FRANK SCHERBAUM
(GERMANY)

FIVE YEARS OF COMPUTATIONAL ANALYSIS OF TRADITIONAL GEORGIAN VOCAL MUSIC: THE GVM PROJECT

1. Introduction

Computational ethnomusicology is still a young, but rapidly growing field of research which involves a number of different disciplines. In this presentation I will look back on five years of collaborative work within the research project „Computational Ethnomusicology of Traditional Georgian Vocal Music (GVM)", which has been funded by the German Research Foundation (DFG) since 2018 and is jointly conducted at the Universities of Erlangen and Potsdam by an international and interdisciplinary team[6]. Specifically, I will share my personal selection of some of our attempts to combine computational approaches with classical ethnomusicological methods. Already before the beginning of the project, a lot of time and ressources went into the generation and processing of data sets. In total, three large digital data sets, particularly suited for computational analysis, were created or re-processed and used during the GVM project. Two of them, the Erkomaishvili dataset (Müller et al., 2017; Rosenzweig et al., 2020; Scherbaum et al., 2017, 2020, 2021), which contains the Tbilisi State Conservatory recordings of Artem Erkomaishvili from the year 1966, and the GVM dataset (Scherbaum et al., 2019; Scherbaum, Mzhavanadze, et al., 2018; Scherbaum & Mzhavanadze, 2018), which mainly contains recordings of Svan music, are already freely accessible through Web-based Audio/Video interfaces at the University of Erlangen[7], through which each of the recorded synchronized tracks can be very conveniently accessed individually or jointly. Regarding the GVM dataset, all the original, ***high-resolution*** audio, video, photographs, and the documentation of the recording sessions, etc. is kept in a long-term archive at the University of Jena which was also funded by the German Research Council[8]. With these three corpora, there now exists a collection of ***high-quality*** recordings of traditional Georgian vocal music which is open for non-commercial use. A key aspect and the main reason why these data are so useful for Computational Ethnomusiclogy is that separate recordings exist for each of the individual voices, either directly recorded with larynx microphones (Kane & Scherbaum, 2016; Scherbaum, 2016; Scherbaum et al., 2015; Scherbaum, Rosenzweig, et al., 2018) or generated through post-processing (Müller et al., 2017). This allows for a variety of interesting and partially new research questions to be addressed. In the following, I will touch upon a few selected topics and some essential lessons which we learned along the way.

---

[1] Project team members in alphabetical order: Reza D. D. Esfahani (Potsdam), Meinard Müller (Erlangen), Nana Mzhavanadze (Potsdam), Sebastian Rosenzweig (Erlangen), Frank Scherbaum (Potsdam), Daniel Vollmer (Potsdam).

[2] https://www.audiolabs-erlangen.de/resources/MIR/2019-GeorgianMusic-Erkomaishvili https://www.audiolabs-erlangen.de/resources/MIR/2017-GeorgianMusic-Scherbaum

[3] https://lazardb.gbv.de/detail/7392

2. The Tonal Organization of Traditional Georgian Music

This has been one of the most controversial topics of scholarly dispute during the last decades (Erkvanidze, 2002, 2016; Tsereteli & Veshapidze, 2014, 2015). Our contribution to this discourse is based on the comprehensive acoustical analysis of the Erkomaishvili- and the GVM dataset (Scherbaum et al., 2020, 2022). This work was technically rather challenging, for example because of the considerable continuous pitch drifts in the Svan repertoire. In this context we had to develop, test and implement, a number of new methods (Müller et al., 2017; Rosenzweig et al., 2019, 2021, 2022).

There are three Main Lessons which I Feel We Learned in the Course of this Work:

1. To understand the Tonal Organisation of the Erkomaishvili- and the GVM dataset requires the understanding of three quantities: first the pitch inventory, in other words what pitches are used as skeleton of a song; second the harmonic interval inventory; and third the melodic step size distribution, which explains how the pitches of the pitch inventory are combined in a melodic phrase.
2. The interval structure of the harmonic interval inventories differ systematically from the interval structure of the pitch inventories, probably because of harmonic intonation adjustments which have already been qualitatively described early in the last century (Nadel, 1933). For example, the harmonic 2nd is nearly always close to 200 cents, while the melodic 2nds are significantly smaller. Part of the reason may be that harmonic 2nds have to appear as side products of the popular Georgian 1-4-5 chord.
3. The melodic step sizes used in actual performances were found to differ widely and scatter around 170 cents. At first glance, this might be taken as evidence for an equi-distant (unitonic) scale. However, this should at best be seen as an approximation. As for the Erkomaishvili dataset (Rosenzweig et al., 2020) for example, the melodic intervals around the most salient pitch group of a chant were often observed to be significantly larger than what would be expected from a unitonic scale (Scherbaum et al., 2020).

