Dept. for Speech, Music and Hearing
Quarterly Progress and
Status Report

Matching the rule parameters of PHRASE ARCH to performances of "Träumerei": a preliminary study

Friberg, A.

KTH Computer Science and Communication

journal: STL-QPSR
volume: 36
number: 2-3
year: 1995
pages: 063-070

http://www.speech.kth.se/qpsr

Matching the rule parameters of PHRASE ARCH to performances of "Träumerei": a preliminary study*

Anders Friberg

Abstract
In music performance a basic principle is the marking of phrases, which often seems to be achieved by means of tone durations. In our grammar for music performance the rule PHRASE ARCH, has been formulated to model this effect. A technical description of the rule is presented. Also presented is an attempt to match the different parameters of this rule to the duration data from 28 performances of the first 9 bars of Robert Schumann's Träumerei as measured by Bruno Repp (1992). The optimisation was based on relative duration measured in percent. On average 44% of the total variation was accounted for by PHRASE ARCH. The discrepancies were mostly at the note level and were mostly associated with small musical gestures.

Introduction

The KTH performance rules translate scores to music performances. These rules have been described in several previous papers (e.g. Friberg, 1995; Sundberg, 1993). We will focus here on the rule PHRASE ARCH. The aim is two-fold: to provide a complete technical description and to see if the rule can be traced in real performances.

PHRASE ARCH is a rule that is particularly suited to romantic classical music, performing the phrases with an initial accelerando and a subsequent ritardando. The definition of the rule was inspired by Neil Todd's phrase model (Todd, 1985, 1989), see also Friberg et al. (1994) and Friberg (1995).

The existing measurements by Repp (1992) on 28 performances of Robert Shumann's "Träumerei" were well suited for testing the validity of the rule. It is clearly a romantic piece with often very pronounced tone duration changes. Many artists and recordings spanning several decades ensured that much of individual plausible variation was exposed in this data. An interesting question was how much of the variation seen in these performances could be explained by varying the parameter values in the rule.

A complete technical description of the rule will be given below which includes both timing and sound-level variations. The following parameter matching is, however, only done with regard to the timing information. In this preliminary study, the matching was limited to the first nine measures of "Träumerei".

* Also appearing in Friberg A & Sundberg J, eds, *Proceedings of the KTH Symposium on Grammars for Music Performance*, May 27, 1995.

PHRASE ARCH

The input is the score complemented with an analysis of the phrase structure in a hierarchical sense. Fig. 2 offers an example of the analysis of "Träumerei." The focus of the current definition of the rule is on tone duration deviations. The sound level is also affected but simply defined as inversely proportional to the duration deviation.

Table 1 lists all available parameters at each phrase level (PhLevel). All phrases at the specified phrase level will be changed according to the parameters. To get a complete phrasing, the rule is supposed to be applied simultaneously at several levels. The effect is additive (as most of the other rules), that is, the current duration is increased with the relative deviation value given by the rule. PHRASE ARCH differs somewhat from previous rules in that many more parameters are available. It may be considered more as a tool controlling the phrasing than a final rule.

Table 1. List of the parameters that are available in the rule PHRASE ARCH.

K	This is the main parameter. It controls the amount of ritardando at the end of each phrase.
Acc	It controls the amount of accelerando, expressed in terms of a factor multiplied by K/2. The default value is 1.
Turn	The position of the turning point between the accelerando and the ritardando. A decimal value between 0 and 1 will be treated as a ratio between the turning point and the phrase length, measured in nominal time (score position). Alternatively, an integer value n specifies the nth note from the beginning of the phrase.
Next	It is used to modify the amount of ritardando for phrases that also terminate a phrase at the next higher level (having a lower PhLevel number). It is expressed as a ratio of K. The default value is 1, which means no change.
2Next	Same as Next for phrases also terminating two levels higher.
Last	It changes the duration of the last note of each phrase. It is a factor multiplied with the final value obtained from all the other parameters. The default value is 1, which means no change.
Power	It determines the shape of the accelerando and ritardando functions. For example, the default value 2 gives a quadratic function. Any positive integer or decimal value is allowed.
Amp	It sets the sound level as a factor multiplied with the default value.

