MUSICOLOGIST
International Journal of Music Studies

Trabzon University State Conservatory © 2017-2024
Volume **8** Issue **1** **June 2024**

Research Article
Musicologist 2024. 8 (1): 28-54
DOI: 10.33906/musicologist.1246886

FRANK SCHERBAUM
Universität Potsdam, Germany
fs@geo.uni-potsdam.de
orcid.org/0000-0002-5050-7331

SIMHA AROM
Unité Eco-anthropologie, MNHN, CNRS, Université Paris Diderot
simha.arom@gmail.com
orcid.org/0000-0002-2129-9190

FLORENT CARON DARRAS
University & Conservatoire National Supérieur de Musique et de Danse de Paris, France
florentcarondarras@gmail.com
orcid.org/0000-0003-3869-1399

ANA LOLASHVILI
ana.lolashvili@tsc.edu.ge
orcid.org/0000-0001-7252-5234

FRANK KANE
kane.frank@gmail.com
orcid.org/0000-0002-9238-8705

On the Classification of Traditional Georgian Vocal Music by Computer-Assisted Score Analysis

ABSTRACT

This paper describes a feasibility study for the computational classification of traditional three-voiced Georgian vocal music, based on characteristic chord sequences extracted from digital scores. We demonstrate that for this purpose the differences between Western 5-line staff notation and a more appropriate heptatonic system for traditional Georgian music can be adjusted for by a simple transformation. A corpus of about 500 digital scores, consisting of labeled ‘song classes’, i.e. subsets of folk songs from different regions and of liturgical songs in different styles, served as a testbed for the development of a classification procedure based on a higher-order Markov model that - in addition to the classification - yields chord progression sequences for each ‘song class’. Their capacity for interpretation was tested by one hundred cross-validation runs, in which randomly selected subsets of ¾ of the size of each song class were used to train classifiers, which were then applied to the remaining ¼ subsets. The sizes of the intersections of the successfully classified songs in all cross-validations are interpreted as direct measures of the degree of representativeness of the songs for their respective ‘song classes’. Based on a second validation experiment, in which we split up the datasets into equally sized subsets of ½ and ¼ of the original subsets, respectively, we estimate that the smallest subset size for an interpretation of the observed chord progression patterns as properties of a ‘song class’ is of the order of 50 songs. Currently, in our corpus, this requirement is only met by the subsets from Svaneti and Shemokmedi.

KEYWORDS

Traditional Georgian Vocal Music

Classification

Computational Ethnomusicology

Received: February 16, 2023; **Accepted:** May 14, 2024

Introduction

More than a decade ago, Arom and Vallejo undertook a first attempt to investigate the chord syntax of traditional three-voiced[1] Georgian vocal music through the manual analysis of a small number of scores (Arom and Vallejo, 2008; 2010). Since then, computational methods have revolutionized the way we do research in general and have led to a wealth of new tools and the emergence of the new research field of 'computational ethnomusicology'. As a consequence, we can now approach the still open key questions posed by Arom and Vallejo (2008) with new tools and a hugely enlarged dataset. Specifically, the present paper describes the computer-assisted extraction of the harmonic chord sequences of traditional Georgian vocal music from transcription-based digital scores in Western 5-line staff notation and the derivation of a workflow for the subsequent analysis of building blocks of their harmonic syntax using a classification algorithm from the field of machine learning (Bernard, 2021). In this context, we have collected a corpus of roughly 500 digital scores consisting of subsets of (folk) songs from different regions and of liturgical songs in different styles. These subsets will be neutrally referred to as 'song classes'. We will illustrate that the characteristic patterns implicitly encoded in the chord progression sequence of a song can be used for its classification, in other words for the identification of its associated 'song class'. For this purpose, we employ the popular n-gram method, a Markov-model-based approach that is commonly used to classify texts, e. g. to identify a language from a snippet of text (Bernard, 2021).

The classification of traditional Georgian music on the basis of symbolic representations in Western 5-line notation seems at first glance to be a contradiction in terms since this notation is based on a 12-tone equal temperament (12-TET) tuning system, which is well known to be inadequate for Georgian traditional music. In addition, the use of 12-TET notation means that these transcriptions have to let key signatures appear, as if this were tonal music. Since all recent acoustical analyses of recordings of traditional Georgian vocal music (Scherbaum et al., 2020; 2022; Tsereteli and Veshapidze, 2014; 2015) indicate a clear heptatonic tuning system, we demonstrate that for the purpose of classification, the differences between the notation of a traditional Georgian song in a diatonic scale derived from the 12-TET system (which in this case is also heptatonic) and what one believes to be a more appropriate heptatonic system for traditional Georgian music can be adjusted

[1] Since all the music in this work is in three voices, we will omit this attribute in what follows.

for by an appropriate transformation. This notwithstanding, we are aware that the task remains very challenging, since we also face the problem that - compared to Western classical music - the corpus available for analysis is rather small (about digital 500 scores) and is not very balanced in terms of regions of origin and/or song styles. To complicate matters further, the durations of the songs also differ substantially. All of these, however, are not uncommon problems in computational ethnomusicological research and can be accounted for by the choice of a Bayesian framework for analysis.

Overall, we view the present work as a feasibility study with the long-term goal of developing building blocks for an optimal workflow for decoding and better understanding the rules underlying the harmonic syntax of traditional Georgian vocal music.

Data Processing

The starting point for our analyses is a corpus of about 500 pieces for three voices, the majority of which (with the exception of the composed songs from urban regions), were created from transcriptions made by ear by various Georgian scholars (cf. Section 3, Dataset).

Cleaning and reduction of scores

Our current processing workflow starts with 'cleaning' the digital scores. This consists of separating the three voices of a song onto different staves, removing extra notes (such as passing and escape tones), *ossias*, or *appoggiaturas*, so that the cleaned scores contain only the notes from the three voices, one per stave, and nothing else. This is a purely technical preprocessing step to simplify the subsequent analysis, which, however, requires great care and can become rather time-consuming.

A particular long-term aspect of our study is to investigate the effect of the reduction of the scores to their presumably essential parts. This way we seek to separate structural aspects from purely ornamental aspects of chord progression sequences: in other words from aspects that are - grammatically speaking - not essential. The motivation for doing so was that we wanted to test if this increases the relative amount of class-specific feature information as compared to the full original score, which usually contains non-specific and purely ornamental, as well as 'essential', traits. Deciding what are ornamental traits, however, is at least partially dependent on the cultural context in which the music has developed and is usually performed. For this reason, we abstained from performing the reduction of the scores only algorithmically. Instead, this work was done by Ana Lolashvili, a graduate of the Chant University, currently studying at the Tbilisi Conservatory. An example is shown in Fig. 1.

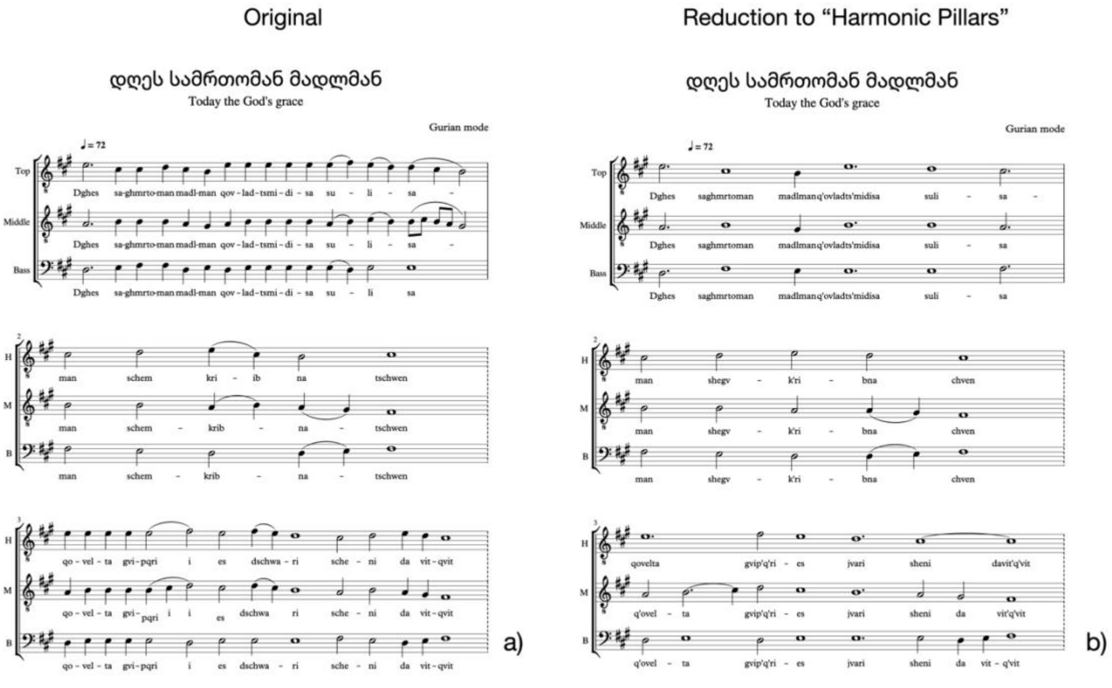

Figure 1. Example of the 'manual' reduction of a score to its harmonic pillars.

