

## Content

**Aim:** Introducing available resources and demonstrating computational analysis tasks for non-Western music traditions.

### Plan:

- Summary of available resources from CompMusic project
- Set up for the session
- Intonation analysis task (case study, Turkish music)
- Rhythm analysis task (case study, Indian music)
- Score analysis task (case study, Turkish music)



Computational approaches for analysis of non-Western music traditions, Serra, Clayton, Bozkurt

1

**Computational models**  
for the discovery of the World's Music

[HOME](#) [DESCRIPTION](#) [TEAM](#) [PUBLICATIONS](#) [CORPORA](#) [SOFTWARE](#) [EVENTS](#) [BLOG](#) [NEWS](#) [RESOURCES](#) [GET INVOLVED](#)

**LATEST NEWS**

[Article by Xavier Serra in Acta Musicologica](#)

In order to reach the musicological community and explain the work done... Georgi Dzhambava defends his PhD thesis

On June 28th 2017, Georgi Dzhambava

[dunya.compmusic.upf.edu](#)

[Info](#) [Legal](#) [Xavier](#) [Logout](#)

**LATEST BLOGS**

[Final report submitted](#)

CompMusic has finished, and our funding agency, ERC, asked us to write a brief report. Here it is. Achievements along the way have been many. The CompMusic project has been a big and long project with many achievements.

[Technology and Multiculturalism](#)

17/04/2018

[Article published in the daily newspaper La Vanguardia on Sunday 17th 2018. English translation of the original text written in catalan.] The article is about the CompMusic project and the developed software tools allow to process the available information in order to study and explore the characteristics of each musical repertoires.

[Explore our collections](#)

[Carnatic](#) [Hindustani](#) [Makam](#) [Jingju](#) [Andalusian](#)

[compmusic.upf.edu](#)

**ERC-Advanced Grant (2011-2017) PI: Xavier Serra** [English](#) [ES](#)

CompMusic is a research project funded by the European Research Council from 2011 to 2017 and coordinated by Xavier Serra from the Music Technology Group of the Universitat Pompeu Fabra in Barcelona, Spain. The project aims to develop methods in the field of music information processing with a domain knowledge approach. The project focuses on five music traditions of the world: Hindustani (North India), Carnatic (South India), Turkish-makam (Turkey), Arab-Andaluzian (Maghreb), and Beijing Opera (China).

[dunya.compmusic.upf.edu](#)

[Info](#) [Legal](#) [Xavier](#) [Logout](#)

**LATEST BLOGS**

[Final report submitted](#)

CompMusic has finished, and our funding agency, ERC, asked us to write a brief report. Here it is. Achievements along the way have been many. The CompMusic project has been a big and long project with many achievements.

[Technology and Multiculturalism](#)

17/04/2018

[Article published in the daily newspaper La Vanguardia on Sunday 17th 2018. English translation of the original text written in catalan.] The article is about the CompMusic project and the developed software tools allow to process the available information in order to study and explore the characteristics of each musical repertoires.

[Explore our collections](#)

[Carnatic](#) [Hindustani](#) [Makam](#) [Jingju](#) [Andalusian](#)

[compmusic.upf.edu](#)

Arab-Andalusian

Turkish-makam

Hindustani

Carnatic

Beijing Opera

Computational approaches for analysis of non-Western music traditions, Serra, Clayton, Bozkurt

3

**CompMusic tutorials, talks, concerts**

CompMusic

1,350 subscribers

[SUBSCRIBE 1.3K](#)

[HOME](#) [VIDEOS](#) [PLAYLISTS](#) [CHANNELS](#) [ABOUT](#)

[Tutorial on Beijing Opera \(JL...](#)

[Nils Sargatz - An evening of...](#)

[CORMA talk](#)

[Ottaped Music LeoDem and...](#)

[CompMusic Seminar, Dec 2...](#)

<https://www.youtube.com/user/CompMusicProject>

Computational approaches for analysis of non-Western music traditions, Serra, Clayton, Bozkurt

4

**CompMusic summary in numbers**

- Corpora of five music traditions:  
**[1400 hours of audio with metadata on MusicBrainz]**
- **24 datasets** developed for specific experiments around these traditions (<https://zenodo.org/communities/mtgupf>)
- ~150 publications, 15 thesis [[open code and data](#)]
- Summary report:  
<http://compmusic.upf.edu/node/350>

5

**Corpora and datasets**

Music repertoire

Research corpus

test datasets

**X. Serra "The computational study of a musical culture through its digital traces".** Acta Musicologica. 2017; 89(1):24-44.

