The (obvious) influence of pitch and dynamics on the perceptual dimensions on the timbre of musical instruments

Saleh Siddiq[1†], Christoph Reuter[1], Isabella Czedik-Eysenberg[1,2], Denis Knauf[2]

[1]Institute of Musicology, University of Vienna, Austria, [2]Vienna University of Technology, Austria
[†]Corresponding Author email: saleh.siddiq@univie.ac.at

Introduction

Assumptions

Reducing musical instruments to only one note, as implicitly suggested by the standard definition [1], is a misconception, because every instrument can produce a broad range of timbres.
Studies with one single note per instrument (e.g. [2]) effectively study single sounds and yield results that perfectly apply to what they investigate: isolated sounds. Unfortunately the results are projected onto the entire instrument.
Studies with several tones per instrument suggest that perceptual dimensions shift with respect to the newly included variable (e.g. [3]).

Objectives

- Taking pitch and musical dynamics into account to provide a reasonable basis for the investigation of musical instruments.
- Identifying the perceptual timbre dimensions together with their respective physical correlates.

Methods

Stimuli selection

Goal
Inclusion of a broad variety of excitation mechanisms.

Restrictions
Stimuli number limited because of listening task duration.

Selection
Five nonpercussive instruments with different excitation mechanisms.

- Bassoon (BS)
- Clarinet (CL)
- Flute (FL)
- Trombone (TM)
- Violoncello (VC)

instrument		octaves of pitch range				
name	excitation	2	3	4	5	6
Flute (FL)	vortex at edge			E	E	E
Clarinet (CL)	reed		E		E	E
Cello (VC)	bowed string	E		E		E
Bassoon (BS)	double reed	E	E		E	
Trombone (TM)	player's lips	E	E	E		

Each note was tested in three different dynamic levels: *pianissimo (pp)*, mezzoforte (*mf*), **fortissimo (*ff*)**.

All sounds were taken from the *Vienna Symphonic Library* (VSL)

Listening task

pairwise comparison
n · (n−1)/2 pairs (n = 45 → 990 pairs)
→ AB-pairs rated, AA and BA omitted
7-point rating scale (7 = max dissimilarity)

Participants

n = 43 (female = 20, male = 22, other = 1)
age: 13–75 yrs. (Ø = 32.86, SD = 14.93)
musical activity: 32 active, 8 formerly, 2 never
experience: 2–60 yrs. (Ø = 20.66, SD = 14.08)

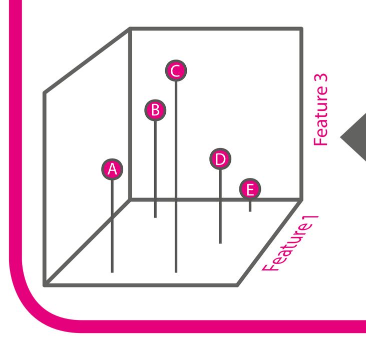
Empirical data

Multidimensional Scaling (MDS):
spatial representation of perceived dissimilarities
ordinal MDS (NMDS) via *smacof* in R

Physial data

Music Information Retrieval (MIR):
computational audio feature extraction
MIRtoolbox in Matlab

Combining empirical and physical data

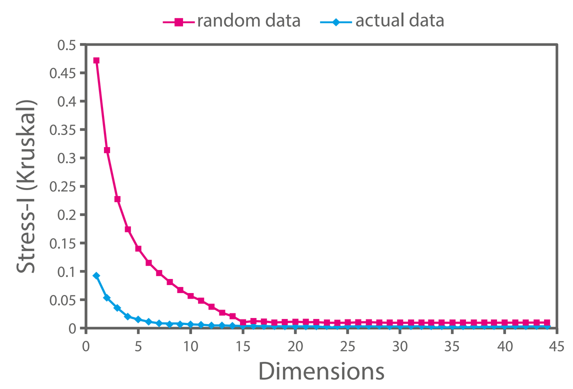


	Feat. 1	Feat. 2	Feat. 3
Sound A	5	2	3
Sound B	2	1	3
Sound C	5	3	5
Sound D	3	4	2
Sound E	1	5	1

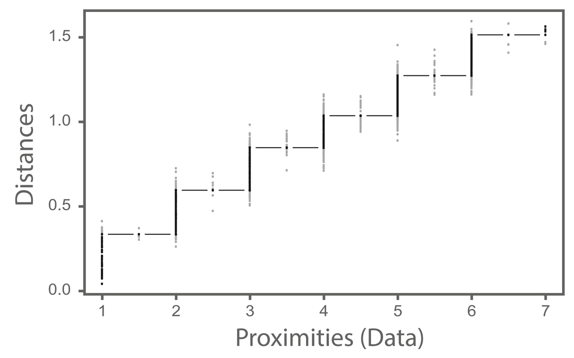
Correlations between feature values and coordinate vectors may explain the main axes of the configuration.

