



Trabzon University State Conservatory © 2017-2021

Volume 5 Issue 1 June 2021

Research Article

Musicologist 2021. 5 (1): 66-82

DOI: 10.33906/musicologist. 875348

FRANK SCHERBAUM

Universität Potsdam, Germany

fs@geo.uni-potsdam.de

orcid.org/0000-0002-5050-7331

NANA MZHAVANADZE

Universität Potsdam, Germany

mzhavanadze@uni-potsdam.de

orcid.org/0000-0001-5726-1656

Svan Funeral Dirges (Zär): Language-Music Relation and Phonetic Properties

ABSTRACT

This paper reports on the results of a follow-up study to Scherbaum and Mzhavanadze (2020) and Mzhavanadze and Scherbaum (2020), which jointly describe the acoustical and musicological properties, respectively, of a new collection of field recordings of three-voiced Svan funeral dirges, known as zär in Svan and zari in Georgian. The focus of the present work is on language-music relation and phonetic properties. It was motivated by the pioneering work of Bolle Zemp (1994; 1997; 2001), who for the first time examined this topic from both an ethnomusicological and a linguistic perspective. We revisit some of Bolle Zemp’s observations and assumptions with a substantially expanded (by a factor of 11) data set and by using computational phonetic analysis tools not available at the time of her study. Her observation of correlations between pitch, duration, and timbre features related to the interjection of woj (wai) is consistent with our analysis for only some of the singers. Therefore, rather than assigning a general semantic meaning to these correlations, we offer an alternative interpretation as a natural consequence of the formant tuning techniques, which we found employed in the vocal production of some of the singers.

KEYWORDS

Svan Funeral Dirges

Language-Music Relation

Phonetic Properties

Received: February 5, 2021; Accepted: June 4, 2021

Introduction

Svaneti in NW Georgia is one of the rare regions at the crossroads of Europe and Asia, where very old (possibly pre-Christian) traditions are still cultivated as part of daily life. Svan funeral dirges, known as *zār* in Svan and *zari* in Georgian, performed exclusively during mourning rituals, occupy a special place within Georgian music. Due to the remote geographic situation, they are still maintained in a comparatively original form and are believed to offer insights into the oldest parts of Georgian vocal music. Svan *zār* have been the core topic of an interdisciplinary study, the results of which have been published or are forthcoming in a series of papers in *Musicologist*. These papers so far cover their acoustical and (ethno)musicological characteristics, as well as the historical and cultural context (Scherbaum & Mzhavanadze, 2020; Mzhavanadze & Scherbaum, 2020) and Mzhavanadze & Scherbaum (2021, forthcoming).

In addition to these topics, we believe that because of the peculiar vocal utterances employed, the analysis of Svan *zār* might also help shed some light onto the language-music relation in the early layers of Georgian musical thinking. The present study was stimulated by the pioneering work of Bolle Zemp (1994; 1997; 2001) on the language-music relation and the phonetic properties of *zār*. We extend her study, which was focused on the audio recording of a single funeral chant from Lat'ali (Lat'li in Svan), to the analysis of eleven different performances of five different variants of *zār*,¹ performed by singers from different villages, which were obtained during an ethnomusicological field expedition in 2016 (Scherbaum and Mzhavanadze, 2018; Scherbaum et al., 2019).² In addition, we want to test if her hypothesis regarding possible correlations between pitch, duration, and timbre features could be tested quantitatively by using a computational acoustical phonetic analysis. Some of the performances in the new corpus were recorded in their natural context at actual funerals, which we were allowed to document. Based on the information we received from our local informants during the 2016 expedition, the present collection represents almost half of the eleven different variants of *zār* believed

¹ Variant 1: K'ala and Ushguli, 2: Mest'ia, 3: Lenjar, 4: Lat'ali, 5: Lower Bal

² All the recordings of the 2016 field expedition have been made publicly available and can be accessed either through the open access long-term archive at the University of Jena, which also hosts the field report and the meta data (<https://lazardb.gbv.de/search>; see Scherbaum et al., 2019 for details), or through the research repository at the University of Erlangen of the GVM project (<https://www.audiolabs-erlangen.de/resources/MIR/2017-GeorgianMusic-Scherbaum>). Contact the first author (fs@geo.uni-potsdam.de) for access information.

to still be practiced today (Scherbaum and Mzhavanadze, 2020). The recording strategy employed during the 2016 field expedition makes the new corpus particularly suitable for modern state-of-the-art analysis. Whenever possible, one singer from each voice group was simultaneously recorded with a high-quality headset microphone and a larynx microphone³. In addition, the whole ensemble was recorded with a high-resolution (4K) video camera, on which a directional microphone was mounted, plus a conventional stereo microphone. The systematic use of larynx microphones allows the undistorted documentation of the acoustical contribution of each singer while all of them are singing together in their natural context (Scherbaum et al., 2015). In addition, larynx microphone recordings have been shown to contain essential information in relation to a singer's voice regarding pitch, intonation, timbre, and voice intensity. This allows the application of computer-based methods to document and analyze vocal music of the oral tradition in new ways. This includes, for example, computerized pitch analysis techniques to determine the fundamental frequency (F0) trajectories and their microtonal structure, to study the tuning systems used by the singers, as well as possible interactions between singers (Scherbaum et al., 2015; Scherbaum, 2016).

³ For reasons of reverence we have refrained from using headset microphones for the recordings which were made at real funerals.

