Analysis of the Tbilisi State Conservatory Recordings of Artem Erkomaishvili in 1966

Frank Scherbaum
University of Potsdam
fs@geo.uni-potsdam.de

Meinard Müller
International Audio Laboratories Erlangen
meinard.mueller@audiolabs-erlangen.de

Sebastian Rosenzweig
International Audio Laboratories Erlangen
sebastian.rosenzweig@audiolabs-erlangen.de

ABSTRACT

In this paper we try to obtain information regarding the musical thinking of Artem Erkomaishvili, one of the last master chanters of traditional Georgian chant. For this purpose, we analyse the recently determined F0-trajectories (Müller et al., 2017) for a set of chant recordings from 1966 in which Artem Erkomaishvili sang all three voices sequentially using two tape recorders in overdubbing mode. The purpose of our study is to determine the tuning of Artem Erkomaishvili's voice and how it compares to the models proposed by various researchers to reflect (in their opinion) the historical Georgian tuning. The analysis of the melodic pitch inventory shows that the sizes of melodic seconds sung by Artem Erkomaishvili vary over a range from approximately 140 to 240 cents, with a peak of the distribution at approximately 180 cents. We do not see evidence for any attempt to precisely use any particular or a small set of melodic interval sizes, as is suggested by some of the proposed tuning models. The harmonic analysis yields an interval distribution which is peaking at justly tuned fifths and octaves at 698 and 1203 cents, respectively. No observational evidence for stretched octaves, as suggested by some models, is seen. Analysing the joint pitch distribution, we find evidence for considerable voice interaction in which Artem Erkomaishvili maintained harmonic intervals despite considerable pitch fluctuations of the individual voices. In short, Artem Erkomaishivli's performance in 1966 seems to reflect a combination of strong harmonic and relaxed melodic thinking.

1. INTRODUCTION

Artem Erkomaishvili (1887-1967) is known today as a key representative of traditional Georgian singing of the 20th century and one of the last grand masters of Georgian chanting (*sruligalobelni*) (cf. Graham, 2015). In 1966, one year before his death, he was asked to perform all voices of a series of chants to save them for posteriority. His performance, part of which was recently remastered (Jgharkava, 2016), was recorded at the Tbilisi State Conservatory using two tape recorders, which were subsequently operated in what is now called overdubbing. The recordings were transcribed by Shugliashvili (2014). Although the use of the overdubbing setup originated from the lack of fellow chanters who could perform the repertoire, it turned into an advantage in view of an analysis of this data. Despite the fact hat polyphonic pitch analysis is still considered an enormous challenge in general situations, the sequential overdubbing considerably simplifies the task of determining the fundamental frequencies F0 (which for simplicity we will also refer to as pitches) for all voice segments. Details of the processing techniques can be found in Müller (2015). The corresponding time-stamped F0-trajectories have been made publicly available[1].

In the present paper, which is a direct follow-up study to Müller et al. (2017), we want to find out what we can learn from this unique set of recordings (respectively analysis results) regarding the characteristics of the tuning system(s) used by Artem Erkomaishvili. The topic of the authentic, historical Georgian tuning system has been a matter of intense and controversial discussion for a number of years, resulting in the proposition of several scale and/or tuning models which have little in common other than the untempered nature of the music (Erkvanidze, 2002; Gelzer, 2002; Westman, 2002; Gogotishvili, 2004 ; Kawai et al, 2010; Tsereteli and Veshapidze, 2014; Erkvanidze, 2016). Based on the analysis of recent field recordings in Svaneti/Georgia, Scherbaum (2016) took a conceptually different perspective on the issue of Georgian tuning systems. Since he found considerable differences in the sequential (melodic) and the concomitant (harmonic) intervals used by traditional singers, he concluded that a single scale/tuning model might not capture the complete tuning characteristics of Georgian vocal music. Instead, in line with Nikolsky (2015), he separately analysed the melodic and the harmonic pitch/interval inventory of the music. In the present study we go one step further and separately analyse the pitch organization in the recordings of Artem Erkomaishvili from a melodic, a harmonic and a voice interaction perspective.

The main part of our study is devoted to the attempt to use the time-stamped F0-trajectories of the individual voices in Artem Erkomaishvili's recordings to determine the associated melodic and harmonic pitch and interval inventories and to investigate how listening to pre-recorded voices affected the tuning of Artem Erkomaishvili's singing. The results are discussed in the context of the predictions of the tuning models suggested by various researchers to reflect (in their opinion) the authentic, historical Georgian tuning practice(s). Our results suggest that voice interaction effects, evidence for which can clearly be seen in the recordings of Artem Erkomaishvili, should be included in the discussions of the tuning systems of traditional Georgian (and possibly other) vocal music. This might require a shift of attention from the purely melodic to the combined melodic-harmonic aspects of the music.

The paper is organized as follows. Following a brief recapitulation of the recording setup and the extraction of the F0-trajectories by Müller et al. (2017), we discuss the determination of the melodic aspects of the performance of Artem Erkomaishvili (Section 2.1). For the top voice segments we show that the individual F0-values, which make up the pitch tracks, exhibit a strong pitch clustering. We interpret the pitch values of the cluster centers (which we determine by k-means cluster analysis) to indicate the pitches of the notes of the mental melodic template Artem Erkomaishvili might have been using during his performance. From the pitch values of the cluster centers for the complete dataset, we determine the set of possible single-step melodic intervals for the complete performance. We compare (as a spot check) the properties of the resulting

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

ISBN: 978-84-697-2303-6

distribution with the results of a note analysis for a single chant using the Tony software (Mauch et al., 2014, 2015), and with the values of the single-step interval sizes from the predictions of some of the published tuning models for Georgian vocal music. Subsequently (Section 2.3), we discuss the harmonic aspects of the tuning used by Artem Erkomaishvili. In this context, we make use of the time-stamps for the individual voice segments determined by Müller et al. (2017) to estimate the F0-values for the concomitantly sung (harmonic) intervals. These also show a strong clustering, the properties of which we interpret to reflect the mental harmonic template Artem Erkomaishvili might have been using during his performance. As final aspect of our analysis, which according to our knowledge has previously been ignored in quantitative investigations of tuning in Georgian vocal music, we investigate (in Section 2.4) the joint pitch distribution of voice combinations for signatures of voice interactions. Considering a single chant, we find several instances in which Artem Erkomaishvili evidently maintained harmonic intervals despite considerable pitch fluctuations of the individual voices. Finally, in section 3, we conclude with a discussion of the main results of our study and their consequences for future work.

