On the Characterization of Expressive Performance in Classical Music
First Results of the *Con Espressione* Game

Carlos Cancino-Chacón[1,4], Silvan Peter[2], Shreyan Chowdhury[2], Anna Aljanaki[3] and Gerhard Widmer[2]

[1]Austrian Research Institute for Artificial Intelligence, Vienna, Austria
[2]Institute of Computational Perception, Johannes Kepler University Linz, Austria
[3]Institute of Computer Science, University of Tartu, Estonia
[4]RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Norway

The *Con Espressione* Game

Research Aims: Find the dimensions of musical expression that can be attributed to a performance, as perceived and described in natural language by listeners

- Web based questionnaire: verbal descriptors of expressive performance.
- Different performances of 9 classical piano pieces (45 performances)
- Dataset enriched with score-to-performance alignments

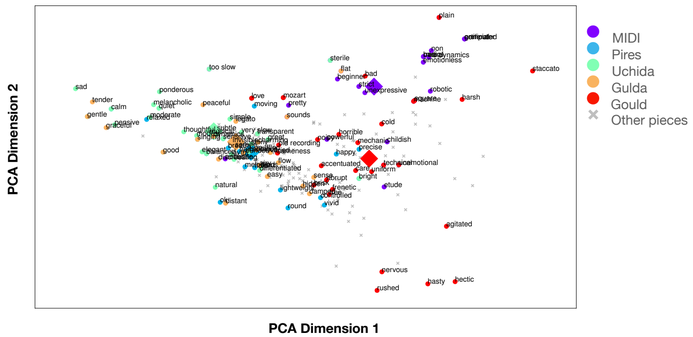
Composer	Piece	#	Pianists
Bach	Prelude No.1 in C, BWV 846 (WTC I)	7	Gieseking, Gould, Grimaud, Kempff, Richter, Stadtfeld, MIDI
Mozart	Piano Sonata K.545 C major, 2nd mvt.	5	Gould, Gulda, Pires, Uchida, MIDI deadpan
Beethoven	Piano Sonata Op.27 No.2 C# minor, 1st mvt.	6	Casadesus, Lazić, Lim, Gulda, Schiff, Schirmer
Schumann	Arabeske Op.18 C major (excerpt 1)	4	Rubinstein, Schiff, Vorraber, Horowitz
Schumann	Arabeske Op.18 C major (excerpt 2)	4	Rubinstein, Schiff, Vorraber, Horowitz
Schumann	Kreisleriana Op.16; 3. Sehr aufgeregt (ex 1)	5	Argerich, Brendel, Horowitz, Vogt, Vorraber
Schumann	Kreisleriana Op.16; 3. Sehr aufgeregt (ex 2)	5	Argerich, Brendel, Horowitz, Vogt, Vorraber
Liszt	Bagatelle sans tonalité, S.216a	4	Bavouzet, Brendel, Katsaris, Gardon
Brahms	4 Klavierstücke Op.119, 2. Intermezzo E minor	5	Angelich, Ax, Serkin, Kempff, Vogt

What are the main dimensions for expressive character?

Principal component analysis (PCA) on the occurrence matrix of the terms and find 4 principal dimensions

Dimension 1				Dimension 2			
positive correlation		negative correlation		positive correlation		negative correlation	
hectic	0.17	sad	-0.20	rushed	0.22	hard	-0.19
staccato	0.15	gentle	-0.18	nervous	0.20	stumbling	-0.18
hasty	0.15	tender	-0.18	too fast	0.17	staccato	-0.17
agitated	0.14	calm	-0.16	bit	0.16	ponderous	-0.14
irregular	0.14	graceful	-0.16	hasty	0.15	monotonous	-0.13

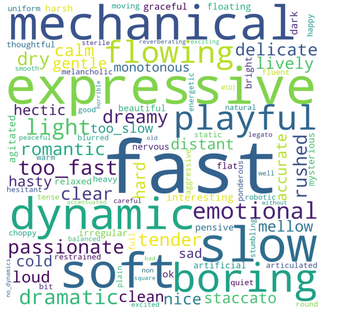
Dimension 3				Dimension 4			
positive correlation		negative correlation		positive correlation		negative correlation	
monotonous	0.22	heavy	-0.14	ok	0.24	cold	-0.15
bad	0.17	graceful	-0.13	happy	0.21	warm	-0.14
warm	0.16	smooth	-0.12	joyful	0.19	floating	-0.14
peaceful	0.16	ponderous	-0.12	free	0.15	blurred	-0.14
beautiful	0.15	soaring	-0.10	breathy	0.14	mysterious	-0.13

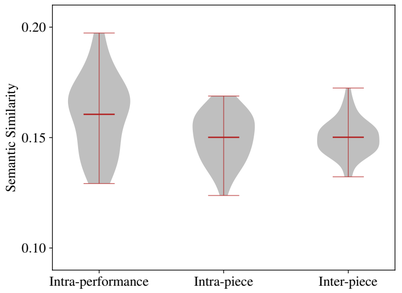


How similarly do listeners describe the performance of a piece?

Distribution of Terms

- 94 participants (on average listened to 4.5 out of 9 pieces)
- 88% had some musical training
- 1,515 individual descriptions, 3,166 terms (45% unique)





Semantic Similarity

- Semantic similarity for short sentences by [Li et al., 2007]
- **Intra-performance**: same piece, same pianist
- **Intra-piece**: same piece, other pianists
- **inter-piece**: other pieces

How do performance features relate to the expressive character dimensions?

Performance Parameters

Performance Parameters (PP)

- *tempo, loudness*

Mid-level Features (MF)

- From [Aljanaki and Soleymani, 2018]: *melodiousness, articulation, rhythmic complexity, rhythmic stability, dissonance, tonal stability, minorness*
- extracted from spectrograms using a CNN from [Chowdhury et al., 2019]

High-level Features (HF)

- 2D emotion space: *arousal and valence*
- Predicted using a CNN + GRU

Multiple Linear Regression to test the position of the pieces (their centroid) in the expressive character dimensions and the performance features

Dimension 1		Dimension 2		Dimension 3		Dimension 4	
PP ($R^2 = 0.24$)		PP ($R^2 = 0.18$)		PP ($R^2 = 0.26$)		PP ($R^2 = 0.24$)	
loudness avg	0.51^{***}	loudness sk	0.45^{**}	loudness std	-0.53^{**}	beat period k	-0.34^{*}
						loudness std	-0.44^{*}
MF ($R^2 = 0.39$)		MF ($R^2 = 0.00$)		MF ($R^2 = 0.00$)		MF ($R^2 = 0.29$)	
rhythmic complexity	-0.74^{*}	minorness	0.15	articulation	-0.15	rhythmic complexity	0.52^{*}
tonal stability	-0.94^{**}					tonal stability	0.84^{***}
articulation	0.46^{*}						
HF ($R^2 = 0.22$)		HF ($R^2 = 0.00$)		HF ($R^2 = 0.36$)		HF ($R^2 = 0.09$)	
valence sk	0.48^{**}	valence avg	0.14	valence k	0.42^{**}	valence k	-0.33^{*}
				arousal avg	-1.24^{***}		
				valence std	0.27^{*}		
				valence avg	-0.82^{*}		

Get the Dataset!

Acknowledgements

This research has received support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 670035 (project "**Con Espressione**") and by the Research Council of Norway through its Centers of Excellence scheme, project number 262762 and the **MIRAGE** project, grant number 287152.