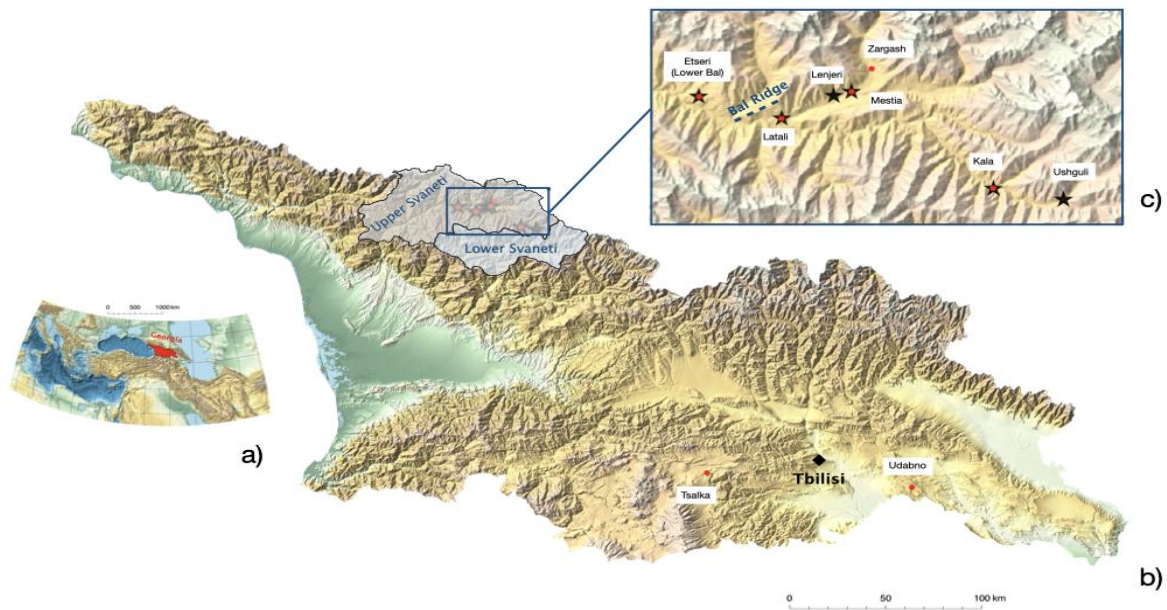


Figure 1. Geographical situation. a) The location of Georgia in its regional context. b) Study area and recordings sites (marked by red solid circles) of the field expedition of 2016. c) Locations of recording sites (red solid circles) and locations of origin of *zār* (black stars) within Upper Svaneti. The Bal ridge (altitude-wise) divides the Lower Bal and Upper Bal regions of Upper Svaneti.

The recording locations and the locations of origin of the different *zār* variants are shown in Figure 1 with solid red circles and black stars, respectively. Apart from the Upper Svaneti area, the field work also covered the eco-migrant Svan communities of Didgori, Tsalka, and Udabno, outside Svaneti, near the capital of Tbilisi (Figure 1).

Language–music relation

The quantitative investigation of language–music relations in *zār* was originally initiated by the Swiss ethnomusicologist Sylvie Bolle Zemp, who discussed the results of her studies in a sequence of papers, most of which are in German (Bolle Zemp, 1994; 1997; 2001). What made her work so innovative at the time is that she examined this topic from both an ethnomusicological and a linguistic perspective. For her analysis, Bolle Zemp had at her disposal the recordings of a single funeral chant from Lat’ali, which she phonetically transcribed using the symbols of the International Phonetic Alphabet (in its 1993 version). Employing the tools accessible at that time, she processed the chant through sonographic images and analyzed the musical content of the verbal ‘text’,

explored all the possible references of the utterances employed in *zār*, and visualized the results in the form of sonograms. Based on the results of her analysis, Bolle Zemp suggested a strong correlation between words and music. She assigned semantic importance to the core utterance *woj*, which she interpreted as an interjection of pain.⁴ She argued that as an utterance of mourning connotation, it lies at the root of several vocal formulas of the chant. In her view, ‘singers’ emphasize human emotions such as pain, dignity, etc. by modification of the sound characteristics of spoken language, e.g. by formation of vowels and consonants in different ways, by manipulation of the interjections, and by stylization of expressions of the spoken language through certain vocal processes (velarization, descending glissandi, nasalization). In her view, the structure of the movement of voices, the duration of a sound, the sequence of concomitants, and the interrelationship of consonant and dissonance chords is greatly conditioned by the ‘text’, which coordinates the musical process (Bolle Zemp, 1997; 2001). In particular, she suspected a correlation between a sequence of high/low/high timbres, high/low/high pitches, and short/long/short durations of vocal formulas. Although the research methodology of Bolle Zemp’s study makes it transparent and reproducible, the analysis of only one example of *zār* limits the generalizability of her results to other variants of the chant. Besides, although the general quality of the recordings of S. Bolle-Zemp is good, the recordings of that time do not allow for the analysis of the acoustical properties of individual voices because of the then yet-unsolved problem of separating individual voices in polyphonic field recordings⁵.

