



# Comparative Analysis of Multiple Musical Performances

Craig Stuart Sapp  
ISMIR; Vienna, Austria  
27 September 2007



# Mazurka Project

AHRC Research

Centre for the History and Analysis of Recorded Music

<http://mazurka.org.uk>

- 2732 recordings of 49 mazurkas by Frédéric Chopin (1810-1849)

= Average of 56 performances/mazurka

*least:* 39 performances of 41/3

*most:* 89 performances of 17/4

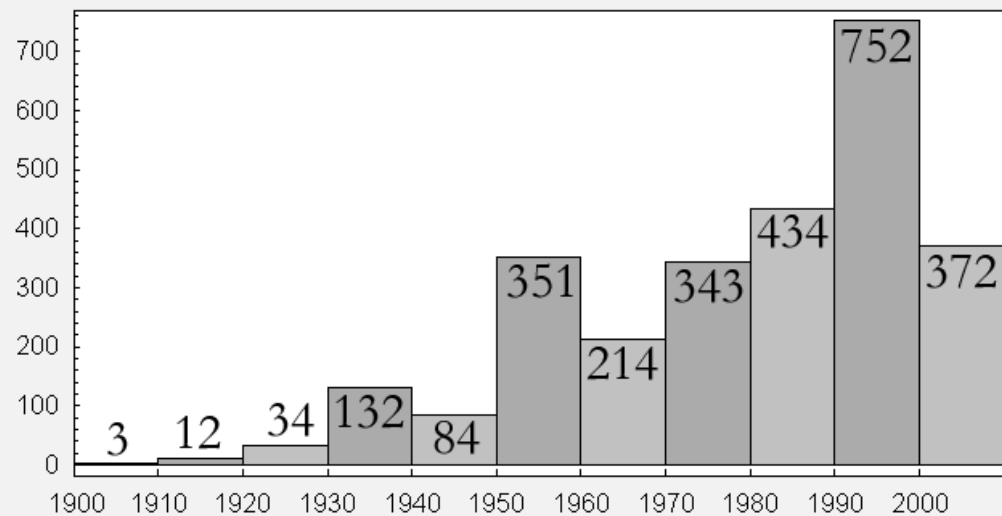
- 157 performers
- on 209 CDs/records
- 123 hours of music
- Earliest performance from 1902 by Alfred Grünfeld: mazurka 67/4



Masako Ezaki, 2006:



## Performance Count by Decade



# Many Performances of Same Composition

## 89 performances of mazurka 17/4

- how to compare and navigate through all the performances?



 (measure 31)

Afanassiev 2001  
Andsnes 1990  
Ashkenazy 1981  
Bacha 1998  
Barbosa 1983  
Beliaevsky 2004  
Ben-Or 1989

Biret 1990  
Blet 2003  
Block 1995  
Brailowsky 1960  
Brunhoff 1963  
Casadesus 1930  
Chiu 1999  
Clidat 1994  
Cohen 1997  
Coop 1987  
Cortot 1951  
Csalog 1996  
Czerny-Stefanska 1949 *live*  
Czerny-Stefanska 1949 *studio*  
Czerny-Stefanska 1989  
Ezaki 2006  
Falvay 1989  
Ferenczy 1958  
Fiorentino 1990  
Flière 1977  
Fou 1978  
François 1956  
Gieseeking 1938  
Ginzburg 1957  
Goldmann 1997  
Guller 1956  
Hatto 1993

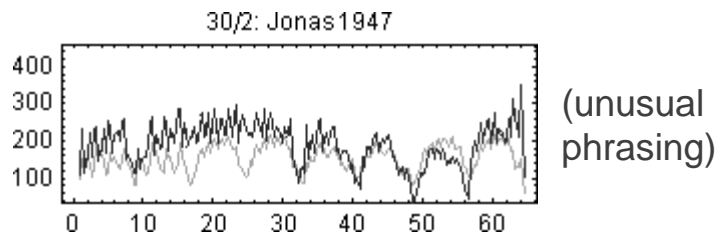
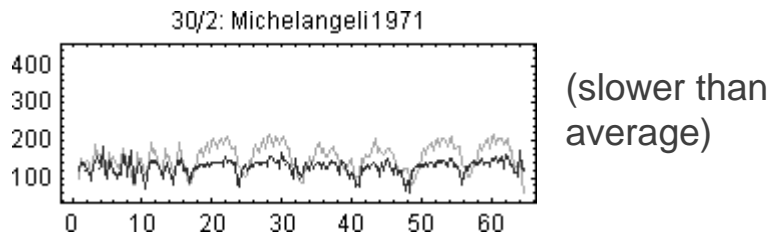
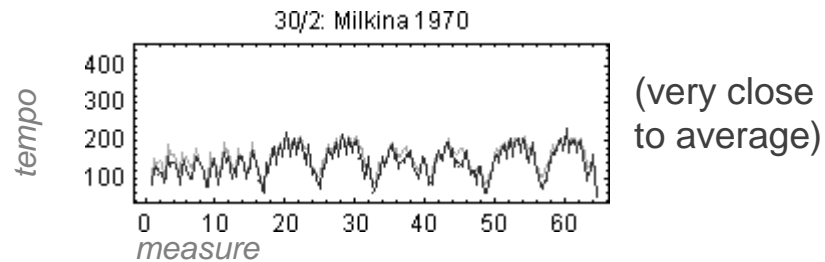
Hatto 2006  
Horowitz 1971  
Horowitz 1985  
Indjic 1988  
Kapell 1951  
Kiepara 1999  
Kilenyi 1937  
Kissin 1993  
Kitain 1937  
Kushner 1990  
Lévy 1951  
Lear 1994  
Lefébure 1950  
Lilamand 2001  
Luisada 1990  
Lushtak 2004  
Lympany 1968  
Lympany 1990  
Magaloff 1977  
Magaloff 1977b  
Magin 1975  
Milkina 1970  
Mohovich 1999  
Nadelmann 1956  
Ohlsson 1999  
Olejniczac 1990  
Olejniczak 1991

Osinska 1989  
Pöntinen 2003  
Paderewski 1912  
Paderewski 1923  
Paderewski 1924 *piano roll*  
Perahia 1994  
Perlemuter 1986  
Poblocka 1999  
Rangell 2001  
Risler 1920  
Rosen 1989  
Rubinstein 1939  
Rubinstein 1952  
Rubinstein 1966  
Rummel 1943  
Shebanova 2002  
Simon 1991  
Smith 1975  
Szpilman 1948  
Sztompka 1959  
Tanyel 1992  
Uninsky 1971  
Vardi 1988  
Wasowski 1980  
Weissenberg 1971  
Zecchi 1942  
Zecchi 1942b

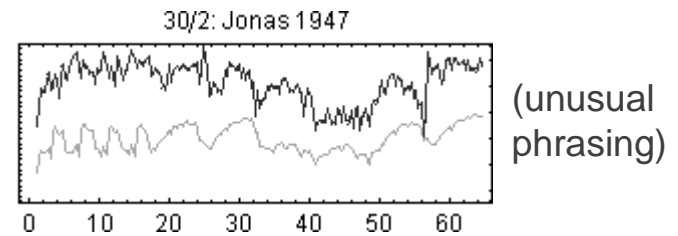
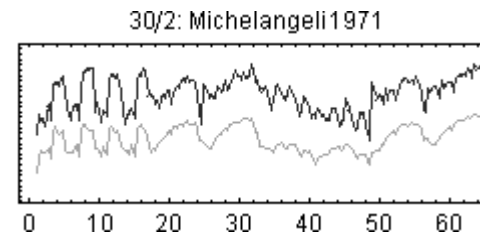
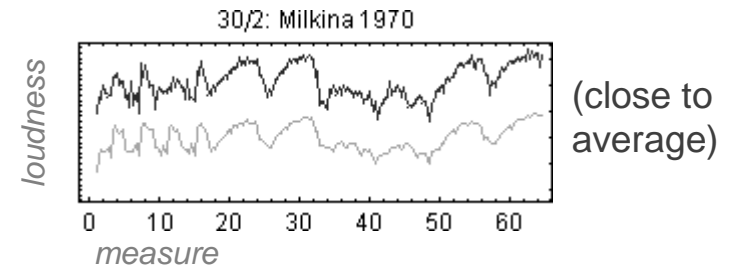
# Tempo/Dynamics Graphs

— Data for particular performance  
— Average of all performances

## Beat tempo:



## Beat dynamics:



Data extracted using Sonic Visualiser / Vamp Plugins developed at C4DM, Queen Mary, U. of London

<http://www.sonicvisualiser.org> & <http://sv.mazurka.org.uk>

# Correlation

*Pearson correlation:*

output range: -1.0 to +1.0

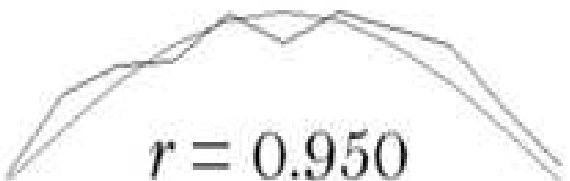
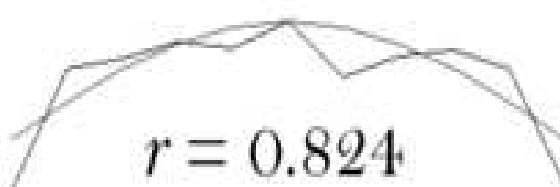
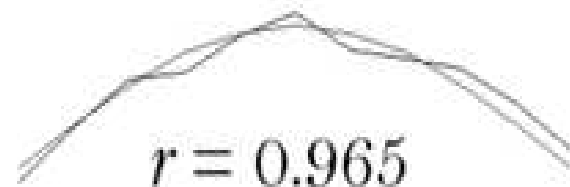
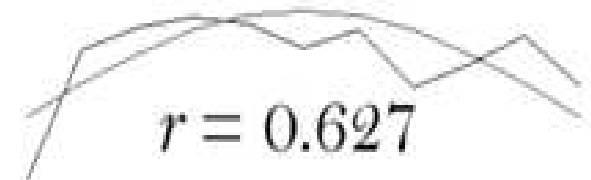
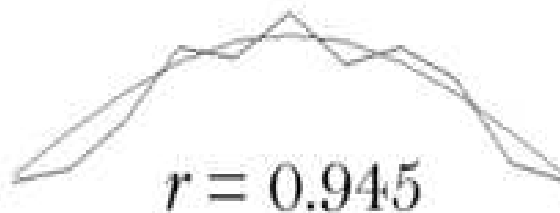
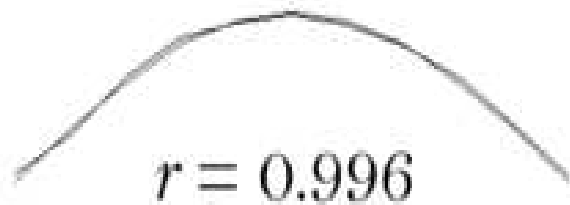
$$\frac{\sum_i (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