Since the 'manual' score reduction is very labor intensive, it was only done on a subset of the complete corpus, consisting of a total of 182 songs. This subset is intended to serve as a reference point for future analysis.

Tuning system transformation

The next step in the processing chain consists of modifying the digital score for the 'distortions' of its tonal content caused by its representation in a fundamentally inappropriate western staff notation. In this context, it should be emphasized that the question of the characteristics of the tonal organization of Georgian traditional music has been one of the most controversial issues in scholarly discussions in recent decades. A review of the related discourse, a considerable part of which has been conducted detached from reproducible observational evidence, can be found in Jordania, 2022.

Starting with the pioneering work of Tsereteli and Veshapidze (2014; 2015) followed by the study of Scherbaum et al. (2020), and most recently by Scherbaum et al. (2022), this discourse is now becoming more and more evidence-based since it shifts to the interpretation of increasingly large and openly accessible data sets of objectively verifiable pitch determinations (in total, an estimate of more than 1 million pitch and interval samples). All acoustical studies show that the melodic pitch inventories used in traditional Georgian vocal music differ significantly from those based on the major and minor systems, in which the melodic scales[2] consist of intervals of half- or whole-note steps. These studies agree that the melodic scale(s) of traditional Georgian music are, in a first approximation, composed of roughly equally spaced intervals. For the larger datasets, in particular the Erkomaishvili dataset (Scherbaum et al., 2020), the interval between the fourth and fifth (scale) degree above the last bass note of a song often corresponds to a whole tone step (200 cents), while the size of the remaining intervals is of the order of 5/6 of a whole tone (167 cents). The size of the melodic 2^{nd} as the most frequent melodic step size of the songs is not fixed, however, but varies significantly around a mean value of approximately 170 cents, whereas the mean value of the harmonic 2nd is significantly larger (Fig. 1). This can possibly be explained, among other reasons, as a consequence of the 1-4-5 chord, very popular in Georgian music, in which a fourth and a fifth are simultaneously intoned as pure intervals above a fundamental. The differences between the results of Tsereteli and Veshapidze (2014; 2015)**,** whose observations suggest a pure unitonic scale with a step size of 6/7 of a whole tone (171 cents), and the results of Scherbaum et al. (2020; 2022) whose synoptic scale model suggests a combination of a whole tone step and six equal sized intervals with a step size of approximately 5/6 of a whole tone (167 cents) are probably due to the different sizes of the analyzed datasets and are irrelevant to the purpose of the present study, especially against the backdrop of the wide spread of the observed melodic 2^{nd}s.

[2] The plural is used here to differentiate the different church modes as individual scales.

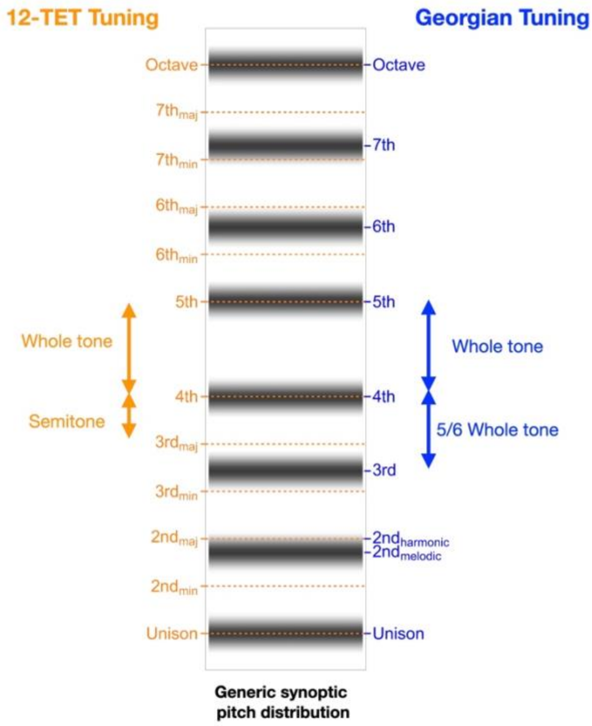

Figure 2. Comparison of the key elements of the 12-TET tuning system (orange labels) with the essential characteristics of the observed tuning systems in traditional Georgian vocal music (blue labels). The generic synoptic pitch distribution shown as a density plot in the middle was generated by combining the key elements of the average scale models derived from the Erkomaishvili dataset (Scherbaum et al., 2020) with the average tuning systems obtained for all Svan ensembles (Scherbaum et al., 2022).

Note that Fig. 2 sketches the key properties of the tuning of traditional Georgian vocal music only in a rough conceptual way, and is not meant to be interpreted as representing an individual dataset. The orange and blue interval labels in Fig. 2 illustrate the interval sizes of the 12-tone equal temperament (12-TET) and the Georgian tuning system, respectively.

What is relevant for the purpose of our present work is that the melodic scale of traditional Georgian vocal music is heptatonic, in other words consists of seven tones per octave. Since this is also the case for all diatonic scales[3], any diatonic scale, even if it is represented in western five-line score notation, can be mapped onto any of the heptatonic tuning systems that have been suggested for Georgian traditional music (Scherbaum et al., 2020; 2022; Tsereteli and Veshapidze, 2014; 2015). In this context it does not matter if one assumes a pure or an approximately equidistant scale with individual larger

[3] A diatonic scale is any heptatonic scale that includes five whole steps (whole tones) and two half steps (semitones) in each octave.

intervals.

Fig. 3 illustrates how the pitches determined from an actual recording of a traditional Georgian song and displayed as pitch and note trajectories in its original tuning (Fig. 3a) and the corresponding transcribed values (Fig. 3d and 3e) are related through the pitch distribution in Fig. 3c.

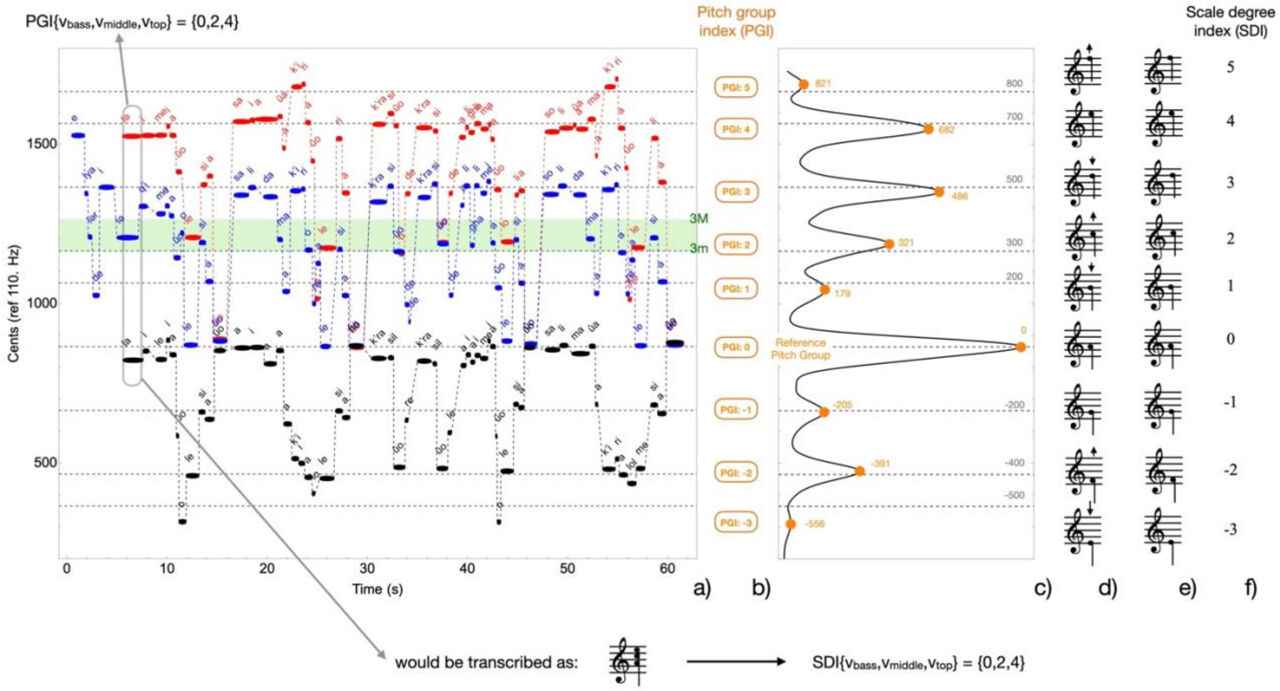

Figure 3. The conceptual relationship between pitch and note trajectories of audio recordings and the tuning system used in transcriptions into to 5-line staff notation. For details, see text.