In this section, we are revisiting some of the issues brought up by Bolle Zemp (1997). In particular, we want to test if her hypothesis regarding possible correlations between pitch, duration, and timbre features could be tested quantitatively by using a computational acoustical phonetic analysis. Since each of the individual singers in our corpus was separately recorded, we can quantitatively exploit the acoustic properties of the individual voices, as well as their interaction for this purpose. For each voice, the fundamental frequency (F0) trajectory, as well as the sequence of individual note objects (containing onset time, pitch, duration and lyrics information), were determined (cf.

⁴ A phonetically related utterance in Georgian is *wai*.

⁵ This problem becomes also acute when focusing on retrospective study of the chant to reconstruct the stages of its development and changes. The quality of older archive recordings are often critically poor and sometimes does not allow even a minimal manual processing.

Figure 3 in Scherbaum and Mzhavanadze, 2020). The text of each note object was identified by a native Georgian speaker (Nana Mzhavanadze), in collaboration with two native Svan speakers (Madona Chamgeliani and Ketevan Margiani), and expressed in Svan language. For this purpose, the Svan character set on the Titus website was used⁶.

In order to obtain an overview of the lyrics used in the complete collection, Figure 2 shows an overview of the relative frequencies of occurrences of the individual note lyrics in the form of so-called word clouds, in which the size of a word corresponds to the relative frequency of its occurrence in the corresponding *zār*. This frequency is not always the same for the same *zār* variant, as can be seen, for example, in Figure 2 a) and b), d) and e), g), and h), and i) – k). In particular, the Lat’ali variant, which was the subject of Bolle Zemp’s work, has considerably different lyrics in the realizations by the singers from Lat’ali and from Udabno, respectively. Figure 2 indicates that overall the most common note texts in our transcriptions are {‘*ჟის*’, ‘*ჟილ*’, ‘*ოო*’, ‘*ოოჰ*’, ‘*ჰის*’}⁷.

⁶<http://titus.uni-frankfurt.de/didact/caucasus/kartlaut.htm>,
<http://titus.uni-frankfurt.de/didact/caucasus/kaukvok.htm>

⁷ In transcription: {‘*ûa*’, ‘*ûo*’, ‘*yi*’, ‘*iyh*’, ‘*ha*’}. It should also be noted that some ‘words’, although they are transliterated slightly differently have the same meaning. For example: ‘*ოო*’ and ‘*ო*’ are the same. So are ‘*ო*’ and ‘*ოლ*’, etc.



Figure 2. Word cloud derived from the set of note texts from the complete collection of *zār* recordings (see Scherbaum and Mzhavanadze, 2020). The size of a word corresponds to the relative frequency of its occurrence in the corresponding *zār*.

The main focus of Bolle Zemp’s (1997) interpretation was on the interjection *woj*, which she assumed to be related to the Georgian interjection of pain, *wai*. In Svan, these two interjections would be written as ‘*ჟადო*’ or ‘*ჟოდო*’, which Bolle Zemp perceived as a single word. In our attempts to transcribe the *zār* lyrics, we almost always perceived ‘*ჟ*’ (‘*ûa*’)- ‘*დო*’ (‘*yi*’) and ‘*ჟო*’ (‘*ûo*’) ‘*დო*’ (‘*yi*’) as two different notes and transcribed them accordingly. It goes without saying that the quality of phonetic transcriptions depends strongly on the training of the transcriber. But even for well-trained transcribers, as demonstrated in section 13.4 of Reetz and Jongman (2009), it is very hard not to be influenced in perception by one’s mother tongue. In the present study, although done by natural speakers and experienced singers, the identification of the note lyrics still turned out to be extremely challenging. For example, it was not always straight forward to distinguish between ‘*ჟ*’ (‘*ûa*’), and ‘*ჟო*’ (‘*ûo*’) sounds, or between the different

realizations of vowel sounds involving ‘*o*’ (‘*i*’). For the subsequent analysis, in order to overcome some of the subjectivity of the phonetic transcriptions, we performed a computational acoustical phonetic analysis.

Phonetic analysis

Since each of the individual voices was separately recorded by either a larynx microphone or by a combination of a headset- and a larynx microphone, we performed a so-called acoustical phonetic analysis, based on the analysis of voice formants. Formants are resonance frequencies of the vocal tract, which are related to its shape and controlled by the position of the tongue, the jaw, the lips, and the velum, all of which are referred to as articulators (cf. Sundberg, 1987; Reetz and Jongman, 2009). The articulatory description of *placement* of consonants or *frontness* or *height* of vowels, or the description of vowels in terms of quality or color, as used by Bolle Zemp (1997), can acoustically be related to the values of the formant frequencies. The *height* of a vowel for example is related to the value of the first formant (F1), while the *frontness* (which is related to the position of the arching tongue) is related to the second formant (F2). The first two formants are essential for the identification of vowel sounds, while the higher formants (F3, F4,..) are believed to be more related to the identity of the singer. Formants are associated with peaks in the smoothed power spectrum of the sound. In the present study, we used the Linear Predictive Coding (LPC) algorithm implemented in Mathematica (Wolfram Research, Inc., 2020), which is also used in the standard phonetic software, Praat (Boersma and Weenink, 2021). Since formant frequencies are controlled by the shape of the vocal tract, they are believed to be independent of the fundamental frequency of the vocal fold oscillations (F0) (Reetz and Jongman, 2009).