In addition, there are fundamental statistical arguments against the use of a small number of melodic step sizes to reconstruct a melodic scale. The fact that the melodic step size distribution peaks at roughly 170 cents and is more or less symmetric, will nearly always result in samples which are taken from the center of this distribution, in other words be close to 170 cents, independent of the pitch distribution. Therefore, I see the analysis of melodic step sizes as the least reliable method to determine the melodic scales.

3. New Dynamic Visual Representations of Traditional Georgian Music

The visual representation of traditional Georgian vocal music is a topic where computational tools can be especially helpful. In this context, we have experimented with different types (perspectives) of video representations of songs[9]. One example in which a cursor follows the pitch and note trajectories while you can listen to the music can be accessed on the project website[10] and is illustrated in Fig. 1.

---

4 https://www.uni-potsdam.de/en/soundscapelab/computational-ethnomusicology/a-new-web-based-video/audio-interface-for-traditional-georgian-vocal-music

5 http://soundscapelab.de/GVMPR/F0.php?varname=3

In this representation, the note lyrics are displayed on top of the "notes". Subtitles for the individual singer's close-up face videos are shown on the left side of the screen. The red cursors ***mark*** the actual position within the song. The green number in the ellipse in the upper right of the plot shows the F0-value of the lowest voice at the cursor position in Cents (relative to the chosen reference frequency of 55 Hz). The numbers displayed as sub- and superscripts give the harmonic interval of the second lowest voice with respect to the bass voice and the harmonic interval of the highest voice with respect to the lowest voice, respectively. Finally, the distribution on the right side of the plot shows the frequency distribution of the F0-values from within all the notes in the song. We refer to this as „pitch distribution" or „pitch inventory". It is one of the ways to display the tuning system used in that song. Each peak corresponds to a scale degree. For a more detailed explanation see (Scherbaum et al., 2020). The tilted numbers in gray between the peaks of this distribution show the intervals between the different scale degrees in cents. As a reminder, 100 cents correspond to a semitone. This way, any tuning system used can be displayed in an undistorted way. For the audio, we made use of the pywebaudioplayer (Pauwels & Sandler, 2018), which allows the control of the volume of the individual voices during playback. This way listeners can generate the own ensemble mix.

4. Svan Funeral Dirges Zär aka Zari

Svan zär is a 3-voiced male chant which is sung on the day of a funeral, before the corpse of the deceased is put into the grave. The singers usually stand outside of the house of the deceased, behind a table with food and drinks. Different Svan villages used to have different variants of zär although now only 11 different variants are said to be actively maintained. Since Svan funeral rituals are believed to very old, our collection of zär recordings is a particular gem in our dataset, for research as well as culturally. The core of our work related to the investigation of Zär has been published in four companion papers (Mzhavanadze & Scherbaum, 2020a, 2021; Scherbaum & Mzhavanadze, 2020, 2021) and a video presentation (Mzhavanadze & Scherbaum, 2020b), which cover different perspectives on this phenomenon: acoustical, musicologicial, phonetic and linguistic, as well as the cultural context. In the part which is touched upon today, we analyzed the musical acoustical properties based on the Zär field recordings in the GVM dataset. The aim of the study was to investigate the tonal organization of the eleven different performances of six different variants of zär, performed by singers from different villages. For some of the performances, we observed a strong gradual pitch rise of up to 100 cents per minute. The intra-variant differences in the performances of different groups of singers were observed to be remarkably different, including the use of significantly different harmonic tuning systems. In contrast, two subsequent performances of the Mest'ia variant of zär by a group of singers recorded in Zargā̈sh were essentially identical. This demonstrates the widespread absence of improvisational elements in these two performances. One of the most interesting results of our analysis is the observation that the musical structure of zär, expressed for example in its ambitus, the complexity of its melodic progression, and its harmonic chord inventory change systematically along the course of the Enguri valley.