The voices of all three singers fluctuate visibly in Fig. 3a, but all notes (horizontal red, blue and black blobs for the top, middle and bass voice, respectively) can still be seen to belong to discrete pitch groups, identifiable by the peaks of the pitch distribution shown in Fig. 3c. In this example we chose the most salient pitch group (the one with the highest peak), which also happens to be the final note for all three voices, as a reference point and assigned positive pitch group indices to the pitch groups above and negative ones to the pitch groups below, respectively (Fig. 3b). The horizontal dashed gridlines (in Fig. 3c) correspond to the pitches of a minor scale anchored at the pitch of the reference pitch group. One can see that the pitch values for some of the pitch groups of the tuning used in this song (shown by the orange labeled numbers close to the solid orange circles indicating the mean value of a pitch group) come very close to pitches of the minor scale, while others show a larger deviation. What is important, however, is that in all cases there exists a unique 1-to-1 match of a note of the minor scale (in our case A-minor, chosen for the convenience of an accidental-free scale) and a pitch group of the tuning system used by the singers. Therefore, if we index (count) the notes of the A-minor scale from the lowest to the highest by integer numbers and choose A as the reference note to which we assign the scale degree index of 0, we can unambiguously transcribe the whole song simply by setting the pitch group index (PGI) (Fig. 3b) of a sung note equal to the scale degree index (SDI) (Fig. 3f) of our chosen A-minor scale. Hence the pitch groups in Fig. 3b would be mapped to the notes in Fig. 3d or Fig. 3e, depending if one cares to indicate the deviations from the 12-TET tunings system by little arrows or not. The former used to be common practice in ethnomusicological transcriptions, but is nowadays rarely seen in digital scores. Since all the notes necessary for the transcription are part of the A-minor scale, one can also proceed in the inverse direction. Knowing the scale degree index of a note, one also knows which pitch group the note actually belongs to. By representing a note through its pitch group index, we lose the precision of the exact cent value but gain in accuracy by being able to correct for the bias of the 12-TET tuning system with respect to the tuning system actually used.

So far so good. But what about a situation in which a score also contains notes from a non-diatonic scale? In this case, the score would still contain accidentals, even if we transpose it to the C-Maj/A-min key, because we have not yet left the 12-TET tuning system. In this case, the properties of the generic pitch distribution shown in Fig. 2, namely that the tuning system is approximately equidistant, suggest a solution. Since in terms of pitch, the pitch groups making up the Georgian sound scale generally lie between the minor and major variants (in the 12-TET system) of the same scale degree, ignoring the accidentals will map all minor and major intervals onto 'neutral' ones and the corresponding scale becomes heptatonic, even if their interval structure is not yet correct.

In the following, we illustrate some of the problems that might arise from trying to represent actual recordings of traditional Georgian songs in a major/minor-based tuning system. The individual pitches belonging to the pitch group with PGI= 3 (in Fig. 3a) fluctuate within the green shaded rectangle indicating the pitch band between a minor and a major 3rd (above the reference pitch). Looking at the first 3-voiced chord in the song (highlighted by the vertical rectangle with rounded edges), one can see that the pitch of the middle voice lies pretty much between the pitch for the minor 3rd and the major 3rd. Hence it would not be surprising if this chord were to be perceived by some transcribers as an A-minor chord {A, C, E} and by others as an A-major chord {A, C#, E}. However, from Fig. 2a it can be seen that is actually neither of the two because the whole concept of minor and major 3rd is inappropriate. In order to resolve this ambiguity we need to interpret both transcribed triads {A, C, E} and {A, C#, E} simply as indicating a chordal build-up of notes from the pitch groups {0, 2, 4}. Technically, this can be achieved by simply dropping the accidental from C# and determining the scale degree indices (Fig. 3f) of the three elements of the triad {A, C, E}, which would result in the list {0, 2, 4}.

Generalizing this idea and dropping all the accidentals of a score, but interpreting the resulting notes simply as indicators of their scale degree indices, allows us unambiguously to map any score in western 5-line staff notation onto a heptatonic tuning system, without having to make any assumptions regarding the details of the interval structure. One can think of the effect of removing all accidentals as combining all pairs of major/minor variants of non-pure intervals into single (neutral) versions. As a consequence, the pair (3rd-min/3rd-maj) will result in a single 3rd, and the same for 2nd, 6th, and 7th. Therefore, removing all the accidentals of a score results in a diatonic tuning system and therefore a heptatonic scale that can be mapped onto the Georgian tuning system by setting the scale degree indices equal to the pitch group indices. Furthermore, because the pitch groups of the generic Georgian tuning for the non-pure intervals (2nd, 3rd, 6th, 7th) are located between their minor and major variants, removing the accidentals results in exactly the mapping that inverts the note assignment during the transcription process.

Therefore, the tuning system modification for the differences between the 12-TET tuning system and any of the suggested Georgian heptatonic tuning systems consists simply of two components. First, the removal of all accidentals from the score and second, the choice of a reference note and the calculation of the scale degree indices for all notes with respect to the chosen reference note. The only assumption that we make in this context is that the 1:1 mapping of the scale degree indices in the accidental-free score to the pitch group indices in the actual Georgian tuning system used makes sense, which seems pretty obvious from Fig. 2.

The choice of the reference note is somewhat arbitrary but very important as to how the different songs of a corpus are quantitatively represented relative to each other. For our study, we chose the final bass note of a song as the reference note, which means that for each song a pitch group index of 1 refers to the first pitch group above the final bass note, a pitch group index of -1 to the first pitch group below the final bass note, and so forth. Choosing the final bass note (the *finalis*) as reference note does not mean that one has to attach a functional meaning to the *finalis*, but it facilitates for example the quantitative comparison of the final cadences of different songs by simply comparing the numerical values of the last few pitch group indices.

Representations for subsequent processing

As a result of the tuning system modification, for each of the three voices in a song we obtain a sequence of 'notes', each represented by a pitch group index and a duration, from which it is a straightforward matter to calculate the corresponding harmonic states, each defined by a concommitance of three pitch group indices and a duration. For the purpose of the subsequent analysis, we ignore the duration of the notes and analyze (for now) only the sequences of pitch group indices. Fig. 4 shows the score (a) and the result of tuning system modification (b) for the song *Kriste Aghdga*.

a)

PGI-T	4	4	4	4	R	5	4	3	4	5	R
PGI-M	3	2	1	2	R	1	0	1	2	3	R
PGI-B	-3	-3	-3	-3	R	-3	-4	-3	-2	-1	R
IDX	1	2	3	4	5	6	7	8	9	10	11

PGI-T	4	3	2	3	4	4	4	3	2	3	4	R
PGI-M	3	1	0	1	2	3	2	1	2	1	0	R
PGI-B	-4	-4	-4	-3	-2	-1	0	-1	-2	-3	-4	R
IDX	12	13	14	15	16	17	18	19	20	21	22	23

PGI-T	5	4	3	2	3	3	4
PGI-M	1	0	-1	0	1	1	0
PGI-B	-3	-4	-3	-2	-2	-1	0
IDX	24	25	26	27	28	29	30

b)

Figure 4. Score of the song *Kriste Aghdga* (a) and the corresponding sequence of harmonic states (chord sequence) expressed by the pitch group indices of the three voices (b). Note that the duration of the notes is ignored and if successive chords have identical pitch values, they are joined into a single harmonic state.

For the subsequent analysis, it turned out to be advantageous to transform the sequence of pitch group indices of the three voices shown in Fig. 4b into different representations. These are shown in Fig. 5.