Figure 3 illustrates the processing sequence of the phonetic analysis. Each of the note objects in Figure 3a) is defined by an onset time, a F0 value, a duration, and a text string representing the lyrics. For a note of interest, the onset time and duration information is used to cut out the corresponding audio segment from the audio file. This is shown in Figure 3b) for a note with the lyrics ‘*ô*’ (‘*û*’) and in Figure 3c) for a note with the lyrics ‘*ô*’ (‘*û*’). The main panels in b) and c) show the spectrograms of the audio segments, the waveforms of which are shown by the gray and brown waveform segments, respectively, on top of the main panels. The small figures to the right of each spectrogram

correspond to the Short-Time Fourier Transform (STFT) of the complete audio segment of the note object. The horizontal red dashed lines indicate the first four formant frequencies (F1, F2, F3, F4), as determined by the Linear Predictive Coding (LPC) algorithm (for details see Wolfram Research, Inc., 2020). Their numerical values are displayed together with the values of the fundamental frequency of the notes in the tables in Figure 3 d) and e).

Finally, F1 and F2 are plotted on a so-called vowel map in Figure 3f), together with the average formant frequencies for the vowels of American English (in magenta), collected by Hillenbrandt et al. (1994), and for German (in blue), by Sendlmeier and Seebode (2006). The positions of the two examples with the lyrics ‘*շն*’ (‘*ûa*’) and ‘*շո*’ (‘*ûo*’) fall roughly on a line between the German ‘*a:*’ and ‘*u:*’.

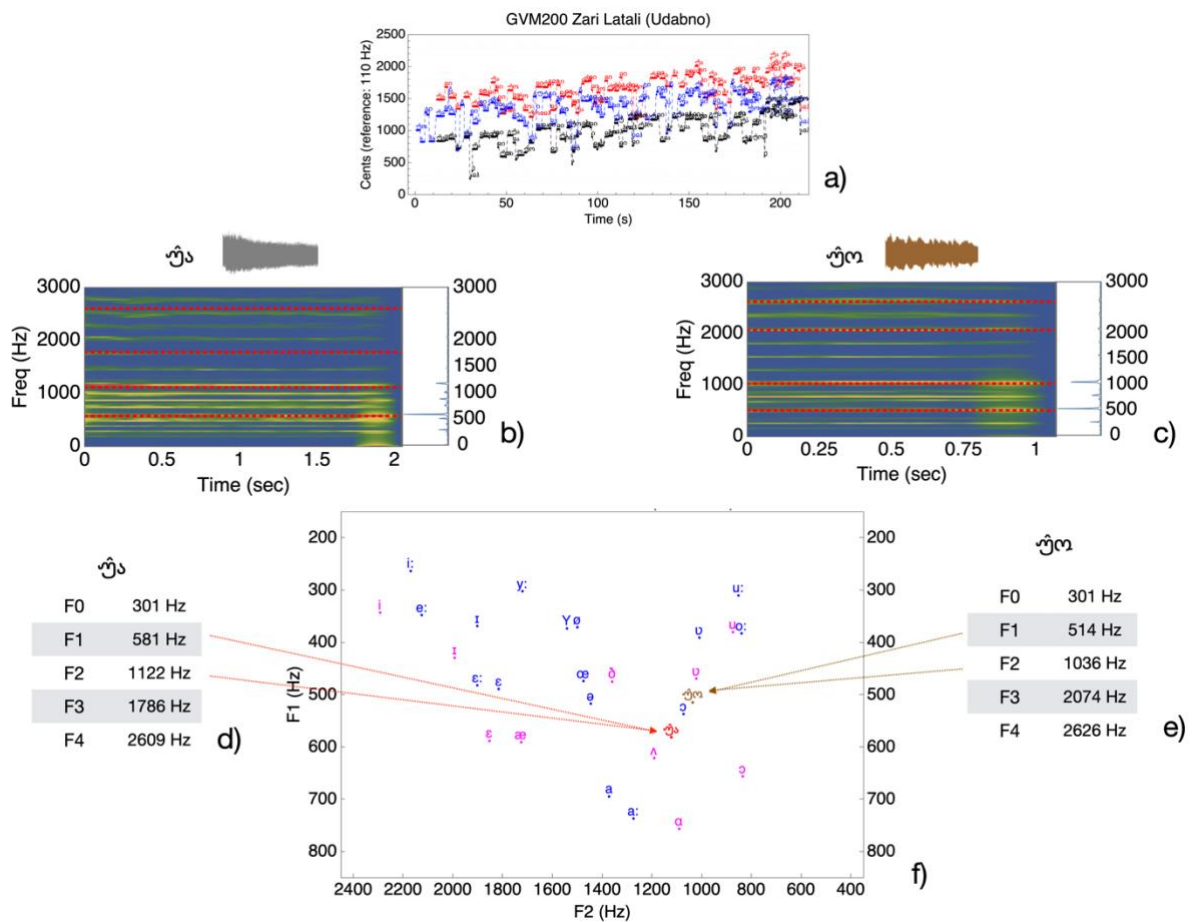


Figure 3. Illustration of the workflow of the phonetic analysis.

Figure 4 shows the vowel map for all notes with lyrics ‘*շն*’ (‘*ûa*’) in red or ‘*շո*’ (‘*ûo*’) in brown in the headset recordings of the Lat’ali *zâr*, sung by the singers from Udabno.

