

# Introduction to Musical Corpus Studies

---

Fabian C. Moss

13 November 2020

Musikwissenschaftliches Seminar - Universität zu Köln - WS 2020/21



Introduction (16:00–17:20)

- I. What are Musical Corpus Studies?
- II. Issues
- III. Examples
- IV. Organization of the course
- V. Questions

— *Break* —

Melody I (17:40–19:00)

# I. What are Musical Corpus Studies

---

## II. Issues

---

### III. Examples

---

Harasim, D., Moss, F. C., Ramirez, M., & Rohrmeier, M. (in press). Exploring the foundations of tonality: Statistical cognitive modeling of modes in the history of Western classical music. *Humanities & Social Sciences Communications*

# Research questions

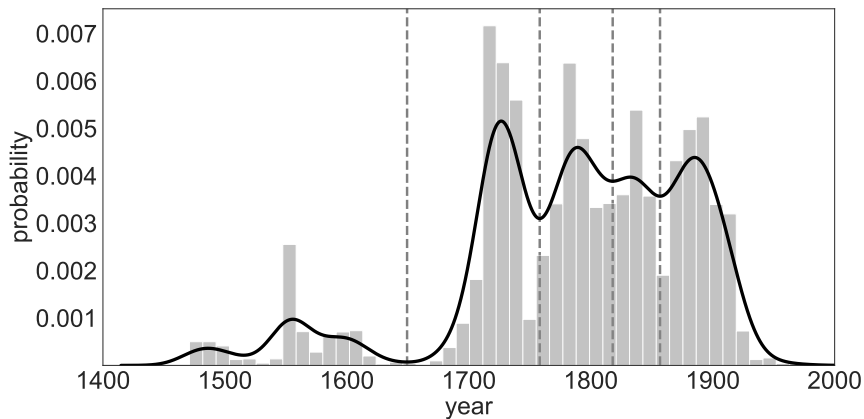
1. How can we find modes automatically?
2. How can the concept of a mode be operationalized?
3. Can we do it without knowing (unsupervised) how many modes there are and what they look like?
4. How do modes change historically?

- 21'000 pieces from <https://classicalarchives.com>
- MIDI format
- user-generated (quality?)
- biases
- metadata: composer names, keys, composition date, ...
- representativeness?
- almost no early music examples → add from other projects
  1. *Citations: The Renaissance Imitation Mass Project* (CRIM)
  2. *The Lost Voices Project*



- 21'000 pieces from <https://classicalarchives.com>
- MIDI format
- user-generated (quality?)
- biases
- metadata: composer names, keys, composition date, ...
- representativeness?
- almost no early music examples → add from other projects
  1. *Citations: The Renaissance Imitation Mass Project* (CRIM)
  2. *The Lost Voices Project*

⇒ in total 13'402 pieces (ca. 55 million notes) with given composition year (but not key)



**Figure 1:** Historical distribution of pieces in the corpus.

# Assumptions

1. pieces can be represented by pitch-class counts
2. enharmonic equivalence
3. transpositional invariance

# Assumptions

1. pieces can be represented by pitch-class counts
2. enharmonic equivalence
3. transpositional invariance

All of these assumptions are highly questionable, especially on a large historical scale!

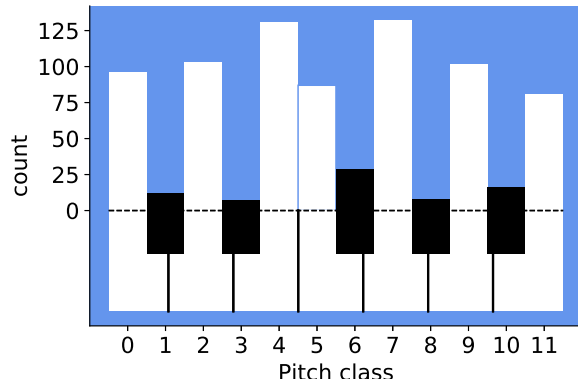
# Assumptions

1. pieces can be represented by pitch-class counts
2. enharmonic equivalence
3. transpositional invariance

All of these assumptions are highly questionable, especially on a large historical scale!