6

**Dunya Server**

- Audio
- Metadata
- Features

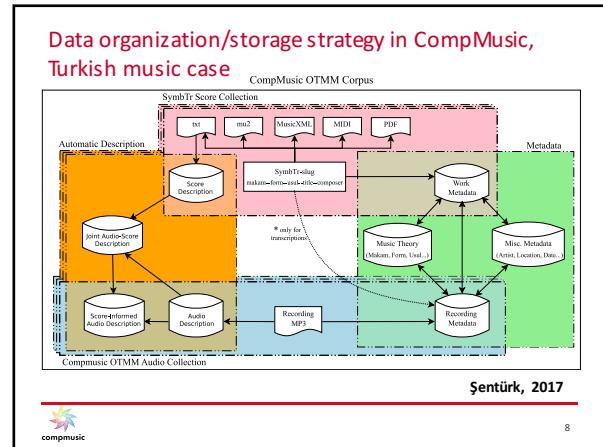
**pycompmusic**

Python tools for accessing audio, annotations and features.

<https://github.com/MTG/pycompmusic>

Features extracted using culture-specific tools designed by MTG researchers

7



**Metadata:** MusicBrainz, <https://musicbrainz.org>

**Tag "makam: hicaz"**

**Recordings**

- Asman Aşılın Hümüslüne Gölge by Hafız Kemal Bey
- Fatih Suresi by Kâni Karabey
- Gâh-i İlahî Pergen Ne Rəsmik Kademe by Gül Göre
- Gördür Yüzünü Gölberimin Nuru Kararı by Hafız Kemal Bey
- Hicaz Mâlik by Hacı Ağa Gündüz Kutbay & Nerdet Yıldız
- Hicaz Peygamber Ağa Gündüz Kutbay & Nerdet Yıldız
- Hicaz Salâti by Kâni Karabey
- Hicaz Sâz by Hacı Ağa Gündüz Kutbay
- Hicaz Sâz Sensesi by Derya Türkân & Özay Özal
- See all 729 recordings

**Works**

- Bahar Zamanı Geddi
- Hicaz Mandra
- Hicaz Sâz
- Hicaz Sûrâ
- Hicaz Sûrâ
- Kâhi-i mahânum nâsi etmez fîgân
- Kâhi-i mahânum gûzâzelîn ve defâne
- Senedde Bir Gün

**Crowdsourcing metadata collection**

**(!! No audio out there)**

9

**CompMusic collections on MusicBrainz:**

<https://musicbrainz.org/user/compmusic/collections>

Collection	Type	Releases
Assalâna	Release	24
Beijing Opera	Release	48
Bootleg Carnatic	Release	76
Carnatic	Release	326
Chingâza	Release	23
Dunya Beijing Opera	Release	91
Dunya Carnatic	Release	246
Dunya Carnatic CC	Release	20
Dunya Hindustani	Release	234
Dunya Hindustani CC	Release	36
Dunya Ottoman-Turkish makam	Release	296
Hindustani	Release	235
Ottoman-Turkish makam	Release	351
Ottoman-Turkish makam excluded	Release	55

10

**Accessing datasets:** <http://compmusic.upf.edu/datasets>

**INDIAN ART MUSIC**

**INDIAN MUSIC TONIC DATASET**

This dataset comprises 597 commercially available audio music recordings of Indian art music (Hindustani and Carnatic music), each manually annotated with the tonic of the lead artist. This dataset is used as the test corpus for the development of tonic identification approaches. <http://compmusic.upf.edu/iam-tonic-dataset>

**CARNATIC VARNAM DATASET**

Carnatic varnam dataset is a collection of 28 solo vocal recordings, recorded for our research on intonation analysis of Carnatic ragas. The collection consists of audio recordings, time aligned tala cycle annotations and swara notations in a machine readable format. <http://compmusic.upf.edu/carnatic-varnam-dataset>

**CARNATIC MUSIC RHYTHM DATASET**

The Carnatic Music Rhythm Dataset is a sub-collection of 176 excerpts (16.6 hours) in four taales of Carnatic music with audio, associated tala related metadata and time aligned markers indicating the progression through the tala cycles. It is useful as a test corpus for many automatic rhythm analysis tasks in Carnatic music. A subset with 118 two minute long excerpts (about 4 hours) is also available with equivalent content. <http://compmusic.upf.edu/carnatic-rhythm-dataset>

...