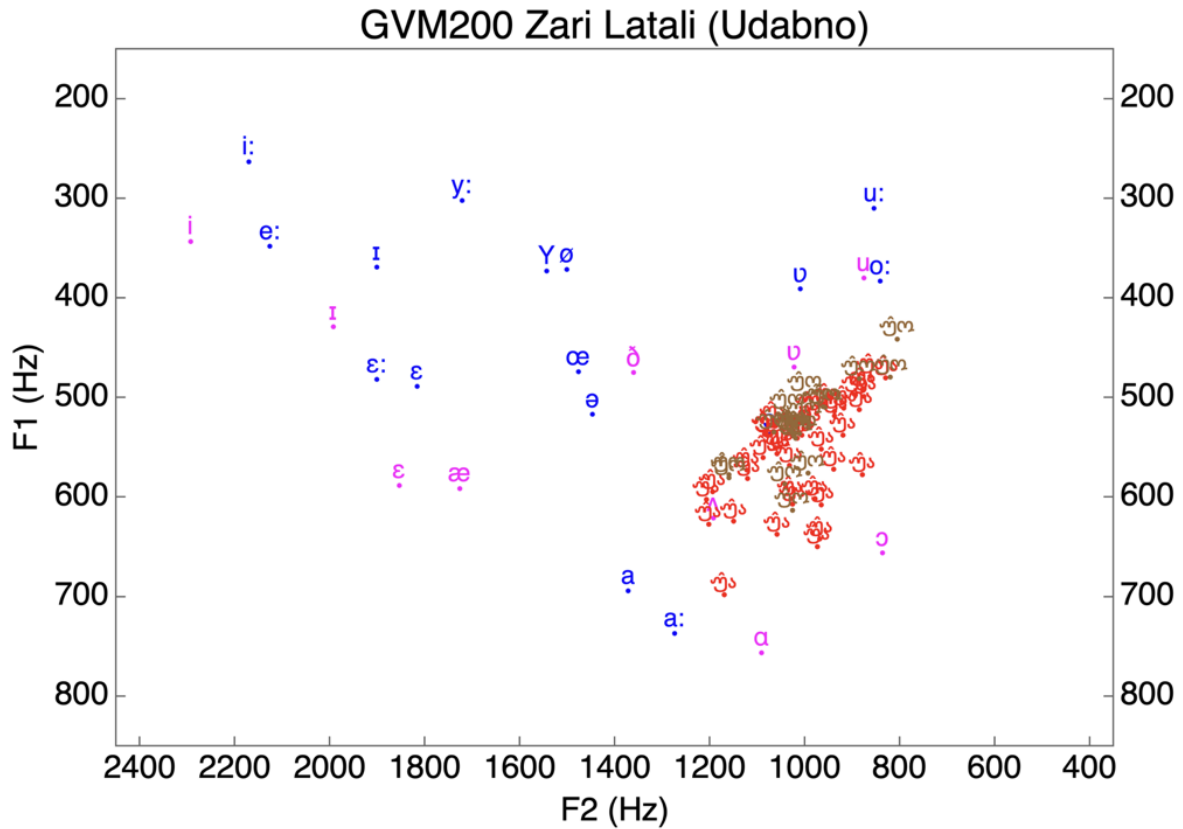


Figure 4. Vowel map for all notes with lyrics ‘*wa*’ (‘*wa*’) in red or ‘*wo*’ (‘*wo*’) in brown in the headset recordings of the Lat’ali *zār*, sung by the singers from Udabno.

It can be seen that the vowel map positions of the individual notes fall again between the German ‘*a*’ and “*u*” sounds, but largely overlap. For once, this is in line with our experience that the two sounds are sometimes difficult to distinguish. The word clouds in Figure 2 show that the syllables ‘*wa*’ and ‘*wo*’ are definitely very present in the phonetic soundscape of all *zār* variants. Comparing the durations of all notes with note lyrics ‘*wa*’ or ‘*wo*’ with the total duration of all notes in the particular *zār* reveals that in terms of durations roughly 50% of all *zār* lyrics in all *zār* variants are spent on either ‘*wa*’ or ‘*wo*’. In about 1/3 of these cases, ‘*wa*’ or ‘*wo*’ appear in combination with either ‘*yi*’ or ‘*o*’ or ‘*iyh*’ as subsequent note.

As a consequence, Bolle Zemp’s conjecture regarding the importance of the interjections *woj* /*wai* seems to be generally applicable to all *zār* variants, if one assumes that they correspond to different combinations of ‘*wa*’ (‘*wa*’) or ‘*wo*’ (‘*wo*’) for the first note text and ‘*yi*’ (‘*yi*’) or ‘*o*’ (‘*o*’) or ‘*iyh*’ (‘*iyh*’) for the second one in our study. But does this mean, as Bolle Zemp (1997) suggested, that these syllables have a strong mourning connotation

and a sematic meaning, which, she suspected, was based on postulated correlations in changes in pitch, duration, and timbre features?