2. PITCH AND INTERVAL ANALYSIS

Figure 1 sketches the three-stage concept used during the recordings of Artem Erkomaishvili in 1966. In the first stage, only the lead (top) voice of a chant was recorded. In the second stage, Artem Erkomaishvili was singing the middle voice while listening to the recording of the lead voice. During the recording of the bass voice, he listened to the overdubbed recordings of the middle and top voice. The extraction of the F0-trajectories was also performed in a segmented way in that the extracted F0-trajectory for the first segment was used as constraint for the extraction of the F0-trajectory of the second segment, and so forth. For details of the analysis see Müller et al. (2017).

To make this audio collection better accessible for musicological research, one important task is to estimate the fundamental frequency (F0) trajectories of the sung pitches from the recordings using automated methods. While this is feasible with standard procedures in the case of monophonic music, the problem becomes much harder in the case of polyphonic music. In Müller et al. (2017), a graphical user interface (GUI) for semi-automatic estimation of F0 trajectories was introduced. The GUI allows a user to specify temporal-spectral constraint regions that guide the estimation process. Furthermore, the GUI provides visual and acoustic feedback mechanisms that can be used to control and refine the estimated results in an interactive fashion. In Müller et al. (2017), we applied this GUI for extracting the F0 trajectories of the sung pitches from the three-voice chant recordings performed by Artem Erkomaishvili. To this end, we first determined the recordings' structures based on the three-stage recording setup (see Figure 1). Subsequently, we determined the F0-trajectories for the lead, middle, and bass voices from the first, second, and third section, respectively. To this end, suitable visualization and sonification functionalities helped us in determining suitable constraint regions to guide the estimation process. All results, including the original recordings, figures of the visual representations, the estimated F0-trajectories, and the sonifications of these trajectories, have been made publicly available.[1] These results serve as an important basis for our subsequent analysis.

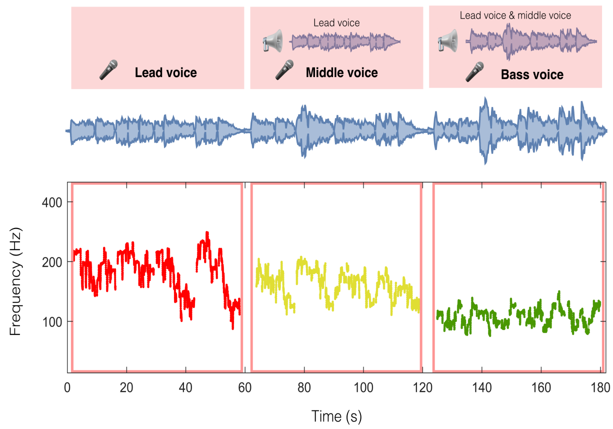

Figure 1. Sketch of the three-stage recording setup (top panel), the recorded waveforms (middle panel), and the F0-trajectories derived for the individual voices of chant no. 2 (Shugliashvili, 2014). The pink rectangles indicate the structure of the three-stage recording process.

2.1 Melodic Analysis

In the first step of our analysis, we determined the melodic pitch inventories of the lead (top) voice segments. These were always sung first and individually. We assume that, in case Artem Erkomaishvili believed that a particular chant should be performed in a specific scale, this will show up as a clustering of pitches around the intended "scale pitches" for this voice segment. This can be seen in Figure 2 for the chant *Aghdgomasa Shensa* (referred to by its chant ID no. 2 in Shugliashvili, 2014).

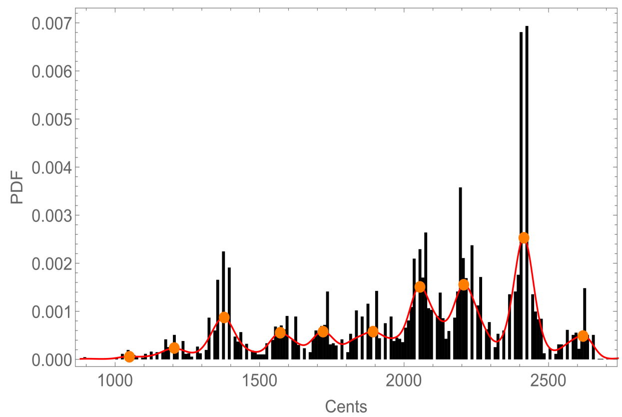

Figure 2. Pitch histogram (vertical axis scaled to match the sample PDF) and smooth kernel distribution (red solid line) of the F0-values in the top voice of chant no. 2. Note the clustering of the pitch samples. The reference note for all absolute cent calculations is A1 (55 Hz).

In order to determine what we assume to be the intended scale pitches quantitatively, we performed a formal cluster analysis (using the k-means algorithm) to determine the locations of the centers of the F0-clusters and the corresponding spreads. Figure 3 shows the resulting separa-

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

ISBN: 978-84-697-2303-6

tion of the pitch set of the top voice of chant no. 2 into 11 pitch clusters.