11

**Software tools**

Many software tools have been developed as part of CompMusic that are freely available:

- Dunya browser** (web browser to explore the Dunya corpora)
- Dunya desktop** (modular and extensible (open source) desktop application to explore the Dunya corpora)
- Sarâqa** (android app to explore and listen a collection of Carnatic and Hindustani music)
- PyCompMusic** (tools to help researchers work with Dunya corpora)
- ...

**Examples for the thesis-work repositories:**

<https://sertansenturk.com/work-research/phd-thesis/>  
<http://compmusic.upf.edu/node/304>  
<http://compmusic.upf.edu/phd-thesis-ajay>  
<http://compmusic.upf.edu/node/333>

12

## Dunya browser



Dunya comprises the music corpora and related software tools that have been developed as part of the CompMusic project. These corpora have been created with the aim of studying particular music traditions and they include audio recordings plus complementary information that describes the recordings. Each corpus has specific characteristics and the developed software tools allow to process the available information in order to study and explore the characteristics of each musical repertoire.

### Explore our collections



compmusic

13

14

## Dunya browser: Turkish music example

<https://www.youtube.com/watch?v=FtY2mZRKeEc> Şentürk, 2017

compmusic

15

## Dunya desktop <https://github.com/MTG/dunya-desktop>



Uses pyCompMusic to access data

compmusic

16

## SARAGA: The App and the data

[https://www.youtube.com/watch?v=z\\_b99xgmwnw](https://www.youtube.com/watch?v=z_b99xgmwnw)

Data: <http://doi.org/10.5281/zenodo.1256127>

compmusic

17

## Content

- Summary of available resources from CompMusic project
- **Set up for the applied session**
- Accessing data using pycompmusic
- Intonation analysis task (case study, Turkish music)
- Rhythm analysis task (case study, Indian music)
- Score analysis task (case study, Turkish music)

**Set up**



```
$ git clone https://github.com/MTG/Ismir2018TutorialNotebooks.git
$ cd Ismir2018TutorialNotebooks/
$ docker-compose up
    ! Mac users should launch Docker before this command
Access http://localhost:8888 using a browser; password: mir
```

Docker image: <https://github.com/MTG/MIR-toolbox-docker>

Contains:

- Common data-science packages: numpy, scikit-learn, pandas, etc.
- MIR tools: mir-eval, music21
- MTG tools: Essentia, freesound-python, pycompmusic
- Jupyter server

compmusic

19

**Content**

- Summary of available resources from CompMusic project
- Set up for the applied session
- Accessing data using pycompmusic**
- Intonation analysis task (case study, Turkish music)
- Rhythm analysis task (case study, Indian music)
- Score analysis task (case study, Turkish music)

Computational approaches for analysis of non-Western music traditions, Serra, Clayton, Bozkurt  
compmusic

20

**Accessing Dunya corpora using the API**

**Demonstration:** [DownloadDataFromDunya\\_no Token.ipynb](#)

This notebook demonstrates downloading data using pycompmusic API.

- downloading a single file using recording's MusicBrainz ID
- downloading files of a CompMusic dataset ([https://github.com/MTG/otmm\\_makam\\_recognition\\_dataset](https://github.com/MTG/otmm_makam_recognition_dataset))

Music Brainz ID for a recording is simply the code you can retrieve from a MusicBrainz link to a recording.

For example:  
the recording <https://musicbrainz.org/recording/e666ec52-b752-492d-9423-24e1c7bffb7>  
has a Music Brainz ID : 'e666ec52-b752-492d-9423-24e1c7bffb7'

compmusic

21

**Accessing features, metadata, annotations****Demonstration:** [downloadAllSARAGAContent.ipynb](#)

This notebook demonstrates the use of Dunya api for downloading Saraga dataset files which includes manual annotation files accompanying recordings with Creative Commons licence(in mp3 format).