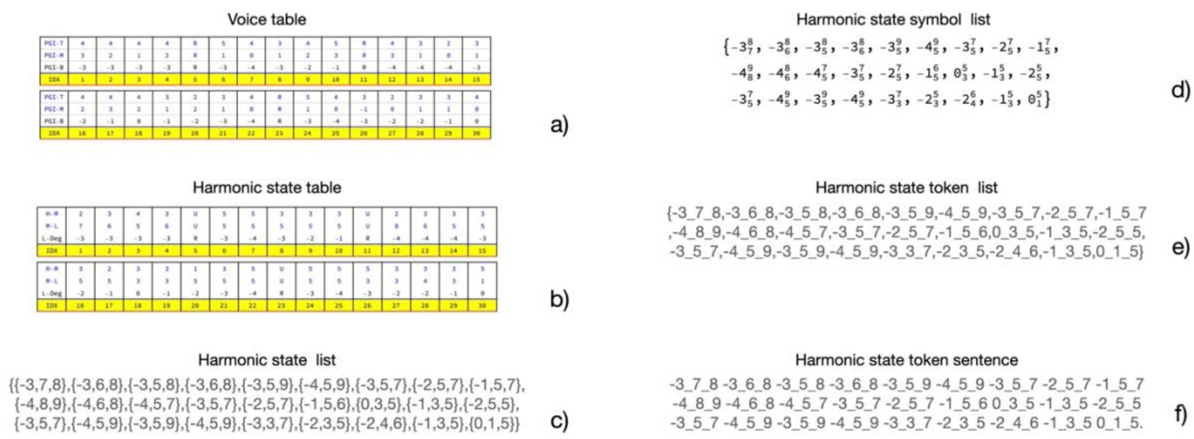

Figure 5. Different representation forms of the sequence of pitch group indices of the three voices shown in Fig. 4b as used in the subsequent analysis. a) as voice table (same as Fig. 4b). b) as harmonic state table. c) as harmonic state list. d) as harmonic state symbol list. e) as harmonic state token list. f) as harmonic state token sentence.

In the representation of 'harmonic state' shown in Fig. 5b, the two top elements in each column contain the intervals between the highest and the intermediate (H-M) and the intermediate and the lowest voice (M-L), respectively. In this context voice crossings, which happen in particular in Gurian songs, were accommodated through correction. The H, M, and L labels in the chord assignments always refer to the highest, the intermediate, and the lowest voice, independent of who sings them. In the case of voice crossings, these will not always correspond to the top, middle, and bass voices. The application of this correction is admittedly a subjective choice. From the structural perspective, one could argue that voice crossings contain information regarding the type of a song or the region of origin. From the acoustic perspective, however, it is only the interval that matters, not who sings a particular voice. This perspective (which is the one we chose) also simplifies the processing because one does not have to deal with signed intervals. The lowest entry in each column (L-Deg) shows the scale degree of the lowest voice with 0 representing the reference scale degree *(*in our case the *finalis)*. The 'harmonic state list' representation in Fig. 5c is simply a representation of the 'harmonic state table' as a list (indicated by the curly brackets {}). This is the computer readable numerical input form used for all our algorithms. The 'harmonic state symbol list' in Fig. 5d, on the other hand, is used to facilitate the 'human' perception of the 'harmonic state list'. The central value of each symbol contains the lowest voice scale degree, while the subscript and superscript contain the intervals between the lowest and the intermediate (M-L) and the lowest and the highest voice (H-L), respectively.

In our analysis we make heavy use of the conceptual similarity of words in a sentence and chords in a chord sequence. To be able to exploit this similarity quantitatively, we transform each numerical element in a harmony into a text string and connect these strings via underscores. This results in the transformation of the numerical representation of a harmonic state into a representation as a character string which is called 'harmonic state token', as shown in Fig. 5e. The benefit of this representation is that it allows the use of algorithms from computer linguistics. Conceptually in this representation a chord becomes the equivalent of a word in an 'unknown language'. The representation as 'harmonic state token sentence', shown in Fig. 5f, completes this transformation conceptually. It allows us to treat a whole song as a sentence of an unknown language. As a final remark on the different types of representation we want to point out that in addition to the representation of a chord as a complete chord in the form of a numerical triplet {L-Deg, M-L, H-L} we have also explored the effect of dropping the bass voice information and only considering the interval pair {M-L, H-L}. The potential advantage of this representation is that it does not depend on the choice of the reference note. The disadvantage, on the other hand, is that we ignore all the melodic information contained in the melody of the lowest voice.

Dataset

For the actual analysis, we have used a set of 493 digital scores, obtained from available song collections (Akhobadze, 1957; Chokhonelidze, 2003; Jordania, 2004; Shugliashvili, 2014) and from the Center of Church Chants of the Patriarchate (2006a; 2006b; 2008); Folklore state Centre of Georgia (2018a; 2018c; 2018b; 2020a; 2020b; 2020c; 2020d); Patarava (2003); Tarkhnishvili (2008); Tbilisi State Conservatoire (2005); Veshapidze (2006); Veshapidze and Kotrikadze (2016). The collections by Shugliashvili (2014) and Jordania (2004) were already in digital form, while the rest had to be manually converted. The original corpus contains liturgical chants from different monasteries and folk songs from different regions of Georgia, some of which, however, were only represented by a few examples. In order to generate a more balanced dataset, we reduced it to only those regions that contributed at least 15 songs. Fig. 6 shows the distributions of songs according to their locations. Fig. 6a shows the breakup of the complete corpus in terms of songs, while Fig. 6b shows the corresponding distribution of harmonic states (chords). The difference in the two pie charts in Fig. 6a and 6b is due to the fact that the lengths of the songs are strongly region-dependent. Fig. 6 shows, for example, that Svan songs are generally shorter than songs from Samegrelo. The imbalance of the distributions will be taken care of in the context of the final analysis processing, as will be described in more detail in that context.

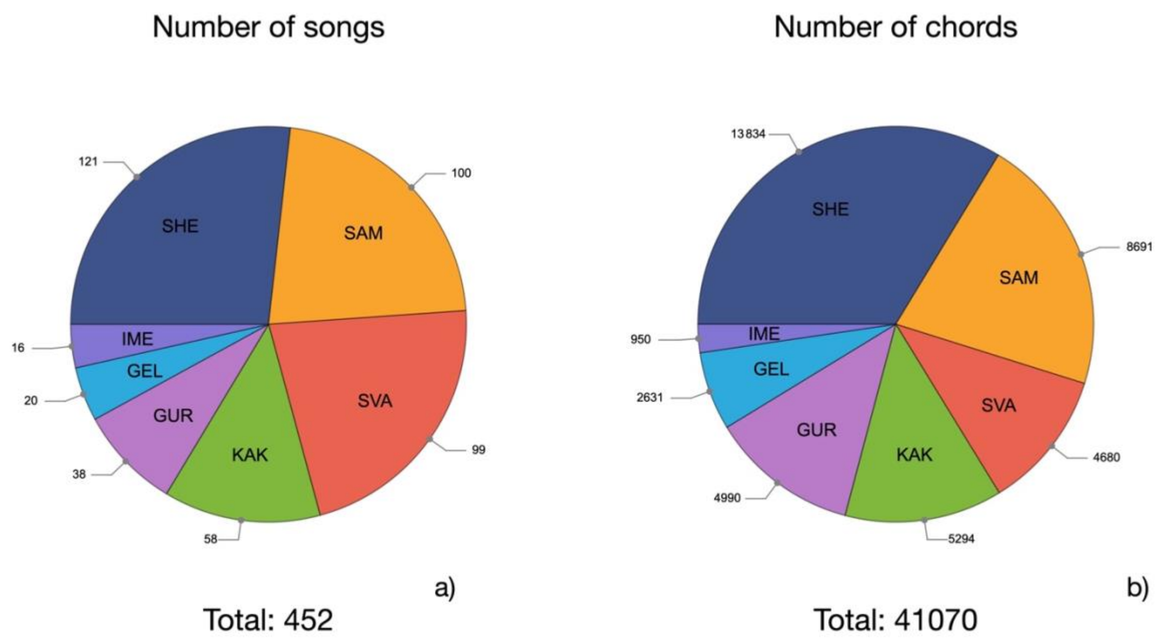

Figure 6. Composition of the corpus by place of origin. a) With reference to the number of songs. b) With reference to the number of chords. The acronyms SAM, SVA, KAK, GUR, GEL, IME correspond to Samegrelo, Svaneti, Kartli-Kakheti, Guria, Gelati, and Imereti, respectively.

Fig. 7, which shows the distribution of the most frequently occurring chords in the different song classes, illustrates that the differences in the composition of the individual subsets are continued at the level of the chord inventories.