• **Measures how well two shapes match:**

$r = +1.0$  is an exact match

$r = 0.0$  means no relation

$r = -1.0$  means upside-down



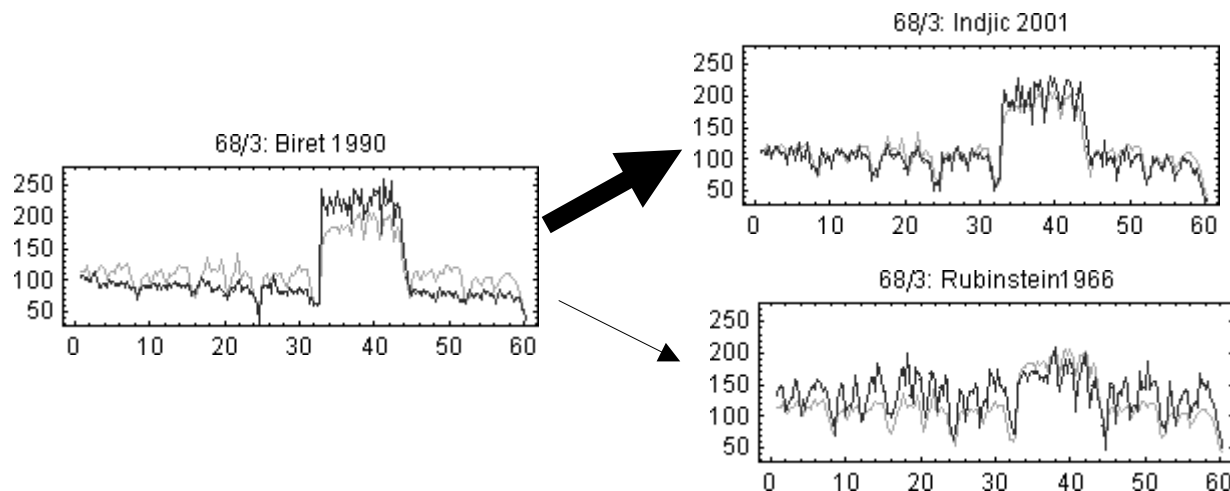
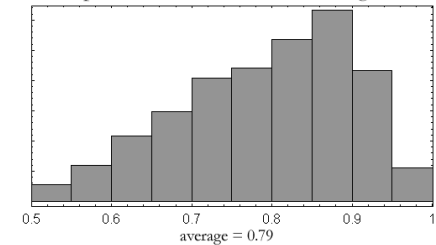
# Performance Correlations

	Bi	Br	Ch	Fl	In	Lu	R8	R6	Sm	Un
Biret	1.									
Brailowsky	0.92	1.								
Chiu	0.81	0.81	1.							
Friere	0.83	0.86	0.86	1.						
Indjic	0.95	0.91	0.86	0.88	1.					
Luisada	0.85	0.84	0.81	0.84	0.88	1.				
Rubinstein 1938	0.62	0.66	0.76	0.73	0.66	0.67	1.			
Rubinstein 1966	0.5	0.55	0.74	0.7	0.59	0.61	0.77	1.		
Smith	0.55	0.65	0.67	0.74	0.63	0.56	0.62	0.59	1.	
Uninsky	0.86	0.85	0.89	0.89	0.9	0.89	0.75	0.69	0.64	1.

Simple Mazurka Correlation Histogram

average = 0.79

Simple Mazurka Correlation Histogram

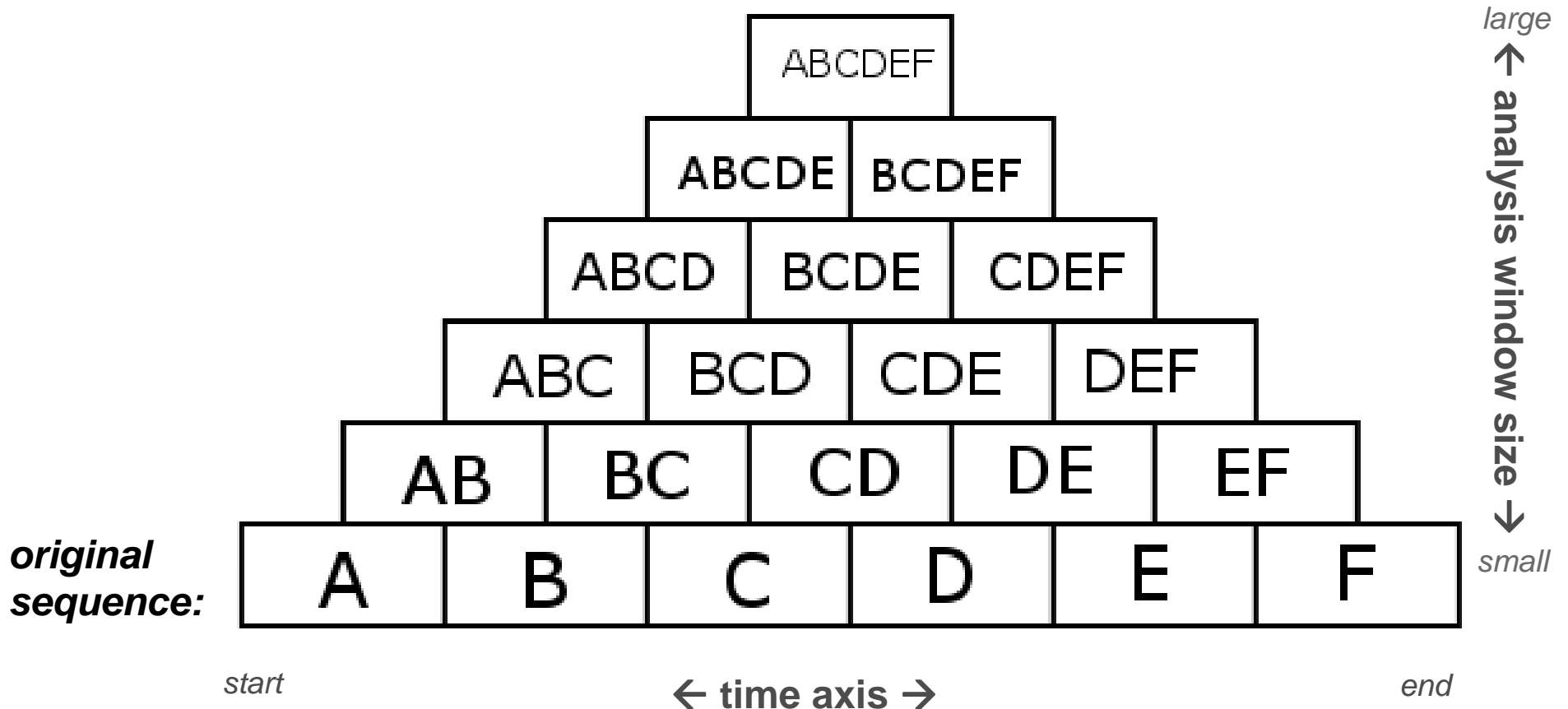


## Highest correlation to Biret: 0.95

**Lowest correlation  
to Biret: 0.50**

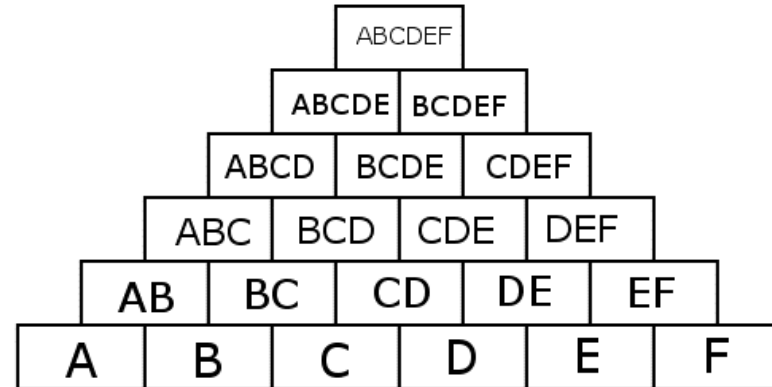
# Scape Plotting Domain

- 1-D data sequence chopped up into all possible  $n$ -grams to form a 2-D plot
- Example of a composition with 6 beats at tempos A, B, C, D, E, and F:

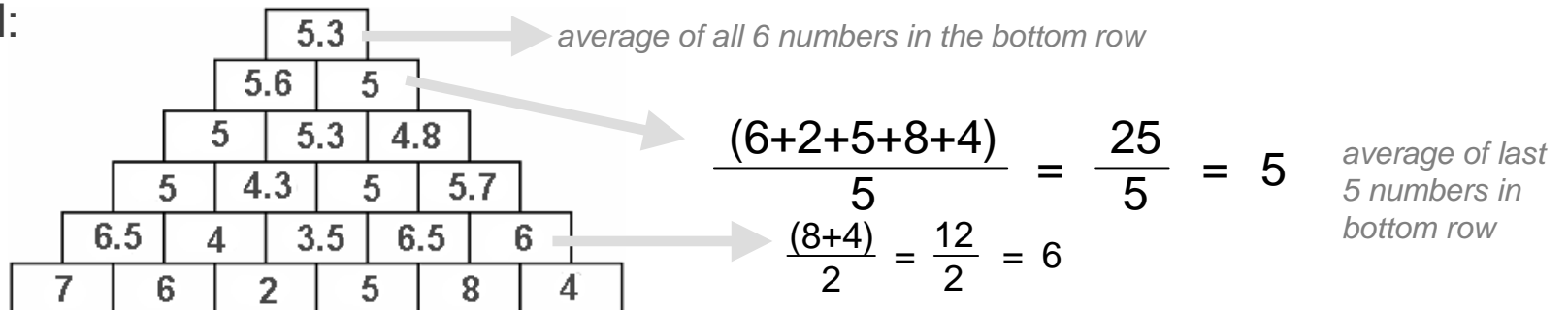


# Scape Plotting Range

- Any operation can be applied to each cell in plotting domain:



- Example using ***averaging*** in each cell:





# Comparative Timescapes

- **Cell operation:** Index of performance which has highest correlation to reference performance (excluding reference).