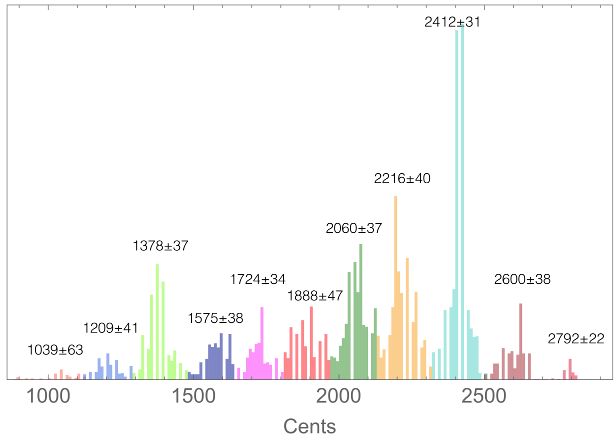

Figure 3. Pitch cluster histograms of the F0-values in the top voice of chant no. 2. The number on top of each cluster shows the cluster mean and the corresponding cluster standard deviation (in cents). The vertical axis is proportional to pitch sample PDF. The labelling of the vertical axis, which is unimportant in the present context, was omitted on purpose on this and similar plots to increase the plot size.

What can be seen in Figure 3 is that the F0-values seem to cluster in such a way that an octave (here e. g. the interval between the cluster at 2412 and the cluster at 1209 cents, which spans 1197 cents) is divided into seven intervals of different sizes. This seems to support the interpretation of the clusters as marking the “scale pitches” of the mental tuning template which Artem Erkomaishvili was using during his performance of the chant. Figure 4 shows the melodic line of the top voice of chant no. 2 as a trajectory through the different pitch clusters.

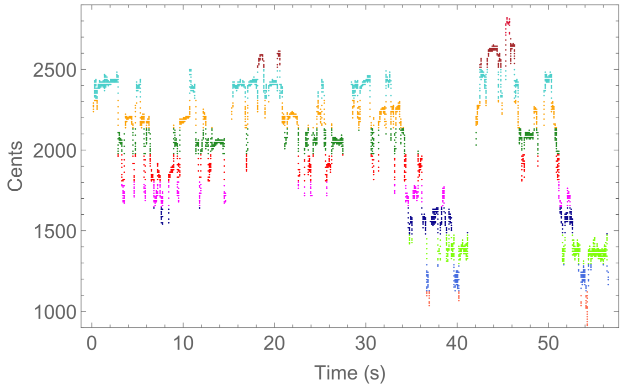

Figure 4. Pitch track of the top voice of chant no. 2, color coded according to cluster membership.

At first glance, one might believe that the spread of the clusters, as indicated by the sample standard deviations in Figure 3, but also the jitter of the pitch track in Figure 4, are rather large since they reach values of one quarter to one half of a semi-tone (25 – 50 cents). However, this is not surprising and must not be seen as a sign of poor pitch control of the singer. For once it is to be expected as an expression of the categorical perception of pitch (e. g. Siegel & Siegel, 1977; Sundberg, 1994). In addition, sliding phases in the beginning of new syllables, breathing, vibrato and consonants all affect the temporal stability of the F0-trajectories. In order to test to what degree these effects, but also the pitch algorithm itself, might affect the determination of the “scale pitches”, we performed an alternative pitch determination using the Tony software (Mauch et al., 2014, 2015). In this context, we visually edited the pitch tracks to remove all what could be considered artifacts of sliding phases in the beginning of new syllables, breathing, vibrato, and consonants. Subsequently, pitch tracks as well as notes, yet another way to determine the pitches for this example, were calculated. The resulting histogram is shown in Figure 5.

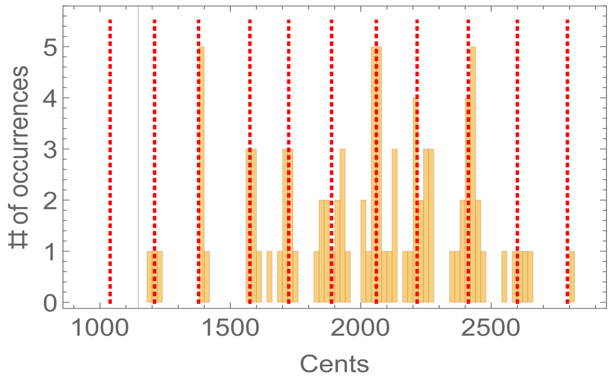

Figure 5. Histogram of note pitches, determined with the Tony software (Mauch et al., 2014, 2015) for the top voice of chant no. 2. The red dotted lines mark the locations of the pitch cluster centers displayed in Figure 1 as determined from the F0-trajectories.

Figure 5 shows that the pitch-cluster mean values determined from the raw F0-trajectories are a reasonable representation of the histogram distribution of the notes, as determined after visual editing of the pitch tracks. As final test of the robustness of the pitch-cluster centers, we performed a k-mean cluster analysis on the individual pre-edited pitch values as determined by the Tony software (based on the PYIN algorithm). The resulting pitch histograms are shown in Figure 6.

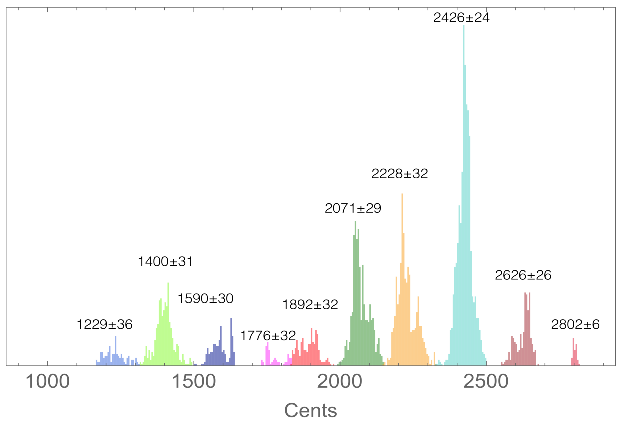

Figure 6. Pitch cluster histograms of the pitch samples in the top voice of chant no. 2 as determined with the PYIN algorithm in the Tony Software (Mauch et al., 2014, 2015. The number on top of each cluster shows the cluster mean and the corresponding cluster standard deviation (in cents).