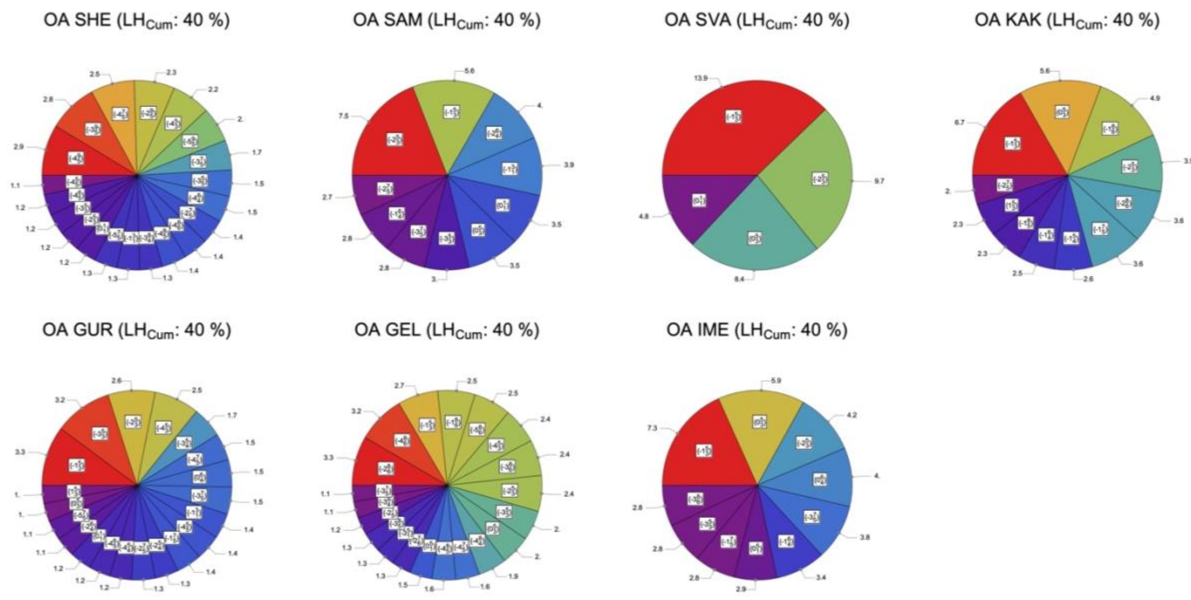

Figure 7. Distribution of most frequently occurring chords in the different song classes, representing 40% of the total likelihood. The numbers in the callouts represent likelihoods in %.

Each of the pie charts covers 40% of the cumulative likelihood (LH). For example, for all the songs from the class SVA (Svaneti) a single, randomly drawn chord has a likelihood of 13.9% to be {-1, 3, 5}, a likelihood of 9.7% to be {-2, 3, 5}, a likelihood of 8.4% to be {0, 3, 5}, and a likelihood of 4.8% to be {0, 1 , 1}. In mathematical terms, which we will need later, this would be written as P({1, 3, 5}| SVA) = .139.

Even without performing quantitative analysis, Fig. 7 already visually indicates significant differences between the individual song classes, as expressed in the occupancy density of the pie charts, namely the number of chords within the top 40% of the likelihood. In the next chapter, we will describe how this information, together with the information contained in Fig. 6b, can be used quantitatively to classify the entire corpus.

Analysis

Theoretical background

The question which we want to address in this section is the following: Given a collection of songs that originate from different song classes, how can we teach a computer to recognize the correct class? Since one cannot expect a definite answer to this question, we weaken it a bit and ask instead about the class that is most likely to be the origin of the song under consideration. This question has a definite answer which fortunately can be obtained by making use of Bayes' theorem, which follows directly from the basic rules of probability theory. The idea behind what is now called Bayesian inference is that one can express the probability for a model M to have produced a set of observations d, which is *posterior probability*, written as $P_{post}(M|d)$, by a very simple formula:

$$P_{post}(M|d) = P(d|M)\ P_{prior}(M)/\text{Factor}_{norm}. \qquad (1)$$

In this context, $P(d|M)$ is the so-called *likelihood* which states the conditional probability for a model M to produce the data d. In our situation, this is essentially what is shown in Fig. 7. The models are the individual song classes, while the data (observations) are the individual chords. The sizes of the individual pieces in the pie charts quantitatively describe the conditional probabilities for a song from a particular song class to have produced a particular chord. As derived above, the probability for a chord from a song from Svaneti to be {1, 3, 5} is P({1, 3, 5}| SVA) = 0.139. Hence its likelihood is 0.139.[4]

The second term, the so-called *prior probability* of the model under consideration $P_{prior}(M)$ is the probability to observe any of the chords from the model (song class) M if we randomly draw a chord from the complete corpus. This, however, is exactly what is expressed by the sizes of the different pieces (one for each song class) in Fig. 6b, which illustrates one of the major benefits of the Bayesian framework, namely that - via the prior probability term - it can deal with imbalanced data sets like ours. To conclude the example from Svaneti, the prior probability for Svaneti, which would be written as $P_{prior}(SVA)$, is given by 4680 (the number of chords from Svaneti), divided by the total number of chords in the whole corpus, which is 41070, which results in $P_{prior}(SVA)$= 0.114.

Hence, the 'posterior probability' for Svaneti to have generated the chord {1,3,5), which is written as P(SVA| {1, 3, 5}), becomes 0.139 x 0.114 = 0.01584 divided by Factor_{norm}. For the solution of the classification problem, it would actually suffice to determine the maximum value of the product of prior probability and likelihood, $P(d|M)\ P_{prior}(M)$. However, the determination of Factor_{norm} is also straightforward, since the sum of the probabilities for all possible models (all seven song classes) has to be 1. Hence, the normalization factor Factor_{norm} is simply obtained by the sum of all $P(d|M_i)\ P_{prior}(M_i)$ for i

[4] In a general situation, the determination of the likelihood term can become technically more challenging since it involves the learning of a probability distribution on categorical data which can contain missing values and/or outliers. This is intrinsically taken care of in the Classify algorithm of Mathematica (Wolfram Research, 2020).

= 1,…N, with N being the number of models (in our case song classes)[5].

Instead of individual chords as above, the 'data' could of course also be chord sequences, which leads directly to the n-gram method (aka Markov model) in which one is interested in the posterior probabilities for a model to have produced a chord sequence of a particular length (n) (Bernard, 2021). For n=1 (unigrams), one is interested in a list of single chords {$chord_1$, $chord_2$, $chord_3$,… }. For n=2 (bigrams) one considers a list of lists of two successive chords, e.g. {{$chord_1$, $chord_2$}, {$chord_2$, $chord_3$}, {$chord_3$, $chord_4$}, … }, etc.

For the determination of the posterior probability of a song class to have produced a whole song (instead of a single chord), one has to simply multiply the posterior probabilities for all the chords of the song. In the context of classifying text this is known as the *bag-of-words assumption* (Bernard 2021). In this context, it is worth emphasizing that *text classification* is conceptually identical to what we are trying to do in the present study if we make use of the transformation of a sequence of chords into a sequence of words, as illustrated in Fig. 5.

Song class identification (classification)

From a practical perspective, the identification of the most likely song class for a given song (or chord sequence) boils down to the determination of the song class with the highest posterior probability. In the present study, we have used the Classify algorithm of Mathematica (Wolfram Research, 2020) for this purpose. Fig. 8 illustrates the quality of the resulting classification in the form of the associated confusion matrices for n-grams with n = 1 to 4. The accuracy values on top of each panel indicate the percentage of the songs which are correctly classified. The implicitly encoded song-class-specific patterns within the chord inventories and the chord-transition-inventories contain enough song-class-specific information to result in accuracies more than three times better than the accuracy baseline (26.8%), namely 80.3 % and 98.9%.

[5] This can be seen, if one divides each $P(d|M_i)\ P_{prior}(M_i)$ by the sum of $P(d|M_i)\ P_{prior}(M_i)$ for all i, the sum of all $P(d|M_i)\ P_{prior}(M_i)$ will become 1.In other words the sum of the posterior probabilities $P_{post}(M|d)$ for all models becomes 1, as it should be.