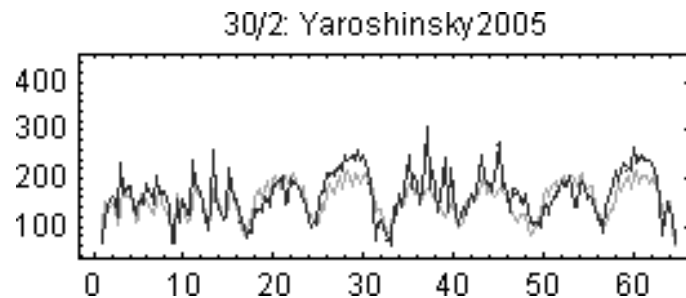
$$\text{cell}_{ij} = \max_{k \neq r} \text{Correlation}(\text{performance}_{ijr}, \text{performance}_{ijk})$$

- **Each performance assigned unique (arbitrary) index color:**

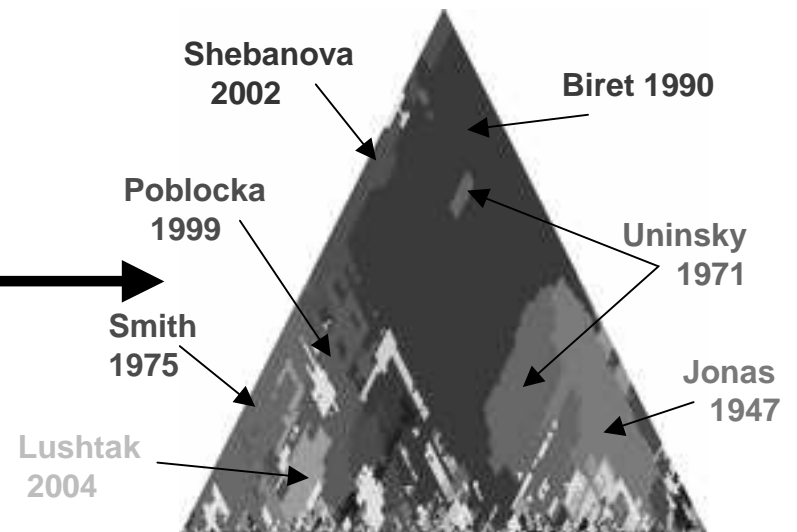
■ performer 1  
■ performer 2  
■ performer 3  
etc.

Resulting plot shows best correlated performance at all timescales:

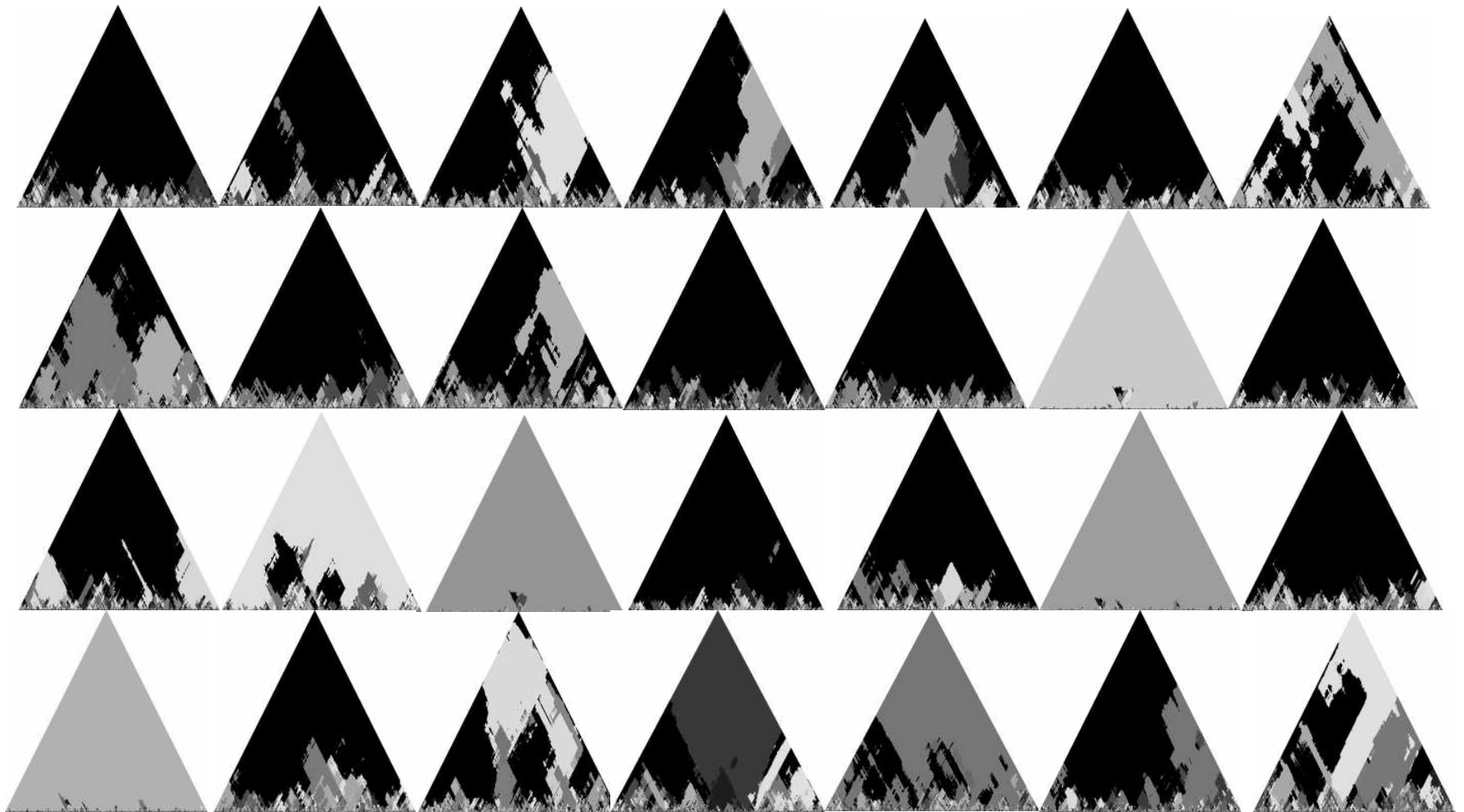
- **Example reference performance:**



Note: causality is not addressed directly in plots



# Performance Correlation Scares



[mazurka.org.uk/ana/pcor](http://mazurka.org.uk/ana/pcor) (tempo)

[mazurka.org.uk/ana/pcor-gbdyn](http://mazurka.org.uk/ana/pcor-gbdyn) (dynamics)

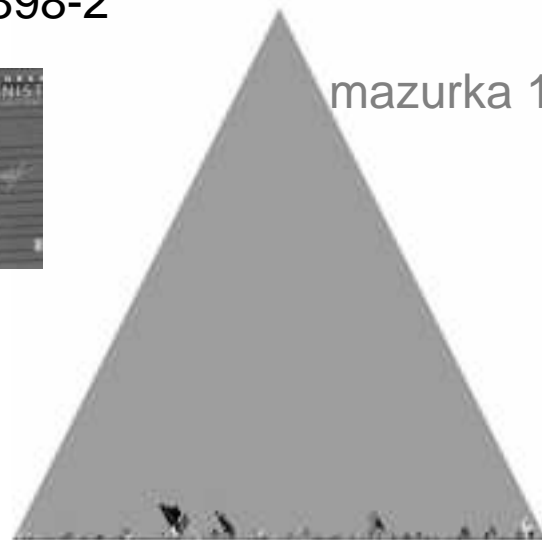
<http://mazurka.org.uk/ana/pcor-joint> (tempo+dynamics)

# Boring Timescape Pictures

Occasionally get over-exposed photographs back from the store,  
and have to throw them in the waste bin.

The same performance by Magaloff on two different CD re-releases:

Philips 456 898-2

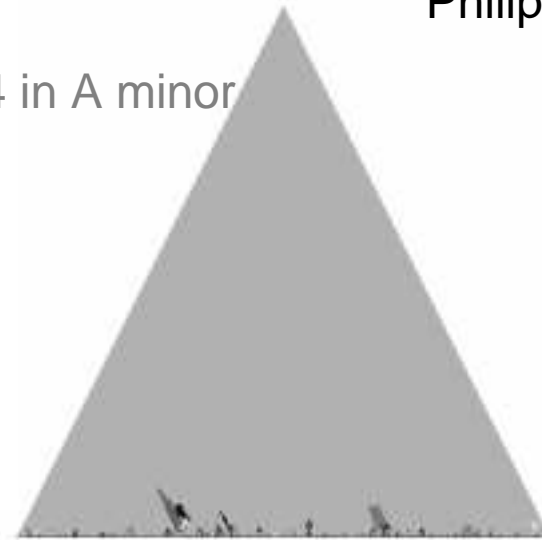


Magaloff 1977



mazurka 17/4 in A minor

Philips 426 817/29-2



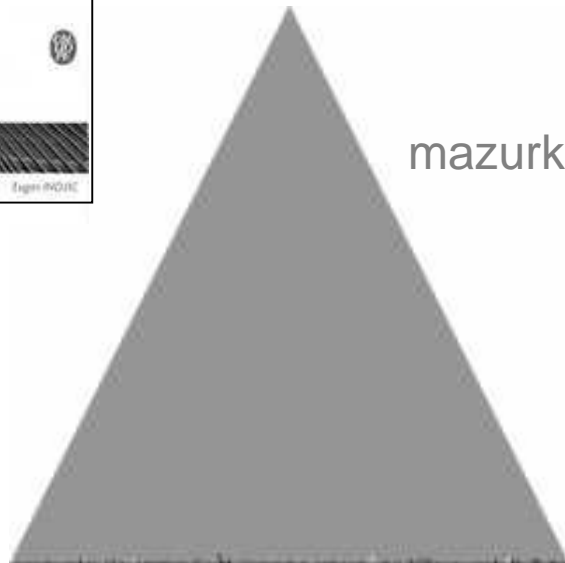
Magaloff 1977b



# Boring Timescape Pictures?

Two difference performances from two different performers on two different record labels from two different countries.