Saraga dataset is also available for direct download on : <https://doi.org/10.5281/zenodo.1256126>

Saraga dataset is composed of two collections:

[Hindustani collection](#)

[Carnatic collection](#)

The notebook creates two subfolders and saves all data in these folders. Each annotation is saved in a separate text file.

compmusic

22

**Creating subsets of data for mode recognition tests**

**Demonstration:** [generateFileLists4Collections.ipynb](#), [formExpSubsets4ModeRecognition.ipynb](#)

This notebook targets accessing Dunya data and collecting file lists for several collections.

The main aim is to create subsets of data for mode and rhythm mode recognition experiments and the process is split into two notebooks. In this first notebook, only the list of files available in meta data is created and saved in a pickle file. For each recording the following information is included:

- Files available for that recording, MusicBrainz id (mbid), Mode information (raga, makam, etc), Rhythm mode information (tala, usul, etc)

Then, the second notebook (formExpSubsets4ModeRecognition.ipynb) reads this file and forms the subsets by grouping recordings with respect to mode or rhythm mode while also checking available files (ex: tonic annotation) for the recording

compmusic

23

**Available mode recognition tools**

The subsets created in the previous step (the *annotations.json* file) can be directly used in running the mode recognition experiments available below.

Using chroma features:

[https://github.com/emirdemirel/Supervised\\_Mode\\_Recognition](https://github.com/emirdemirel/Supervised_Mode_Recognition)

Using pitch distributions:

<https://github.com/altugkarakurt/morty>

<https://github.com/furkanyesiler/pitchdistamr>

compmusic

24

## Content

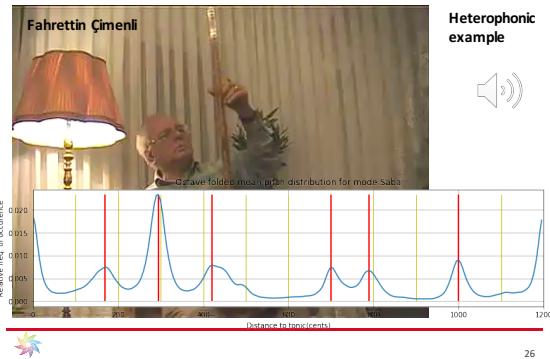
- Summary of available resources from CompMusic project
- Set up for the applied session
- Accessing data using pycompmusic
- **Intonation analysis task (case study, Turkish music)**
- Rhythm analysis task (case study, Indian music)
- Score analysis task (case study, Turkish music)



*Computational approaches for analysis of non-Western music traditions, Sema, Clayton, Bozkurt*

25

## Microtonality in Turkish music



26

## Tuning analysis, Turkish makam music

**Demonstration:** `tuningAnalysis_SingleRecording.ipynb`  
`tuningAnalysis_SetOfRecordings.ipynb`

### Tuning analysis for a single recording

Here, we demonstrate tuning analysis of a single recording in makam Huseyni (a commonly used makam with a scale close to western A dorian with a characteristic second tone lower in pitch resulting in two microtonal intervals around it). The following steps are carried:

- We first download a specific recording: <https://musicbrainz.org/recording/9d195cbe-8ef9-47ef-83f0-e7c0b8af1697>, Dunya direct-link: <http://dunya.com/music.uof.edu/makam/recording/9d195cbe-8ef9-47ef-83f0-e7c0b8af1697>. It includes automatically aligned scores. You can first listen to the recording through Dunya.
- Then pitch analysis and pitch distribution computation is carried out.
- Tonic intonation is accessed from this and further interval distributions are computed (with respect to tonic).
- Octave folding and automatic peak picking applied to detect scale degree's' distance to the tonic.
- Interval list stored in the Scala format (<http://www.nuygens-lotker.org/scala/>) to be able to sonify the intervals using computer keyboard.



27

## Content

- Summary of available resources from CompMusic project
- Set up for the applied session
- Accessing data using pycompmusic
- Intonation analysis task (case study, Turkish music)
- **Rhythm analysis task (case study, Indian music)**
- Score analysis task (case study, Turkish music)



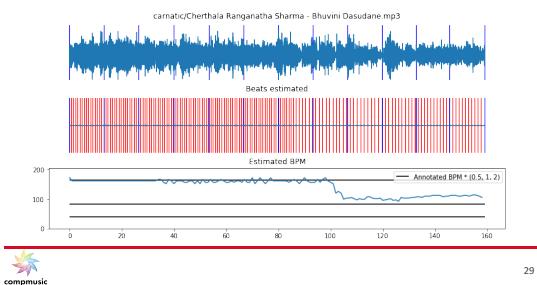
*Computational approaches for analysis of non-Western music traditions, Sema, Clayton, Bozkurt*

28

## Rhythm analysis task (case study, Indian music)

**Demonstration:** `visualizeAnnotations.ipynb`

This notebook aims at demonstrating some of the contents of Saraga database annotations: sections, typical phrases, sama and tempo annotations.