5. Singer interaction

One more aspect which I want to briefly mention is some very recent and still ongoing work about the non-verbal interaction of singers during performances of polyphonic music. In July 2019, together with Lasha Chkhart'ishvili (bass voice), Guram Guntadze (middle voice), and Mamuka

Siradze (top voice) from the trio Khelkhvavi in Ozurgeti, we conducted an experiment to monitor the singers‘ heartbeat rates during singing. So in addition to the usual headset- and larynx microphones, each singer got equipped with a pulse sensor taped to his index finger Fig. 2). We wanted to make sure that the singers were already well tuned to each other. In addition, we wanted to record a song which is sufficiently complex to set it aside from similar experiments conducted for mantra singing. These two conditions let us choose the Gurian folk song Chven Mshvidoba, which was sung close to the end of the session.

Without going into any detail, which can be found in (Scherbaum & Müller, 2022), we observed that the hearts of Guram and Mamuka (middle and top voice singers) accelerate and decelerate more or less simultaneously during the performance. This may be caused by a synchronization of their respiration and a coupling between respiration and heart beat via a phenomenon called respiratory sinus arrythmia (RSA), but there may also be other factors contributing. These results are quite encouraging and suggest that there is a lot of new and relevant information to be gained from the joint use of different sensor types during ethnomusicological recording sessions.

6. Conclusions

The selected results demonstrate some of the potential benefits of computational tools for the analysis of Georgian and other non-western oral music traditions and the benefits of extending ethnomusicological field recordings to new types of multimedia, and multi-channel recordings (including video, audio, muscle vibrations, heartbeat, and respiration sensors). In the long run, I hope that such an approach will allow shifting the research focus from the pure documentation of “what is sung?” to more complex questions such as “how does the music work?” and “how do singers interact?”.

Acknowledgements

I want to close with a big thank you to everyone who contributed to the success of this project, in particular to all the singers who shared their music with us. It has been an incredibly enriching experience for me, personally as well as scientifically. This work was supported by the German Research Foundation (DFG MU 2686/13-1, SCHE 280/20-1).

References

Erkvanidze, M. (2002). On georgian scale system. In Rusiko Tsurtsumia & Joseph Jordania (Eds.), *The First International Symposium on Traditional Polyphony* (pp. 178–185). Tbilisi: Tbilisi State Conservatoire

Erkvanidze, M. (2016). The Georgian Musical System. In Beauguitte, P., Duggan, B., Kelleher, J. (Eds.), *6th International Workshop on Folk Music Analysis* (pp. 74–79). Dublin: Technological University Dublin

Kane, F., & Scherbaum, F. (2016). Using body vibrations for teaching, visualization and analysis of traditional Georgian singing. In Rusiko Tsurtsumia & Joseph Jordania (Eds.), *8th International Symposium on Traditional Polyphony* (pp.190–197). Tbilisi: Tbilisi State Conservatoire.

Müller, M., Rosenzweig, S., Driedger, J., & Scherbaum, F. (2017). Interactive Fundamental Frequency Estimation with Applications to Ethnomusicological Research. *Conference on Semantic Audio,* 8. Erlangen, Germany. http://www.aes.org/e-lib/browse.cfm?elib=18777

Mzhavanadze, N., & Scherbaum, F. (2020a). Svan Funeral Dirges (Zär): Musicological analysis. *Musicologist,* 4(2), 168–197. https://doi.org/10.33906/musicologist.782185

Mzhavanadze, N., & Scherbaum, F. (2020b). Zär, polyphonic group laments from Svaneti/Georgia. Society of Ethnomusicology (SEM) Annual Meeting.

Mzhavanadze, N., & Scherbaum, F. (2021). Svan Funeral Dirges (Zär): Cultural Context. *Musicologist,* 5(2), 133–165. https://doi.org/10.33906/musicologist.906765

Nadel, S. F. (1933). *Georgische Gesänge.* Lautabt, Leipzig: Harrassowitz in Komm.

Pauwels, J., & Sandler, M. B. (2018). pywebaudioplayer: Bridging the gap between audio processing code in Python and attractive visualisations based on web technology. Web Audio Conference WAC-2018, Berlin, Germany.

