Polyphonic Singing Datasets for MIR Research

27.10.2020, Singing & AI Group

Sebastian Rosenzweig, Helena Cuesta

AUDIO LABS

upf. Music Technology Group

About me

Since 2017:
PhD Student @ Music Technology Group, UPF
Supervisor: Dr. Emilia Gómez

Research interests:

- (Multiple) fundamental frequency estimation
- Analysis of polyphonic vocal music
  - Intonation
  - Performance
  - Unisons
- Machine learning for music signals

Icons made by Freepik from www.flaticon.com

About me

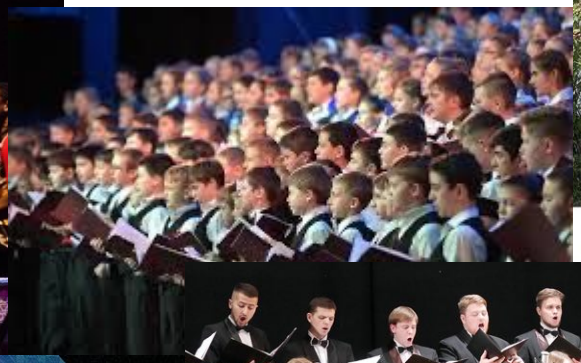
Since 2017:
PhD Student @ International Audio Laboratories Erlangen
Supervisor: Prof. Dr. Meinard Müller

Research interests:

- Analysis of polyphonic vocal music
  - Intonation
  - Performance
- Close-up microphones
- Fundamental Frequency Estimation
- Time-Scale Modification

Icons made by Freepik from www.flaticon.com, Photo by Paul Träger

Polyphonic Singing



Polyphonic Singing

Many facets are yet to be explored and understood:

- Choir tuning
- Singer interaction, pitch adjustment
- Cultural differences

⇒ In order to address these questions using MIR techniques, we need suitable datasets!

Polyphonic Singing Datasets - An Overview

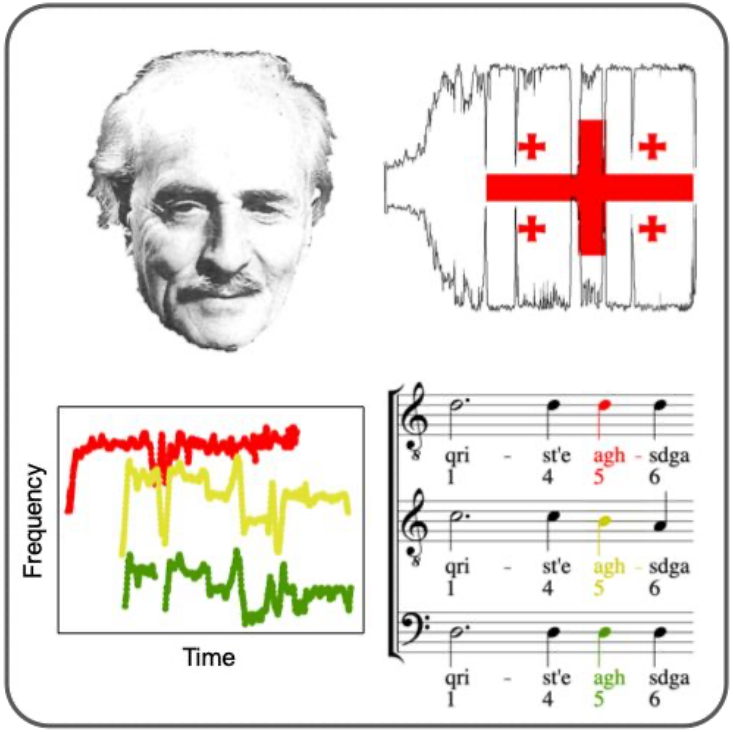
Name/Author	Multitrack	Annotations	Publicly Available	# Recordings	Duration (hh:mm:ss)
Su et al. (2016)	No	MIDI	On Request	5 excerpts	00:02:11
Barbershop Quartets[10]	Yes	MIDI	No	22 songs	00:42:10
Bach Chorales[11]	Yes	MIDI	No	*26* songs	00:58:20
Scherbaum et al. (2019)	Yes	–	On Request	216 songs	06:04:40
Choral Singing Dataset (CSD) (Cuesta et al., 2018)	Yes	MIDI, F0, Notes	Yes	3 songs	00:07:14

[10] https://www.pgmusic.com/barbershopquartet.htm
[11] https://www.pgmusic.com/bachchorales.htm

Polyphonic Singing Datasets - An Overview

Name/Author	Multitrack	Annotations	Publicly Available	# Recordings	Duration (hh:mm:ss)
Su et al. (2016)	No	MIDI	On Request	5 excerpts	00:02:11
Barbershop Quartets[10]	Yes	MIDI	No	22 songs	00:42:10
Bach Chorales[11]	Yes	MIDI	No	*26* songs	00:58:20
Scherbaum et al. (2019)	Yes	–	On Request	216 songs	06:04:40
Choral Singing Dataset (CSD) (Cuesta et al., 2018)	Yes	MIDI, F0, Notes	Yes	3 songs	00:07:14
Erkomaishvili Dataset (Rosenzweig et al., 2020)	No	Structure, F0, Score, Onsets	Yes	101 songs	07:05:00
Dagstuhl ChoirSet (DCS) (Rosenzweig and Cuesta et al., 2020)	Yes	MIDI, F0, Beats	Yes	2 songs, exercises	00:55:30

[10] https://www.pgmusic.com/barbershopquartet.htm
[11] https://www.pgmusic.com/bachchorales.htm

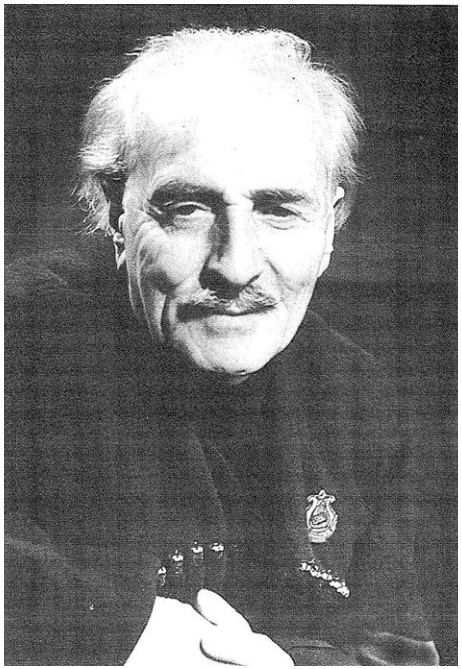


Erkomaishvili Dataset

Sebastian Rosenzweig, Frank Scherbaum, David Shugliashvili, Vlora Arifi-Müller, and Meinard Müller
Erkomaishvili Dataset: A Curated Corpus of Traditional Georgian Vocal Music for Computational Musicology
Transactions of the International Society for Music Information Retrieval (TISMIR), 3(1): 31–41, 2020.

