Orchestral Manoeuvres in the Light

Formants and their Movements in a Dynamic Formant Map

Christoph Reuter[1], Saleh Siddiq[1], Isabella Czedik-Eysenberg[1,3], Daniel Muzzulini[2], Michael Oehler[3]

[1] Musicological Department, University of Vienna
[2] Zürcher Hochschule der Künste, Zurich
[3] Musicological Department, University of Osnabruck

See Timbre Spaces and Formant Maps at https://muwiserver.synology.me/dynamic/timbremaps_e.htm

21st Quinquennial
IMS Congress, Athens
22.-26. August 2022

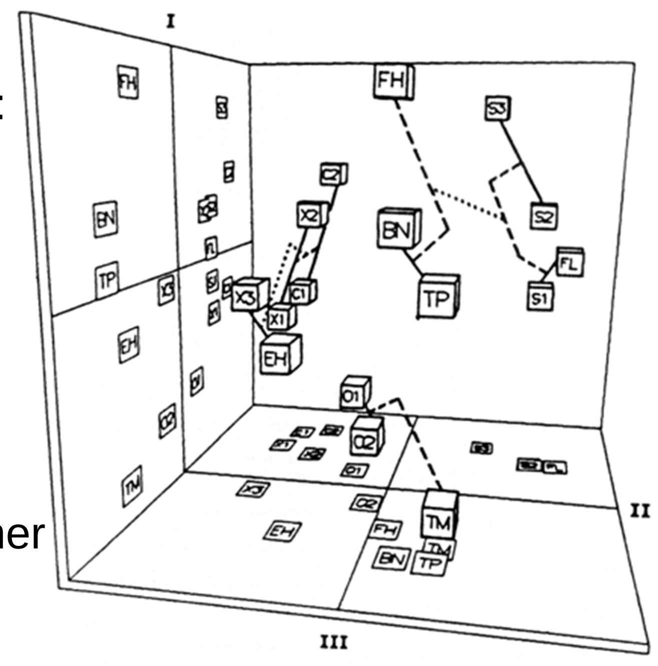
Popular Methods of Timbre Description

In the last 50 years, three types of **timbre description** in particular have become established:

Timbre Spaces

First known **timbre space** in which the **similarities of timbres** are plotted along three spatial dimensions: (since Grey 1975)

- **I:** Spectral Energy Distribution
- **II:** Attack transients and the synchronicity of higher partials there
- **III:** Fluctuations and inharmonicity



Timbre Space (Grey 1975, S. 62)

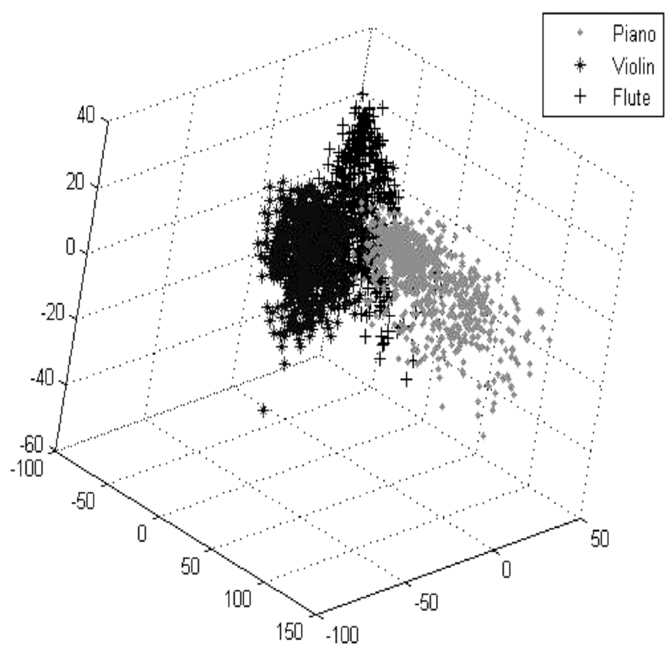
However, a comparison of the three most popular timbre spaces in one **Meta Timbre Space** showed that timbres of the **same stimuli set** were **more similar** to each other than timbres of the same instrument:
Timbre Spaces are only conditionally suitable for describing timbre, since they can **neither be compared with each other nor generalized**. (Siddiq et al. 2018).

Mel Frequency Cepstral Coefficients (MFCC)

Initially a computational method **for automatic speaker recognition** or **speech similarity estimation** (since Davis & Mermelstein 1980)

- **Pro:** well suited for calculating **similarities** in speech/music/instrumental timbres
- **Cons:** not very intuitive, numerical output is **difficult to interpret**.