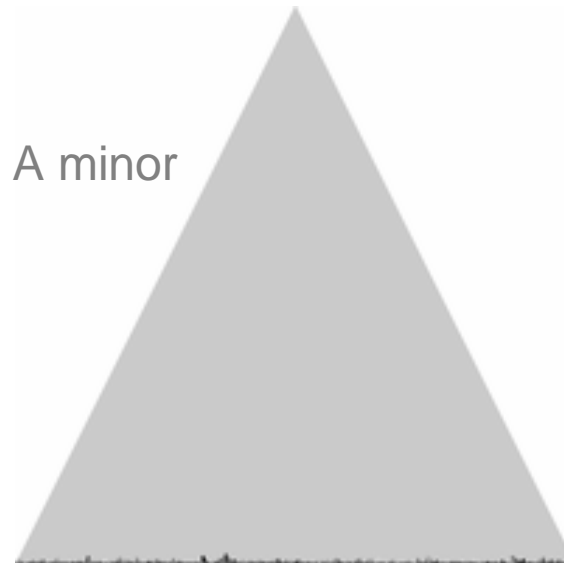
Calliope 3321



■ Indjic 1988



Concert Artist 20012



■ Hatto 1997



see: [http://www.charm.rhul.ac.uk/content/contact/hatto\\_article.html](http://www.charm.rhul.ac.uk/content/contact/hatto_article.html)

# Hatto Hoax

~ 100 CDs of Joyce Hatto performances issued on the Concert Artist label in 2003-2006. (also 70 cassettes in 80's and 90's)

Indjic mazurka performances first borrowed for a Hatto cassette release in 1993.

- Origins of ~65 CDs on other commercial recordings have been identified (23 in the first week after story broke)

[http://en.wikipedia.org/wiki/Joyce\\_Hatto](http://en.wikipedia.org/wiki/Joyce_Hatto)

<http://www.farhanmalik.com/hatto/cdlist.html>

- Borrowed performances from at least 70 pianists

<http://www.farhanmalik.com/hatto/pianistslist.html>

---

LETTER FROM ENGLAND  
**FANTASIA FOR PIANO**  
*Joyce Hatto's incredible career.*  
BY MARK SINGER



September 17, 2007  
The New Yorker

# Fiorentino Fakes

Sergio Fiorentino



Con. Artist:  
CACD 9200-2  
(2003)

Mazurka 6/2  
**Mazurka 7/1**  
Mazurka 7/2  
**Mazurka 7/3**  
**Mazurka 7/4**  
Mazurka 7/5  
Mazurka 17/2  
**Mazurka 17/4**  
Mazurka 24/1  
Mazurka 24/2  
Mazurka 24/3  
**Mazurka 24/4**  
Mazurka 30/1  
Mazurka 30/2  
Mazurka 30/3  
**Mazurka 30/4**  
Mazurka 33/1  
Mazurka 33/3  
Mazurka 33/4  
Mazurka 50/2  
Mazurka 50/3  
Mazurka 56/3  
Mazurka 59/1  
Mazurka 59/2  
Mazurka 59/3  
Mazurka 63/2

**Mazurka 7/1**  
Mazurka 7/2  
**Mazurka 7/3**  
**Mazurka 7/4**  
Mazurka 7/5  
Mazurka 68/1  
Mazurka 68/2  
Mazurka 68/3  
Mazurka 68/4  
Mazurka 24/1  
Mazurka 24/2  
Mazurka 24/3  
**Mazurka 24/4**  
Mazurka 17/1  
Mazurka 17/2  
Mazurka 17/3  
**Mazurka 17/4**  
Mazurka 30/1  
Mazurka 30/2  
Mazurka 30/3  
**Mazurka 30/4**  
Mazurka 41/2  
Mazurka 67/4

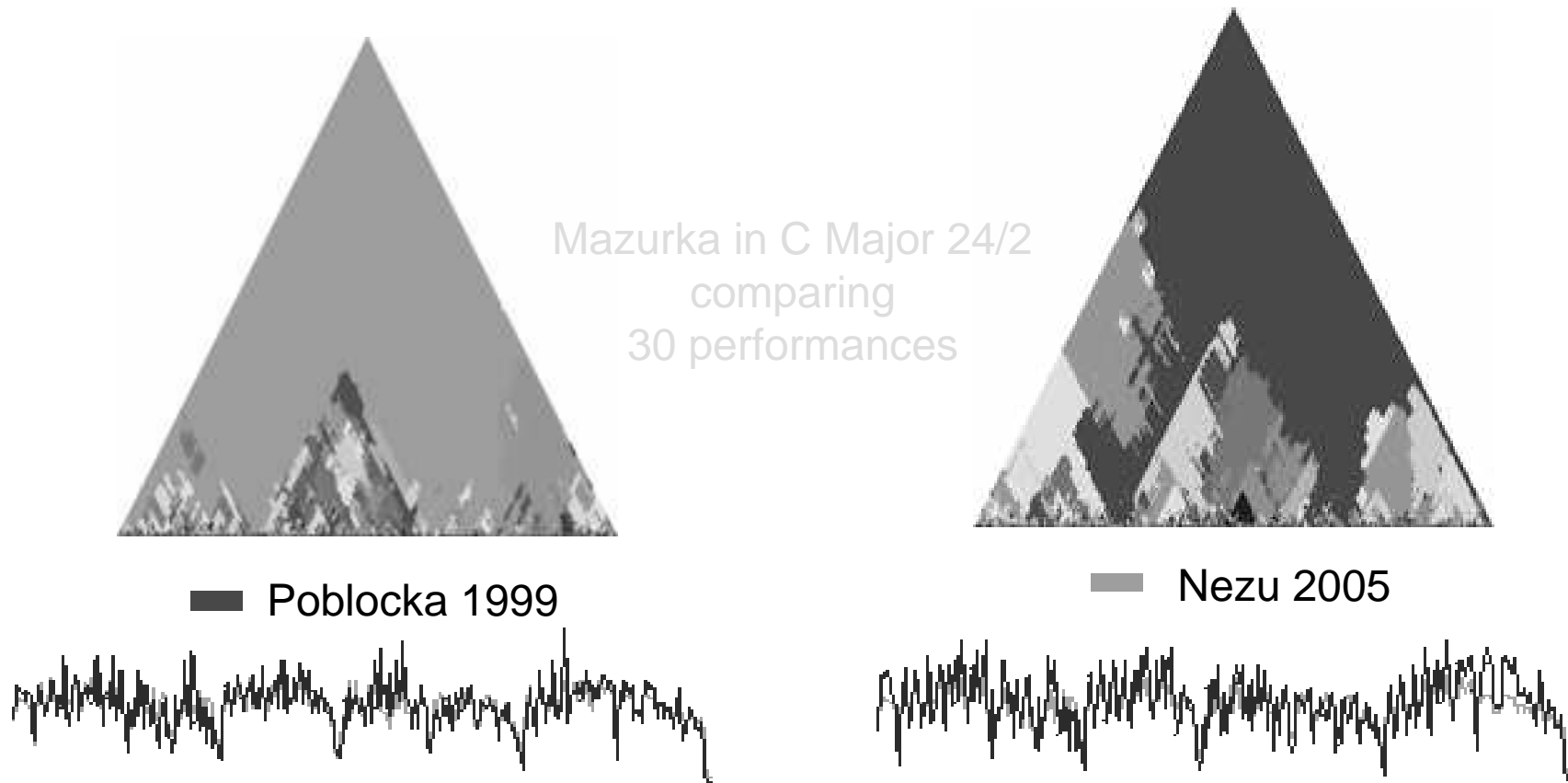
Janusz Olejniczak



Naïve/OPUS 111:  
OP20002 (1991)

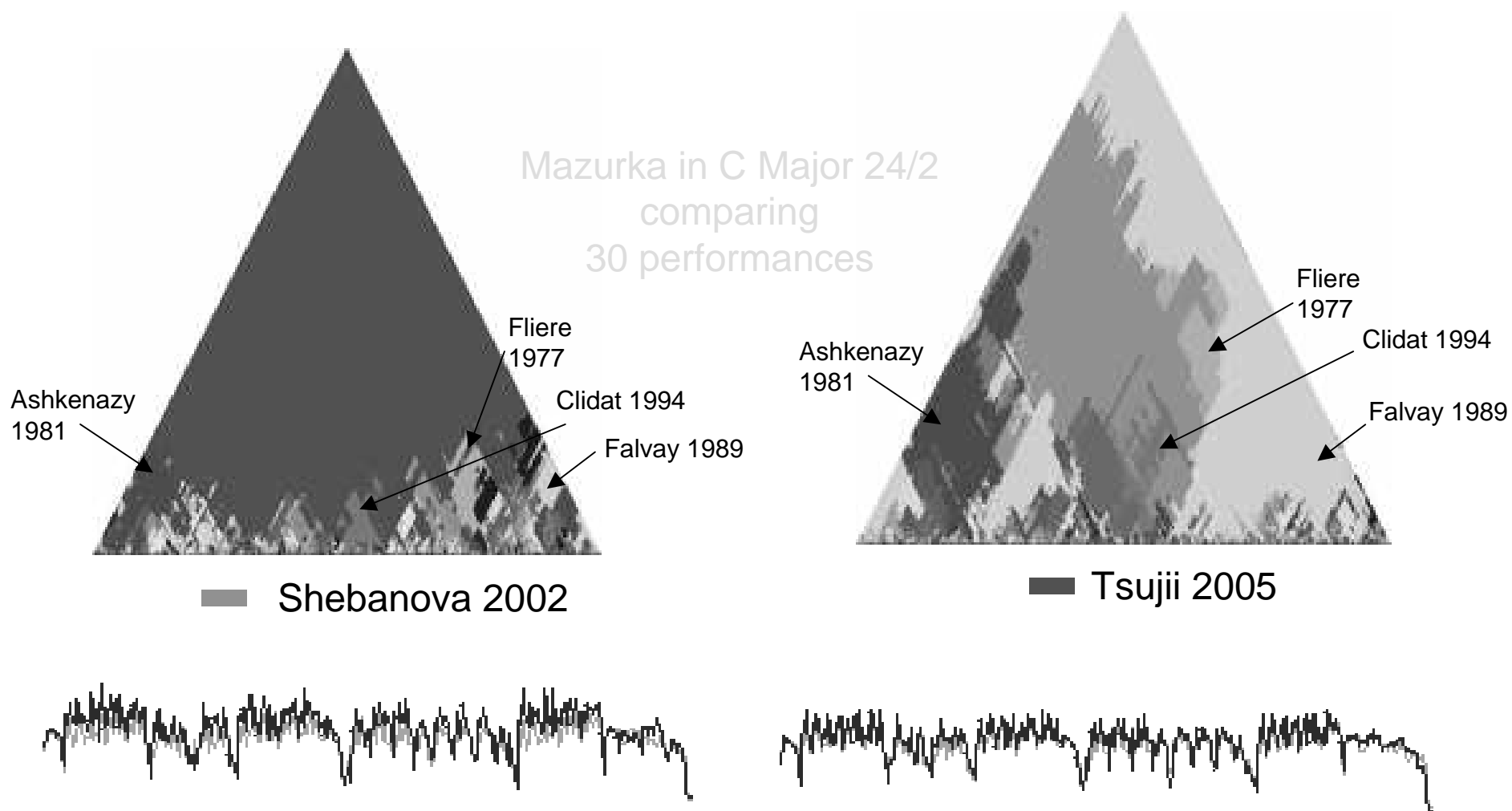
# Strong Interpretive Influences

- Timescapes were developed to examine “soft plagiarism” in performances  
-- more interesting than Con. Artist’s copyright infringement