Overall, comparing the F0-distributions in Fig. 3 and Fig. 6, the set of F0-cluster means in Fig. 3 is shifted by approximately 20 cents towards lower values with respect to the set of cluster means calculated from the pitch samples determined with the Tony software in Fig. 6. This shift might be due to the preprocessing of the pitch trajectories and the removal of presumed artifacts, e. g. glissandi at the beginning of new syllables which tend to start from sometimes rather low pitch values (cf. Figure 4). If

ISBN: 978-84-697-2303-6

one would remove this constant shift, eight of the ten corresponding peaks in the two sets of cluster centers would be less than 10 cent apart, one 15 cent, and one (the smallest cluster in Figure 6 between 1700 and 1800 cents) 33 cent. Based on this result, we might assume that the vast majority of interval sizes, calculated as differences between the cluster centers of neighboring pitch clusters, carry an average uncertainty of less than 10-15 cent.

From the analysis of all chants for which the pitch range of the top voice covers more than an octave (58 out of 101), we obtain 467 intervals between neighboring cluster centers. Their histogram distribution is shown in Figure 7.

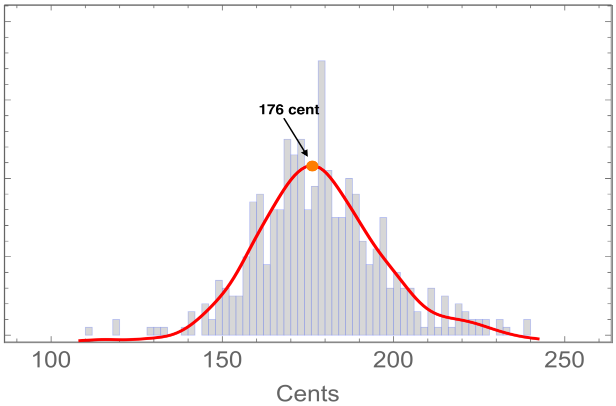

Figure 7. Distribution of possible single-step melodic interval sizes, determined from the top voices of all 58 chants for which for which the pitch range covers more than an octave. The red line corresponds to the 5-point smooth kernel distribution calculated from the histogram data.

No particular preference is seen for any of the interval sizes characterizing seconds in the scale models by Erkvanidze (2002, 2016), by Tsereteli and Veshapidze (2014), or in the tempered diatonic model, respectively. Erkvanidze (2002) suggests that the authentic Georgian tuning system uses interval sizes for single melodic steps that can take values of either 154, 172, or 204 cents. One can see that none of these interval sizes is incompatible with Figure 6, but so are many other intervals between 140 and 220 cents. Tsereteli and Veshapidze (2014) on the other hand suggest a seven-interval scale of equal interval size of 1200/7 = 172 cents. This value, which is also one of the interval sizes in the Erkvanidze model, is actually very close to the peak value of the distribution (176 cents), but there is no visual evidence that Artem Erkvanidze has intentionally tried to achieve this with any precision. Finally, western tempered diatonic scales assume single melodic step sizes of either 100 (semitone) or 200 cents, the first of which is completely absent in Figure 7.

One has to note, however, that Figure 7 does not show the frequency distribution of single melodic steps of notes which were actually sung in a particular chant. It provides an overall view of the possible melodic single-step sizes for the whole corpus. In order to perform a simple test of how different these may be, we used the Tony software (Mauch et al., 2014, 2015) to determine the sung notes (instead of the pitch track centers) in chant no. 2 and to calculate the melodic intervals based on them. The results are shown in Figures 8 and 9.

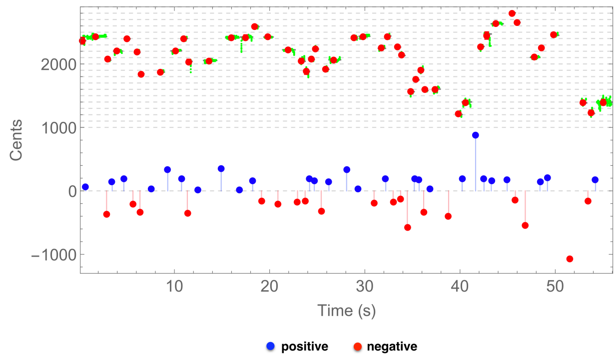

Figure 8. Notes (red blobs plotted at the note pitches in the upper part of the figure, superimposed by the pitch track segments in green) and melodic step sizes (vertical lines in the lower part of the figure). The melodic steps are color coded according to their direction (up- blue, down- red).

In Figure 9a) the distributions of the melodic step sizes is shown independent of direction while in Figure 9b) and c) the distributions are split up according to upwards (b) and downwards (c) movements. As a note on the side, we want to mention that the upwards steps taken by Artem Erkomaishvili in this example seem to be little larger on average than the downward steps. In field observations of traditional village singers in Upper Svaneti (Scherbaum, 2016) this was observed as a systematic feature, which might point to a more general performance element of Georgian vocal music which deserves further study.

It can be seen in Figure 9a) that the central body of the step size distribution for single melodic steps in this example covers a similar range of approximately 140 - 220 cent than the distribution for all possible step sizes shown in Figure 7. In other words, even in a single chant, the variability of melodic seconds is not found to be reduced.