MFCCs are **standard features** when it comes to calculating timbral similarity.



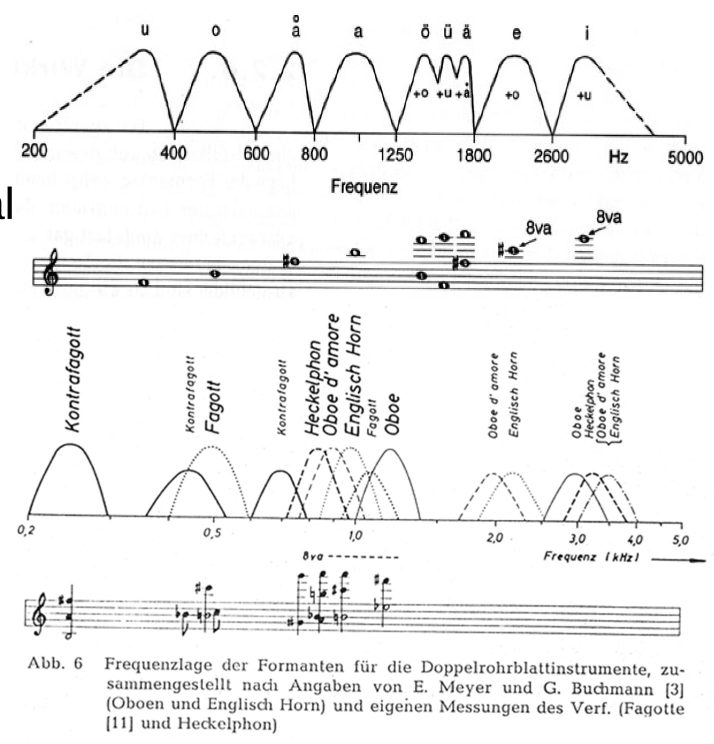
Automatische Klangfarbentrennung im MFCC-Space (Loughran et al. 2008, p. 3)

Formants

Pitch-independent stable maxima in the spectrum of instrument timbres, which - like vocal formants - have a **characteristically timbre-defining** effect.
(since Schumann 1929)

In the German-speaking world, formants are often used to describe the **timbre of musical instruments** based on their **vowel character**.

Similar **formant ranges** lead to similar **timbre sensations**. Instruments with similar formant ranges blend better sonically than instruments with different formant ranges.



Vocal formants and their frequency ranges (top)
Formants of double-reed instruments (bottom)(Meyer 2015, p. 33 and 63)

Formant Analysis in Common Signal Analysis Libraries?

Why are there almost no formant analysis functions in libraries for the signal analysis of musical sounds?

Algorithms for formants **are completely missing** in the common libraries for musical signal analysis (e.g. in **MIRToolbox** (Lartillot, Toiviainen 2007), **Timbre Toolbox** (Peeters et al. 2011), **Essentia** (Bogdanov et al. 2013), **Yaafe** (2013), **Meyda** (Rawlinson 2014), **LibROSA** (McFee et al. 2015), **JS-Xtract** (Jillings 2016), **Aubio** (2017), **MiningSuite** (Lartillot 2019) etc.)

Possible **reasons**:

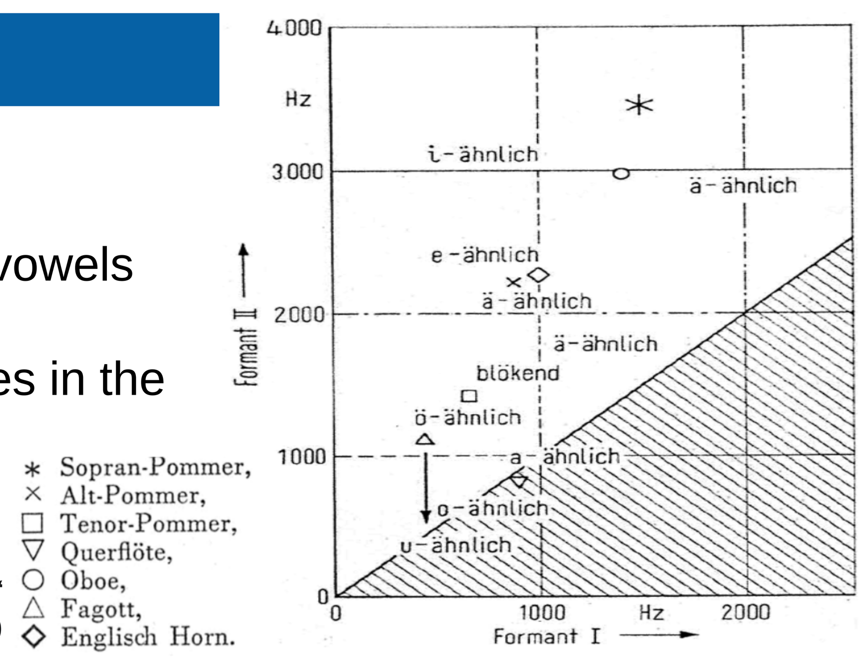
- The research of Stumpf, Schumann etc. **has not been translated** into English, the ideas of these authors were **not present** in English-language literature.
- Formants were discovered "**too early**", long before there was any idea of predictable timbre features at all.
- Formant estimation via LPC is **strongly dependent** on the algorithms default settings.
- Formants tend to be attributed to phonetics ("vowels"), not to musical acoustics.