- Nezu did graduate studies with Poblocka

# Strong Interpretive Influences (2)

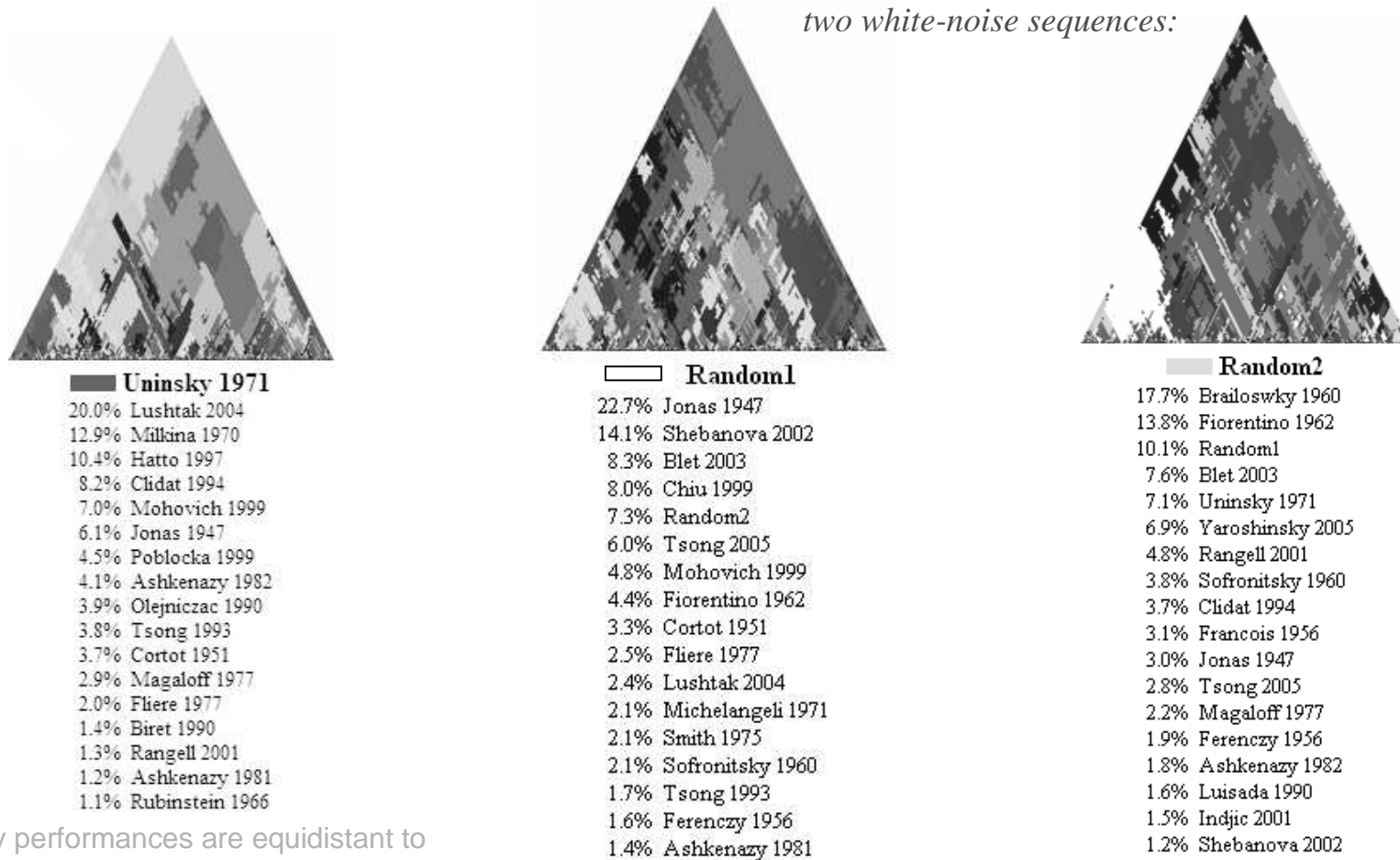


- note parallel colors: same neighboring pianists (clustering in performance space)



# Purely Random Matches

- Plots have to show some match at all points  
-- not necessarily a good one
- Small color regions, inverted triangles & broken borders = poor matches

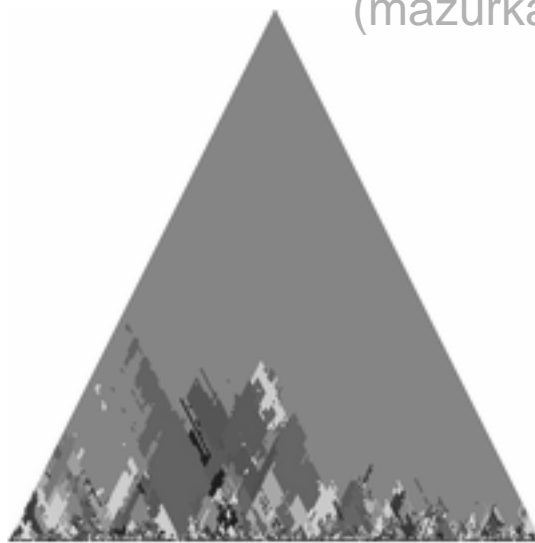


- Many performances are equidistant to Uninsky performance, none probably particularly similar to his performance.

(mazurka 30/2; 36 performances)

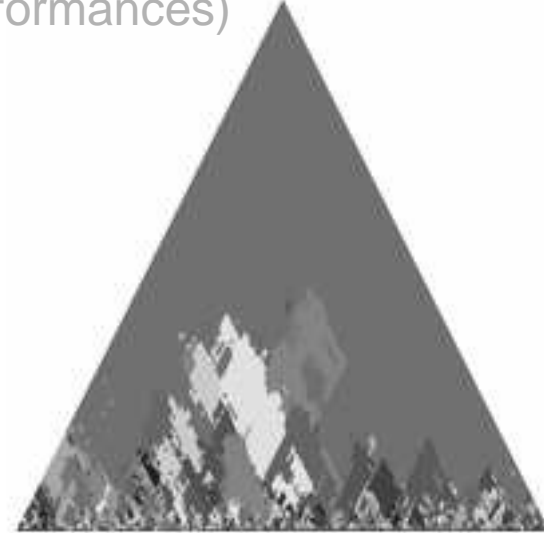
# Same Performer over Time

(mazurka 63/3; 60 performances)



■ Uninsky 1932

74.1% Uninsky 1971  
4.9% Moravec 1969  
2.8% Gierzod 1998  
2.8% Harasiewicz 1955  
1.7% Schilhawsky 1960  
1.2% Czerny 1949  
1.1% Rabcewiczowa 1932



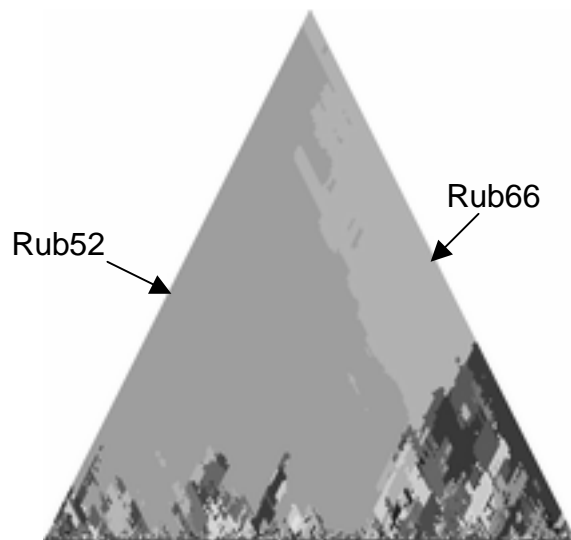
■ Uninsky 1971

66.1% Uninsky 1932  
6.0% Indjic 1988  
5.2% Boshniakovich 1969  
2.6% Czerny 1949  
2.6% Gierzod 1998  
2.1% Schilhawsky 1960  
1.9% Milkina 1970  
1.7% Harasiewicz 1955

- 40 years between recordings
- 78 rpm recording / 33.3 rpm recording
- France in 1932 / Texas in 1971

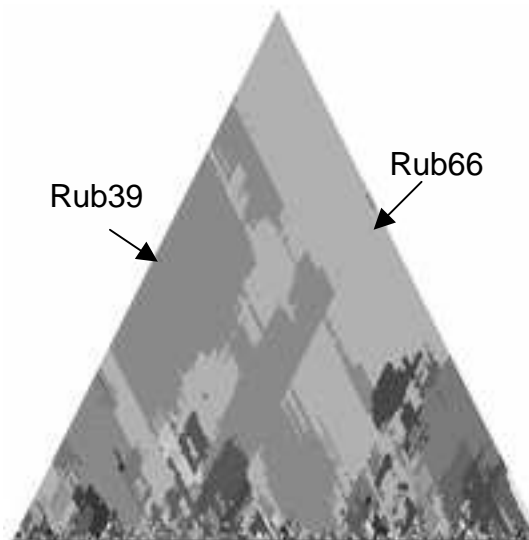
# Same Performer over Time (2)

(Mazurka 63/3; 60 performances)



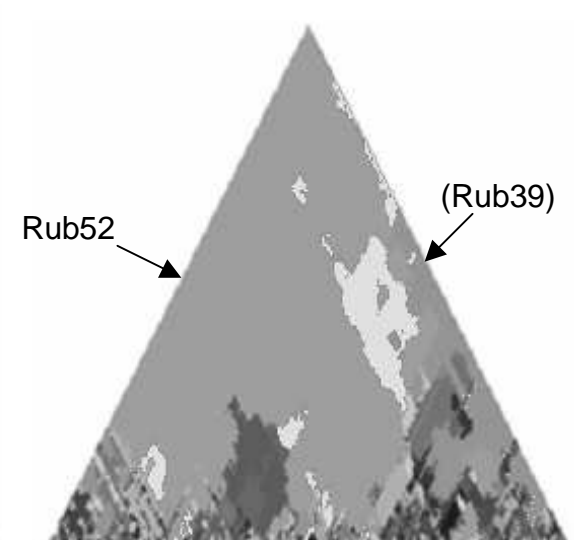
## Rubinstein 1939

57.6% Rubinstein 1952  
20.5% Rubinstein 1966  
2.4% Zak 1937  
1.9% Milkina 1970  
1.9% Luisada 1991  
1.8% Rachmaninoff 1923  
1.3% Czerny 1949  
1.2% Ashkenazy 1981



