



**Predicting the similarity between
expressive performances of music
from measurements of tempo and
dynamics**

By Renee Timmers

Presented by Isaac Schankler

ISE 575 / Spring 2010 / University of Southern California



Introduction

Measurements of performances (with emphasis on piano) have become an important means of understanding musical expression:

- Have shown extensive subtle variations in tempo, timing, articulation, and dynamics (Seashore 1938)
- Consistency of variations in repeated performances (Seashore 1938)
- Controllability of variations (Kendall and Carterette 1990; Palmer 1989, 1996)
- Diversity of interpretations greater among professional musicians than among piano students (Repp 1990, 1992a)
- Diversity tends to be smaller at phrase level than below phrase level (Repp 1990, 1992a)



Introduction (continued)

To what extent do measurements reflect the psychological reality of performers and listeners?

- Professional musician's remarks translated into concrete rules for variations in tempo, timing, dynamics and intonation (Sundberg, Friberg, and Fryden 1989; Friberg et al. 1991)
- ... and quality of synthesized performances with variations was judged to be higher than those without (Thompson et al. 1989)

However:

- Perceptual length of a timing perturbation is dependent on its position within the phrase structure (Repp 1992b, 1998)
- Short, empty durations systematically overestimated by a constant interval (Najakima 1987)



Introduction (continued)

So which quantitative representations of tempo and dynamics come closest to the perceptual representation?

- Timing as duration or tempo variation (Friberg and Sundberg 1999)
- Different time scales: note, beat, bar, or phrase levels (Bengtsson and Gabrielson 1987, 1988)
- Normalization, i.e. variations are expressed relative to the mean (Gabrielson 1987, 1988)
- Absolute values, e.g. ms, bpm (Repp 1992a, 1992b; Langner and Goebel 2003)
- Derivatives (rates of change) of tempo and dynamics (Kronman and Sundberg, 1987)
- Second derivative or "change of change" (Gjerdingen 1988)
- Time scale and representation unit has considerable effect on characteristics of the data (Timmers and Honing 2002)



Aims

1. How well does measured data represent perceptually salient characteristics of performances?
2. What data representation comes closest?
 - time span (local vs. global)
 - unit (absolute vs. normalized)
 - derivative (absolute, first, or second)
 - relative salience of tempo and dynamics
 - compound measure of interaction between tempo and loudness



Method

Two experiments with same aim and musical material (experiment 2 has stricter procedure):

- 20 participants rate similarity between pairs of performances
- Chopin prelude and Mozart sonata
- Tempo and loudness measured at the beat level
- Distance in tempo and loudness calculated using different representations
- Regression analyses of measures attempt to explain similarity ratings

How much each representation accounts for variance in similarity ratings = its ability to capture salient characteristics of performances

Musical Material

Chopin's Prelude Op. 28, No. 17

mm. 1-10 (Ch)

Five performances:

- Argerich
- Harasiewicz
- Kissin
- Pollini
- Rubinstein

17. Allegretto.

17. Allegretto.

17. Allegretto.

17. Allegretto.

Procedure (Experiment 1)

Participants:

- 7 women and 13 men, aged 26-45
- 15 experienced musicians (10+ yrs.)
- 5 nonmusicians (3 years or less)
- 6 pianists and 9 nonpianists

	Ref					
Comparisons	1	2	3	4	5	
7 (very similar)	0	0	0	0	0	
6	0	0	0	0	0	
5	0	0	0	0	0	
4 (neutral)	0	0	0	0	0	
3	0	0	0	0	0	
2	0	0	0	0	0	
1 (very dissimilar)	0	0	0	0	0	OK/save

FIG. 1. User interface for the similarity rating in experiment 1.