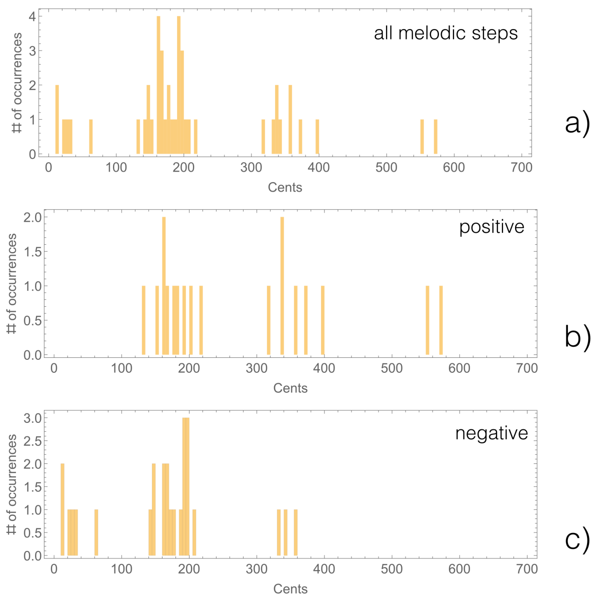

Figure 9. Intervals between notes in the top voice of chant no. 2 (vertical lines) and corresponding histogram. Fig. 9a) shows all steps sizes independent of direction, while these are split up according to positive (b) and negative (c) in the two lower panels.

ISBN: 978-84-697-2303-6

Figure 7 to 9 suggest that the mental tuning templates Artem Erkomaishvili might have been using during the performance of the chants do not seem to be very rigid regarding the single-step melodic interval sizes, which can also be referred to as "melodic seconds". For the whole set of 58 chants analysed, the values range approximately between 140 and 240 cent, which corresponds to a semitone in the 12-tone equal tempered scale. There is no visual evidence for any intention to precisely sing any particular melodic scale.

Several questions arise in this context. Does the lack of evidence for precision of the melodic seconds tell us anything regarding the validity of any of the scale models proposed for Georgian music? Is it unintentional or intentional, in other words does it characterize uncertainties or is it actually an important feature of the music and serves a particular purpose? Before we get back to these questions in the discussion section, we are going to look at the other voices and their interaction with each other.

2.2 Harmonic Analysis

To analyze the harmonic tonal organization in the recordings, we realigned the individual voice segments to a common start time. The start and end times for the individual segments were obtained manually and are publiccly available at the website[1] accompanying Müller et al. (2017). First, we aligned all voice segments to a common zero start time. Subsequently, we selected only those F0-samples for which all three voices are active (with valid F0-values). For chant no. 2 this results in the F0-trajectories shown in Figure 10.

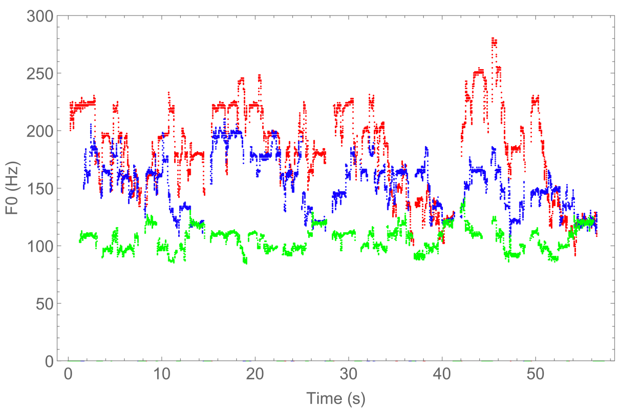

Figure 10. F0 trajectories for the aligned voices of chant no. 2. The top, middle and bass voices are plotted in red, blue, and green, respectively.

Subsequently, we determined the F0-values for all the concomitant pitches from which we calculated the harmonic intervals. These were again subjected to a cluster analysis to quantitatively determine the harmonic structure of the chant (Figure 11).

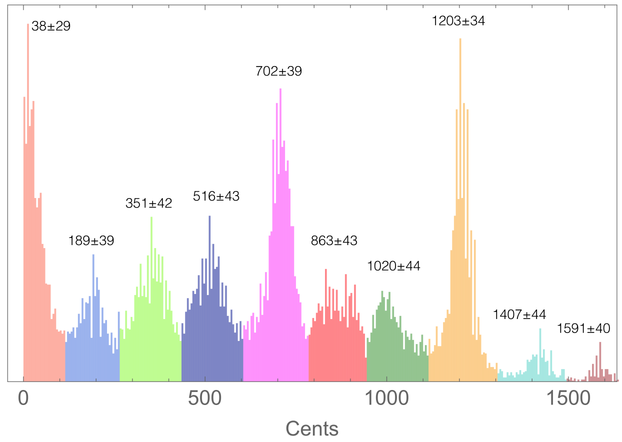

Figure 11. Distribution of concomitant (harmonic) intervals in chant no. 2 and derived clusters thereof. The numbers indicate the cluster means and standard deviations.

One can see in Figure 11 that the most prominent harmonic pitch cluster centers occur around 38 cents, 702 cents (a perfectly justly tuned fifth) and at 1203 cents (a perfectly tuned octave). The harmonic thirds are close to neutral with a cluster center at 351 cents, while the fourths at 516 cents appears sharper than a justly tuned fourth (which would be at 498 cents). Performing the same kind of analysis to all 44394 harmonic intervals in the analysed corpus results in the distribution shown in Figure 12. The general picture remains very similar to Figure 11, except that the harmonic seconds get closer to the tempered value of 200 cents, moving farther away from the distribution of the melodic intervals (cf. Figure 7). Overall, the fifth is the most frequent harmonic interval occuring in the complete corpus.

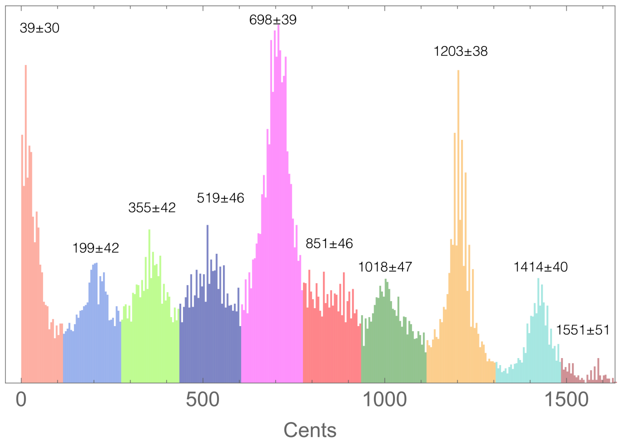

Figure 12. Distribution of all 44394 concomitant (harmonic) intervals in all 58 chants of the corpus for which the top voice covers a range of more than one octave, separated into pitch clusters. The numbers indicate the cluster means and standard deviations.