## Rubinstein 1952

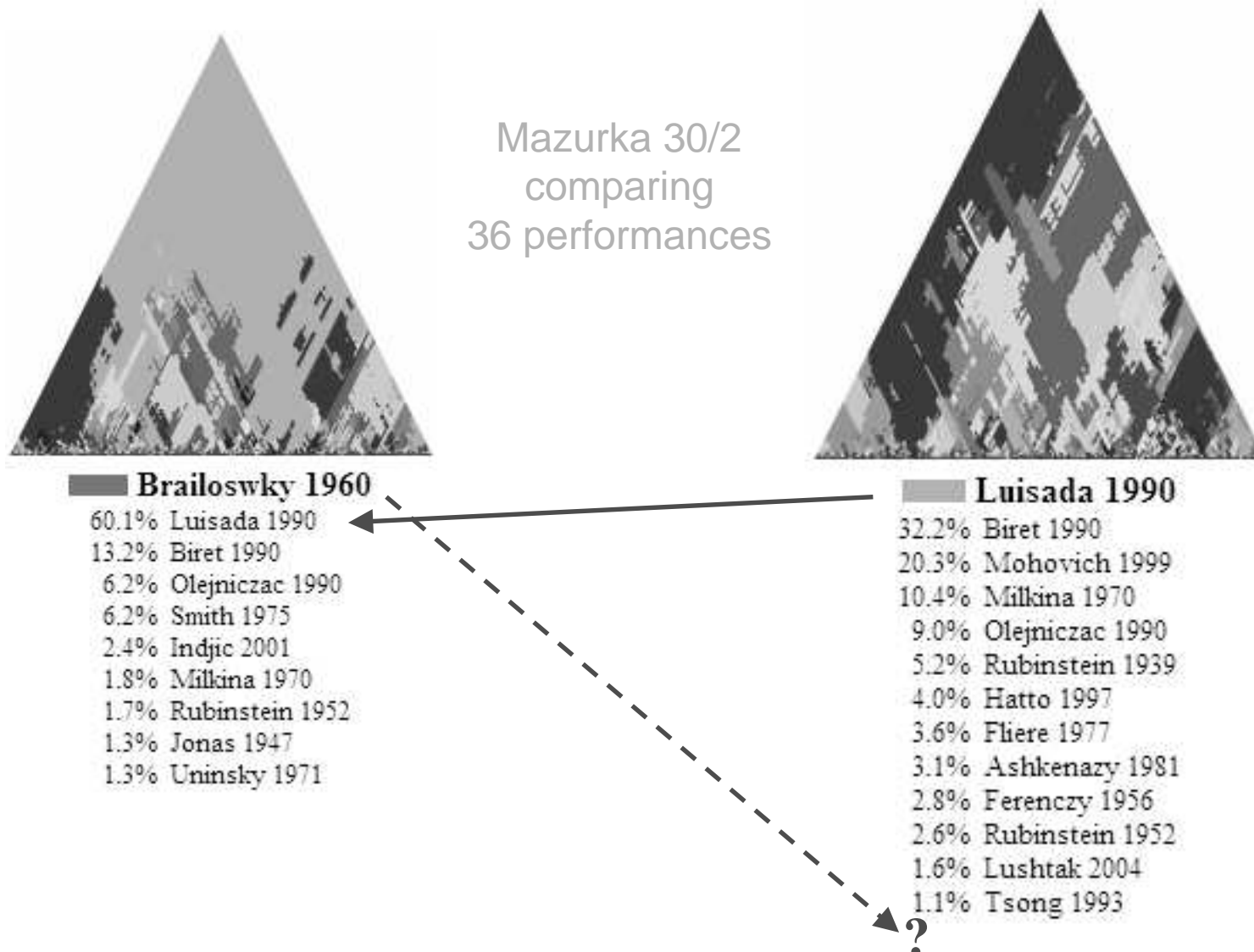
37.5% Rubinstein 1966  
35.2% Rubinstein 1939  
4.3% Falvay 1989  
2.5% Poblocka 1999  
2.5% Rachmaninoff 1923  
2.2% Mohovich 1999  
1.6% Kapell 1951  
1.4% Afanassiev 2001  
1.2% Shebanova 2002  
1.2% Milkina 1970



## Rubinstein 1966

59.3% Rubinstein 1952  
8.3% Kushner 1989  
8.2% Rubinstein 1939  
3.2% Osinska 1989  
2.2% Kapell 1951  
2.1% Hatto 1988  
1.8% Chiu 1999  
1.5% Milkina 1970  
1.3% Zak 1937  
1.1% Falvay 1989  
1.1% Lushtak 2004

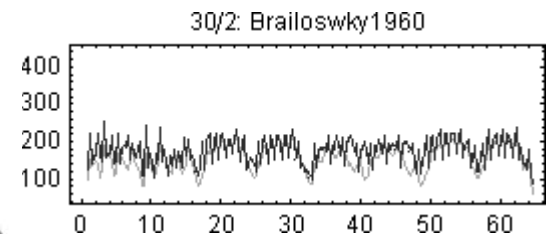
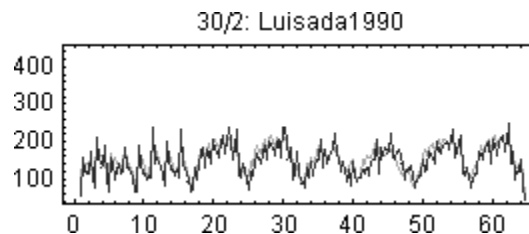
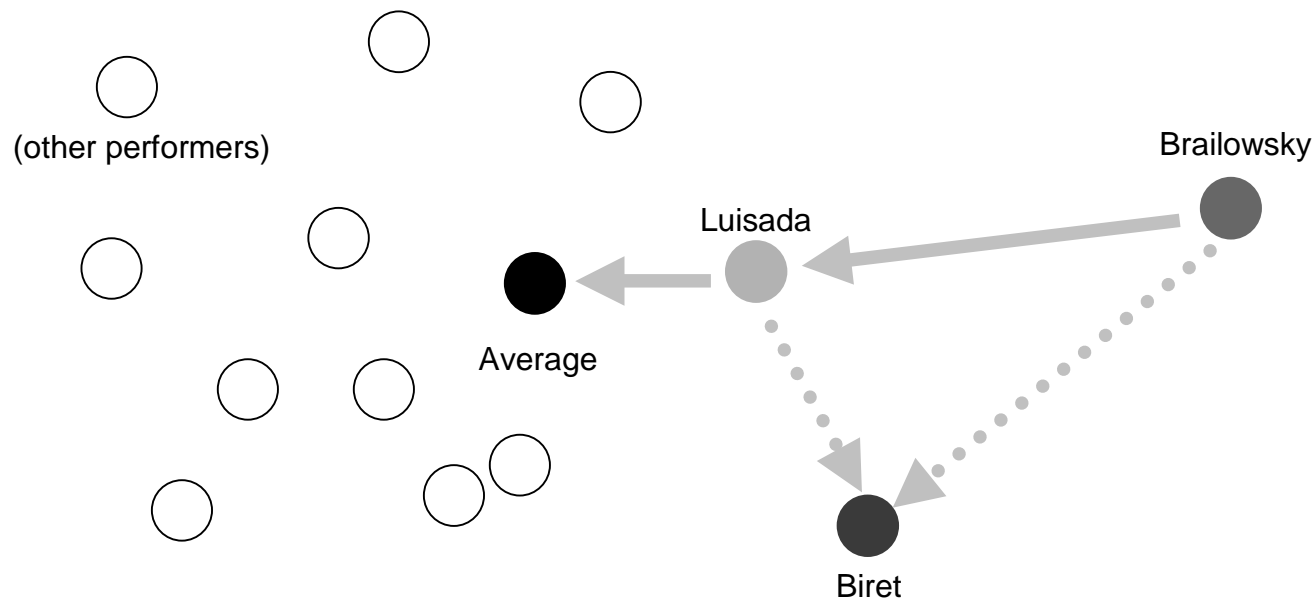
# Non-mutual Matching