Current algorithms for **formant calculation/estimation** could easily be taken from already existing **phonetics libraries** (a.g. **Colea** (Loizou 1998/9), **Praat** (Boersma, Weenink, ab 2001), **Parselmouth** (Jadoul et al. 2018), **Formant Tracker** (Kamath 2021), **Formant Estimation** (Rabiner et al. 2021))

Formant Maps

1974: First Formant Map

Matching **double-reed instruments** to vowels based on their first two formant ranges, determined from the strongest amplitudes in the spectrum in each case.

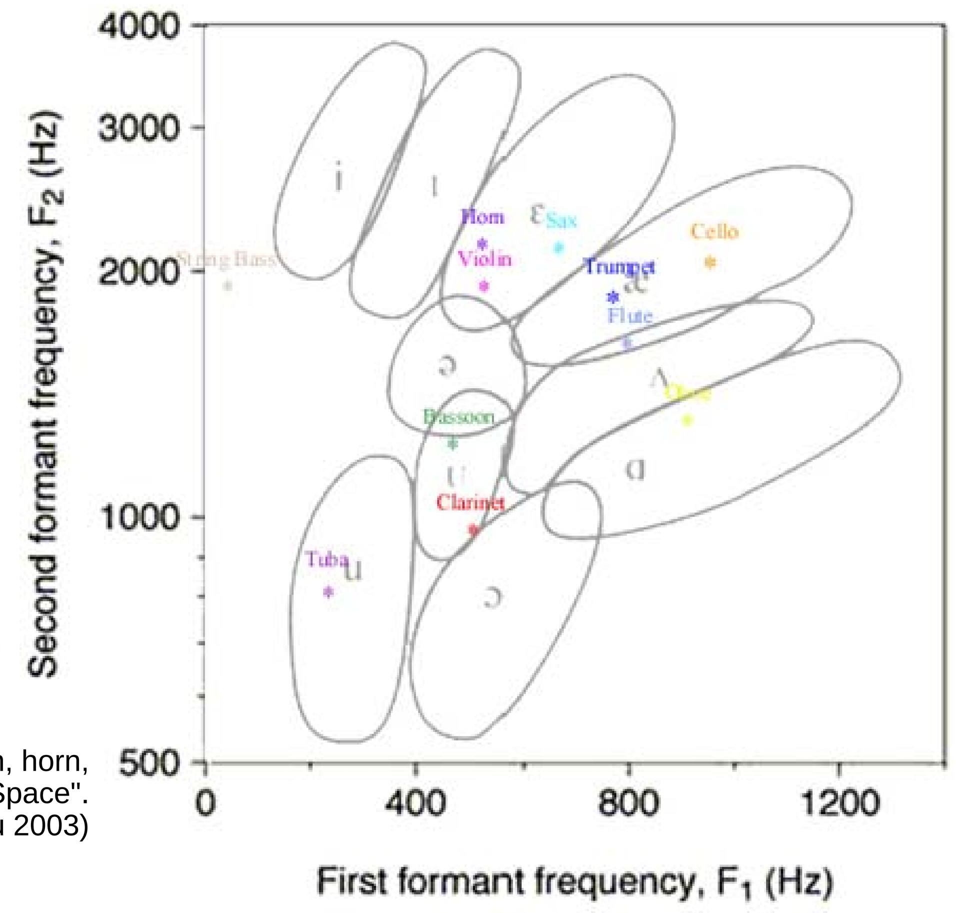


„ Perceptual space for timbres with two-formants" (Sirker 1974, p. 52)

Formant Maps

2003: Vowel Space: Arrangement of the timbres of common orchestral instruments on a (per instrument differing) pitch in a **vocal formant map** by John McCarty (calculated via Colea).