To account for two phrase levels, two strategies are available: either applying the rule separately at both levels, or only apply it at the lower level and use the parameter Next. The difference is that in the former case the higher level will be performed with an arch over the whole phrase, whereas in the latter case, the higher level will be performed with, for example, an increased ritardando in the end (as illustrated in Fig. 1 below).

Using the parameter Last with values less than one is important for the impression of continuation to the next phrase. It can also be used to compensate for the case when the phrase is already composed with a long final note.

The phrase is divided in two parts: the accelerando and the ritardando. The length of each part is determined by the Turn parameter. Let x_i be the normalised score position for note i in one part ($0 \le x \le 1$, $0 \le i \le N$).

For the accelerando, let

$$\Delta DR_i = 0.1K * Acc * (1 - x_i)^{Power} \quad (1)$$

and for or the ritardando, let

$$\Delta DR_i = 0.2K * x_i^{Power} \quad (2)$$

If the conditions for Next or 2Next are true, the ΔDR_i values for the ritardando will be multiplied with these parameters.

The last note is finally modified

$$\Delta DR_N <= \Delta DR_N * Last \quad (3)$$

The new duration of each note is given by

$$DR_i <= DR_i * (1 + \Delta DR_i) \quad [ms] \quad (4)$$

The sound level is computed similarly.
For the accelerando, let

$$\Delta L_i = 0.5K * Acc * Amp * (1 - x_i)^{Power} \quad (5)$$

and for the ritardando, let

$$\Delta L_i = 0.1K * Amp * x_i^{Power} \quad (6)$$

If the conditions for Next or 2Next are true, the ΔL_i values for the ritardando will be multiplied with these parameters. The new sound level of each note is given by

$$L_i <= L_i + \Delta L_i \quad [dB] \quad (7)$$

The sound level of the last note is not modified.

An example of the duration deviation produced by the rule is given in Fig. 1. There are two phrases where the second also terminates a phrase at the next higher level as indicated in the phrase structure in the figure.

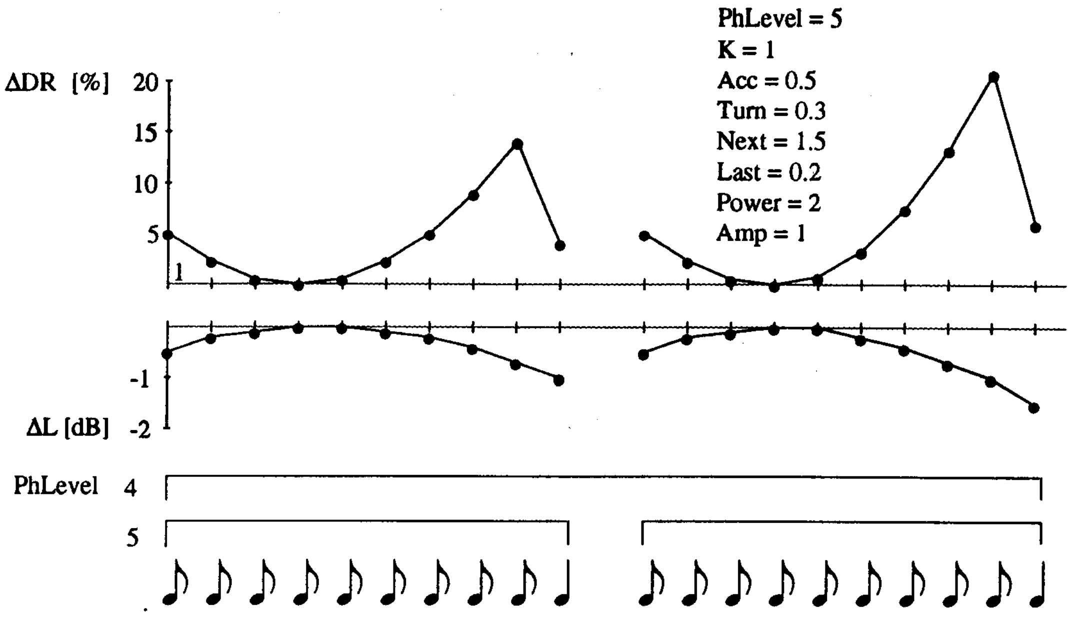

Figure 1. Example of the result of the rule PHRASE ARCH, *applied with the parameters listed in the figure.*

Parameter matching

Phrase analysis

The input of PHRASE ARCH is the score together with the phrase analysis. This analysis can be considered part of the interpretation since it can differ among artists. Consequently, the analysis used in the experiment was derived from performance notes done independently by our musical expert Lars Frydén. This analysis is shown in Fig. 2. The phrase structure of the piece is rather complex with partly overlapping phrases occurring in different voices. This interplay between the phrases was avoided by reducing the analysis to an overall one. Observe that two phrases are starting and ending on the same note in measure 5 and 9. In such an overlap region, the values of the second phrase override those of the first.