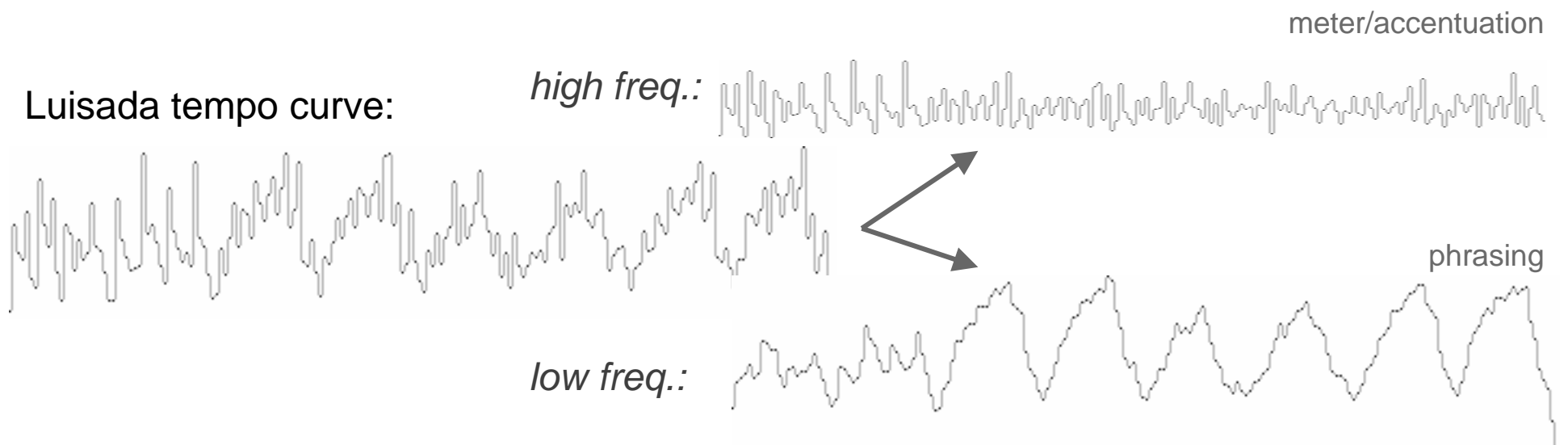
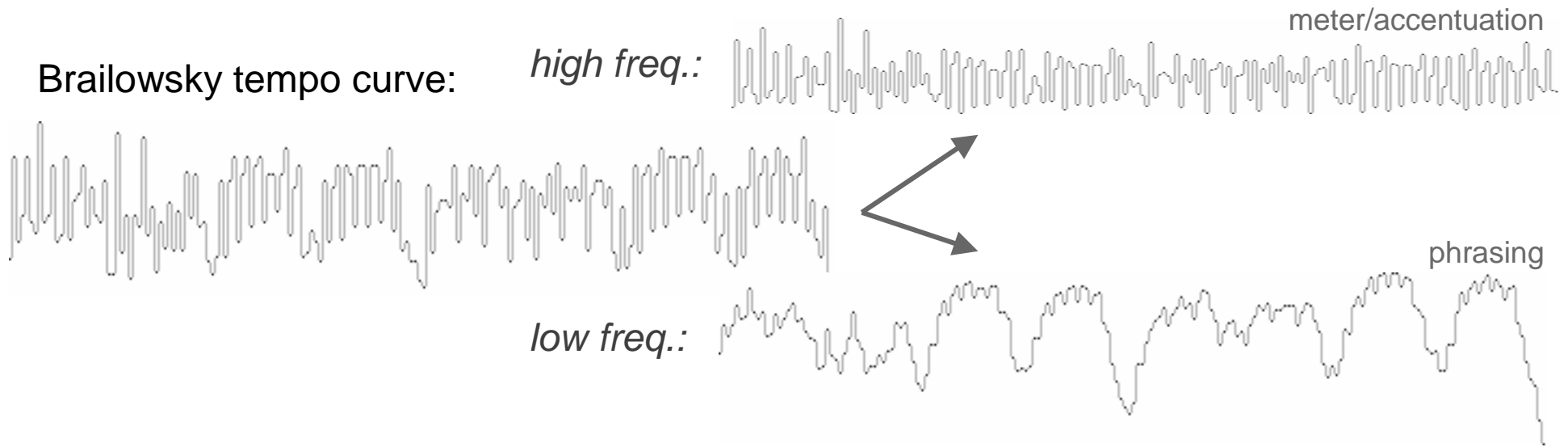
- Looking at both performers pictures helps interpreting “significance” of matches

# Performance Map Schematic

- Brailowsky has the strongest mazurka meter pattern
- Luisada has the second strongest mazurka meter pattern



# Tempo Components



# Segregated Tempo Components

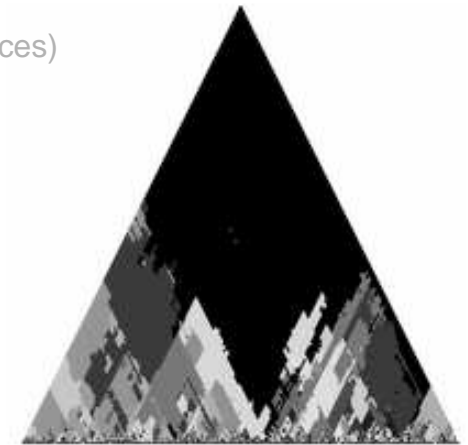
(mazurka 30/2; 36 performances)

High frequency  
tempo components:  
*accentuation*  
(metrical pattern)



Brailoswky 1960

57.8% Luisada 1990  
12.1% Biret 1990  
6.7% Average



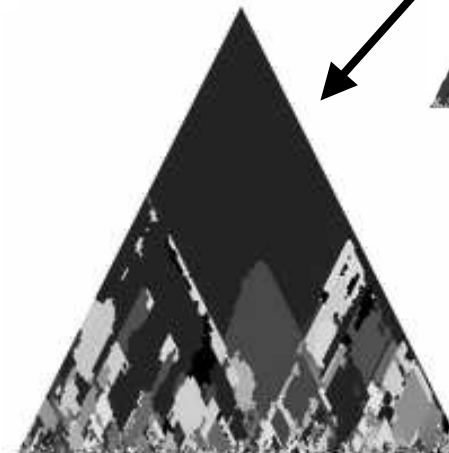
Luisada 1990

52.9% Average  
14.4% Biret 1990  
6.7% Brailoswky 1960

Raw  
Beat Tempo

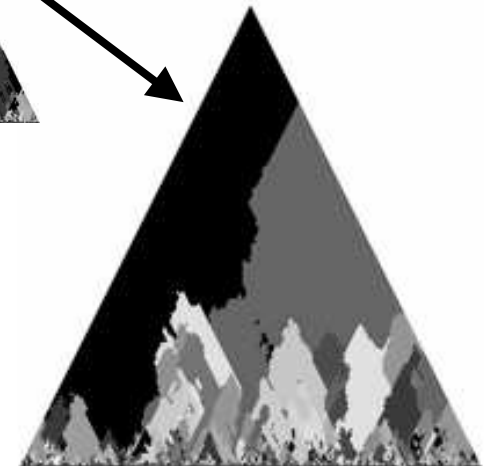


Low frequency  
tempo components:  
*phrasing (accel., rit.)*



Brailoswky 1960

54.8% Magaloff 1977  
7.1% Yaroshinsky 2005  
3.3% Shebanova 2002



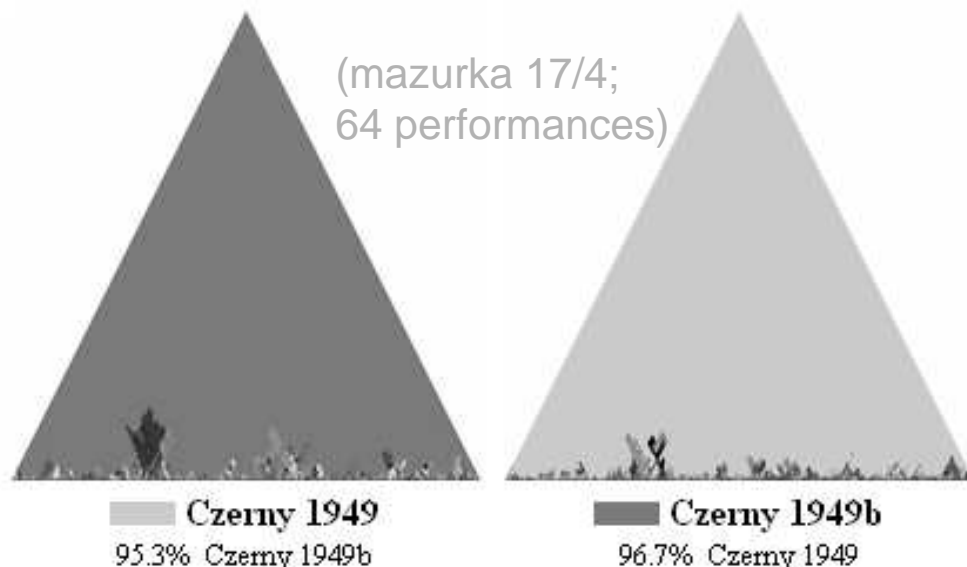
Luisada 1990

36.8% Average  
28.0% Mohovich 1999  
6.1% Cortot 1951

# More Same Performers

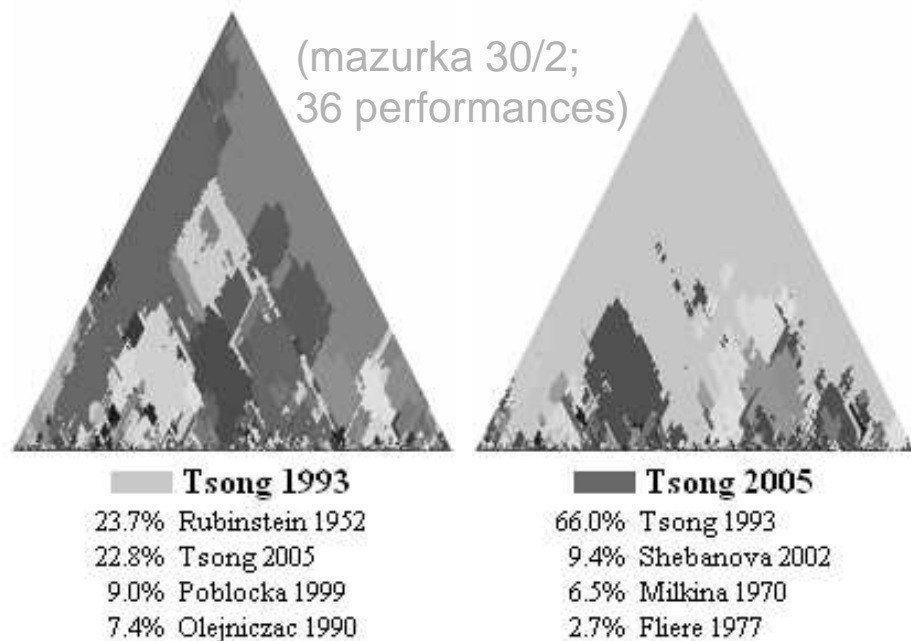
- **Most self-similar**

- Recorded in same year
- One performance live in concert hall (Chopin competition)
- Other recorded in studio.