"Now I can say that the tuba's timbre has a 'U' sound or (oo) as in the word who, or the trumpet has an 'ae' timbre like the work actor."
(McCarty, Stanford.edu 2003)

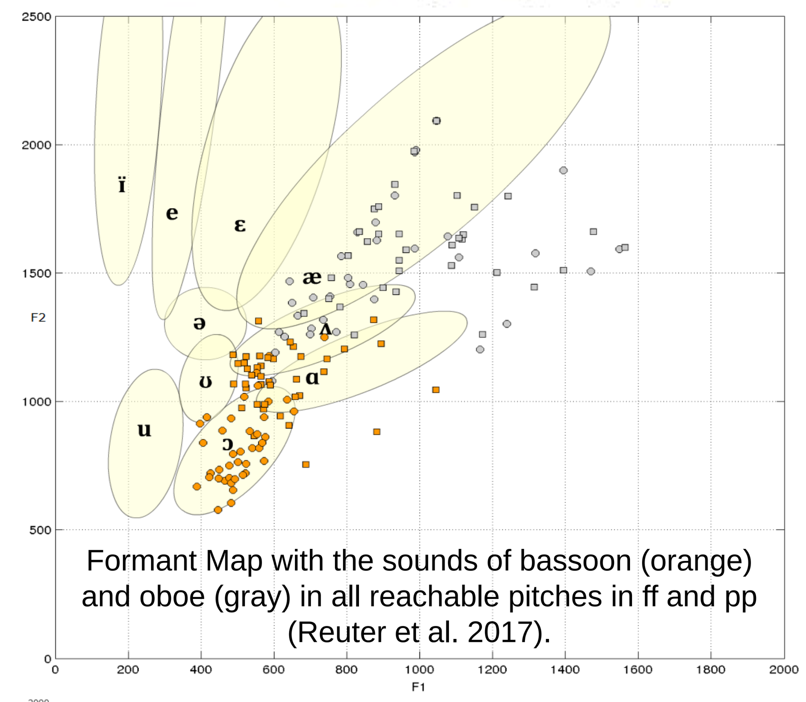


Sounds of saxophone, flute, oboe, clarinet, bassoon, horn, trumpet, tuba, violin, cello, bass in the "Vowel Space". (McCarty, Stanford.edu 2003)

2016: Interactive Formant Map for the common woodwind and brass instruments from the VSL library in **all attainable pitches** and **two dynamic levels** (ff and pp) by Reuter et al. (593 single sounds).

Mean values (circle center) and **standard deviation** (diameter of ellipses) can be displayed per instrument and register.

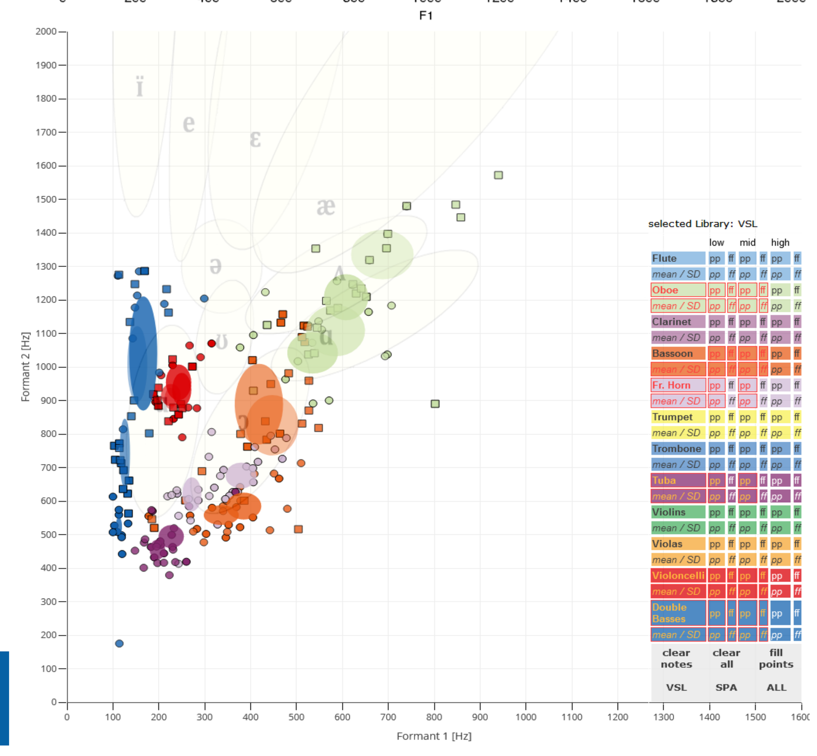
Overlapping areas sound **similar**, while **widely separated** areas sound **dissimilar.**.



Formant Map with the sounds of bassoon (orange) and oboe (gray) in all reachable pitches in ff and pp (Reuter et al. 2017).