2.3 Voice Interaction

When Artem Erkomaishvili was singing the middle voice, he was listening to the top voice played back to him from one of the tape recorders. Similarly, he would listen to the recording of the overdubbed top and middle voices when singing the bass. Can one tell from the F0-trajectories, if hearing another voices affects his singing?

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

ISBN: 978-84-697-2303-6

If we look at the individual F0-distributions for the different voices shown in Fig. 13, all one can see is that the pitch clusters seem to be pretty much in phase with a similar spread.

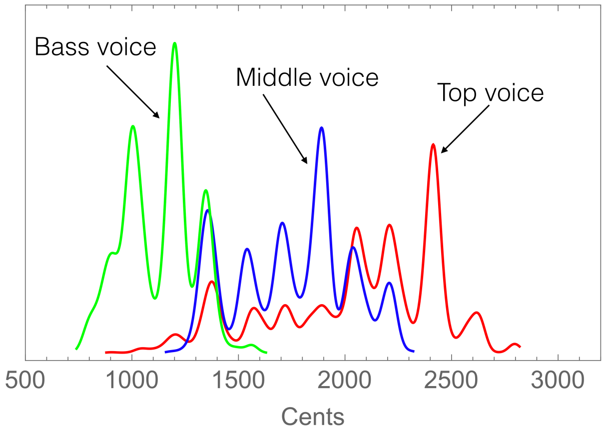

Figure 13. Smooth kernel distributions of the F0 values for the top voice (red), middle voice (blue), and bass voice (green) for chant no. 2.

One way to identify possible voice interactions is by studying the joint distributions of concomitant pitches. These are shown in Figs. 14 to 16 for the middle-top voice, the bass-middle voice, and the bass-top voice pairs, respectively. Each dot represents a pair of simultaneously sung pitches. Jointly sung notes will appear in this plot as a two-dimensional cluster of dots. The x- and y- coordinates of a note cluster should be close to one of the cluster centers for the individual voices shown in Fig. 13. For example the x- coordinates of any of the clusters in Figure 14 (middle against top voice) should be close to one of the peaks of the middle voice (blue curve) in Figure 13, while the corresponding y-coordinates should be close to one of the peaks of the red curve representing the top voice pitches. The reason for this is simply that mathematically speaking the blue and red distributions in Figure 13 are the marginal distributions to the joint distribution of pitch pairs shown in Figure 14. The tilted lines in Figure 14 correspond to different harmonic intervals between the top and the middle voice. The solid black line indicates unisone. So if the two voices are in perfect unisone, the corresponding pitch dot would plot exactly on the solid black line. If the top voice would be exactly 200, 350, 500, or 700 cents above the middle voice, the corresponding pitch dot would plot on the dashed orange, the dashed green, the dashed blue or the solid red line, respectively.

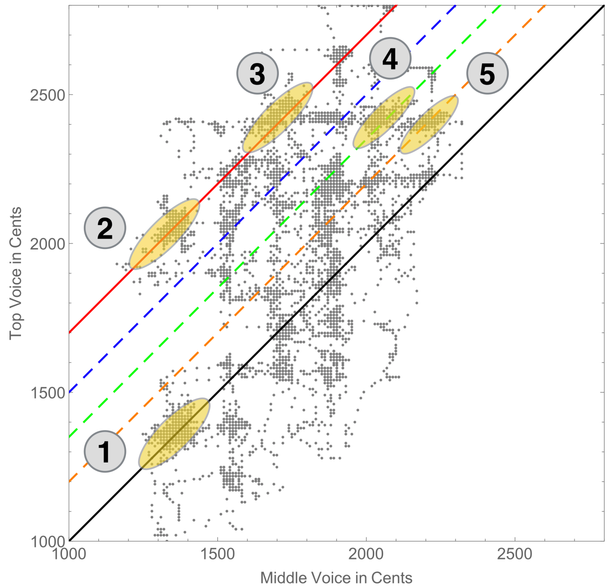

Figure 14. Concomitant middle-top voice pitch pair sample distribution of chant no. 2.

It is the shape of the two-dimensional clusters which tells us if the pitch of the middle voice is influenced by the pitch of the top voice heard. Lets assume, for example, that the top voice sings a note in which the mean pitch is at 1300 cents and fluctuates within a range of ± 20 cents. If the middle voice wants to sing the same note it will also produce a range of pitch values fluctuating by some amount, lets say also ± 20 cents. If the two fluctuations will be compleley independent of each other, say Gaussian, the two-dimensional distribution of pitch pairs will be a two-dimensional Gaussian distribution which would be visible as a distribution around the center point which looks similar in all direction (circular). If, on the other hand, the middle voice would be absolutely stable (no fluctuation at all), one would see a vertical alignment of the two dimensional pitch cluster for that note. If the top voice is stable, but the middle voice fluctuates, then the alignment of the cluster should be horizontal. If, however, the top voice fluctuates by some amount and the middle voice wants to maintain a particular harmonic interval, it must sing in such a way that the middle voice will fluctuate in phase with the top voice by exactly the same amount. In such a case, the note cluster would show an alignment of exactly 45 degrees. In Fig. 14 we can identify several of these structures labeled by numbers. Note cluster 1 in Figure 14, for example, represents a situation in which top and middle voice maintain unisone despite the fact that the voices fluctuate by a considerable amount (by roughly 100 cents). Note clusters 2 and 3 represent situations in which the middle voice sings a stable 5th below the top voice while both voices fluctuate by approximately 100 cents. Note cluster 4 and 5 indicate similar situations for a harmonic neutral third and a harmonic major second, respectively.