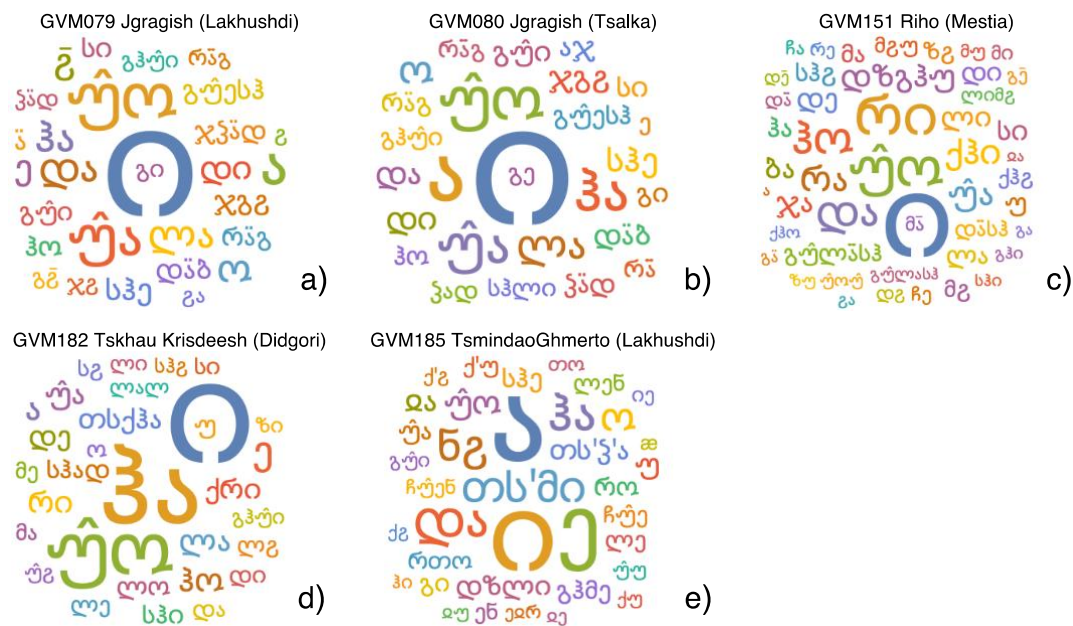


Figure 5. Word cloud derived from the set of note texts from a selection of other Svan song recordings.

As one can see in Figure 5, the syllables ‘*ჴა*’ (‘*ûa*’) or ‘*ჴო*’ (‘*ûo*’) are present in significant amounts in other Svan songs as well. This raises doubts regarding their connotation being limited to mourning.

Since the first formant (F1) is related to the *height* of a vowel, while the second formant (F2) is related to the position of the arching tongue, and therefore, to the amount of velarization, one can check for correlations between changes in pitch, duration, and timbre features by calculating pairwise correlations between F0, F1, F2 and note durations. For all three voices in the Lat’ali *zâr*, sung by the Udabno singers, and for all notes with lyrics ‘*ჴა*’ or ‘*ჴო*’, these so-called Trellis plots⁸ are shown in Figure 6.

⁸ A Trellis plot is a group of smaller plots arranged in a grid in which each subplot is conditioned on a different variable.

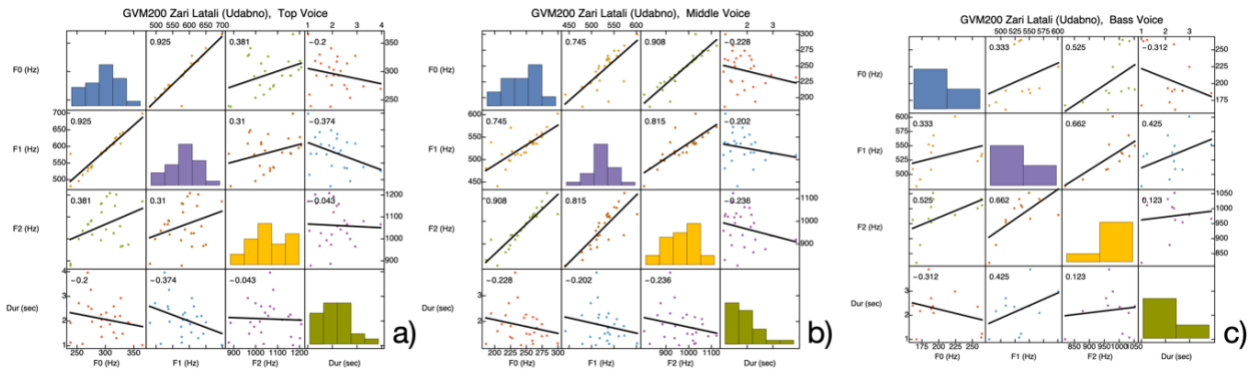


Figure 6. Trellis plot of F0, F1, F2, and note duration of notes with lyrics ‘ژب’ or ‘ژر’ in the Lat’ali zār, sung by the Udabno singers.

With correlation coefficients above 0.9 in both cases, Figure 6 a) and b) show strong correlations between F0 (pitch) and F1 (in relation to vowel height) for the top voice and between F0 (pitch) and F2 (in relation to tongue position) for the middle voice, respectively. For the bass voice, the correlations are much weaker, and hence not really convincing. Closer inspection revealed that the top voice singer shifts his first formant towards $2 \times F0$ (cf. dotted line in Figure 7a), while the middle voice singer shifts his second formant towards the fourth harmonic of the sound spectrum, which has a frequency of $4 \times F0$ (cf. dashed line in Figure 18b). This “formant tracking” (Bozeman, 2013) gives particular harmonics, in the present case the second (H2 at $2 \times F0$) and fourth (H4 at $4 \times F0$), an energy boost, makes them appear louder, rougher, and gives them a ‘ringing’ quality. It is a vocalization strategy well known for both classical and non-classical singing (Bozeman, 2013; Sundberg et al., 2013 and references therein).

