c). Friberg & Sundberg (1999) found the final ritard alluding to human movement: the pattern of runners' deceleration. Such a deceleration pattern can be defined by a model where tempo v is defined as a function of (normalized) score position x, with q for curvature (varying from linear to convex shapes; see Figure 1a and 1b) and w denoting a non-zero final tempo:

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

The rationale for these models of physical motion is that constant braking force (q=2; see Figure 1b) or constant braking power (q=3; see Figure 1a) are types of movement the listener is quite familiar with, and consequently facilitate for prediction of the actual end, the final stop of the performance.

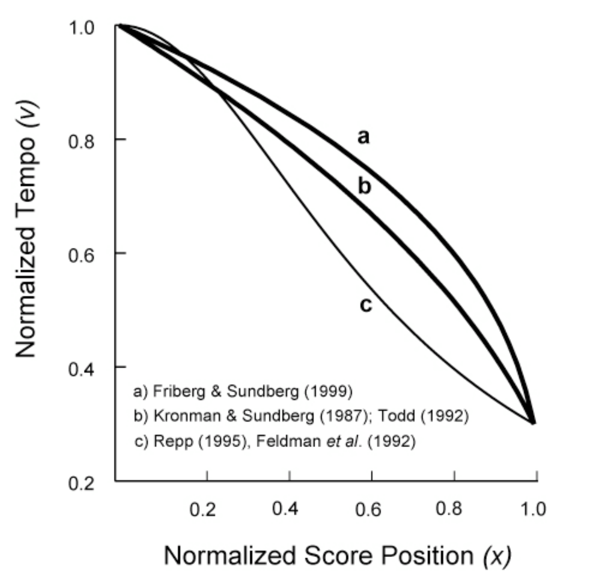

Figure 1: Predictions by a number of models of the *final ritard*. Tempo and score position are normalized.

2. IMPLICIT PREDICTIONS MADE BY KINEMATIC MODELS

However, it has to be noted that these kinematic models predict the shape of the final ritard solely based on the (relative) position in the score (x) and the final tempo (w). While these models intend to describe the common characteristics of final ritards as measured in a music performance, they in fact also state that the shape of a ritard is independent of (1) the number of events (or note density), (2) the rhythmic structure (i.e. differentiated durations), and (3) the overall tempo of the performance; These aspects have no effect on the predicted shape of the ritard. However, for all three aspects it can be argued that they should influence the overall curvature, i.e. the. prediction of the model (an argument in part put forward in Desain & Honing, 1994; Honing, 2003):

1. With regard to the effect of note density one would expect, for example, that a ritard of many notes can have a deep *rubato*, while one of only a few notes will be less deep (i.e. less slowing down), simply because there is less material to communicate the change of tempo to the listener (Gibson, 1975).

2. With regard to the effect of the rhythmical structure (i.e. a musical fragment with differentiated durations) one would expect, for example, a difference in the shape of a ritard for an isochronous as opposed to a rhythmically varied musical fragment. Empirical research in rhythmic categorization has shown that the expressive freedom in timing –the amount of timing that is allowed for the rhythm still to be interpreted as the same rhythmic category (i.e. the notated score)– is affected by the rhythmic structure (Clarke, 1999). Simple rhythmic categories such as 1-1-1 or 1-1-2 allow for more expressive freedom –they can be varied more in timing and tempo before being recognized as a different rhythm by the listener– than relatively more complex rhythms, like 2-1-3 or 1-1-4 (Desain & Honing, 2003).

3. And finally, for the kinematic models the overall tempo of the performance has no effect on the predicted shape of the ritard (tempo is normalized). However, several authors have shown that global tempo influences the use of expressive timing (e.g., Desain & Honing, 1994; Friberg & Sundström, 2002) – at different tempi different structural levels become salient and this will have an effect on the expressive freedom and variability observed (Clarke, 1999).

As a way to show the importance of these three factors (not considered relevant by kinematic models), the predictions of a number of existing models of rhythm perception will be investigated below.