Rosenzweig, S., Scherbaum, F., & Meinard Müller. (2022). Computer-Assisted Analysis of Field Recordings: A Case Study of Georgian Funeral Songs. *ACM Journal on Computing and Cultural Heritage,*16(1), 1–16. https://doi.org/10.1145/3551645

Rosenzweig, S., Scherbaum, F., & Müller, M. (2019). Detecting Stable Regions in Frequency Trajectories for Tonal Analysis of Traditional Georgian Vocal Music. *20th International Society for Music Information Retrieval Conference* (352–359). Delft, The Netherlands. https://archives.ismir.net/ismir2019/paper/000041.pdf

Rosenzweig, S., Scherbaum, F., & Müller, M. (2021). Reliability assessment of singing voice F0-estimates using multiple algorithms. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings* (261–265). https://doi.org/10.1109/ICASSP39728.2021.9413372

Rosenzweig, S., Scherbaum, F., Shugliashvili, D., Arifi-Müller, V., & Müller, M. (2020). Erkomaishvili Dataset: A Curated Corpus of Traditional Georgian Vocal Music for Computational Musicology. *Transactions of the International Society for Music Information Retrieval*, 3(1), 31–41. https://doi.org/10.5334/tismir.44

Scherbaum, F. (2016). On the benefit of larynx-microphone field recordings for the documentation and analysis of polyphonic vocal music. *Proc. of the 6th International Workshop Folk Music Analysis* (80–87). Dublin/Ireland,

Scherbaum, F., Loos, W., Kane, F., & Vollmer, D. (2015). Body vibrations as source of information for the analysis of polyphonic vocal music. *Proceedings of the 5th International Workshop on Folk Music Analysis,* 5, 89–93, Paris, France: University Pierre and Marie Curie.

Scherbaum, F., & Mueller, M. (2023). From Intonation Adjustments to Synchronization of Heart Rate Variability: Singer Interaction in Traditional Georgian Vocal Music. *Musicologist,* 7(2), 155–177. https://doi.org/10.33906/musicologist.1144787

Scherbaum, F., Müller, M., & Rosenzweig, S. (2017). Analysis of the Tbilisi State Conservatory Recordings of Artem Erkomaishvili in 1966. *7th International Workshop on Folk Music Analysis,* 29–36. Malaga, Spain. https://www.researchgate.net/publication/320716492–Analysis–of–the–Tbilisi–State–Conservatory–Recordings–of–Artem–Erkomaishvili–in–1966

Scherbaum, F., & Mzhavanadze, N. (2018). *A new archive of multichannel-multimedia field recordings of traditional Georgian singing, praying, and lamenting with special emphasis on Svaneti.* LaZAR-Database. https://www.uni-potsdam.de/fileadmin/projects/soundscapelab/PapersMusic/2017/ArchiveDescription–V5small.pdf

Scherbaum, F., & Mzhavanadze, N. (2020). Svan Funeral Dirges (Zär): Musical Acoustical Analysis of a New Collection of Field Recordings. *Musicologist,* 4(2), 138–167. https://doi.org/10.33906/musicologist.782094

Scherbaum, F., & Mzhavanadze, N. (2021). Svan Funeral Dirges (Zär): Language-Music Relation and Phonetic Properties. *Musicologist*, 5(1), 66–82. https://doi.org/10.33906/musicologist

Scherbaum, F., Mzhavanadze, N., Arom, S., Rosenzweig, S., & Müller, M. (2020). Tonal Organization of the Erkomaishvili Dataset: Pitches, Scales, Melodies and Harmonies. In Scherbaum, F. (Ed.), *Computational Analysis of Traditional Georgian Vocal Music (Issue 1)* (pp. 1–64). Postdam: Universitätsverlag Potsdam https://publishup.uni-potsdam.de/frontdoor/index/index/docId/47614

Scherbaum, F., Mzhavanadze, N., Arom, S., Rosenzweig, S., & Müller, M. (2021). Analysis of Tonal Organization and Intonation Practice in the Tbilisi State Conservatory Recordings of Artem Erkomaishvili of 1966. *Sixth Analytical Approaches to World Music Conference,* Musée de l'Homme, Paris/France.