Performance pairs grouped into blocks, with 4 or 5 performances compared to a reference performance

Pairs rated from 7 (very similar) to 1 (very dissimilar)

Session ended when all performances had been the reference once, resulting in 20 comparisons for Chopin and 30 for Mozart

Each performance pair was rated twice (once with one performance as the reference, once with the other performance as the reference)



Procedure (Experiment 2)

Participants:

- 11 women and 9 men, aged 21-48
- 15 experienced musicians (10 years or more)
- 5 nonmusicians (3 years or less)
- 9 pianists and 6 nonpianists

Each participant rated all pairs of performances of a fragment once

Order of presentation of fragments and pairs was randomized over subjects (no reference performance)

Order of performances within a pair was counterbalanced over participants



Similarity Predictions

- Beat-tracking algorithm was used and output hand-corrected (Dixon 2001; Dixon and Goebel 2001)
- Location of beat defined to coincide with onset of melody note
- Max amplitude level around each beat selected (Langner and Goebel 2003)
- Level in dB converted to sones, an approximation of perceived loudness (Pampalk et al. 2002)
- Equations given for each representation (analogous equations for loudness implied)
- T = vector consisting of a tempo indication for each interbeat interval ($T = [t_1, t_2, \dots, t_n]$, where t stands for local tempo at interbeat interval n)
- Multiple and stepwise regression analyses used to compare predictive power of different representations



Similarity Ratings

To test if effect of performance pair is systematic over participants and presentations

Experiment 1:

- Repeated measures ANOVA in SPSS10 with pair and order as independent within-subject variables and similarity rating as dependent variable
- Separate ANOVA for each fragment
- For M1 and Ch, main effect of pair was only significant effect
- For M2, all effects were significant, indicating that similarity rating may depend on which performance was the reference

Experiment 2:

- Repeated measures ANOVA in SPSS10 with pair as independent within-subject variable, order as between-subjects variable and similarity rating as dependent variable
- Separate ANOVA for each fragment
- For all fragments, main effect of pair was only significant effect