3. NOTE DENSITY

Models of perceived periodicity (or *tempo trackers* for short), try to capture how a listener distinguishes between small timing variations and those deviations that account for a change of tempo: when is a note just somewhat longer and when does the tempo change? These models can be used to show the effect of note density on the ability to track the tempo intended by the performer.

For the simulations summarized below a tempo track model was used based on the notion of coupled oscillators (Large & Kolen, 1994). This model is elaborated in several variants (see Toivainen, 1998) and validated on a variety of empirical data (see, e.g., Large & Jones, 1999). Such a model can make precise predictions on how well a certain *ritardando* can be tracked as a function of note density and depth of the ritard (see Figure 2).

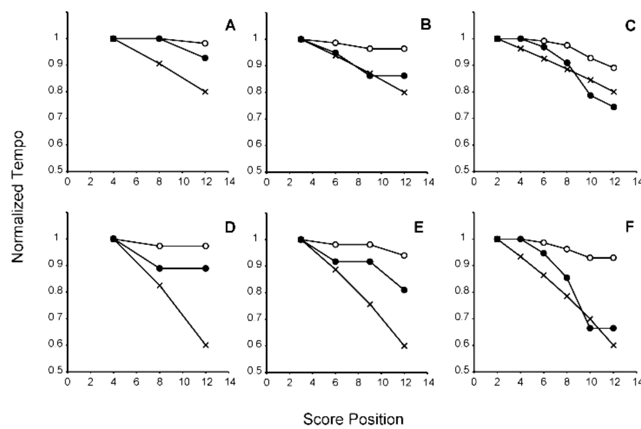

Figure 2: Influence of note density and curvature on a model of tempo tracking. Crosses indicate the input data, filled circles the predicted period by the Large & Kolen model with optimal parameters (in the top row η_ϕ, η_p and γ are 0, 1, and .5; in the bottom row 1, 1, and .5), and open circles show the results for the same model with default parameters (η_ϕ, η_p and γ are 1, 0.4, and 3, respectively, from Large & Palmer, 2002).

Figure 2 shows that the more notes (shown here for 3, 4 and 6 IOIs), the better the tempo tracker (filled circles) will be able follow the tempo. Furthermore, a performance with a deeper rubato (Figure 2, bottom row) is more difficult to track than one that is less deep (Figure 2, top row), as expected. As such this model predicts which changes in timing are relevant to the perception of tempo: it shows the constraints of the perception of *tempo rubato*.

4. RHYTHMIC STRUCTURE

Tempo trackers, however, are relatively insensitive to the microstructure of expressive timing. They focus on the (expected) beat and generally ignore the musical material within the expected beats. Models of rhythmic categorization (or *quantizers* for short) might therefore be more appropriate to study the possible influence the rhythmic structure might have on the shape of the ritard. These models can make precise predictions on the amount of expressive freedom that is allowed before a certain rhythm is perceived as a different rhythm (or category). Three well-known quantizers were used in the simulations summarized below. These are a symbolic (Longuet-Higgins, 1987), a connectionist (Desain & Honing, 1989) and a traditional quantizer (Dannenberg & Mont-Reynaud, 1987). They were applied to isochronous (see Figure 3a and 3c) and rhythmically varied (see Figure 3b and 3d) artificial performances with ritards of different depth.

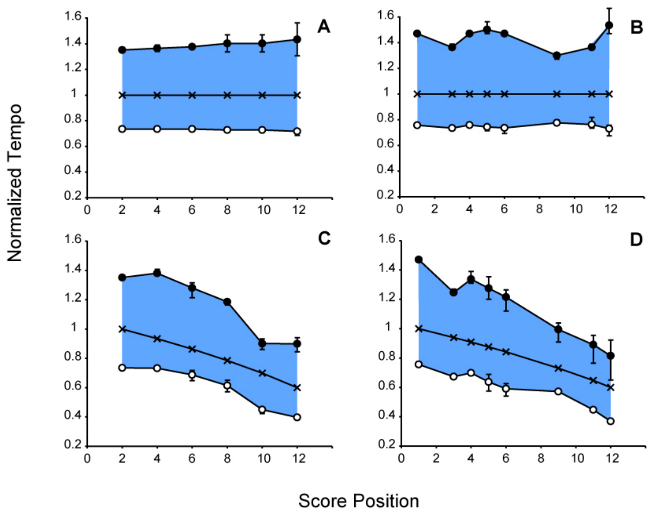

Figure 3: The influence of rhythmic structure on the amount of expressive freedom allowed, as predicted by three quantizers. Crosses indicate the input data, filled circles the average upper boundary, and open circles the average lower boundary. Error bars indicate the minimum and maximum boundary predicted by the three models.