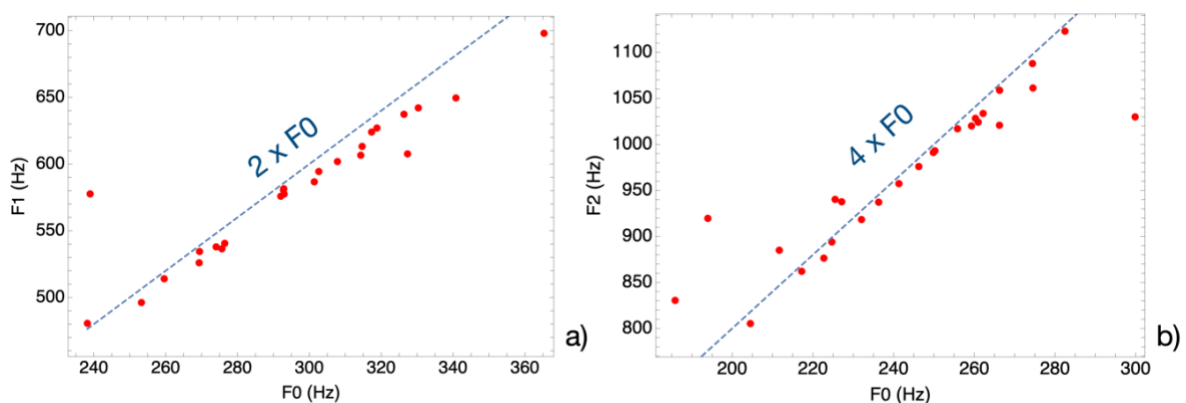


Figure 7. Evidence for formant tuning in the top voice (a) and middle voice (b) recordings of the Lat’ali zār, sung by the Udabno singers.

Men are known in particular to tune their second formant to higher harmonics, if they want a real powerful note, or when their voice turns over (Bozeman and O'Connor, 2017). When harmonics pass through the first formants, one can hear this as a closed timbre ('*voce chiusa*'); the more harmonics are below the first formant, the more the timbre is called open ('*voce aperta*') (Bozeman and O'Connor, 2017)

Since in the present example the correlations in changes in pitch and timbre features (we could not detect a correlation with durations) differ in style between the different voices or are not even used at all, as in the case of the bass voice, we interpret them more as an expression of personal taste of the singers, than as having a semantic meaning. In the latter case, we would expect them to be similar for all three voices.

Conclusions

In this paper, we have investigated the topics of language–music relation and phonetic characteristics of Svan *zār*, first described and analyzed in the pioneering study by Bolle Zemp (1997). As a new contribution to the discourse on these issues, we present a computational analysis of a new set of field recording of Svan *zār*, which comprises audio and video recordings of eleven different performances of five different *zār* variants (Ushguli/K'ala, Mest'ia, Lenjar, Lat'ali, Lower Bal), sung by ensembles from seven different locations (in alphabetical order: Etseri, Kala, Latali, Lenjeri, Mestia, Tsalka, Udabno). The data for our study were audio and larynx microphone recordings of 33 individual vocal tracks, already separated by voice groups during the recording. The substantially expanded scope and the special recording setup used made it possible to address a whole range of questions from a variety of perspectives. One of the goals of our study was to compare the results of the personal analyses of Bolle Zemp (which were based on the recordings of a single performance of the Latali variant only) with the computational analysis of the new and expanded data set.

The particular results and conjectures from the Bolle Zemp studies, which we wanted to revisit were: a) the importance of the interjections *woj /wai* in *zār*, b) the possible mourning connotation of these syllables, and c) the postulated correlations in changes in pitch, duration, and timbre features. Since all voices are available separately in digital form in the new dataset, we were able to address the last question quantitatively, by using

computer-aided acoustic-phonetic analysis. The findings we have obtained with regard to the various aspects can be summarized as follows.

As to a). Although the Lat'ali variant, which was the subject of Bolle Zemp's work, has considerably different lyrics in the realizations by the singers from Lat'ali and from Udabno, we found that overall Bolle Zemp's observation of the frequency of occurrence of the interjections *woj* (*wai*) in *zār* is consistent with the expanded dataset, if one assumes that they correspond to different combinations of 'ჟა' ('wa') or 'ჟო' ('wo') for the first note text and 'ოი' ('yi') or 'ო' ('i') or 'ოჟ' ('iyh') for the second one in our study.

As to b). What we could not confirm, since the syllables 'ჟა' ('ûa') or 'ჟო' ('ûo') are present in significant amounts in other Svan songs unrelated to mourning, is their exclusive mourning connotation.

As to c). Based on a computational phonetic analysis, we could detect the existence of correlations in changes in pitch and timbre features related to the interjection of *woj* (*wai*) for some, but not for all of the singers. If it would have a semantic meaning, as suggested by Bolle Zemp, we would expect it to be employed by everyone. Closer inspection of singers, whose recordings showed these correlations, revealed that they employed formant tuning, a vocal technique, which is often employed by singers to make their voice appear louder, and to give it a particular rough sound quality. Therefore, in contrast to Bolle Zemp, we do not assign a semantic meaning to these correlations, but interpret them as a natural consequence of the formant tuning technique and an expression of the personal taste of the singers.