ISBN: 978-84-697-2303-6

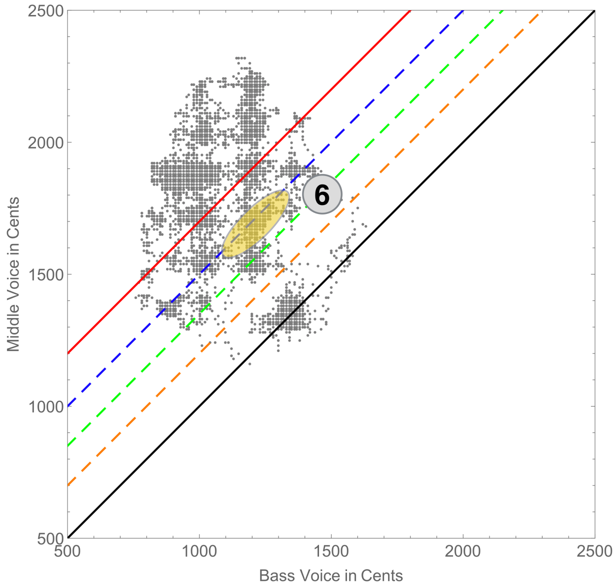

Figure 15. Concomitant bass-middle voice pitch pair sample distribution of chant no. 2.

In the bass-middle voice pitch distribution shown in Figure 15, one can identify more note clusters which are either vertically or horizontally aligned, meaning that there was no or little voice tuning of the bass voice. There is one structure (labeled 6), however, in which a harmonic fourth is attempted to be maintained.

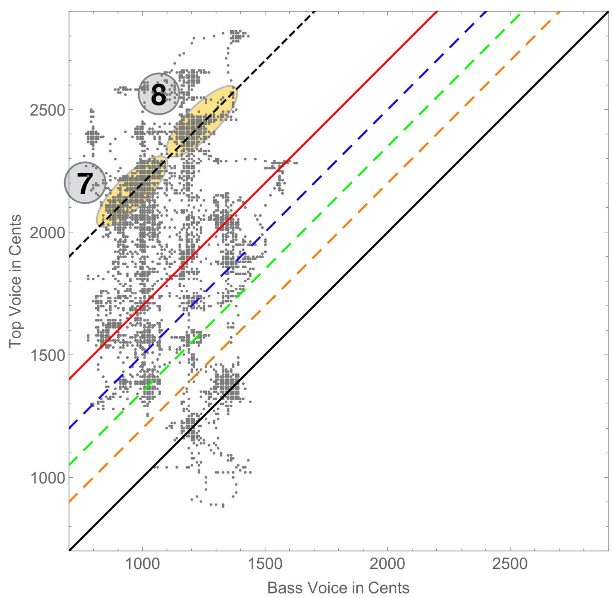

Figure 16. Concomitant bass-top voice pitch pair sample distribution of chant no. 2.

Finally in the bass-top voice distribution, one can see at least two note clusters in which the bass voice tried to maintain an octave to the top voice. Overall, it looks like Artem Erkomaishvili, when singing the bass voice, was switching his attention between the middle and the top voice. When singing the middle voice, on the other hand, he only heard the top voice, so this was his only audible reference which he could relate to. This may explain why Figure 14 shows more note clusters with evidence for voice interactions than Figs. 15 and 16.

3. DISCUSSION AND CONCLUSIONS

With the present study we want to make a contribution to a better understanding of the musical thinking of Artem Erkomaishvili, one of the last master chanters of traditional Georgian chants. Based on a unique set of recordings which was obtained at the Tbilisi State Conservatory in 1966 and for which the F0-trajectories were determined in a prior study by Müller et al. (2017), we investigated the pitch inventories of all three voices of 58 chants separately and jointly. In addition we took a first step at investigating possible signatures of voice interactions between different voices, making use of the special recording setup. In this context it needs to be mentioned, however, that the use of the overdubbing technique, although initiated by Artem Erkomaishvili himself, was new to him (pers. communication by Anzor Erkomaishvili, grandson of Artem Erkomaishvili, 2017). We do not know if this has been influencing the recordings in any way. In any case, there is still much more to be done in the context of trying to understand the influence of voices on each other, in particular on the structural and temporal context in which this happens (cf. Graham, 2013), but this is the objective of a separate study.

Our main results of the analysis of the melodic pitch inventory show that the sizes of melodic seconds vary over a large range from approximately 140 to 240 cents with a peak of the distribution at approximately 180 cents. We do not see a preference for any of the interval sizes characterizing seconds in the scale models by Erkvanidze (2002, 2016), by Tsereteli and Veshapidze (2014), or in the tempered diatonic model, respectively. Loosely speaking, one could characterize the 1966 performance of Artem Erkomaishvili as relaxed regarding the precisons of single melodic steps. In contrast, we observe a high precision when it comes to the harmonic structure of the performance. The harmonic analysis yields a distribution in which precisely justly tuned fifths at 698 cents and octaves at 1203 cents appear as the most frequently intervals. The key to this “melodic flexibility” and “harmonic precision” may lie in the interaction of the voices for which we see clear evidence in the results of the analysis of the joint pitch distributions. There are several cases in which Artem Erkomaishvili maintained particular harmonic intervals despite considerable pitch fluctuations of the individual voices. Therefore, maintaining harmonic precision seems to go hand in hand with the relaxation of melodic precision, which in turn allows for rapid retuning of the voice to maintain an intended harmonic interval.

Relaxing the aim for melodic precision while at the same time aiming at harmonic precision may also relate to the way chants were documented in the past using neumes, which by principle do not allow to document a melody at a very high precision. As discussed in detail in the dissertation of John A. Graham (Graham, 2015), Artem Erkomaishvili used his own neume system, but only for the documentation of the top voice. He is quoted of having told his grandson Anzor Erkomaishvili that “the other voice parts would remember their parts by ear, following the first voice`s lead” (from Graham, 2015). Naturally, if the middle and bass voice are developed by ear from the lead voice, harmonic precision is an asset.