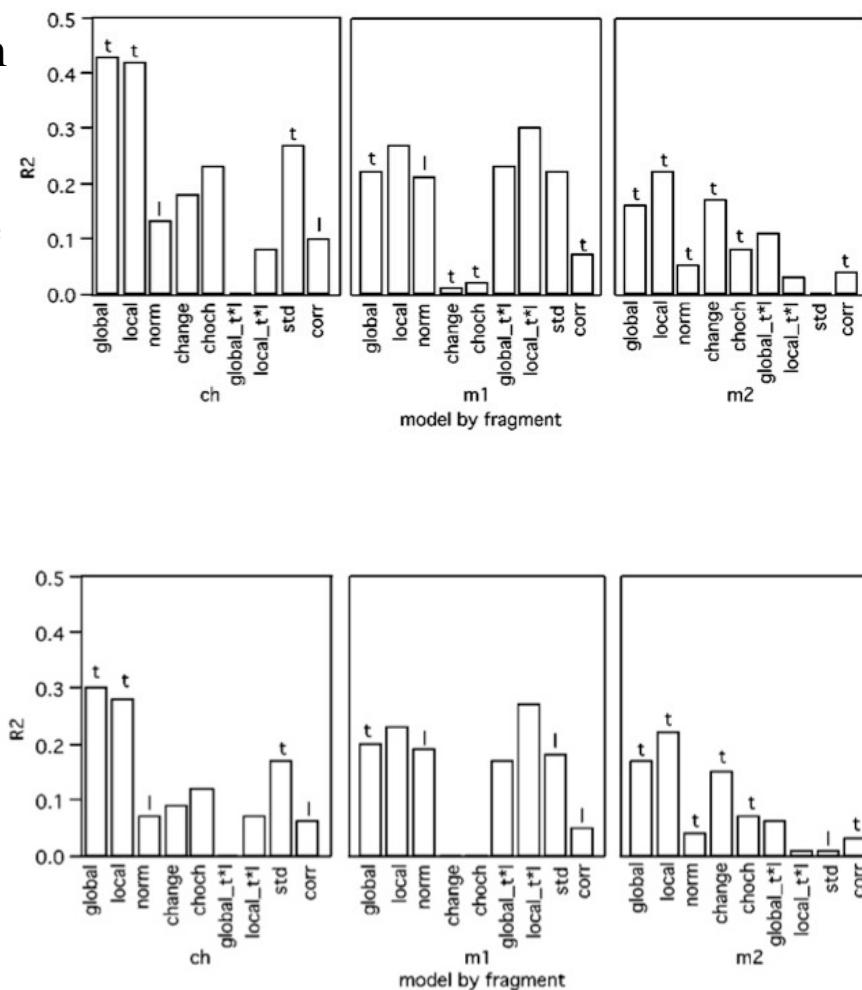
Results: Multiple Regression Analysis

Experiment 1:

- Variance is higher for the Chopin than the Mozart
- Tempo is most important factor
- Local tempo is more important for the Mozart
- Absolute better than normalized
- Derivatives are not as good
- Local tempo and loudness as compound measure works best for M1, but not others
- Correlation does not explain similarity well

Experiment 2:

- Very similar to Experiment 1, but overall variance is lower





Results: Stepwise Regression Analysis

Components are sorted in order of explained variance -- no overlap in explained variance

Experiment 1 (Tables I+II):

- Local or global tempo strongest component of Ch and M2
- Tempo times loudness is strongest component of M1
- Differences in results for musicians and nonmusicians

Experiment 2 (Tables IV+V):

- Similar to Experiment 1, especially for musicians

Results: Stepwise Regression Analysis

TABLE I. Results of the stepwise regression analysis for musicians ($N = 15$). Parameters in order of entrance of the stepwise regression model; the total explained variances for each step, and the F and p value for the full model.

Fragment	Parameters	R^2	F value	p value
Ch	Global t	0.45	$F(1,148) = 120$	<0.0001
M1	Local t^*l	0.27	$F(2,222) = 46.3$	<0.0001
	Local t	0.29		
M2	Local t	0.23	$F(3,221) = 43.4$	<0.0001
	Global t^*l	0.35		
	Change l	0.37		

TABLE IV. Results of the stepwise regression analysis for musicians ($N = 15$) for experiment 2. Parameters in order of entrance of the stepwise regression model; the total explained variances for each step, and the F and p value for the full model.

Fragment	Parameters	R^2	F value	p value
Ch	Global t	0.42	$F(1,148) = 106$	<0.0001
M1	Local t^*l	0.33	$F(2,222) = 58.2$	<0.0001
	Local t	0.34		
M2	Local t	0.25	$F(3,221) = 33.4$	<0.0001
	Global t^*l	0.30		
	Change l	0.32		

TABLE II. Results of the stepwise regression analysis for nonmusicians ($N = 5$). Parameters in order of entrance of the stepwise regression model; the total explained variances for each step, and the F and p value for the full model.

Fragment	Parameters	R^2	F value	p value
Ch	Local t	0.52	$F(2,47) = 30.0$	<0.0001
	Global t^*l	0.56		
M1	Local t^*l	0.40	$F(1,73) = 47.8$	<0.0001
M2	Local t	0.22	$F(2,72) = 17.3$	<0.0001
	Global t^*l	0.32		

TABLE V. Results of the stepwise regression analysis for nonmusicians ($N = 5$) for experiment 2. Parameters in order of entrance of the stepwise regression model; the total explained variances for each step, and the F and p value for the full model.

Fragment	Parameters	R^2	F value	p value
Ch	Local t^*l	0.11	$F(2,47) = 4.8$	<0.02
	Global t	0.17		
M1	Global t^*l	0.20	$F(1,73) = 18.4$	<0.0001
M2	Local t	0.14	$F(3,71) = 11.5$	<0.0001
	Global t^*l	0.26		
	Corr l	0.33		



Conclusions

- Salience of global tempo in line with other studies (Hevner 1937; Levitin and Cook 1996)
- Role of local tempo opens up unexplored area of investigation
- Integrated tempo and loudness also largely unexplored
- Study not conclusive about absolute strength of representations
- What about articulation, timbre, fluency, quality, structure, etc.?