Results

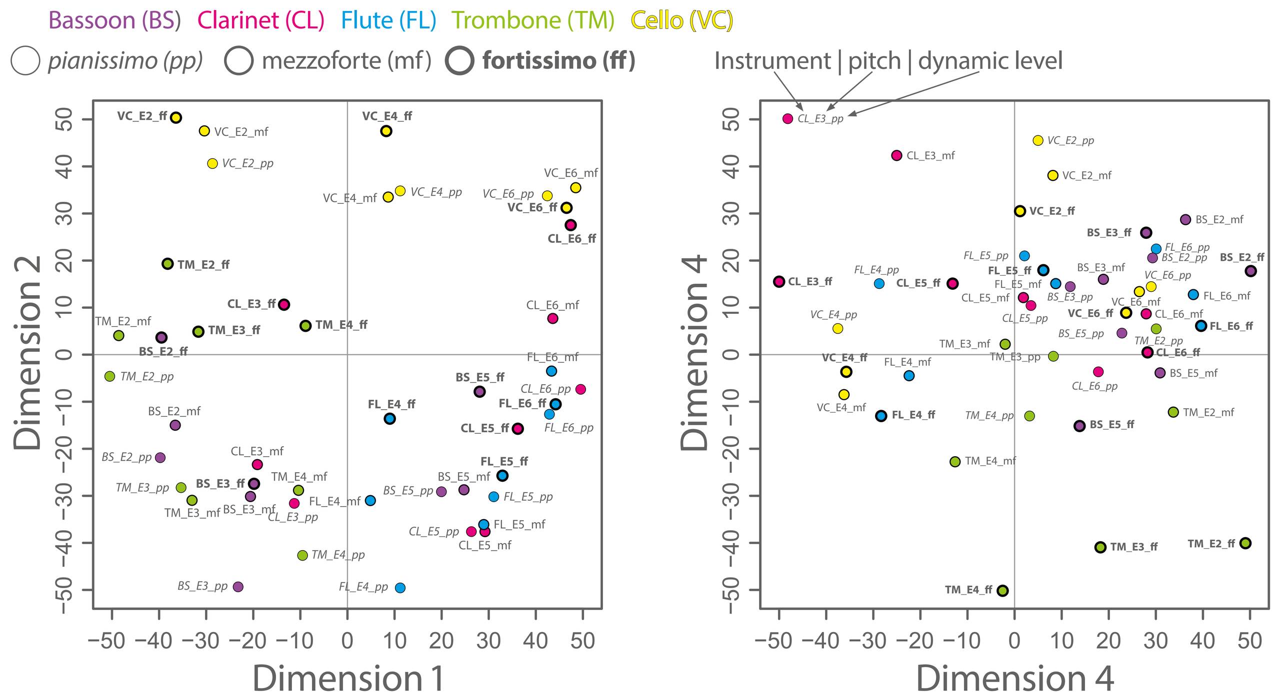
The spatial MDS configuration



Stress value for 4D configuration:
Stress-I (Kruskal): 0.03202913



Correlation of distances and proximities: **r = 0.97101671**



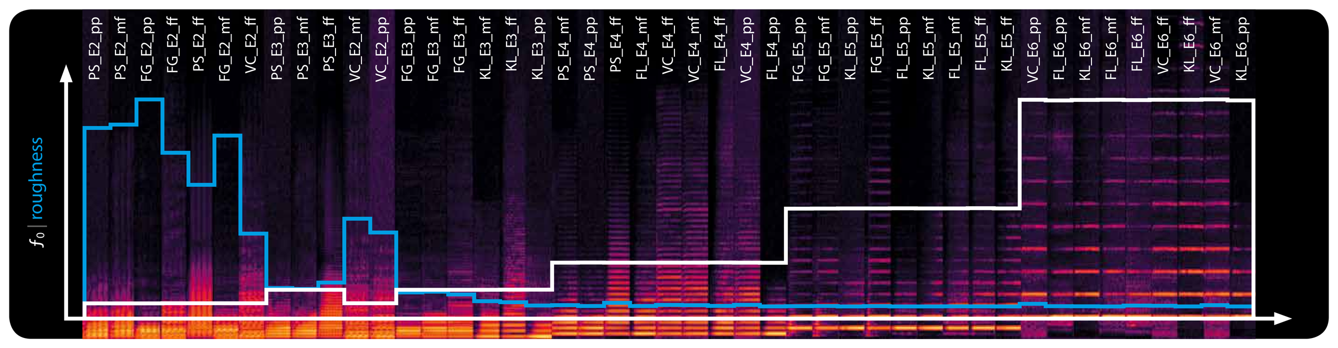
Three main takeaways

- Besides the cello along dimension 2, there are no instrument clusters.
- Instruments strew widely across the space.
- Similarities do not indicate that tones stem from the same instrument.