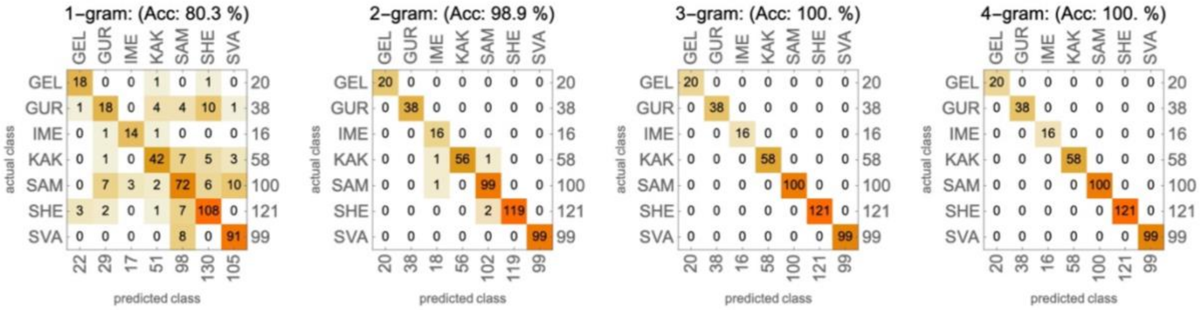

Figure 8. Confusion matrices, which illustrate which and how many actual song classes are mispredicted and how, for n-grams with n=1-4. The accuracy baseline (accuracy if predicting the commonest class) is 26.8%.

For n=3 and beyond, the classification accuracy becomes 100%. This also means that it does not need the reduced versions to achieve a very good classification accuracy[6]. In other words, all it takes to uniquely identify the song class of a song is to know a sufficient number of sequences of three subsequent chords (3-grams). In contrast, the information contained in the chord inventories (1-grams) and the chord transition inventories (2-grams) is obviously not specific enough to allow a unique identification of a song class. Different song classes can obviously share the same chords and chord transitions, but do so rarely if ever with n-grams of lengths larger than three. This is illustrated in Fig. 9 for the chant *Holy is the Lord our God* from the Gelati (GEL) subset of our corpus.

[6] For the determination of the chord syntax, the differences of the results between the original and the reduced version are expected to be much stronger.

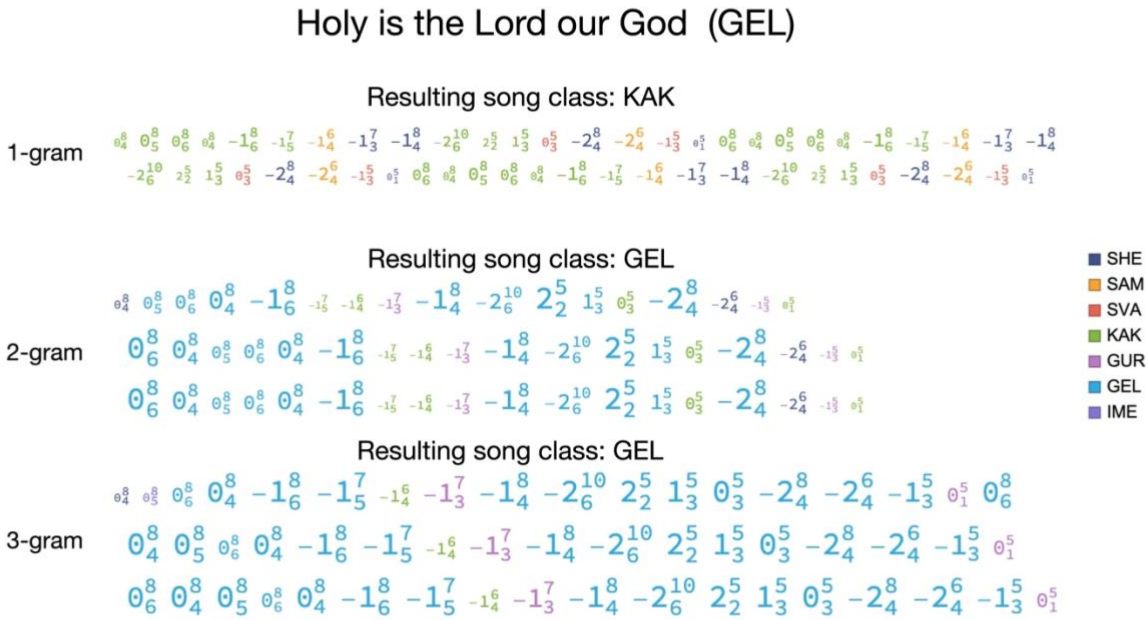

Figure 9. Visualization of the change in posterior probabilities in the course of a song for different n-gram lengths.

The visualization type was inspired by Bernard (2021), who used it in the context of language classification. The color of the last element of each n-gram indicates the song class for which the posterior probability for the n-gram has the maximum value of all the song classes. The sizes of the chord symbols correspond to the actual probability values. The larger the size, the more certain is the algorithm about the n-gram's song class. The assumed song class for the whole song is the product of all n-gram probabilities. The uppermost panel in Fig. 9 shows that from the 1-gram perspective (the chord inventory), the song contains characteristic chords from a variety of song classes, with the majority of the chords color-coded in green (KAK). In terms of the chord transitions (2-grams), however, the situation changes, and the algorithm becomes more confident (indicated by the larger symbol sizes) that the assigned song class should be GEL. As to the 3-grams, the choice of GEL as suggested song class is further emphasized, as indicated by the still increasing symbol sizes and the reduction in the number of n-grams which suggest a song class other than GEL.

It must be emphasized that we do not consider the solution to the classification problem as an end in itself. Our main long-term goal is to obtain a better understanding of the harmonic syntax of the music. However, in this context, the components of the classification algorithm described above provide important information, e.g., if we apply equation (1) not to the n-grams in a song, but to all n-grams in a given song class. To make this more precise, in Tables 1-3 we show the 20 most likely 1-grams, 2-grams, and 3-grams, respectively, for each song class together with the corresponding likelihood values.

If one would extend these tables to all n-grams for each song class, one could perform the classification of a song as a simple bookkeeping exercise, which simply would involve:

For each song class: multiplying the posterior probabilities for all n-grams in the song (which can be derived from the corresponding likelihood table and the prior probability value), and finally choosing the song class with the largest resulting value.

Table 1. Twenty most likely 1-grams for each song class together with the corresponding likelihood values in %.