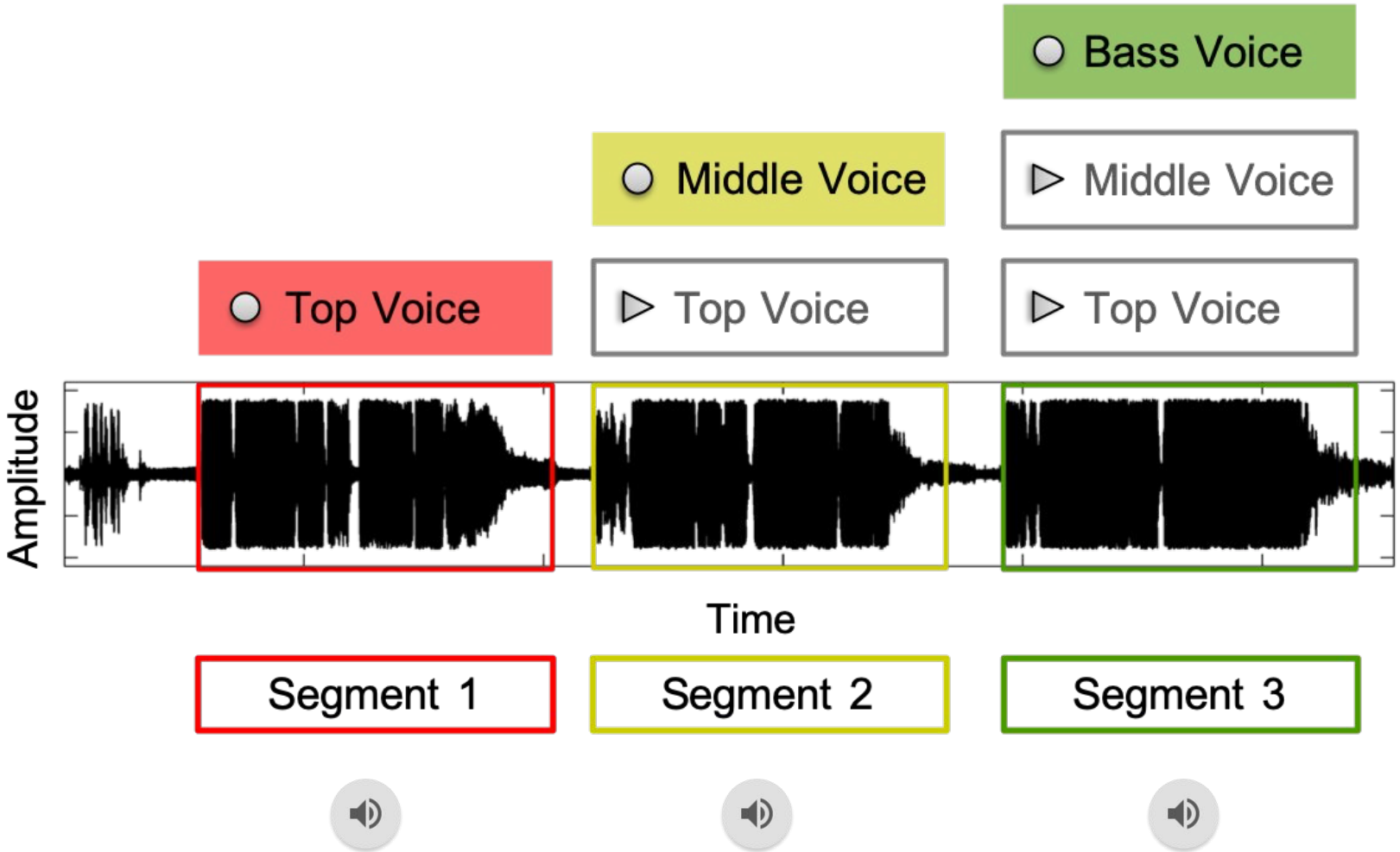
Erkomaishvili Dataset

- Collection of traditional three-voice **Georgian songs**
- Performed by the former Georgian master chanter **Artem Erkomaishvili** (1887-1967)
- Recorded using tape recorders in 1966

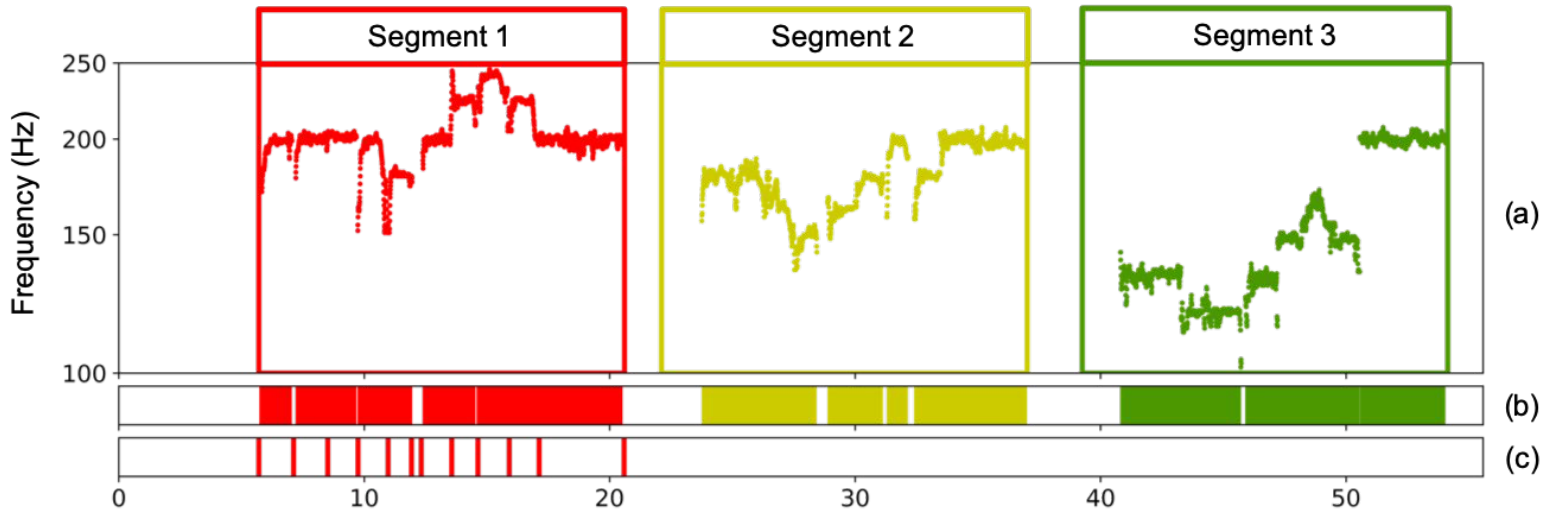


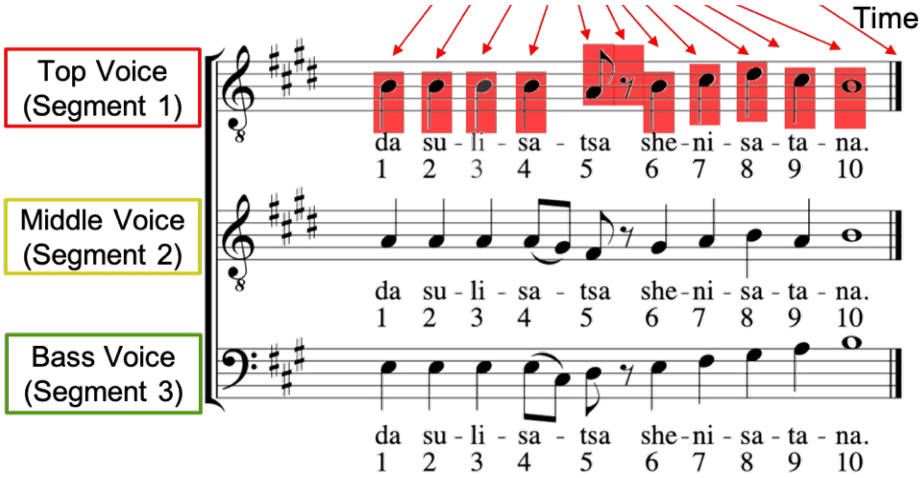
"*original masterpieces of Georgian musical thinking*"
(Shugliashvili, 2014)