Dimension 1: pitch

Dimension 1 correlates strongly with features that are (very) closely related to pitch.

f_0: r = .97 (p < .000001)
roughness: r = −.86 (p < .000001)
Zero crossing rate: r = .76 (p < .000001)



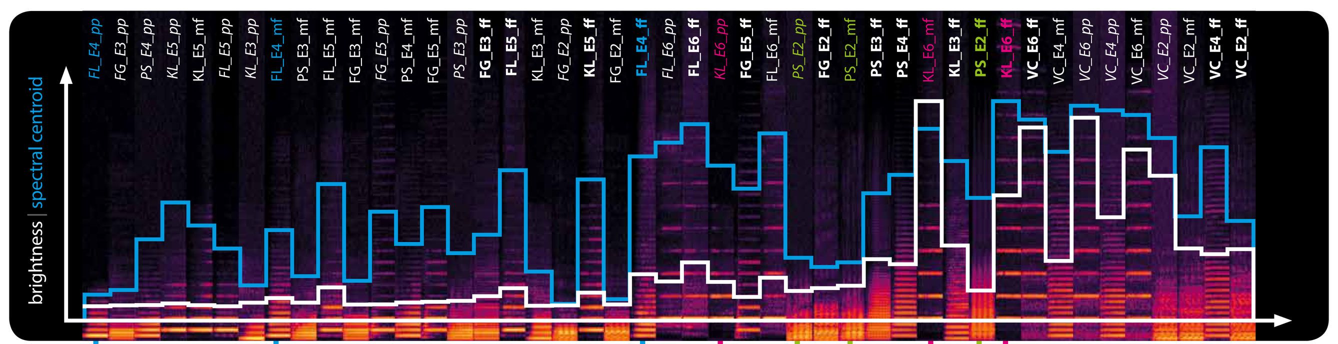
The graph shows a spectrogram of all the sounds along dimension 1. Superimposed are histograms of the strongest correlating features from the table above. Roughness (cyan) decreases as pitch (fundamental frequency, white) increases.

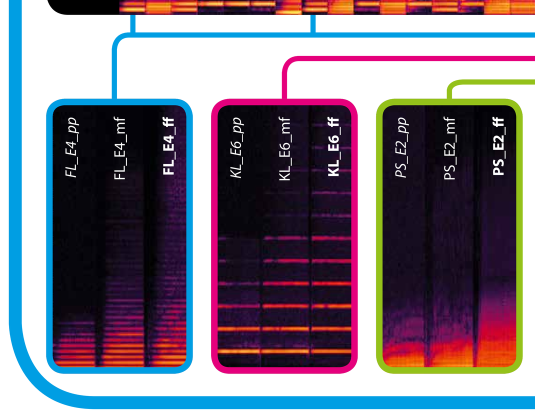
Dimension 2: brightness

Dimension 2 correlates strongly with features that are very closely related to the spectral energy distribution.

Brightness: r = .86 (p < .000001)
Spectral rolloff: r = .72 (p < .000001)
Spectral centroid: r = .62 (p < .00001)
Dynamic levels: r = .56 (p < .0001)

The graph below shows a spectrogram of all sounds along dimension 2. The histograms show the tendencially increasing values of brightness (white) and the spectral centroid (cyan). Labels are set set in italic (dynamic value pp), regular (mf) or bold (ff) for better visual orientation concerning dynamics. Colored labels relate to the exemplifying sounds below the main graph.



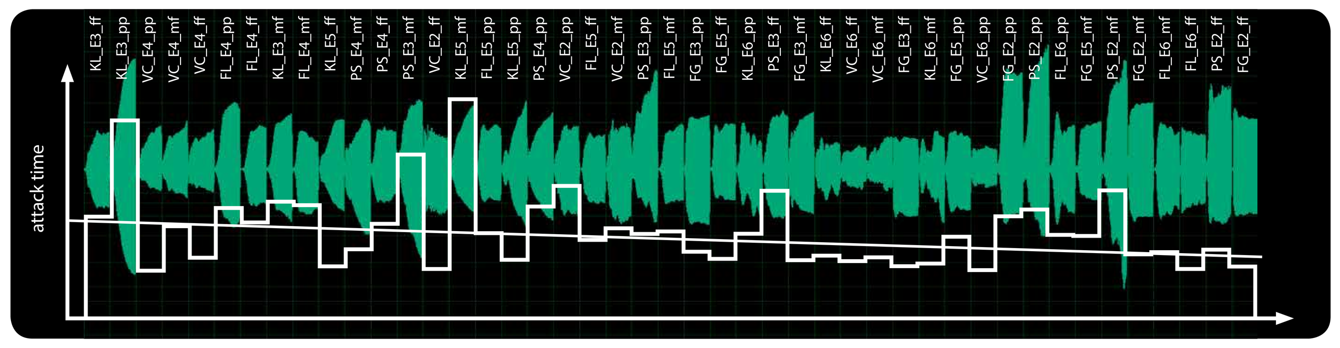


Dynamics hidden in Dimension 2
Within one note (e.g. FL E4), sounds are ordered by dynamic level in almost every case. The cello is the odd exception. Avg. correlations of dim2 and dynamics: **cello r ø = .38 | rest r ø = .88**

Dimension 3: temporal structure?