SHE		SAM		SVA		KAK		GUR		GEL		IME	
$\{-2^2_2, -1^3_1, 0^1_1\}$	0.86	$\{-2^5_3, -1^3_1, 0^1_1\}$	1.13	$\{0^5_3, -1^5_3, -2^5_3\}$	2.03	$\{-1^5_4, -1^5_3, 0^5_1\}$	0.75	$\{-2^4_2, -1^3_1, 0^1_1\}$	0.85	$\{-4^8_4, -4^8_5, -4^8_6\}$	0.54	$\{0^6_4, 0^5_3, -1^5_3\}$	1.63
$\{-2^3_1, -2^2_2, -1^3_1\}$	0.40	$\{-2^7_5, -2^6_4, -2^5_3\}$	0.75	$\{0^5_4, 0^5_3, -1^5_3\}$	1.07	$\{-1^8_4, -1^8_5, -1^8_6\}$	0.75	$\{-2^3_3, -2^4_2, -1^3_1\}$	0.33	$\{-2^8_6, -2^7_6, -1^5_5\}$	0.50	$\{0^7_5, 0^6_4, 0^5_3\}$	1.53
$\{-4^7_5, -4^6_5, -4^5_5\}$	0.31	$\{0^5_3, -1^5_3, -2^5_3\}$	0.69	$\{0^6_4, 0^5_3, -1^5_3\}$	1.00	$\{-1^7_5, -1^6_4, -1^5_3\}$	0.58	$\{-2^4_3, -2^4_2, -1^3_1\}$	0.26	$\{-2^2_2, -1^3_1, 0^1_1\}$	0.39	$\{-1^6_4, -1^5_3, 0^5_1\}$	1.09
$\{-3^4_2, -2^2_2, -1^3_1\}$	0.29	$\{-2^4_2, -1^3_1, 0^1_1\}$	0.68	$\{-2^5_3, -1^3_1, 0^1_1\}$	0.89	$\{-1^8_6, -1^8_7, -1^8_6\}$	0.54	$\{2^4_2, 3^3_1, 4^1_1\}$	0.24	$\{-1^8_6, -1^9_6, -1^8_6\}$	0.39	$\{-1^7_5, -1^6_4, -1^5_3\}$	0.98
$\{-3^4_2, -2^3_1, -2^2_2\}$	0.28	$\{-2^6_4, -2^5_3, -1^3_1\}$	0.53	$\{-2^5_3, -1^5_4, -1^5_3\}$	0.89	$\{-2^8_4, -2^8_5, -2^8_6\}$	0.54	$\{-4^5_5, -4^5_4, -4^5_3\}$	0.22	$\{0^8_4, 0^8_5, 0^8_6\}$	0.35	$\{0^6_4, 0^7_5, 0^6_4\}$	0.87
$\{-5^7_4, -4^5_3, -3^5_1\}$	0.27	$\{-1^7_5, -1^6_4, -1^5_3\}$	0.49	$\{-1^5_3, -2^5_3, -1^3_1\}$	0.65	$\{0^7_5, 0^6_4, 0^5_3\}$	0.52	$\{1^5_4, 2^4_3, 2^4_2\}$	0.22	$\{-3^8_4, -3^8_5, -3^8_6\}$	0.35	$\{3^3_1, 3^4_2, 3^5_3\}$	0.87
$\{-2^7_4, -1^5_3, 0^5_1\}$	0.26	$\{-2^5_4, -2^5_3, -1^3_1\}$	0.42	$\{0^5_4, 0^6_4, 0^5_3\}$	0.65	$\{-2^8_6, -2^8_7, -2^8_6\}$	0.46	$\{2^4_3, 2^4_2, 3^3_1\}$	0.22	$\{-1^5_3, 0^5_1, -1^8_6\}$	0.31	$\{0^5_3, -1^5_3, -2^5_3\}$	0.87
$\{-3^7_4, -2^5_3, -1^5_1\}$	0.26	$\{0^6_4, 0^7_5, 0^6_4\}$	0.41	$\{-1^5_3, -2^5_3, -1^5_3\}$	0.56	$\{-1^6_4, -1^5_3, 0^5_1\}$	0.44	$\{-4^6_5, -4^5_5, -4^5_4\}$	0.20	$\{-2^6_4, -1^5_3, 0^5_1\}$	0.31	$\{3^6_4, 3^5_3, 2^5_3\}$	0.76
$\{-5^4_3, -4^3_1, -3^1_1\}$	0.26	$\{-1^6_4, -1^7_5, -1^6_4\}$	0.38	$\{-3^7_5, -2^5_3, -1^3_1\}$	0.51	$\{-1^8_6, -1^7_5, -1^6_4\}$	0.42	$\{0^5_4, 0^6_4, 0^7_4\}$	0.20	$\{-2^8_6, -2^9_6, -2^8_6\}$	0.31	$\{-3^8_6, -3^7_5, -3^6_4\}$	0.76
$\{-2^9_5, -3^9_5, -4^9_5\}$	0.25	$\{0^6_4, 0^5_3, -1^5_3\}$	0.37	$\{0^5_3, 0^6_4, 0^5_3\}$	0.49	$\{0^5_3, 0^5_4, -1^5_4\}$	0.41	$\{0^7_4, 0^6_4, 0^5_4\}$	0.20	$\{-2^5_4, -1^5_3, 0^5_1\}$	0.31	$\{-3^9_7, -3^8_6, -3^7_5\}$	0.76
$\{-2^7_5, -2^6_5, -2^5_5\}$	0.24	$\{-1^5_3, 0^5_3, -1^5_3\}$	0.34	$\{-1^5_3, -1^3_1, 0^1_1\}$	0.47	$\{0^5_4, -1^5_4, -1^5_3\}$	0.41	$\{-2^4_4, -2^4_3, -2^4_2\}$	0.20	$\{-2^7_6, -1^5_5, 0^5_4\}$	0.31	$\{-2^5_3, -1^5_3, -1^6_4\}$	0.65
$\{-3^5_5, -3^5_4, -3^5_3\}$	0.24	$\{-2^8_6, -2^7_5, -2^6_4\}$	0.33	$\{-2^5_3, -1^6_4, -1^5_3\}$	0.45	$\{-1^5_5, -1^5_4, -1^5_3\}$	0.41	$\{0^5_5, 1^5_4, 2^4_3\}$	0.20	$\{-5^5_4, -4^5_3, -3^5_1\}$	0.27	$\{3^5_3, 2^5_3, 1^5_3\}$	0.65
$\{-6^9_8, -6^8_8, -5^7_6\}$	0.24	$\{0^7_5, 0^6_4, 0^7_5\}$	0.31	$\{-2^6_4, -2^5_3, -1^5_3\}$	0.45	$\{-1^6_4, -1^7_5, -1^6_4\}$	0.41	$\{-4^5_3, -4^5_4, -3^4_4\}$	0.18	$\{-5^{10}_6, -5^{10}_7, -5^{10}_8\}$	0.27	$\{-1^5_3, -2^5_3, -3^5_3\}$	0.65
$\{-6^8_8, -5^7_6, -5^7_4\}$	0.24	$\{-2^5_3, -2^4_2, -1^3_1\}$	0.29	$\{-1^5_3, 0^6_4, 0^5_3\}$	0.40	$\{-1^7_5, -1^6_4, -1^7_5\}$	0.41	$\{-4^7_5, -4^6_5, -4^5_5\}$	0.18	$\{-2^7_5, -2^6_4, -1^5_3\}$	0.27	$\{-3^5_3, -3^6_4, -3^7_5\}$	0.65
$\{-5^7_6, -5^7_4, -4^6_4\}$	0.24	$\{-3^7_5, -2^5_3, -1^3_1\}$	0.29	$\{-3^5_3, -2^5_4, -2^5_3\}$	0.40	$\{-2^{10}_8, -1^8_6, -1^8_7\}$	0.39	$\{-1^5_3, 0^5_4, 0^6_4\}$	0.18	$\{-1^8_6, -2^8_6, -2^7_6\}$	0.27	$\{0^5_3, 0^6_4, 0^5_3\}$	0.54
$\{-5^9_5, -4^9_5, -3^9_5\}$	0.24	$\{-2^5_3, -2^6_4, -2^5_3\}$	0.28	$\{-2^5_4, -2^5_3, -1^5_4\}$	0.40	$\{-1^5_3, 0^5_3, -1^5_3\}$	0.39	$\{-4^5_3, -3^5_3, -3^5_4\}$	0.18	$\{-1^5_5, 0^5_4, -2^8_6\}$	0.27	$\{-2^5_3, -3^5_3, -3^6_4\}$	0.54
$\{-1^9_5, -2^9_5, -3^9_5\}$	0.23	$\{0^5_3, 0^6_4, 0^7_5\}$	0.28	$\{0^5_3, 0^5_4, 0^5_3\}$	0.40	$\{0^5_3, -1^5_3, 0^5_3\}$	0.39	$\{-5^8_7, -5^8_6, -5^8_7\}$	0.18	$\{0^5_4, -2^8_6, -4^{10}_8\}$	0.27	$\{-3^7_5, -3^6_4, -3^5_3\}$	0.54
$\{-5^4_4, -5^4_3, -4^3_1\}$	0.21	$\{0^7_5, 0^6_4, 0^5_3\}$	0.28	$\{-2^6_4, -2^5_3, -1^5_4\}$	0.40	$\{-2^5_3, -2^5_4, -2^5_3\}$	0.37	$\{-3^5_3, -3^5_4, -2^4_4\}$	0.18	$\{-2^8_6, -4^{10}_8, -3^{10}_6\}$	0.27	$\{-3^7_5, -2^5_3, -1^3_1\}$	0.54
$\{-2^9_5, -1^9_5, -2^9_5\}$	0.19	$\{-2^6_4, -2^7_5, -2^6_4\}$	0.27	$\{-2^5_3, -2^5_5, -1^5_4\}$	0.40	$\{-1^8_6, -2^{10}_8, -1^8_6\}$	0.37	$\{-3^5_4, -2^4_4, -2^4_3\}$	0.18	$\{-2^9_6, -2^8_6, -2^7_6\}$	0.27	$\{-1^5_4, -1^5_3, 0^5_1\}$	0.54
$\{-4^9_5, -3^9_5, -2^9_5\}$	0.19	$\{-3^8_6, -3^7_5, -2^5_3\}$	0.27	$\{0^6_4, 0^5_3, 0^6_4\}$	0.40	$\{0^5_3, -1^5_3, -1^5_4\}$	0.37	$\{-2^7_5, -2^6_5, -2^5_5\}$	0.18	$\{-1^7_3, 0^5_1, -1^8_5\}$	0.23	$\{-3^8_5, -3^8_7, -3^8_6\}$	0.54

Table 2. Twenty most likely 2-grams for each song class together with the corresponding likelihood values in %.