Recording Process



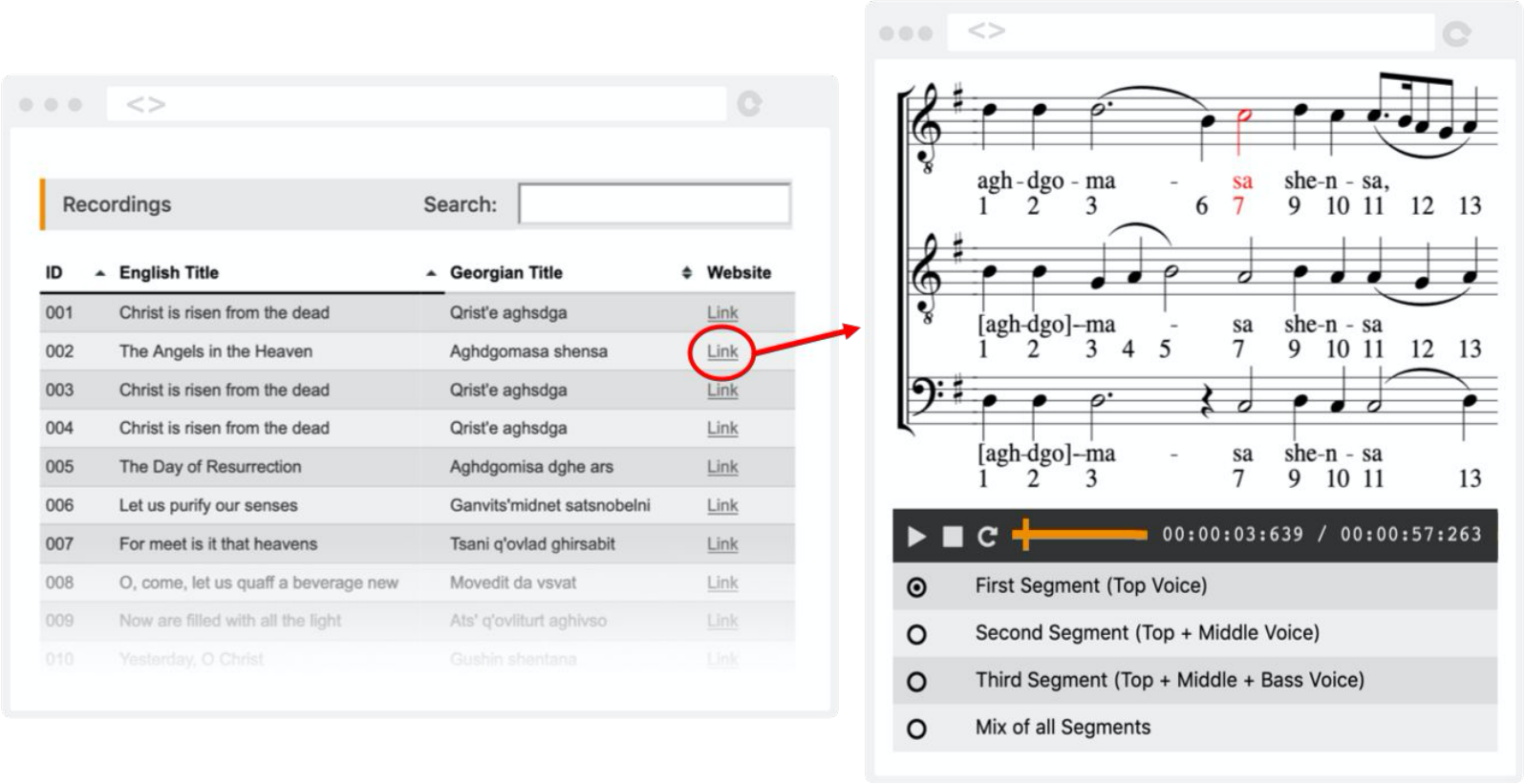
Available Annotations





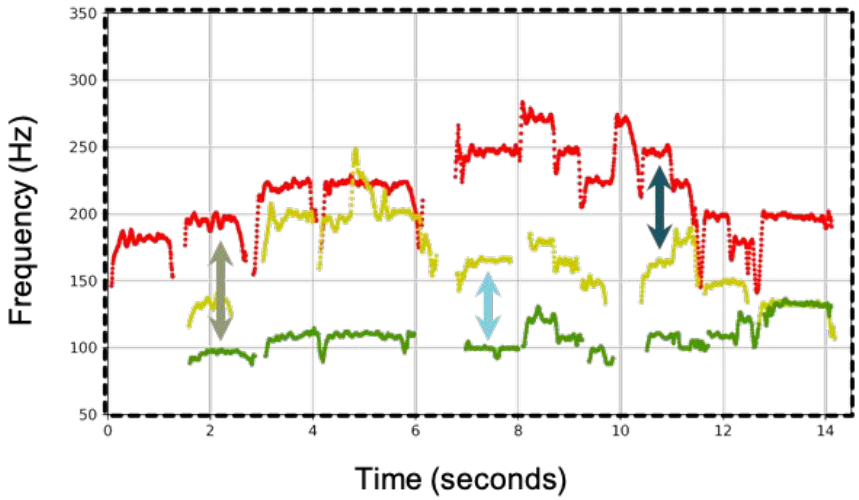
- F0-annotations
- Recording structure annotations
- Onset annotations
- Digital sheet music