ISBN: 978-84-697-2303-6

In conclusion, Artem Erkomaishvili's performance in 1966 seems to be characterized by a combination of harmonic and melodic thinking rather than by the single aim for melodic precision.

If this interpretation is correct and if it is valid for traditional Georgian vocal music in general, it would raise the fundamental question whether the concept of a single scale (whatever its parameters are) is appropriate to describe the characteristics of Georgian vocal music. Our results are at odds with any melodic scale model which requires a very high precision in singing the melodic intervals. Since the melodic and the harmonic structure of vocal music does not have to be identical, it seems more appropriate to consider the tonal organization of vocal music as an at least two-dimensional property connecting melodic and harmonic aspects. A very stimulating in-depth discussion of this topic, in particular on the properties of melodic and harmonic seconds, can be found in the paper by Nikolsky (2015).

In a recent paper, Erkvanidze (2016) emphasizes the importance of studying the properties of the old audio recordings of professional master chanters as a means to understand the old Georgians musical system. In his thesis, John Graham writes "any theory must account for both the tuning system heard in the 1966 Erkomaishvili recordings and evidence from earlier singers and other regional chant systems seen in the transcription record." (Graham, 2015). We fully agree with both statements and want to emphasize that in this paper we do not propose any new tuning model. The main aim of the present study is to analyse those acoustical characteristics of the 1966 Erkomaishvili recordings which seem relevant as boundary conditions for model building and provide them for discussion.

Since science usually benefits most from a healthy competition of different ideas and perspectives, we also invite other researchers to test their models (or develop new ones) using the F0-trajectories of the Erkomaishvili recordings, which for this purpose have been made publicly available[1].

4. REFERENCES

Erkvanidze, M. (2002). On Georgian Scale System. In The First International Symposium on Traditional Polyphony: 2-8 September, 2002, Tbilisi, Georgia (pp. 178–185).

Erkvanidze, M. (2016). The Georgian Musical System. In 6th International Workshop on Folk Music Analysis, Dublin, 15-17 June, 2016. (pp. 74–79).

Gelzer, S. (2002). Testing a scale theory for Georgian folk music. In The First International Symposium on Traditional Polyphony: 2-8 September, 2002, Tbilisi, Georgia (pp. 194–200).

Gogotishvili, V. (2004). On authentic and plagal types of monotonic (non-octave) scales in Georgian traditional vocal polyphony. In Proc. of the Second International Symposium on Traditional Polyphony, 23-27 September, 2004, Tbilisi, Georgia (pp. 218–226).

Graham, J. A. (2013). Unity and Variety in Orthodox Music: Theory and Practice. In I. M. and M. Takala-Roszczenko (Ed.), *Proceedings of the Fourth International Conference on Orthodox Church Music University of Eastern Finland*. Joensuu, Finland: The International Society for Orthodox Church Music (ISOCM).

Graham. J. A. (2015). The transcription and transmission of Georgian Liturgical chant. PhD thesis, Princeton University.

Jgharkava, I. (2016). Pearls of Georgian Chant, CD produced by the Georgian Chanting Foundations & Tbilisi State Conservatoire.

Kawai, N., Morimoto, M., Honda, M., Onodera, E., & Oohashi, T. (2010). Study on sound structure of Georgian traditional polyphony. Analysis of its temperament structure. In The Fifth International Symposium on Traditional Polyphony, 4-8 October, 2010, Tbilisi, Georgia (Vol. 1, pp. 532–537).

Mauch, M., Cannam, C., Bittner, R., Fazekas, G., Salamon, J., Dai, J., Dixon, S. (2015). Computer-aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency. In *Proceedings of the First International Conference on Technologies for Music Notation and Representation* (p. 8). Retrieved from https://code.soundsoftware.ac.uk/projects/tony/

Mauch, M., Cannam, C., & Fazekas, G. (2014). Efficient Computer-Aided Pitch Track and Note Estimation for Scientific Applications. In *SEMPRE*. Retrieved from http://code.soundsoftware.ac.uk/projects/tony

Müller, M. (2015). Fundamentals of Music Processing — Audio, Analysis, Algorithms, Applications. Springer Verlag, ISBN: 978-3-319-21944-8, 2015

Müller, M., Rosenzweig, S., Driedger, J., & Scherbaum, F. (2017). Interactive Fundamental Frequency Estimation with Applications to Ethnomusicological Research. Submitted to *Conference on Semantic Audio, 2017 June 22 – 24, Erlangen, Germany* (8 pages).

Nikolsky, A. (2015). Evolution of tonal organization in music mirrors symbolic representation of perceptual reality. Part-1: Prehistoric. *Frontiers in Psychology*, *6*(OCT), 1–36. http://doi.org/10.3389/fpsyg.2015.01405

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,15 - 17 June, Dublin/Ireland, 80–87.

Shugliashvili, D. (2014). Georgian Church Hymns, Shemokmedi School (pp. XXIII–XXIX).

Siegel, J. A., & Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception & Psychophysics*, *21*(5), 399–407.

Sundberg, J. (1994). Perceptual aspects of singing. *Journal of Voice*, *8*(2), 106–122.

Tsereteli, Z., & Veshapidze, L. (2014). On the Georgian traditional scale. In The Seventh International Symposium on Traditional Polyphony: 22-26 September, 2014, Tbilisi, Georgia (pp. 288–295).

Westman, J. (2002). On the problem of the tonality in Georgian polyphonic songs: The variability of pitch, intervals and timbre. In The First International Symposium on Traditional Polyphony: 2-8 September, 2002, Tbilisi, Georgia (pp. 212–220).

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

ISBN: 978-84-697-2303-6