Scherbaum, F., Mzhavanadze, N., & Dadunashvili, E. (2018). A Web-Based, Long-Term Archive of Audio, Video, and Larynx-Microphone Field Recordings of Traditional Georgian Singing, Praying and Lamenting with Special Emphasis on Svaneti. In Tsurtsumia, R., & Jordania, J. (Eds.), 9*th International Symposium on Traditional Polyphony* (pp. 174–180). Tbilisi: Tbilisi State Conservatoire.

Scherbaum, F., Mzhavanadze, N., Rosenzweig, S., & Müller, M. (2019). Multi-media recordings of traditional Georgian vocal music for computational analysis. *9th International Workshop on Folk Music Analysis*. Birmingham. https://www.researchgate.net/publication/341680587–MULTI-MEDIA–RECORDINGS–OF–TRADITIONAL–GEORGIAN–VOCAL–MUSIC–FOR–COMPUTATIONAL–ANALYSIS

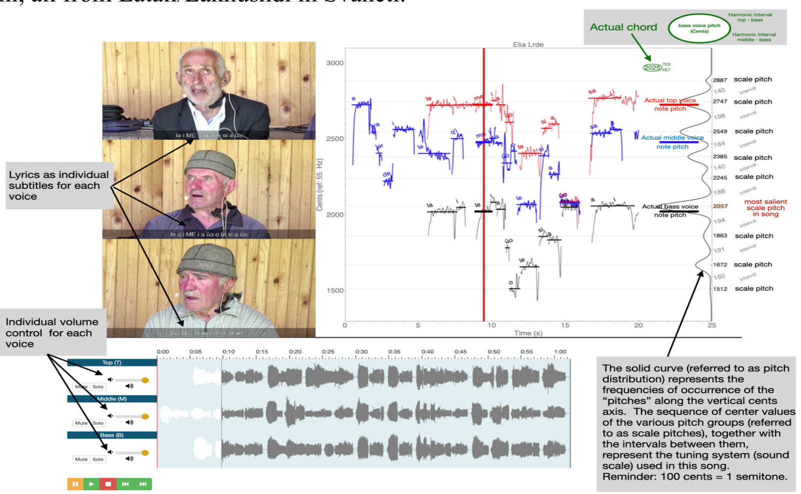
Scherbaum, F., Mzhavanadze, N., Rosenzweig, S., Müller, M. (2022). Tuning Systems of Traditional Georgian Singing Determined from a New Corpus of Field Recordings. *Musicologist,* 6(2), 142–168. https://doi.org/10.33906/musicologist.1068947

Scherbaum, F., Rosenzweig, S., Müller, M., Vollmer, D., & Mzhavanadze, N. (2018). Throat Microphones for Vocal Music Analysis. *Demos and Late Breaking News of the International Society for Music Information Retrieval Conference (ISMIR).*

Tsereteli, Z., & Veshapidze, L. (2014). On the Georgian traditional scale. In Rusiko Tsurtsumia & Joseph Jordania (Eds.), 7*th International Symposium on Traditional Polyphony* (pp. 288–295). Tbilisi: Tbilisi State Conservatoire.

Tsereteli, Z., & Veshapidze, L. (2015). Video of the presentation "The empirical research of a Georgian sound scale". 2015 IAML/IMS Congress. https://youtu.be/TFncneafovI

სურათი 1. დინამიური აუდიო/ვიდეო ინტერფეისის ეკრანული ანაბეჭდი, რომელიც შეიქმნა GVM პროექტის განმავლობაში. მომღერლები ზემოდან ქვემოთ: გიგო ჩამგელიანი, მურად ფირცხელანი და გივი ფირცხელანი, ყველანი ლატალი/ლახუშდიდან, სვანეთი.
Figure 1. Screendump of the dynamic audio/video interface which was developed during the GVM project. Singer from top to bottom are Gigo Chamgeliani, Murad Pirtskhelani, and Givi Pirtskhelani, all from Latali/Lakhushdi in Svaneti.



სურათი 2. ტრიო „ხელხვავი“ ოზურგეთიდან. მარცხნიდან მარჯვნივ: ლაშა ჩხარტიშვილი (ბანი), გურამ გუნთაძე (შუა ხმა) და მამუკა სირაძე (მაღალი ხმა).
Figure 2. Trio Khelkhvavi from Ozurgeti. From left to right: Lasha Chkhart'ishvili (bass voice), Guram Guntadze (middle voice), and Mamuka Siradze (top voice).