Web-Based Interface

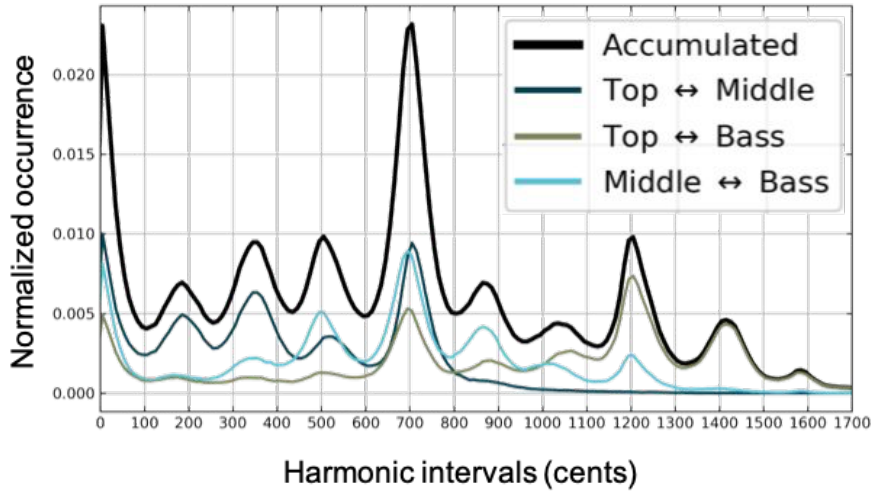


https://www.audiolabs-erlangen.de/resources/MIR/2019-GeorgianMusic-Erkomaishvili

Harmonic Interval Analysis

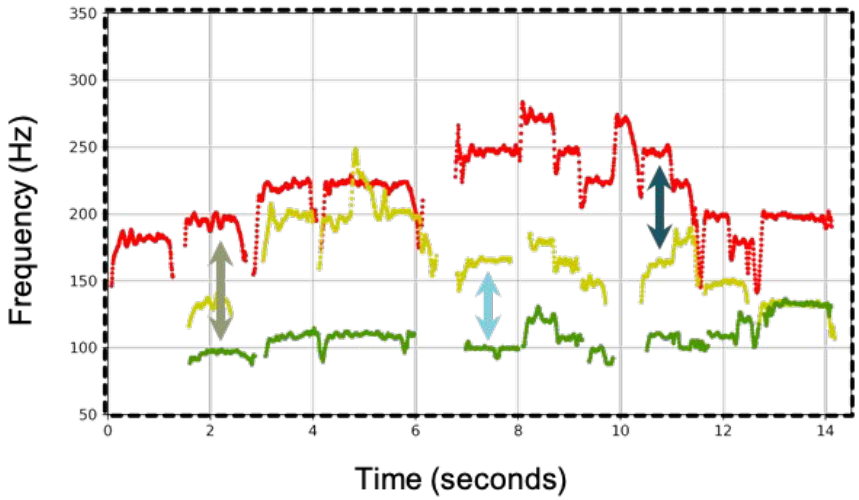


Snippet of superimposed F0-trajectories

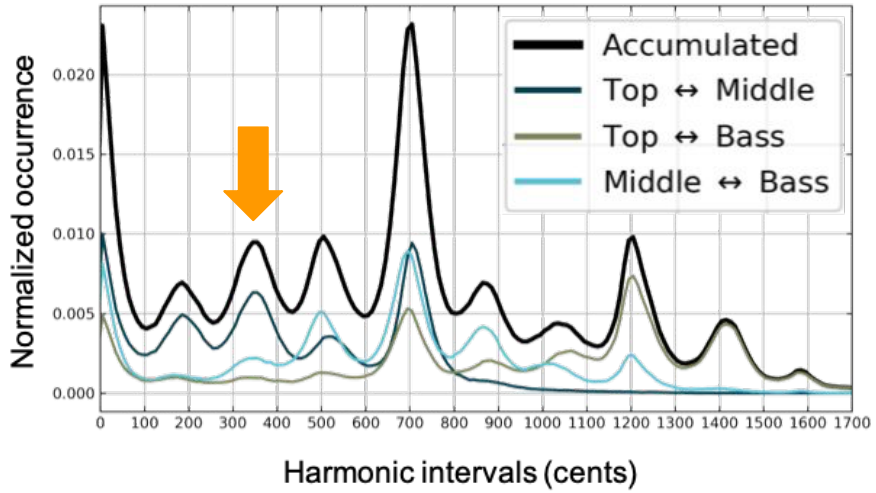


Averaged over all recordings

Harmonic Interval Analysis



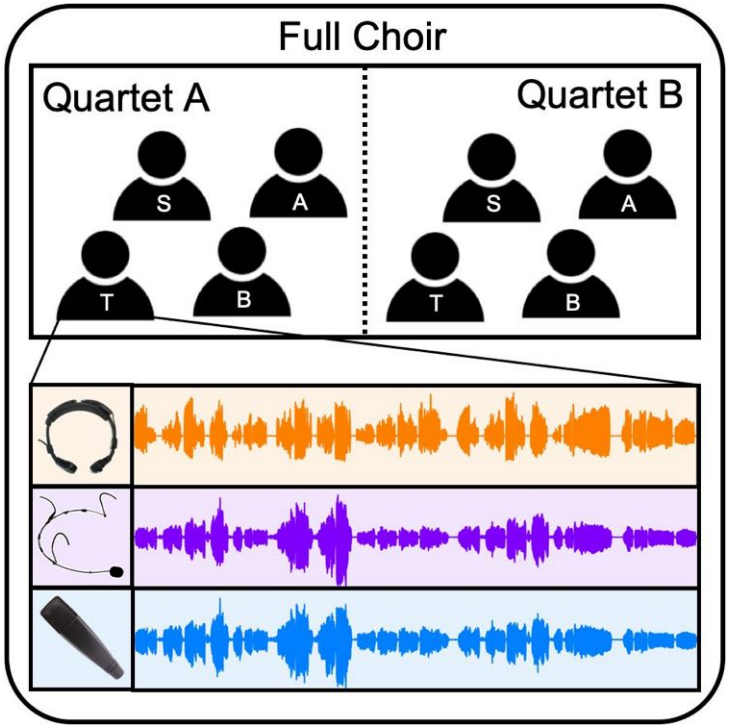
Snippet of superimposed F0-trajectories



Averaged over all recordings

Future Work: From 1966 to 2016

- Analyze and compare to recent field recordings from Georgia (Scherbaum et al., 2019)



Dagstuhl ChoirSet

Sebastian Rosenzweig, Helena Cuesta, Christof Weiß, Frank Scherbaum, Emilia Gómez, and Meinard Müller
Dagstuhl ChoirSet: A Multitrack Dataset for MIR Research on Choral Singing
Transactions of the International Society for Music Information Retrieval (TISMIR), 3(1): 98–110, 2020.

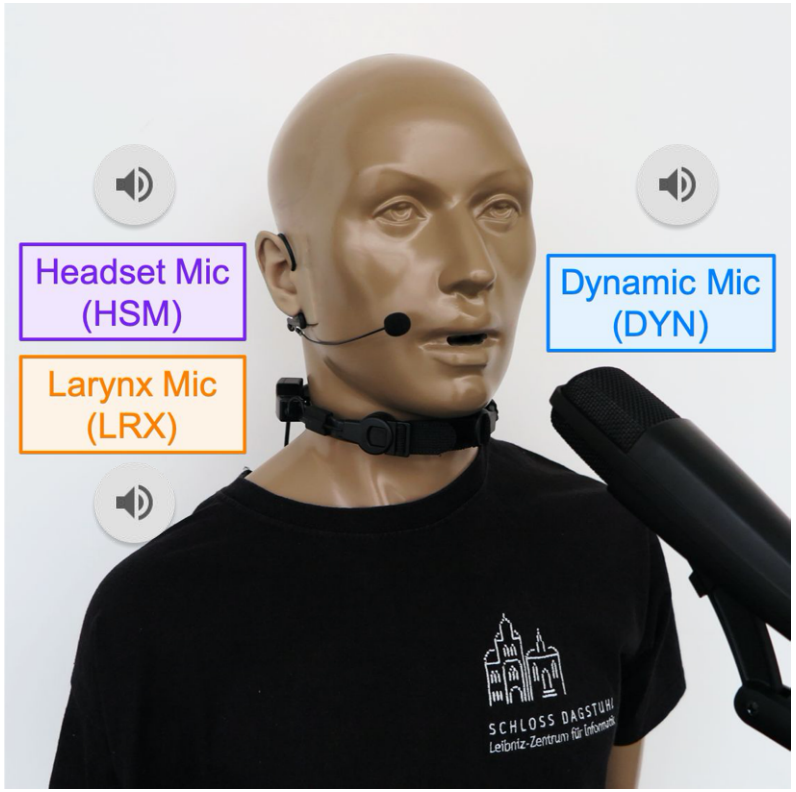
Dagstuhl Recordings

SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

- MIR seminar on “Computational Methods for Melody and Voice Processing in Music Recordings”
- Ensembles consisting of MIR researchers (mostly amateur singers)

Multitrack Recordings

- Two pieces
  - *Locus Iste*, Anton Bruckner
  - *Tebe Poem*, Dobri Hristov
  - Systematic Exercises
- Multiple takes
- Different settings:
  - **Full choir** (13 Singers, 2S 2A 4T 5B)
  - Two SATB **quartets** (4 singers each)
- ORTF stereo microphone (STM)
- Singers equipped with close-up microphones



Annotations

- Manual **beat annotations**
- Automatically extracted **fundamental frequency** (F0) trajectories: CREPE, pYIN
- Time-aligned **score representations** (from beats + score)

Web-based interface: **https://www.audiolabs-erlangen.de/resources/MIR/2020-DagstuhlChoirSet**
Toolbox: **https://github.com/helenacuesta/DCStoolbox**

Annotations

- Manual **beat annotations**
- Automatically extracted **fundamental frequency** (F0) trajectories: CREPE, pYIN
- Time-aligned **score representations** (from beats + score)

To validate the F0-trajectories:

- Manual annotations for two quartet recordings
- Evaluate automatic vs. manual annotations.

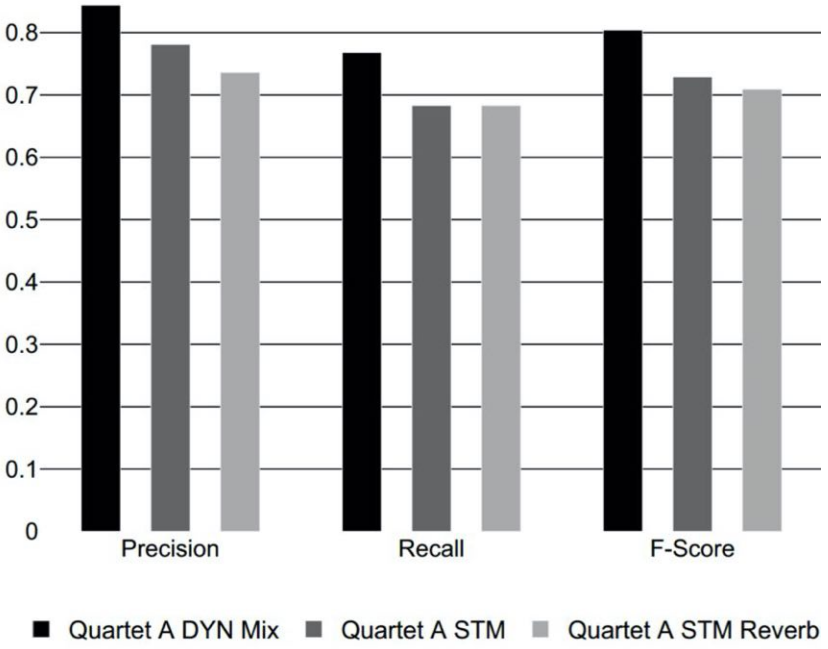
Web-based interface: **https://www.audiolabs-erlangen.de/resources/MIR/2020-DagstuhlChoirSet**
Toolbox: **https://github.com/helenacuesta/DCStoolbox**

Multi-F0 Estimation: a case study

In the paper, we use DeepSalience (Bittner et al., 2017) as a multi-F0 estimation model.