Acknowledgements

First and foremost, our gratitude goes to all the people, who allowed us to be part of and record their rituals during the 2016 field expedition. We are thankful to Meinard Müller and his team for the stimulating collaboration within the framework of the project "Computational Analysis of Traditional Georgian Vocal Music (GVM)" (DFG MU 2686/13-1, SCHE 280/20-1), as well as for hosting the web-based repository of the GVM data. We would also like to thank Adamantios Gafos for stimulating discussions and his encouragement during the initial stages of this study.

REFERENCES

- Boersma, Paul and Weenink, David. (2021). Praat: doing phonetics by computer [Computer program]. (Version 6.1.48), Retrieved from <http://www.praat.org/>. (17 June 2021)
- Bolle Zemp, Sylvie. (1994). Géorgie: Polyphonies de Svanétie (Georgia: Polyphony of Svaneti.) [Liner notes], le Chant du Monde: Diffusion Harmonia Mundi.
- Bolle Zemp, Sylvie. (1997). "Mehrstimmige Wehrufe? Ein Begräbnisgesang aus Swanetien" (Multi-Voiced Laments? A Funeral Chant From Svaneti). *Georgica*. 20: 134-148.
- Bolle Zemp, Sylvie. (2001). Khmovnebi da Ak'ordebi. Simghera Zemo Svanetshi (Vowels and Chords. Singing in Zemo Svaneti). *Sasuliero Da Saero Musikis Mravalkhmianobis Problemebi* (Problems of Polyphony in Sacred and Secular Music), Ed. Tsurtssumia, Rusudan: pp. 292–303. Tbilisi: Tbilisi State Conservatoire.
- Bozeman, Kenneth W. (2013). *Practical vocal acoustics*. Hillsdale, NY: Pendragon Press.
- Bozeman, Kenneth W.; O'Connor, Karyn (Writers). (2017, 17.06.2021). *Formant Tuning, Vowel Modification, Male and Female Passaggio, and Registration*. [Video File]. Retrieved from <https://www.youtube.com/watch?v=dxttfvX-FRs>
- Hillenbrand, James; Getty, Laura A.; Wheeler, Kimberlee and Clark, Michael J. (1994). "Acoustic characteristics of American English vowels" *The Journal of the Acoustical Society of America*. 95: 2875-2875.
- Mzhavanadze, Nana; Scherbaum, Frank. (2020). "Svan Funeral Dirges (Zär): Musicological analysis" *Musicologist*. 4(2): 168-197.
- Mzhavanadze, Nana; Scherbaum, Frank. (2021). "Svan Funeral Dirges (Zär): Cultural Context", accepted for publication in *Musicologist*.
- Reetz, Henning; Jongman, Allgard. (2009). *Phonetics. Transcription, Production, Acoustics, and Perception*. Massachusetts: Wiley-Blackwell.

Scherbaum, Frank. (2016). "On the benefit of larynx-microphone field recordings for the documentation and analysis of polyphonic vocal music" *Proceedings of the 6th International Workshop Folk Music Analysis* [Paper presented at the 6th International Workshop Folk Music Analysis on June 15-17, 2016, Dublin, Ireland.]. Beaguitte, Pierre; Duggan, Bryan; Kelleher, John D. (Eds.), (pp. 80-87). Retrieved from <https://arrow.tudublin.ie/fema/8/>

Scherbaum, Frank; Loos, Wolfgang; Kane, Frank and Vollmer, Daniel. (2015). Body vibrations as source of information for the analysis of polyphonic vocal music. Paper presented at the *5th International Workshop on Folk Music Analysis, June 10-12, 2015, University Pierre and Marie Curie*.

Scherbaum, Frank; Mzhavanadze, Nana. (2018). "A new archive of multichannel-multimedia field recordings of traditional Georgian singing, praying, and lamenting with special emphasis on Svaneti" *LaZAR-Database*. Retrieved from <https://lazardb.gbv.de/>

Scherbaum, Frank; Mzhavanadze, Nana. (2020). "Svan Funeral Dirges (Zär): Musical Acoustical Analysis of a New Collection of Field Recordings" *Musicologist*. 4(2): 138-167.

Scherbaum, Frank; Mzhavanadze, Nana; Arom, Simha; Rosenzweig, Sebastian and Müller, Meinard. (2020). "Tonal Organization of the Erkomaishvili Dataset: Pitches, Scales, Melodies and Harmonies" *Computational Analysis of Traditional Georgian Vocal Music*. 1: 1-64.

Scherbaum, Frank; Mzhavanadze, Nana; Rosenzweig, Sebastian and Müller, Meinard. (2019). Multi-media recordings of traditional Georgian vocal music for computational analysis. Paper presented at the *9th International Workshop on Folk Music Analysis* Birmingham, England, 2-4 July 2019.

Sendlmeier, Walter F; Seebode, Julia. (2006). "Formantkarten des deutschen Vokalsystems" (Format maps of the German vowel system). Retrieved from https://www.kw.tu-berlin.de/fileadmin/a01311100/Formantkarten_des_deutschen_Vokalsystems_01.pdf

Sundberg, Johan. (1987). *The science of the singing voice*. Illinois: Northern Illinois University Press.

Sundberg, Johan; Lã, Filipa M.B. and Gill, Brian P. (2013). "Formant tuning strategies in professional male opera singers" *Journal of Voice*. 27: 278-288.

Wolfram Research, Inc. (2020). "Mathematica (12.3)" [Computer program]. Champaign, Ill., USA: Wolfram Research, Inc.