Introduction to Musical Corpus Studies

Release 0.0.1

Fabian C. Moss

Jun 29, 2020

CONTENT:

1 Schedule **3**

2 Credits **5**

3 Background **7**
- 3.1 What are Musical Corpus Studies? . . . 7
- 3.2 Epistemological goals . . . 7
- 3.3 Issues . . . 7
- 3.4 MCS and traditional musicology . . . 7

4 Solos in the Weimar Jazz Database **9**
- 4.1 The Project . . . 9

5 Harmony in Beethoven's String Quartets **11**
- 5.1 Access the data . . . 11
- 5.2 Harmonic Annotations . . . 11
- 5.3 Chord Transitions . . . 11

Warning: These pages are (heavily) under construction!

In the last two decades *Musical Corpus Studies* evolved from a niche discipline into a veritable research area. The growing availability of digital and digitized musical data as well as the application and development of modern methodologies from computer science, machine learning, and data science cast new light on old musicological questions and generate entirely novel approaches to empirical music research.

Moreover, the general methodological and epistemological approach of Musical Corpus Studies allows to transcend traditional intra-musicological boundaries between its sub-disciplintes (historical/systematic/ethnological/...) without sacrificing the respective specific viewpoints and perspectives.

This course offers a fundamental and practical introduction into these topics. It demonstrates, explores, and critically reflects central thematic areas and methods by means of a number of case studies. Among the contents are:

- Beethoven's string quartets
- 19th century piano music
- Popular music charts
- Electronic music 1950-1990
- Brazilian Choro
- Malian percussion music
- Jazz solos

In the engagement with these topics the course also introduces elementary methods from natural language and music processing, as well as statistics, data analysis and visualization.

CHAPTER

ONE

SCHEDULE

The following table outlines the schedule and summarizes the contents of this course.

No.	Date	Time	Topic	Corpus	Methods
1	Fr., 13.11.2020	16:00-17:20 Uhr	Introduction / Background		
2		17:40-19:00 Uhr	Folk Songs, Melodies, Pitches and Intervals	Essen Folk Song Collection	frequencies, mean, variance
3	Sa., 14.11.2020	09:00-10:20 Uhr	Jazz Solos, Melodies	Weimar Jazz Database	Regular Expressions
4		10:40-12:00 Uhr	Beethoven's string quartets Harmony	Annotated Beethoven Corpus	n-grams, Markov models
		12:00-13:00 Uhr	Lunch Break		
5		13:00-14:20 Uhr	Pop Charts Billboard 100, harmony,	McGill Billboard Dataset	Hidden Markov Models
6		14:40-16:00 Uhr	Free group work		
7	Fr., 11.12.2020	10:00-11:20 Uhr	Brazilian Choro, harmony, form,	Choro Songbook Corpus	Context-Free Grammars
8		11:40-13:00 Uhr	19th century piano music, harmony	DCML Piano Corpus	Probabilistic CFGs
9	Sa., 12.12.2020	09:00-10:20 Uhr	Malian Percussion Music, rhythm, meter	Interpersonal Entrainment in Music Performance: Malian Jembe	
10		10:40-12:00 Uhr	Electronic Music 1950-1990	Curated Corpus of Historical Electronic Music	
		12:00-13:00 Uhr	Lunch Break		
11		13:00-14:20 Uhr	Free group work		
12		14:40-16:00 Uhr	Recapitulation and conclusion		

CHAPTER

TWO

CREDITS

Ich gehe in der Seminarplanung von 12 Semesterwochen à 2 SWS aus, für das gesamte Blockseminar also 24 SWS. Das Seminar wird mit 3 CP bewertet, was 90 Stunden aktiver Arbeit entspricht. Davon entfallen 24 SWS an die Präsenzzeit im Seminar plus 48 SWS an Vor- und Nachbereitung der Seminarsitzungen. Die verbleibenden 18 SWS sind für die Lektüre der Fachliteratur vorgesehen.

CHAPTER

THREE

BACKGROUND

3.1 What are Musical Corpus Studies?

tbc…

3.2 Epistemological goals

tbc…

3.3 Issues

tbc

3.4 MCS and traditional musicology

tnc

References

1. Cook (2006). Border Crossings: A Commentary on Henkjan Honing's "On the Growing Role of Observation, Formalization and Experimental Method in Musicology". Empirical Musicology Review 1(1), 7-11.
2. Honing (2006). On the Growing Role of Observation, Formalization and Experimental Method in Musicology. Empirical Musicology Review 1(1), 2-6.
3. Huron (2013). On the Virtuous and the Vexatious in an Age of Big Data. Music Perception: An Interdisciplinary Journal. 31(1), 4-9.
4. Marsden (2016). Music Analysis by Computer: Ontology and Epistemology. In: David Meredith (ed.) Computational Music Analysis. Springer.
5. Neuwirth & Rohrmeier (2016). Wie wissenschaftlich muss Musiktheorie sein? Chancen und Herausforderungen musikalischer Korpusforschung. Zeitschrift der Gesellschaft für Musiktheorie 13(2), 171-193.
6. Pugin (2015). The Challenge of Data in Digital Musicology. Frontiers in Digital Humanities 2(4), 1-3.
7. Schaffer (2016). What is Computational Musicology? https://medium.com/@krisshaffer/what-is-computational-musicology-f25ee0a65102.

8. Temperley & VanHandel (2013). Introduction to the Special Issue on Corpus Methods. Music Perception: An Interdisciplinary Journal. 31(1), 1-3.

CHAPTER

FOUR

SOLOS IN THE WEIMAR JAZZ DATABASE

The first project we will have a look at is the Jazzomat project

4.1 The Project

References

1. Pfleiderer et al. (2017). Inside the Jazzomat: New Perspectives for Jazz Research.

CHAPTER

FIVE

HARMONY IN BEETHOVEN'S STRING QUARTETS

5.1 Access the data

The data lies on the GitHub repository DCMLab/ABC. Either download the .tsv file directly and open it in pandas or load it from the URL as follows:

```
import pandas as pd

df = pd.read_csv("https://github.com/DCMLab/ABC/corpus.tsv", sep="\t")
```

The corpus is now stored in the variable *df*.

5.2 Harmonic Annotations

- regular expressions

5.3 Chord Transitions

- n-grams