2020: Update of the Formant Map:

- Switch from Flash to **Javascript**
- **Expansion of instruments** to include violin, viola, cello and bass
- Integration of an **additional orchestra library** (Spitfire Audio)
- **1100 individual sounds** in total



Formants, their mean values and standard deviations of oboe, bassoon, horn, tuba, cello and bass in low and middle registers

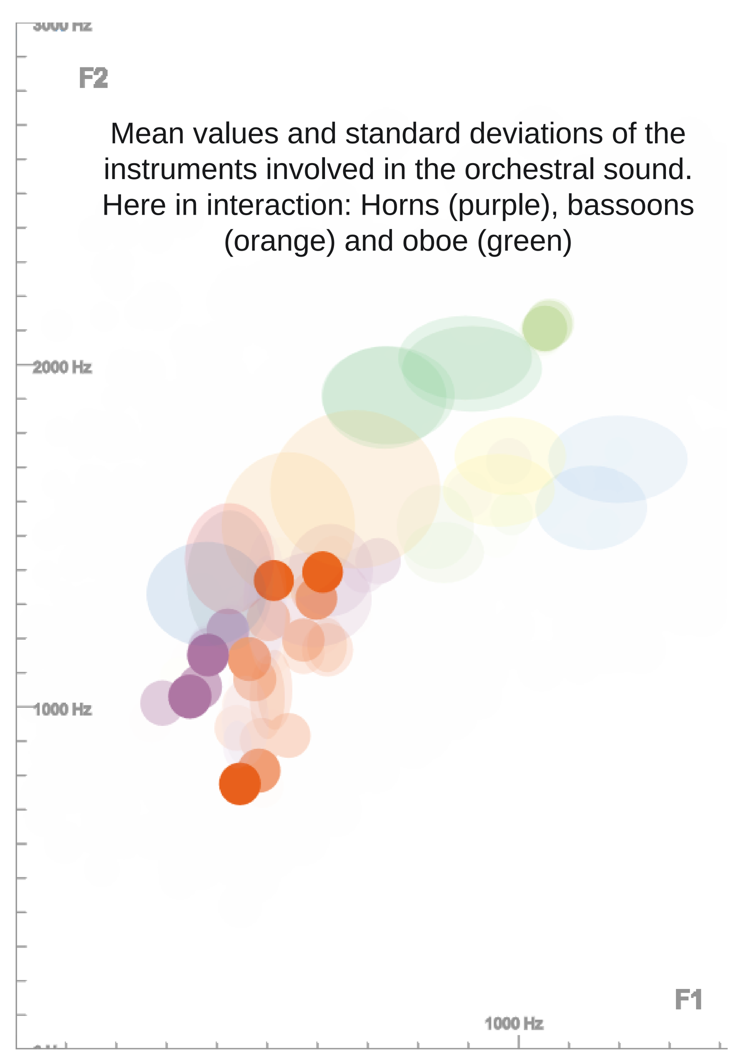
Dynamic Formant Map

Via Parselmouth, the formants of the individual musical instruments in the **1st movement of Beethoven's 7th Symphony** were tracked.
(Recordings of the single tracks from Pätynen et al. 2008).

Via Plotly and P5, the collected values were synchronized with the music and transferred to an **interactive dynamic formant map**, so that the behavior of instrumental formants in "the wild" in specific " territories" or areas can be visualized.

Here, the arrangement of the instruments in typical areas, already known from previous formant maps, becomes apparent.

Similar results can be obtained with **MFCCs**. For comparison, a visualization in **two dimensions of the Timbre Space** was also created.



Mean values and standard deviations of the instruments involved in the orchestral sound. Here in interaction: Horns (purple), bassoons (orange) and oboe (green)

Summary

- Although both **MFCCs** and **formants** for describing timbral properties have been adopted from **phonetics** into **musical acoustics**, so far **only MFCCs** exist as **timbre descriptors** in common signal analysis libraries for music or timbre.
- The **formant extraction algorithms** currently found in Python and Matlab scripts **could easily be adopted** for sound/music feature extraction in the corresponding **signal analysis libraries** as well.
- At least since the **1970s**, there have been approaches to **map musical instruments** to the **vowel trapezium** of phonetics or the **formant map** based on it
- With a sufficiently **high number of individual sounds** in all attainable pitches and different dynamic levels, **clear ranges per register** become visible, in which the instrumental timbres are located in the **formant map**.
- The **tracked formants** of the interacting instruments in the formant map **visualize comprehensibly** the **formant behavior** during pitch and dynamic changes as well as **timbre similarities**.