Compare three scenarios of the same performance:

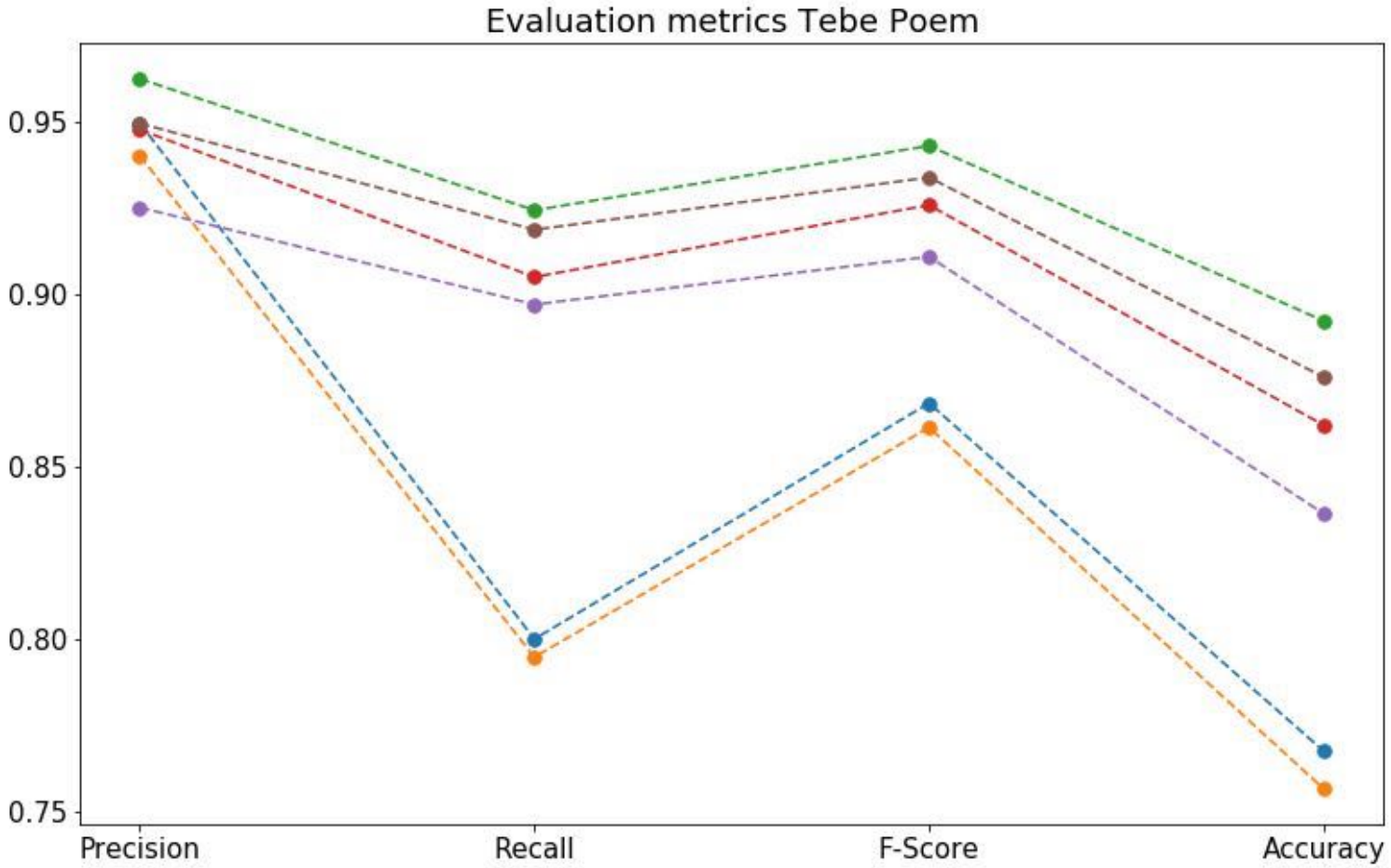
- Mixture of 4 DYN microphones (SATB).
- Stereo recording of the mixture.
- Stereo recording of the mixture + reverb.



Multi-F0 Estimation: a case study

- New experiment using **Tebe Poem** from **DCS**.
- Multi-F0 estimation model from Cuesta et al. (2020), trained with 4-part polyphonic vocal music.
- Evaluate all takes of Tebe Poem. Two main quartet scenarios:
  - Mixture of DYN microphones (SATB, from FC).
  - Stereo recording of the mixture (SATB quartet).