⇒ explicit modeling

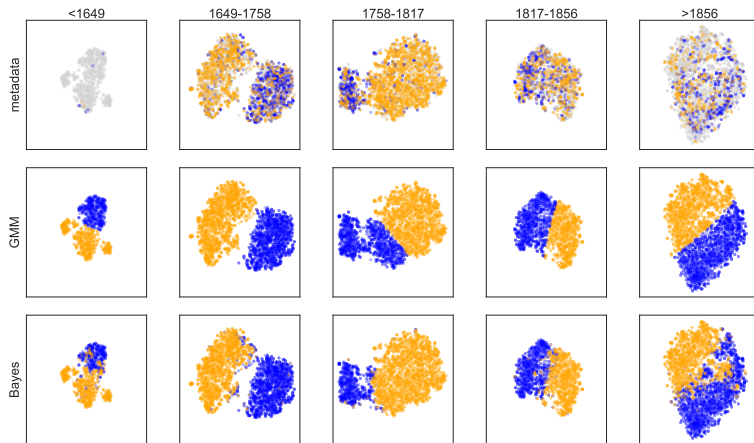
## An example



**Figure 2:** Pitch-class counts of an example piece in C major.

# The model

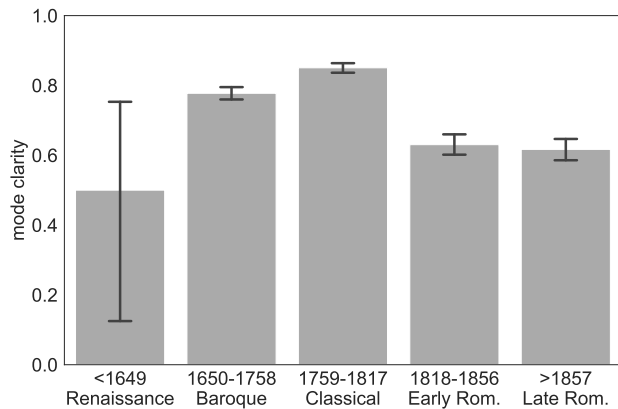
# Automatically finding modes



**Figure 3:** Three models for automatic mode finding.



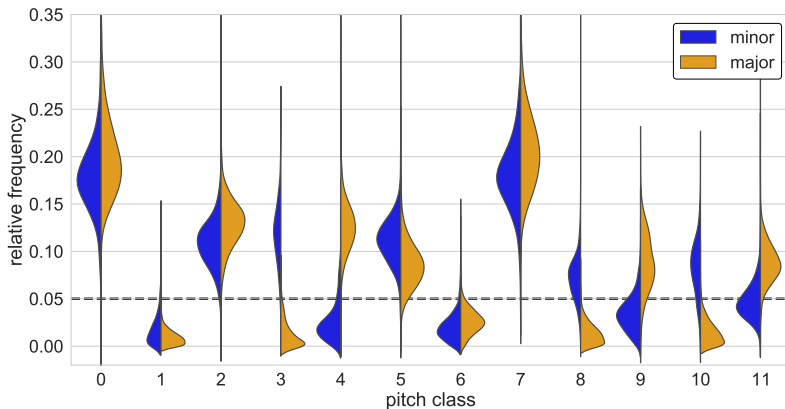
## Quality of the model



**Figure 4:** Accuracy scores of our model in five historical periods.

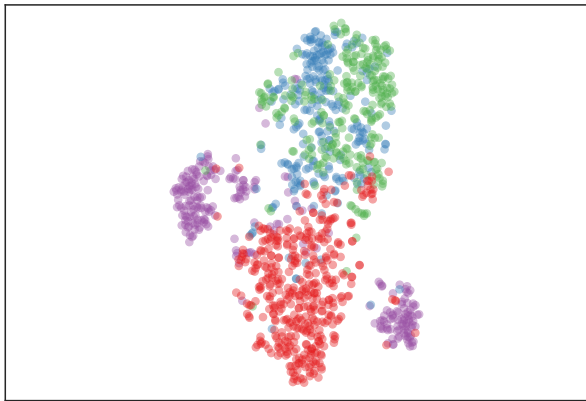
# The major and minor modes

Pitch-class distributions of all pieces in the Baroque and Classical periods:



**Figure 5:** Pitch-class distribution of the major and minor modes.

## Modes in the Renaissance



**Figure 6:** Clustering into four modes in the Renaissance.

# Modes in the Renaissance

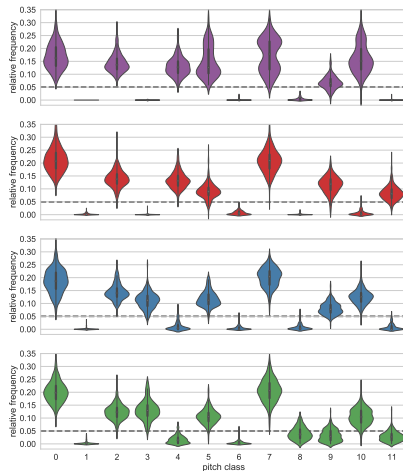


Figure 7: Pitch-class distribution of Renaissance modes.

1. ...

## IV. Course organization

---

- main organization via ILIAS
- literature
- forum
- Zoom link (you are all here)
- external website: <https://fabianmoss.github.io/intro-corpusmus>
  - general info
  - course materials (updated after each session)
- HfMT students: by your group members

- 3 CPs = 90 SWS
  - 24 SWS presence in seminar
  - 24 SWS preparation of and follow-up on course materials
  - 42 SWS reading of literature and writing of report



## Group work

- you will meet with your group in the breakout rooms
- discussions
- exercises

- report due on **31 January 2021, 23:59h**
- 6–8 pages
- suggested structure
  1. Introduction
  2. Discussion
  3. Issues
  4. Various
  5. Contributions

## References

---

Harasim, D., Moss, F. C., Ramirez, M., & Rohrmeier, M. (in press). Exploring the foundations of tonality: Statistical cognitive modeling of modes in the history of Western classical music. *Humanities & Social Sciences Communications*.