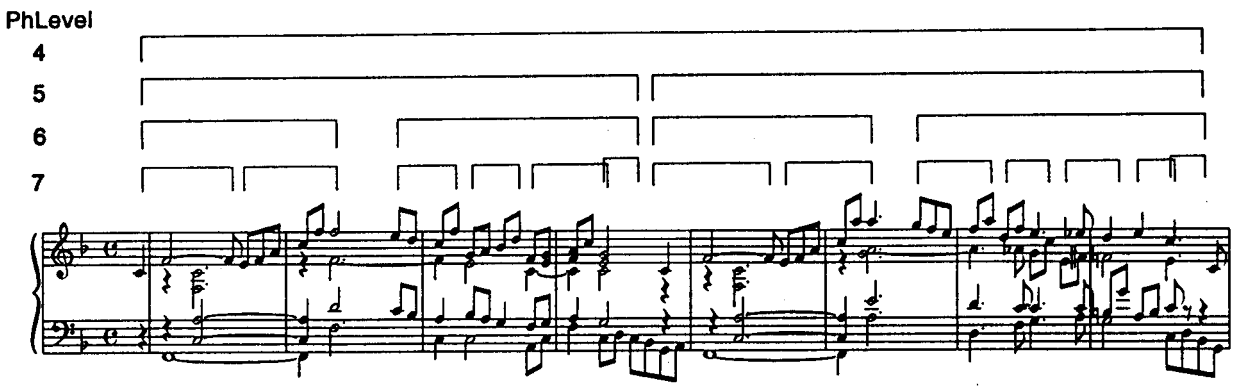

Figure 2. The phrase analysis of "Träumerei," measures 1-9.

Model

Initially the matching was performed at each phrase level individually. The results showed that all phrase levels (4 to 7) were relevant with a significant amount at least by some artists. Consequently, the final model was a combination of all four phrase levels, using 18 parameters totally, see Table 2. The parameters Next and 2Next were included at the lowest level to account for the relatively large ritardandi in the second and eighth phrase that the artists often performed. In this preliminary study, the Power parameter was set at a fixed value of two. This was a reasonable initial value since quadratic functions was found by Repp (1992) to fit very well with the durations of the second and eighth phrase at level seven.

It was necessary to redefine the rule slightly for the optimisation in order to avoid the influence of K in the other parameters. Also, all parameter values were expressed in terms of percentage deviation. Formulas (1) to (3) were modified as follows:

$$\Delta DR_i = Acc * (1 - x_i)^{Power} \quad (1b)$$

$$\Delta DR_i = K * x_i^{Power} \quad (2b)$$

$$\Delta DR_N = Last \quad (3b)$$

If Next and 2Next were used, formula (2b) was replaced with

$$\Delta DR_i = Next * x_i^{Power} \quad \text{or} \quad \Delta DR_i = 2Next * x_i^{Power} \quad (8)$$

After the application of the rules, all durations were adjusted proportionally so that the total duration was the same both for the rule generated durations and the measured durations.

Table 2. The four PHRASE ARCH rules used and the corresponding parameters that were varied (marked with X).

Rule #	PhLevel	K	Acc	Turn	Next	2Next	Last	Power	Amp
1	4	X	X	X	-	-	X	2	-
2	5	X	X	X	-	-	X	2	-
3	6	X	X	X	-	-	X	2	-
4	7	X	X	X	X	X	X	2	-

Optimisation method

The distance between the rule generated and measured performance was estimated by the average relative difference of all durations:

$$\text{Distance} = \frac{\sum_{i=1}^{N} \left| \frac{DR_{measured} - DR_{rule}}{DR_{measured}} \right|}{N},$$

where N is the total number of notes, disregarding rests. The choice of relative deviation as the basis was motivated by experiments showing that the just noticeable difference in duration of one note in an isochronous sequence is constant in relative terms (Friberg & Sundberg, in press).