Figure 3 shows the average upper and lower category boundaries as predicted by the three quantizers. They indicate the amount of expressive timing (tempo change or variance) a performed note can exhibit before being categorized as a different duration category (according to the three models). When a certain input duration *IOI* is categorized as C (e.g., quarter note), the upper border (filled circles) indicates the tempo boundary at which an input duration will be categorized as $C \times 3/4$ (dotted eighth note), the lower border (open circles) the boundary at which it would be categorized as $C \times 4/3$ (double dotted quarter note). The error bars indicate the maximum and minimum values for the three models. (Note that this just one possible category boundary.)

Figure 3a and 3b show a metronomic performance of an isochronous and a rhythmically varied fragment. In Figure 3c and 3d the same rhythmic fragments are shown but with a final ritard applied. In Figure 3a, for instance, it can be seen that the more context the more expressive freedom allowed (i.e. a wider area at end than at the beginning). The overall results show that the rhythmic pattern constraints the expressive freedom allowed according to these rhythmic categorization models. And one could argue that a performer, while applying *rubato*, would, in general, not cross these borders as it would mean that the intended rhythm would not be perceived by the listener.

5. GLOBAL TEMPO

A final aspect that was argued to have an influence on the rubato is the global tempo. To show this effect the same models and input data were used as in Figure 3d, but here the overall tempo of the input data is scaled with a factor 1.25 (Figure 4a) and .8 (Figure 4b). Where in the kinematic approach tempo is normalized (see Figure 1), in Figure 3 it can been seen that, according to these models of rhythm perception, the overall tempo also constraints the expressive freedom: the contours are different for different tempi.

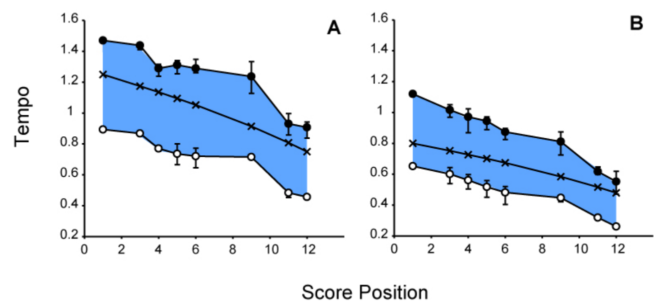

Figure 4: The effect of overall tempo. A slightly faster (*left*) and slower (*right*) version of the input data shown in Figure 3d. Crosses indicate the input data, filled circles the average upper boundary, and open circles the average lower boundary. Error bars indicate the minimum and maximum boundary predicted by the three quantizers.

6. CONCLUSION

The simulations presented in this paper show that computational models of rhythm perception could potentially be a perception-based alternative to the kinematic approach to modeling expressive timing. The former approach has the added characteristic that it is sensitive to note density, rhythmical structure and global tempo, and yields constraints on the shape of a ritardando (restrictions not made by kinematic models). While a final ritard might coarsely resemble a square root function (according to a kinematic model), the predictions made by perception-based models are also influenced by the perceived temporal structure of the musical material that constraints possible shapes of the ritard. It might therefore be considered a potentially stronger theory than one that only makes a good fit (cf. Roberts & Pashler, 2000). However, the theoretical predictions made by the combination of a quantization and tempo track model still needs a systematic empirical study to see how precisely the structural and temporal factors mentioned constrain a musical performance. Next to these empirical issues, theoretical issues of how best to evaluate this type of cognitive models on empirical data (cf. Pitt & Myung, 2002) will be further explored.