Multi-F0 Estimation: a case study

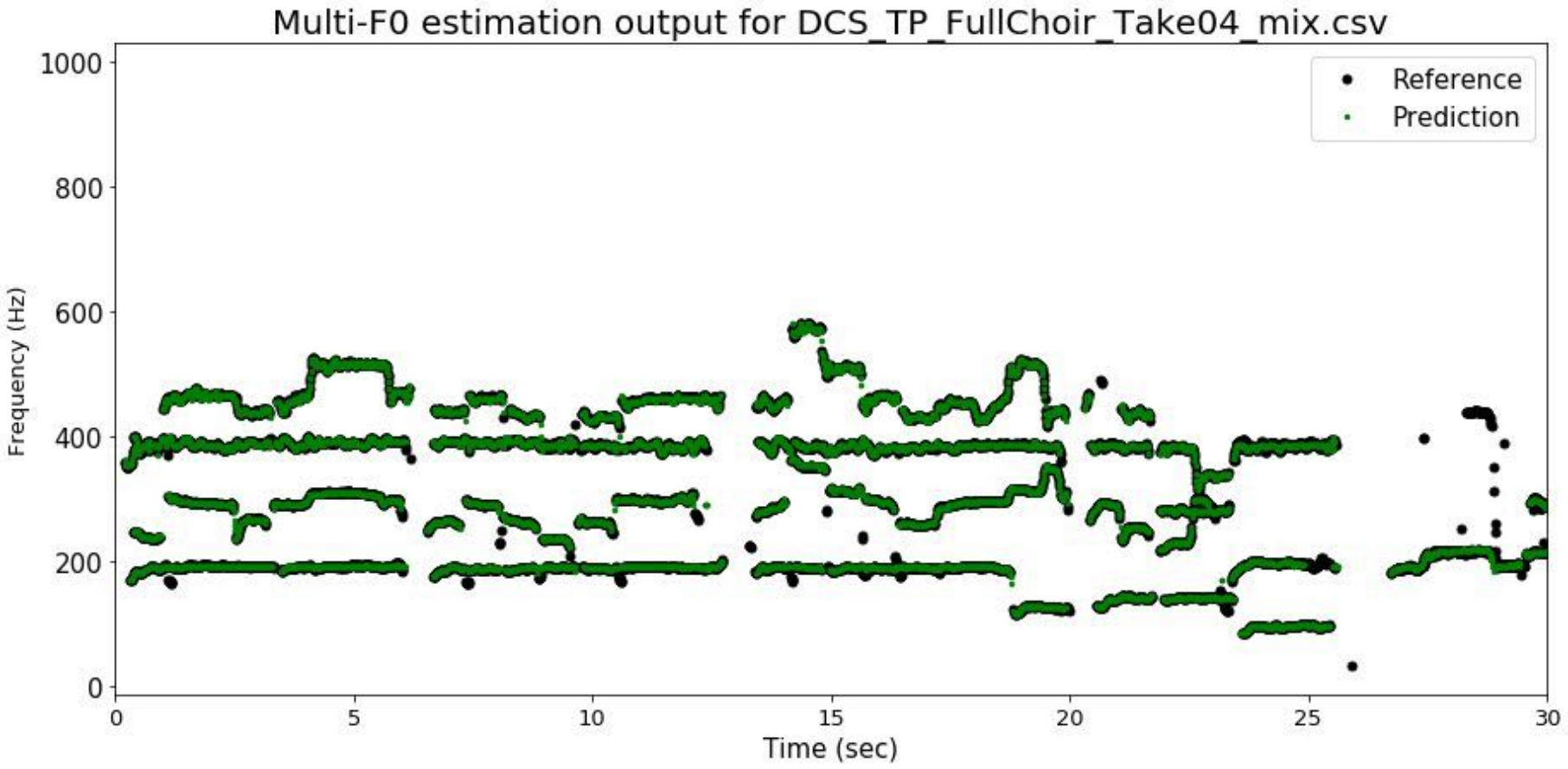


Stereo mic quartet

Mix of DYN mics

Multi-F0 Estimation: a case study

Mix of DYN mics

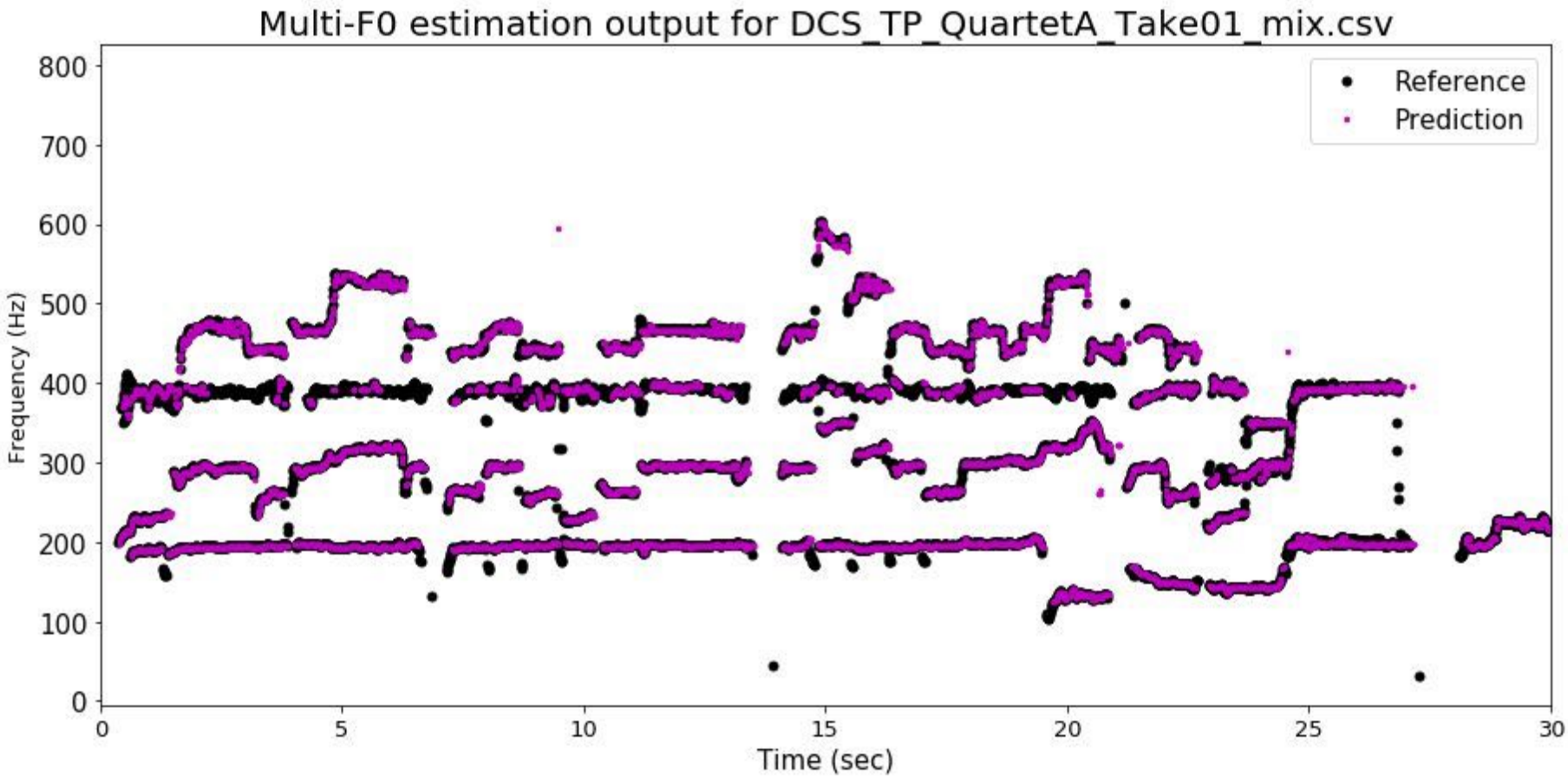


Original mix

Synth output

Multi-F0 Estimation: a case study

Stereo mic quartet

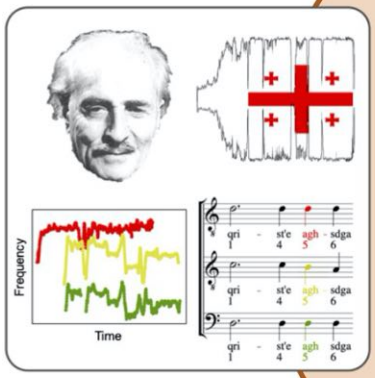


Original mix

Synth output

Summary

Erkomaishvili Dataset

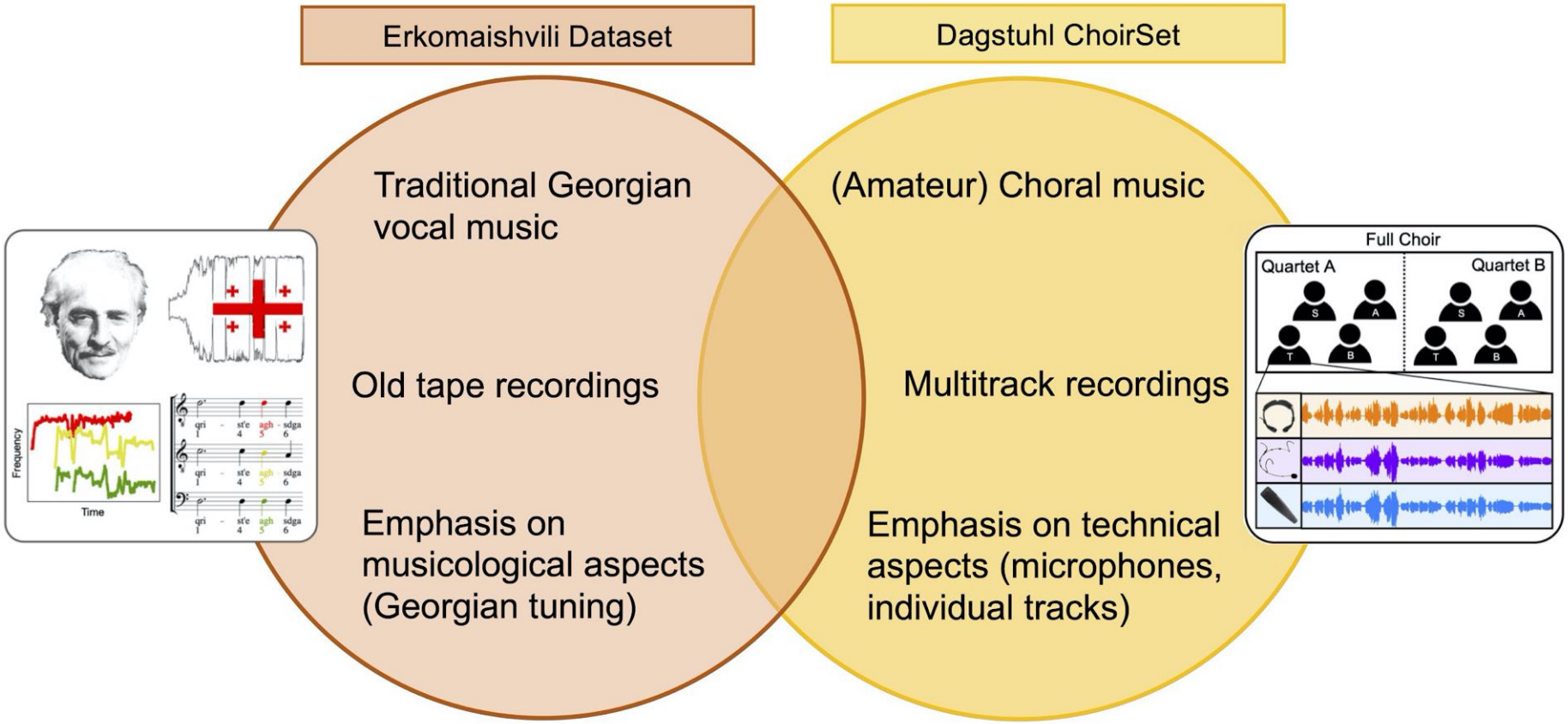


Traditional Georgian vocal music

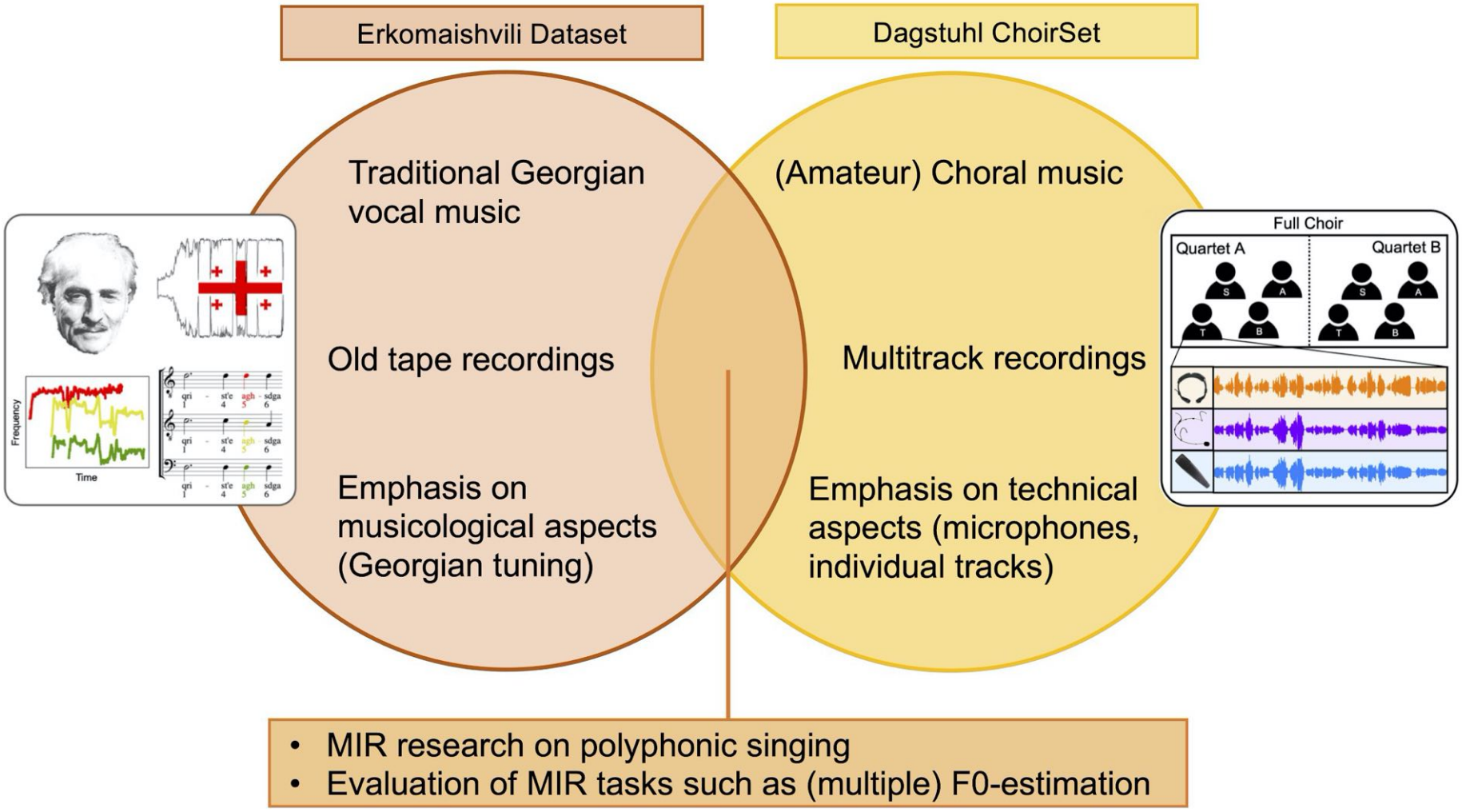
Old tape recordings

Emphasis on musicological aspects (Georgian tuning)

Summary


Erkomaishvili Dataset
Dagstuhl ChoirSet
Traditional Georgian vocal music
(Amateur) Choral music
Old tape recordings
Multitrack recordings
Emphasis on musicological aspects (Georgian tuning)
Emphasis on technical aspects (microphones, individual tracks)
Frequency
Time
Full Choir
Quartet A
Quartet B
S
A
T
B

Summary


Erkomaishvili Dataset
Dagstuhl ChoirSet
Traditional Georgian vocal music
(Amateur) Choral music
Old tape recordings
Multitrack recordings
Emphasis on musicological aspects (Georgian tuning)
Emphasis on technical aspects (microphones, individual tracks)
Frequency
Time
Full Choir
Quartet A
Quartet B
S
A
T
B
• MIR research on polyphonic singing
• Evaluation of MIR tasks such as (multiple) F0-estimation

References

Bittner, R. M., McFee, B., Salamon, J., Li, P., & Bello, J. P. (2017). **Deep salience representations for F0 tracking in polyphonic music.** In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), pages 63–70. Suzhou, China.

Cuesta, H., Gómez, E., & Chandna, P. (2019). **A framework for multi-f0 modeling in SATB choir recordings.** In Proceedings of the Sound and Music Computing (SMC) Conference, pages 447–453.

Kim, J. W., Salamon, J., Li, P., & Bello, J. P. (2018). **Crepe: A convolutional representation for pitch estimation.** In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 161–165. Calgary, Canada. DOI: https://doi.org/10.1109/ICASSP.2018.8461329.

Mauch, M., & Dixon, S. (2014). **pYIN: A fundamental frequency estimator using probabilistic threshold distributions.** In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 659–663. Florence, Italy. DOI: https://doi.org/10.1109/ICASSP.2014.6853678.

Scherbaum, F., Mzhavanadze, N., Rosenzweig, S., & Müller, M. (2019). **Multi-media recordings of traditional Georgian vocal music for computational analysis.** In Proceedings of the International Workshop on Folk Music Analysis, pages 1–6. Birmingham, UK.