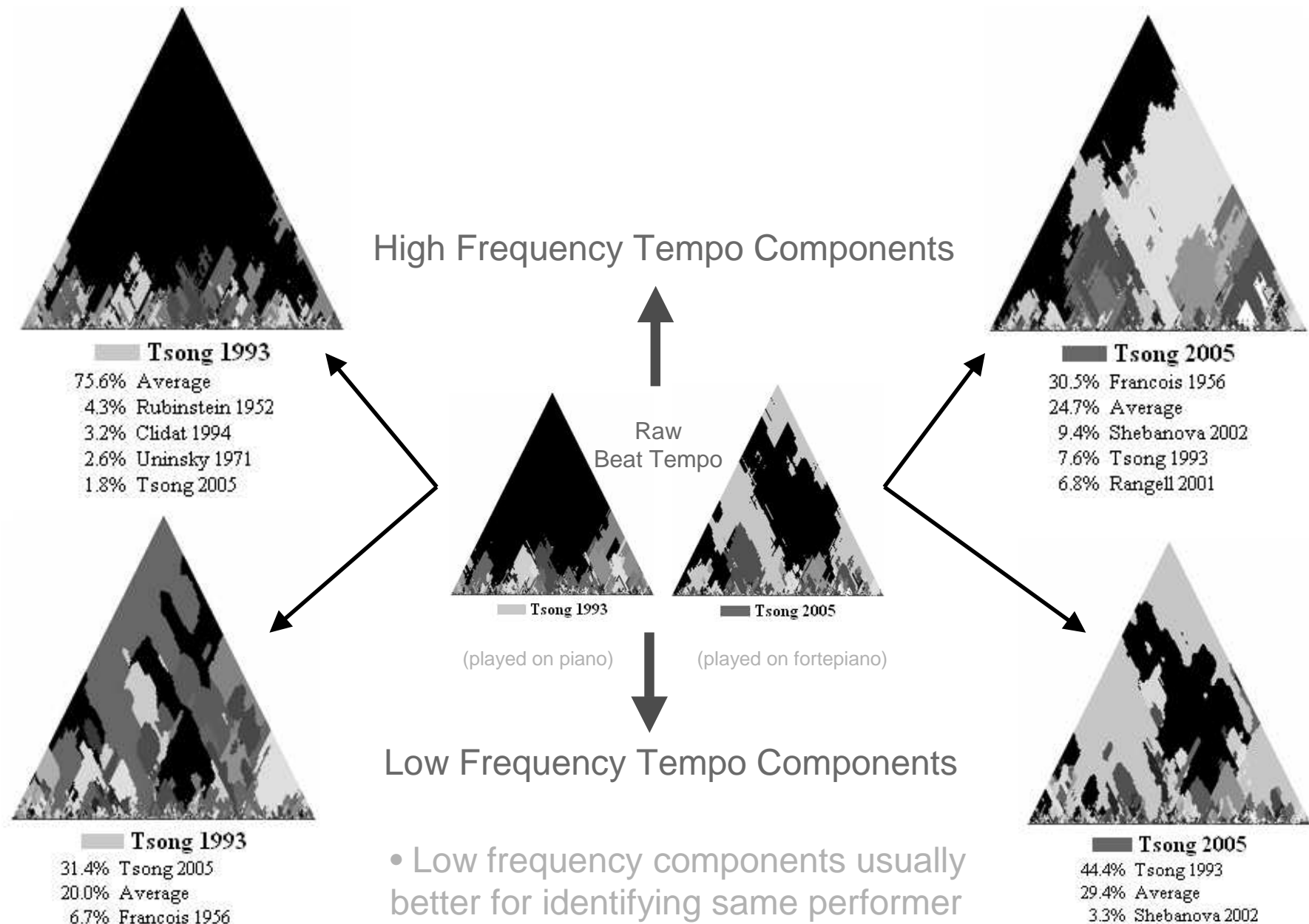
- **Least self-similar**

- Recorded 22 years apart
- Later recording on a pianoforte
- Significant reinterpretation of some phrases (towards Shebanova)





# Characteristic Features of a Performer



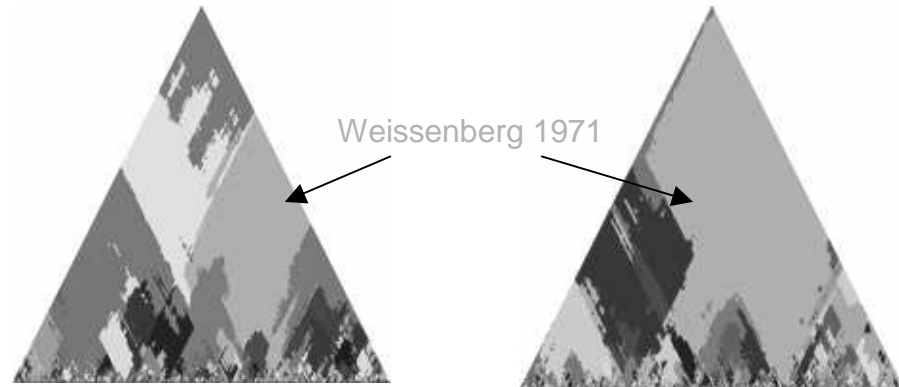
# Dynamics + Tempo scapes

(mazurka 17/4; 64 performances)

■ Chiu 1999

■ Falvay 1989

***timescape:***



***dynascape:***

(dynamics usually more variable than tempo between performances)



***dymescape(?):***



# Future Work

- More detailed performance features

LH/RH synchrony; off-beat rhythms; ornaments, *etc.*

- More composite/segregated feature analysis

separating features: Brailowsky/Luisada example for 24/2

joining features: Chiu/Falvay example for 17/4

- Quantitative measurements of performance similarity

timescapes are primarily qualitative, but not exclusively

- Authenticity analysis of Con.

Artist's Cortot mazurkas

and more on Fiorentino...





# Extra Slides

# Prasing Comparisons

Luisada



Brailowsky



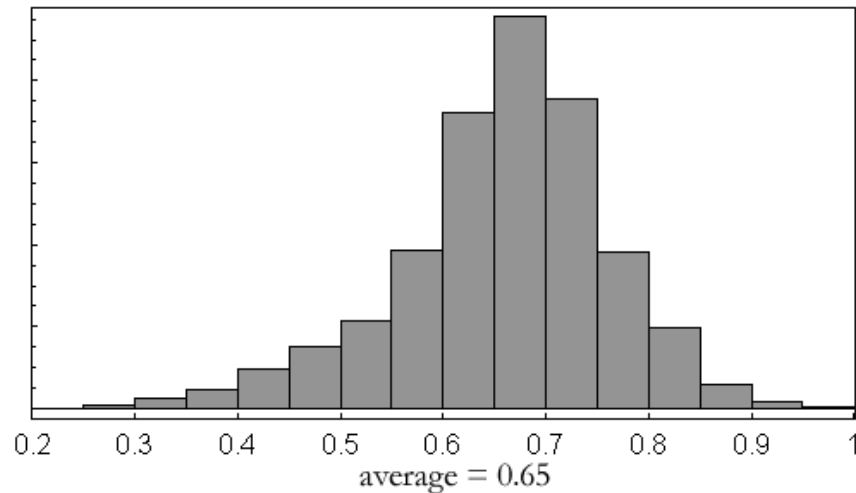
both are Russians  
& studied/worked in Paris...

Magaloff

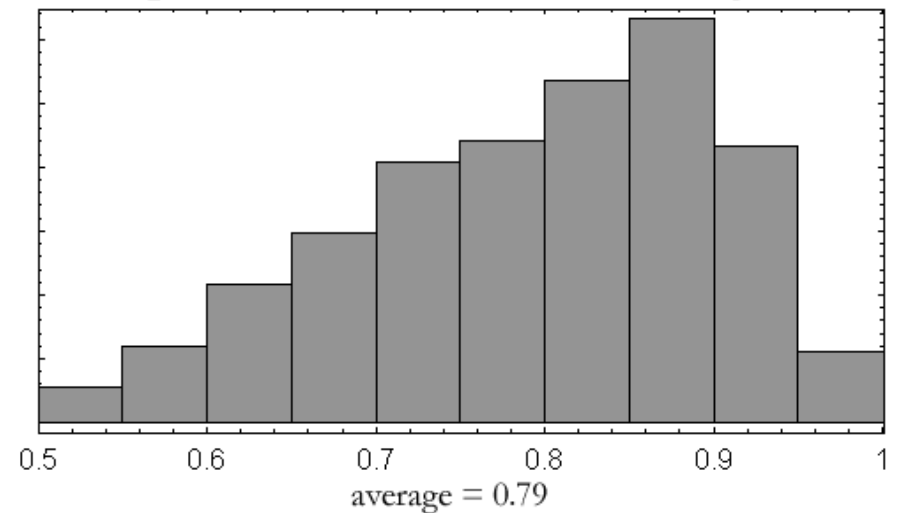


# Expected Correlation Values

Complex Mazurka Correlation Histogram



Simple Mazurka Correlation Histogram



- Different mazurkas have different correlation value distributions
- Simpler/shorter mazurkas have higher average correlations
- Complex/longer mazurkas have lower average correlations

# Hatto Ghost Performers

Ashkenazy, Vladimir	Duchable, François-René	Marshev, Oleg	Szokolay, Balázs
Aspaas, Tor Espen	Frith, Benjamin	Matsuzawa, Yuki	Tateno, Izumi
Babayan, Sergei	Gindin, Alexander	Moreira-Lima, Arthur	Thiollier, François-Joel
Banowetz, Joseph	Grante, Carlo	Muraro, Roger	Tipo, Maria
Baselga, Miguel	Gutierrez, Horacio	Nagy, Peter	Tomsic, Dubravka
Bellucci, Giovanni	Haebler, Ingrid	Nicolosi, Francesco	Trzeciak, Joanna
Benoit, Prisca	Hamelin, Marc-André	Nojima, Minoru	Wodnicki, Adam
Biret, Idil	Hegedüs, Endre	O'Connor, John	Zarafiants, Evgeny
Bloch, Boris	Heisser, Jean-François	Ogawa, Noriko	Zilberstein, Lilya
Bronfman, Yefim	Hiseki, Hisako	Ohlsson, Garrick	
Browning, John	Hobson, Ian	Okashiro, Chitose	
Brownridge, Angela	Indjic, Eugene	Pagny, Patricia	
Budiardjo, Esther	Jandó, Jenő	Raekallio, Matti	
Campanella, Michele	Kim, Paul	Rahkonen, Margit	
Chen, Pi-hsien	Kissin, Evgeny	Ránki, Dezső	
Collard, Jean-Philippe	Kramreiter, Tomás	Reyes, Alberto	
Dalberto, Michel	Kuzmin, Leonid	Scherbakov, Konstantin	
Didenko, Yuri	Long, Beatrice	Simon, Lazlo	
Du Plessis, Herbert	Malikova, Anna	Sterczynski, Jerzy	



# How time + dynamics are mixed

**Correlation:**

$$\frac{\sum_i (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

**t\_n = (t1, t2, t3, t4, t5, t6, t7, t8, ..., tn)**

*original tempo sequence*

**d\_n = (d1, d2, d3, d4, d5, d6, d7, ... dn)**

*original dynamic sequence*

**J\_n = (Jt1, Jd1, Jt2, Jd2, Jt3, Jd3, ..., Jtn, Jdn)**     *joint sequence*

original time sequence is  
unaltered:

*original dynamic sequence is scaled  
to match tempo sequence's mean  
and standard deviation:*

$$J_{t,n} = t_n$$

$$J_{d,n} = \sigma_t \left( \frac{d_n - \mu_d}{\sigma_d} \right) + \mu_t$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

# Peeling Back the Layers