The distance function was minimised using a simple algorithm that moved one step up or down in each parameter if a smaller distance was obtained. This procedure was repeated until all parameters were kept at the same value for two consecutive cycles. The step length was constant for each parameter. A fixed set of initial values was used for all performances. The optimisation procedure did not necessarily find the lowest minimum. It found, however, minima that were reasonably close to the lowest minima; only small improvements were obtained by trying other initial values.

Results and discussion

Table 3 shows the resulting parameter values averaged over all 28 performances. The values for the parameter Last are difficult to interpret since they are dependent on each other at the different levels and hence will not be further discussed. If we assume that the just noticeable deviation is about 5% (Friberg, 1995, p. 26), we see that the rule produced clearly audible effects in most cases. The exception is the accelerando terms that are comparatively small in all cases except at phrase level five.

Table 3. The average, standard deviation (SD), min and max values for the parameters obtained after the matching to the 28 performances. All values in percentage deviation of tone duration except Turn that is percentage of phrase length.

PhLevel		K	Acc	Turn	Next	2Next	Last
4	mean	**28**	**8**	**67**	-	-	43
	SD	27	12	22	-	-	43
	min	0	-10	8	-	-	-10
	max	112	40	96	-	-	182
5	mean	**29**	**38**	**53**	-	-	-10
	SD	20	22	13	-	-	14
	min	0	4	30	-	-	-40
	max	66	100	84	-	-	20
6	mean	**26**	**3**	**55**	-	-	31
	SD	20	8	26	-	-	18
	min	0	-10	4	-	-	8
	max	70	20	100	-	-	78
7	mean	**40**	**4**	**15**	**68**	**-30**	5
	SD	24	6	11	36	20	8
	min	12	-10	0	14	-80	-10
	max	86	20	42	168	34	24

The difference between the measured performance distance to deadpan and the final distance after optimisation can be considered as a measure of how much of the variation that was explained by the rule. On average a little less than half of the variation could be accounted for by the rule (average = 44%, standard deviation = 7%, max = 59%, min = 27%). This average may be considered a rather high value since one would expect a performer to use considerably more than only one strategy in a performance.

Three of the best fits are shown in Fig. 3. We can see that the long term variation is quite well modeled and that the discrepancies are mostly at the lowest phrase level (PhLevel=7). The differences in the end of measure five indicate that our phrase analysis was different from the phrase analysis used by these performers. They seemed to extend the last phrase in measure five to the pickup in the melody. The often used cliché to lengthen the note before an important note, also usually appearing together