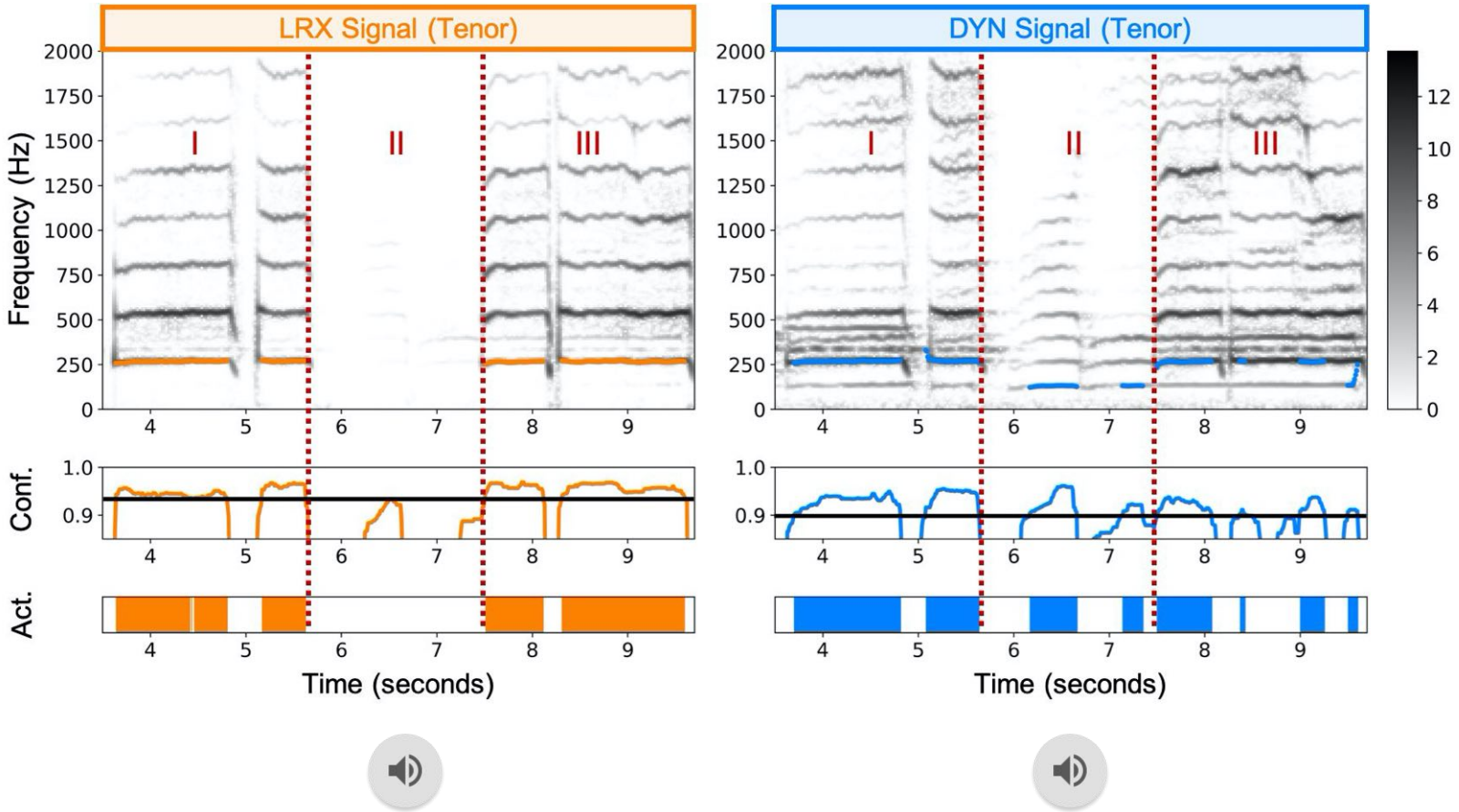
Shugliashvili, D. (2014). **Georgian Church Hymns**, Shemokmedi School. Georgian Chanting Foundation.

References

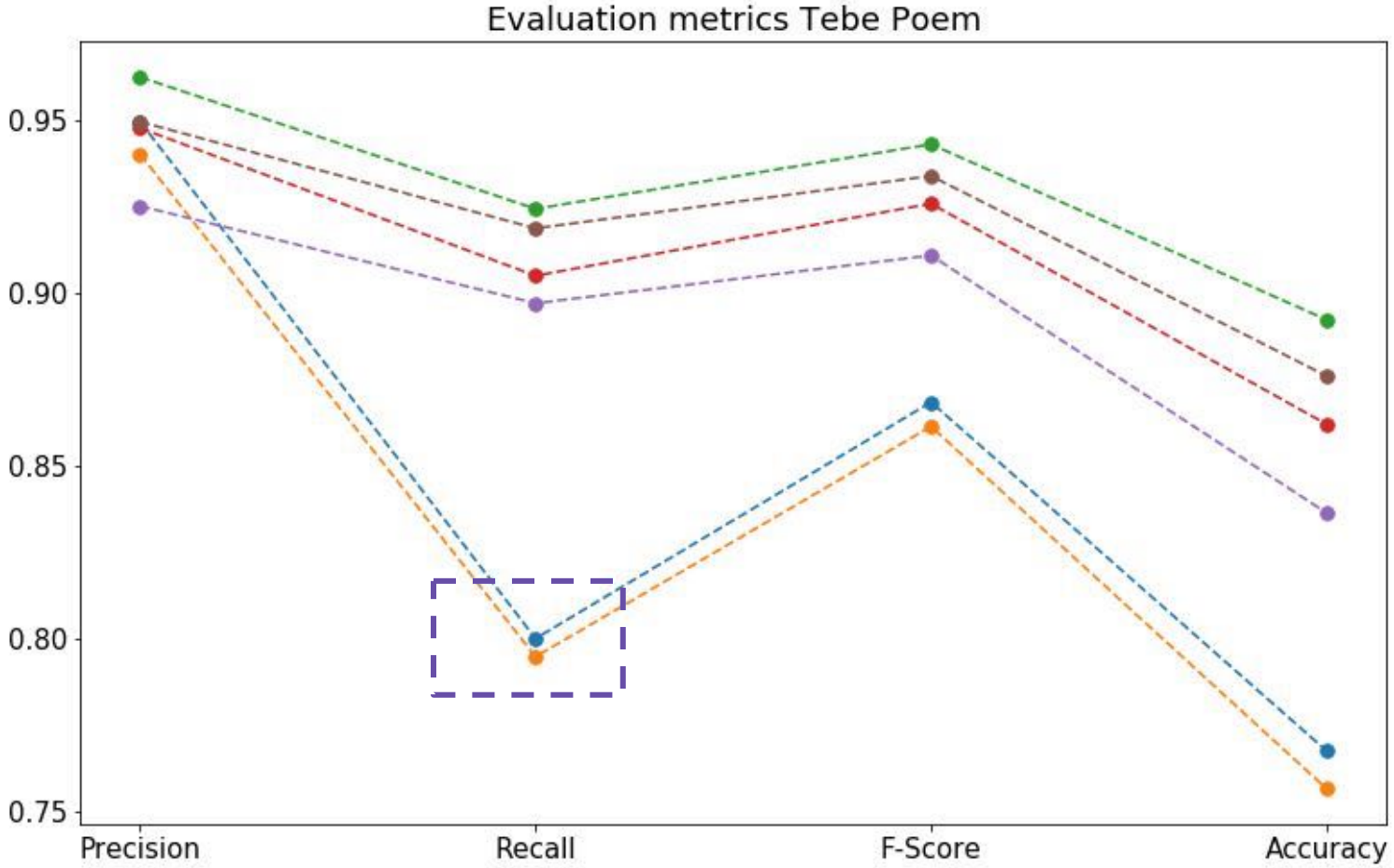
Su, L., Chuang, T.-Y., & Yang, Y.-H. (2016). **Exploiting frequency, periodicity and harmonicity using advanced time-frequency concentration techniques for multipitch estimation of choir and symphony.** In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), pages 393–399. New York City, USA.

Cuesta, H., McFee, B., & Gómez, E. (2020). **Multiple F0 Estimation in Vocal Ensembles using Convolutional Neural Networks**. In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), pages 302-309. Montreal, Canada (virtual).

Close-Up Microphones



Multi-F0 Estimation: a case study



Stereo mic quartet

Mix of DYN mics