Dimension 3 correlates weakly and negatively but significantly with the attack time.

attack time: r = −.32 (p < .05)

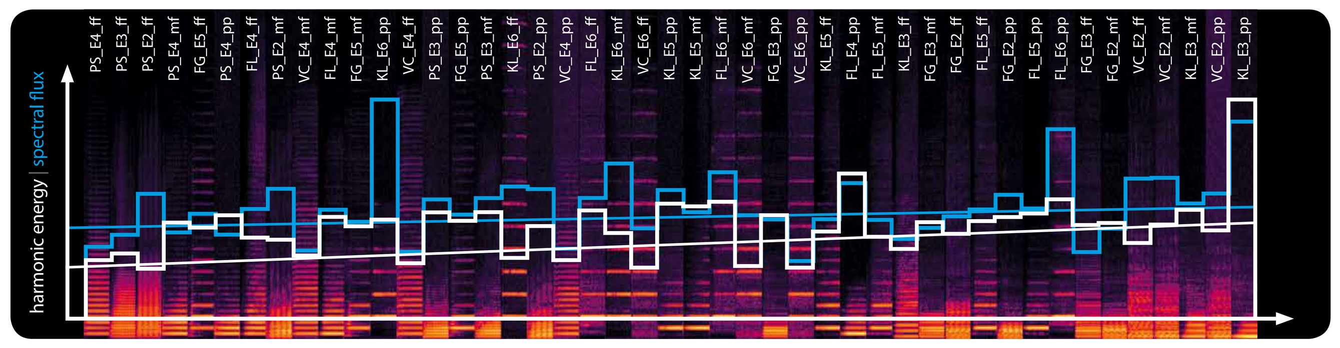


The graph shows the waveform of all sounds along dimension 3. The waveform was chosen because it's better suited to visualize temporal structures. Superimposed is a histogram of the attack time. Tendency is shown via regression line.

Please note: The attack time is only measured until the first peak. Thus, any irregularities force the algorithm to return lower values than there should be. With that in mind, it can be seen from the waveforms that correlations should be even higher.

Dimension 4: noise components?

Dimension 4 shows only weak and insignificant correlations. However, examining the spectrograms as well as listening to the sounds indicates an increasing dominance of noise components along dim4.



The graph shows a spectrogram of all sounds along dimension 4. The superimposed plots show histograms of the strongest correlating features along with the respective regression lines. Spectral fluctuations (cyan) increase as well as the harmonig energy (white).

Increasing fluctuation falls in line with bowing or blowing noises but an increase in harmonic energy is odd. But (1.) strong harmonics, of course, do not require low noise components and (2.) any horizontal line in the spectrum might be interpreted as a harmonic partial by the algorithm.

Takeaways

Conclusions

Although three pitch and dynamic levels are still a rather weak resolution, the inclusion of pitch and dynamics as such already provides a much clearer view on the timbre of musical instruments in the following aspects:

- the timbre of an instrument varies with pitch and dynamics changes
- timbral spectra of musical instruments are quite heterogeneous
- properties like pitch may override similarities based on excitation mechanisms

This study provides empirical evidence

- of Carl Stumpf's concept of timbre in a narrow sense (spectrum) and timbre in a broader sense (noise and temporal structure) [4].
- that the entity „musical instrument" might not be a psychoacoustical but rather a psychological phenomenon.

References

[1] American National Standards Institute (1960). American Standard. Acoustical Terminology (Including Mechanical Shock and Vibration). New York, NY: ASA

[2] Grey, J. M. (1975). *An exploration of musical timbre using computer based techniques for analysis, synthesis and perceptual scaling*. PhD dissertation, University of Stanford.

[3] Handel, S. & Erickson, M. L. (2004). Sound Source Identification: The Possible Role of Timbre Transformations, Music Perception 21(4), 587–610.

[4] Stumpf, C. (1890). Tonpsychologie [Tone Psychology] (Vol. 2). Stuttgart: Hirzel.

Acknowledgements

Supported by funds of the *Österreichische Nationalbank* (OeNB, Anniversary Fund, project number: 16473). Special thanks to the *Vienna Symphonic Library* for supporting our work with two free "Vienna Super Package" licenses.