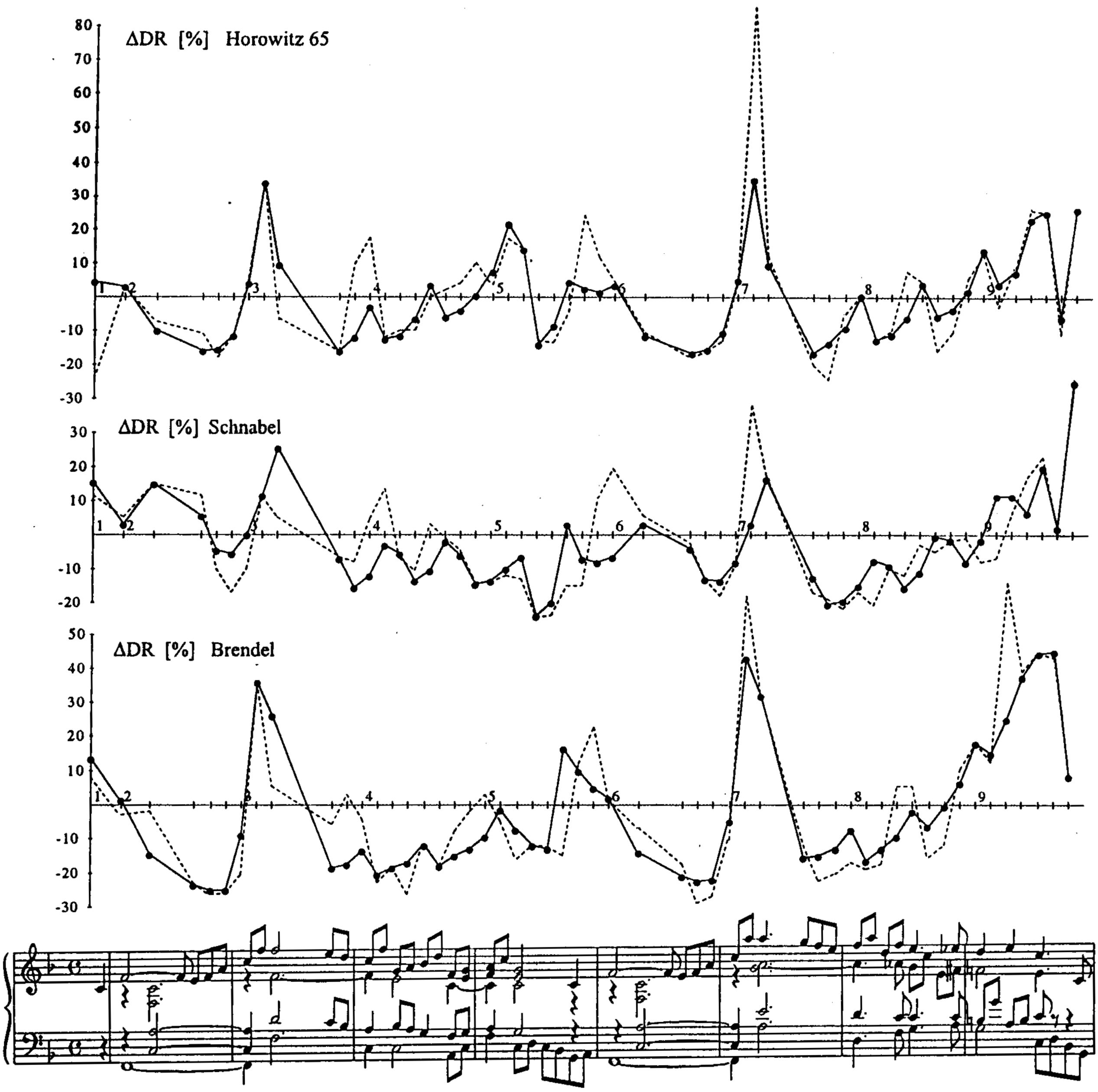

Figure 3. Three examples of the resulting parameter optimisations. The solid lines are the rule generated performances and the dotted lines measured performances by Horowitz, Schnabel and Brendel, respectively.

with a new chord, is also seen in these examples (the second note in measure three and seven). As this case is not included in the present model there are some problems at these points. The extreme lengthening of one note in measure seven by Horowitz may not be predicted from the structure and consequently not possible to generate by a rule.

There were 53 durations and 18 parameters totally. With so many parameters compared to the number of notes, one may assume that almost any model would perform quite well. However, that would imply that the resulting parameter values would have been more arbitrarily than they were. Most of the resulting values were well within the expected range. For example, there were very few negative values for the accelerando or the ritardando parameter. Still it would be desirable to reduce the number of parameters in the future. One way this could be done is to use more of the parameters Next and 2Next instead of applying the rule at all phrase levels. It would also be possible to use a longer music example. One complication in "Träumerei" that we tried to avoid, is that there are a notated ritardando and a fermata in the later part, which affects the phrasing.

Acknowledgements

I would like to thank Bruno Repp for kindly supplying the data. Without his data this investigation would not have been possible. Lars Frydén and Johan Sundberg have been listening to the rule and adjusting the rule parameters; their comments and suggestions are gratefully acknowledged. This work was partially supported by a grant from The Swedish Council for Research in the Humanities and Social Sciences (HSFR).

References

Friberg A & Sundberg J. Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America.* (In press).

Friberg A (1995). *A Quantitative Rule System for Musical Expression*, doctoral dissertation, Royal Institute of Technology, Sweden.

Friberg A, Sundberg J & Frydén L (1994). Recent musical performance research at KTH. In: Sundberg J, ed, *Proceedings of the Aarhus Symposium on Generative Grammars for Music Prformance 1994,* 7-12.

Repp BK (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Träumerei". *Journal of the Acoustical Society of America.* 92/5: 2546-2568.

Todd N (1985). A model of expressive timing in tonal music. *Music Perception* 3: 33-58.

Todd N (1989). A computational model of rubato. *Contemporary Music Review* 3: 69-88.