29

## Content

- Summary of available resources from CompMusic project
- Set up for the applied session
- Accessing data using pycompmusic
- Intonation analysis task (case study, Turkish music)
- Rhythm analysis task (case study, Indian music)
- **Score analysis task (case study, Turkish music)**



*Computational approaches for analysis of non-Western music traditions, Sema, Clayton, Bozkurt*

30

### Turkish Makam Music Symbolic Data Collection

<https://github.com/MTG/SymbTr/releases/tag/v2.0.0>

**2200 pieces:** 150 distinct makams, about 100 usuls, 50 forms, 865,000 musical notes, 80 hours nominal playback time.

Scores in formats: Midi, musicXml, mu2, pdf

Karaosmanoğlu, M. K. (2012). A Turkish makam music symbolic database for music information retrieval: SymbTr. Proc. Int. Society for Music Information Retrieval (ISMIR)

M. K. Karaosmanoğlu, Barış Bozkurt, Andrie Holzapfel and Nilgün Doğrusöz Dışaçık, "A symbolic dataset of Turkish makam music phrases", Proceedings of FMA, 2014.



31

### SymbTr

#### Gönlümü Düşür Eden Bu Hale Hep

Hicezkar Şarkı

Beste: Sevki Bey (1880 - 25/6/1891)

Gülte: ?

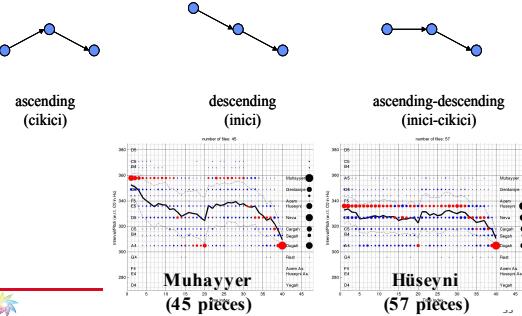


Index	Code	NoteS3	NoteAE	CommasS3	CommasAE	Num	Denum	Ms	LNS	VelOn	Syll
1	9	La4	A4	305	305	1	4	909	95	96	Ben
2	9	Do5#4	C5#4	322	322	1	4	909	95	96	gam
3	9	Si4b4	B4b4	310	310	1	8	455	95	96	li
4	9	Do5#4	C5#4	322	322	1	8	455	95	96	ha
5	9	La4	A4	305	305	3	8	1364	95	96	zan
6	53			0	0	0	0	0	0	0	
7	9	Re5	D5	327	327	1	4	909	95	96	sen
8	9	Do5#4	C5#4	322	322	1	8	455	95	96	se
9	9	Re5	D5	327	327	1	8	455	95	96	ba



32

### 'Seyir'(melodic progression) concept in Turkish makam music



### Content

- Summary of available resources from CompMusic project
- Set up for the applied session
- Accessing data using pycompmusic
- Intonation analysis task (case study, Turkish music)
- Rhythm analysis task (case study, Indian music)
- Score analysis task (case study, Turkish music)

THANK YOU FOR YOUR ATTENTION



Computational approaches for analysis of non-Western music traditions, Sema, Clayton, Bozkurt

35

### Melodic progression

#### Demonstration:

`symbolicDataPro_symbTripynb`

#### Symbolic data processing, Turkish music case

In this notebook, the SymbTr data (Turkish Makam Music Symbolic Data Collection) <http://compmusic.ucl.ac.uk/node/140>, <https://github.com/MTG/SymbTr> is used. Here we demonstrate accessing pieces in a specific form and makam, reading a specific a section of the form and plotting the melodic curves. The makams are chosen to apply the same scale but different melodic progressions ('seyir').

The following steps are carried:

- Downloading SymbTr data from the github repo
- Printing a list of makams and forms sorted with respect to the number of files in those categories
- Plotting melodic curves from the first sections in *saz-semaisi* form in two makams: *rust* and *mahir*.



34