Acknowledgements

Thanks to Ed Large, Aline Honingh, Eleni Lapidaki, Else Buitenhuis and Peter Hannewijk for their comments on an earlier version of this paper.

* This is research in progress (March 2004). A more complete version of this paper (including an elaborate method section) can be found in Honing (*submitted*) and Honing (*in preparation*); see www.hum.uva.nl/mmm under 'Publications'.

7. REFERENCES

1. Clarke, E. F. (1999). Rhythm and Timing in Music. In D. Deutsch (Ed.), *Psychology of Music, 2nd Edition* (pp. 473-500). New York: Academic Press.

2. Dannenberg, R. B. & Mont-Reynaud, B. (1987). Following a jazz improvisation in real time. In *Proceedings of the 1987 International Computer Music Conference*. San Francisco: International Computer Music Association, 241-248.

3. Desain, P., & Honing, H. (1989). The quantization of musical time: a connectionist approach. *Computer Music Journal,* 13(3), 56-66.

4. Desain, P., & Honing, H. (1994). Does Expressive Timing in Music Performance Scale Proportionally with Tempo? *Psychological Research*, 56, 285-292.

5. Desain, P., & Honing, H. (2003) The formation of rhythmic categories and metric priming. *Perception*. 32(3), 341-365.

6. Epstein, D. (1994) *Shaping time*. New York: Schirmer.

7. Feldman, J., Epstein, D., & Richards, W. (1992) Force Dynamics of Tempo Change in Music. *Music Perception*, 10(2), 185-204.

8. Friberg, A., & Sundberg, J. (1999) 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), 1469-1484.

9. Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19(3), 333-349.

10. Gibson, J. J. (1975) Events are perceivable but time is not. In *The Study of Time*, 2, edited by J.T. Fraser & N. Lawrence. Berlin: Springer Verlag

11. Honing, H. (2003) The final ritard: on music, motion, and kinematic models. *Computer Music Journal*, 27(3), 66-72.

12. Kronman, U., & Sundberg, J. (1987) Is the musical ritard an allusion to physical motion? In A. Gabrielsson (ed.) *Action and Perception in Rhythm and Music*. Royal Swedisch Academy of Music. No. 55, 57-68.

13. Large, E. W., & Jones, M. R. (1999). The dynamics of attending: how we track time varying events. *Psychological Review*, 106 (1), 119–159.

14. Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*, 177–208.

15. Large, E. W., & Palmer, C. (2002). Temporal responses to music performance: Perceiving structure in temporal fluctuation. *Cognitive Science, 26, 1-37.*

16. Longuet-Higgins, H. C. (1987). *Mental Processes*. Cambridge, Mass.:MIT Press.

17. Longuet-Higgins, H. C., & Lisle, E. R. (1989) Modelling music cognition. *Contemporary Music Review*. 3, 15-27.

18. Pitt, M. A., & Myung, I. J. (2002) When a good fit can be bad. *Trends in Cognitive Science*, 6(10), 421- 425.

19. Repp, B. H. (1992) Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's Träumerei. *Journal of the Acoustical Society of America*. 92, 2546-2568.

20. Roberts, S., & Pashler, H. (2000) How persuasive is a good fit? A comment on theory testing. *Psychological Review*, 107 (2), 358–367.

21. Shove, P., & Repp, B. H. (1995). Musical motion and performance: Theoretical and empirical perspectives. In J. Rink (Ed.), *The practice of performance*. Cambridge, U.K.: Cambridge University Press, 55-83.

22. Sundberg, J., & Verillo, V. (1980) On the anatomy of the ritard: A study of timing in music. *Journal of the Acoustical Society of America*. 68, 772-779.

23. Todd, N. P. M. (1992) The dynamics of dynamics: a model of musical expression. *Journal of the Acoustical Society of America*, 91(6), 3540-3550.

24. Todd, N. P. M. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*, 97(3), 1940-1949.

25. Toiviainen, P. (1998). An interactive MIDI accompanist. *Computer Music Journal,* 22(4), 63-75.