SHE		SAM		SVA		KAK		GUR		GEL		IME	
$\{-1^3_1, 0^1_1\}$	1.25	$\{-1^3_1, 0^1_1\}$	2.54	$\{0^5_3, -1^5_3\}$	4.10	$\{-1^5_4, -1^5_3\}$	1.59	$\{-1^3_1, 0^1_1\}$	1.07	$\{-1^5_3, 0^5_1\}$	0.96	$\{-1^5_3, 0^5_1\}$	2.89
$\{-2^2_2, -1^3_1\}$	0.85	$\{-2^6_4, -2^5_3\}$	1.98	$\{-1^5_3, -2^5_3\}$	3.60	$\{0^5_3, -1^5_3\}$	1.47	$\{-2^4_2, -1^3_1\}$	0.93	$\{-4^5_3, -3^5_1\}$	0.96	$\{0^6_4, 0^5_3\}$	2.46
$\{-4^5_3, -3^5_1\}$	0.81	$\{-2^5_3, -1^3_1\}$	1.72	$\{0^6_4, 0^5_3\}$	2.38	$\{-1^5_3, 0^5_1\}$	1.36	$\{-4^5_3, -3^5_3\}$	0.69	$\{-2^8_6, -2^7_6\}$	0.88	$\{0^5_3, -1^5_3\}$	2.14
$\{-4^9_5, -3^9_5\}$	0.69	$\{0^5_3, -1^5_3\}$	1.29	$\{-1^3_1, 0^1_1\}$	2.27	$\{-1^7_5, -1^6_4\}$	1.18	$\{-1^5_3, 0^5_1\}$	0.65	$\{-1^3_1, 0^1_1\}$	0.77	$\{0^7_5, 0^6_4\}$	1.82
$\{-2^9_5, -3^9_5\}$	0.63	$\{-1^6_4, -1^5_3\}$	1.26	$\{-1^5_4, -1^5_3\}$	2.27	$\{-1^8_5, -1^8_6\}$	1.09	$\{-4^5_5, -4^5_3\}$	0.46	$\{-4^8_4, -4^8_5\}$	0.69	$\{-2^5_3, -1^5_3\}$	1.82
$\{-2^5_3, -1^5_1\}$	0.62	$\{-1^5_3, -2^5_3\}$	1.22	$\{-2^5_3, -1^3_1\}$	1.99	$\{-1^5_3, -1^5_4\}$	1.07	$\{-3^5_3, -2^5_3\}$	0.44	$\{-4^8_5, -4^8_6\}$	0.65	$\{0^5_3, 0^6_4\}$	1.18
$\{-3^9_5, -2^9_5\}$	0.57	$\{-2^7_5, -2^6_4\}$	1.12	$\{0^5_4, 0^5_3\}$	1.83	$\{-1^8_6, -1^7_5\}$	1.05	$\{-2^5_3, -1^5_3\}$	0.44	$\{-3^5_3, -2^5_1\}$	0.61	$\{-3^8_6, -3^7_5\}$	1.18
$\{-1^5_3, 0^5_1\}$	0.56	$\{-2^5_4, -2^5_3\}$	1.08	$\{-2^5_3, -1^5_3\}$	1.66	$\{-1^6_4, -1^5_3\}$	1.03	$\{-1^5_3, 0^5_4\}$	0.42	$\{-2^5_3, -1^5_3\}$	0.54	$\{-1^7_5, -1^6_4\}$	1.07
$\{-3^6_5, -3^5_5\}$	0.49	$\{-2^5_3, -1^5_3\}$	0.95	$\{-1^6_4, -1^5_3\}$	1.62	$\{-1^8_7, -1^8_6\}$	1.01	$\{-4^5_5, -4^5_4\}$	0.40	$\{-2^8_6, -2^7_5\}$	0.54	$\{-1^6_4, -1^5_3\}$	1.07
$\{-5^9_5, -4^9_5\}$	0.48	$\{-1^7_5, -1^6_4\}$	0.94	$\{-2^6_4, -2^5_3\}$	1.55	$\{-1^8_6, -1^8_7\}$	0.94	$\{-3^5_3, -3^5_4\}$	0.40	$\{-1^8_6, -1^7_5\}$	0.50	$\{3^5_3, 2^5_3\}$	1.07
$\{-3^9_5, -4^9_5\}$	0.48	$\{0^6_4, 0^5_3\}$	0.92	$\{-2^5_3, -1^5_4\}$	1.42	$\{0^5_3, 0^5_4\}$	0.94	$\{-4^6_5, -4^5_5\}$	0.38	$\{-3^5_5, -2^5_3\}$	0.50	$\{-1^5_3, -2^5_3\}$	0.96
$\{-4^8_5, -4^7_5\}$	0.46	$\{-1^5_3, 0^5_3\}$	0.78	$\{-2^5_4, -2^5_3\}$	1.29	$\{-1^5_3, 0^5_3\}$	0.94	$\{-4^5_5, -3^5_3\}$	0.36	$\{-2^7_6, -1^5_5\}$	0.50	$\{-3^7_5, -2^5_3\}$	0.96
$\{-4^9_5, -5^9_5\}$	0.45	$\{-2^4_2, -1^3_1\}$	0.73	$\{-3^7_5, -2^5_3\}$	1.00	$\{-2^5_4, -2^5_3\}$	0.92	$\{-5^8_7, -5^8_6\}$	0.36	$\{-1^9_6, -1^8_6\}$	0.50	$\{0^6_4, 0^7_5\}$	0.86
$\{-4^7_5, -4^6_5\}$	0.44	$\{0^7_5, 0^6_4\}$	0.73	$\{0^5_4, 0^6_4\}$	0.94	$\{0^5_4, 0^5_3\}$	0.86	$\{-4^8_5, -4^7_5\}$	0.36	$\{0^5_4, -2^8_6\}$	0.46	$\{3^3_1, 3^4_2\}$	0.86
$\{-2^3_1, -2^2_2\}$	0.43	$\{-3^7_5, -2^5_3\}$	0.71	$\{-1^5_3, -1^5_4\}$	0.94	$\{-1^8_4, -1^8_5\}$	0.86	$\{-2^6_4, -1^5_3\}$	0.34	$\{-1^7_3, 0^5_1\}$	0.42	$\{3^4_2, 3^5_3\}$	0.86
$\{-5^9_5, -4^7_5\}$	0.42	$\{-2^8_6, -2^7_5\}$	0.62	$\{0^5_3, 0^6_4\}$	0.94	$\{-2^8_5, -2^8_6\}$	0.86	$\{-2^7_5, -2^6_5\}$	0.34	$\{0^8_4, 0^8_5\}$	0.42	$\{-2^5_3, -3^5_3\}$	0.86
$\{-6^9_5, -5^9_5\}$	0.40	$\{0^6_4, 0^7_5\}$	0.59	$\{-1^5_3, 0^1_1\}$	0.92	$\{-2^8_6, -2^8_7\}$	0.86	$\{-4^5_3, -3^5_4\}$	0.32	$\{-5^8_6, -4^8_4\}$	0.42	$\{-3^7_5, -3^6_4\}$	0.86
$\{-2^7_5, -2^6_5\}$	0.39	$\{-4^5_3, -3^5_3\}$	0.59	$\{-1^3_3, 0^1_1\}$	0.79	$\{0^6_4, 0^5_3\}$	0.84	$\{-2^3_3, -2^4_2\}$	0.32	$\{-3^8_5, -3^8_6\}$	0.42	$\{-1^5_3, -1^6_4\}$	0.75
$\{-4^6_5, -4^5_5\}$	0.39	$\{0^5_3, 0^6_4\}$	0.58	$\{0^1_1, 0^5_3\}$	0.79	$\{-2^{10}_8, -1^8_6\}$	0.80	$\{-4^5_3, -4^5_4\}$	0.30	$\{-2^7_5, -2^6_4\}$	0.42	$\{-1^6_4, -1^7_5\}$	0.75
$\{-6^9_8, -6^8_8\}$	0.38	$\{-1^5_3, 0^5_1\}$	0.58	$\{-1^5_4, -1^6_4\}$	0.74	$\{-1^5_3, -2^5_3\}$	0.80	$\{-4^7_5, -4^6_5\}$	0.30	$\{-2^6_4, -1^5_3\}$	0.42	$\{3^6_4, 3^5_3\}$	0.75

Table 3. Twenty most likely 3-grams for each song class together with the corresponding likelihood values in %.

SHE		SAM		SVA		KAK		GUR		GEL		IME	
$\{-2^2_2, -1^3_1, 0^1_1\}$	0.86	$\{-2^5_3, -1^3_1, 0^1_1\}$	1.13	$\{0^5_3, -1^5_3, -2^5_3\}$	2.03	$\{-1^5_4, -1^5_3, 0^5_1\}$	0.75	$\{-2^4_2, -1^3_1, 0^1_1\}$	0.85	$\{-4^8_4, -4^8_5, -4^8_6\}$	0.54	$\{0^6_4, 0^5_3, -1^5_3\}$	1.63
$\{-2^3_1, -2^2_2, -1^3_1\}$	0.40	$\{-2^7_5, -2^6_4, -2^5_3\}$	0.75	$\{0^5_4, 0^5_3, -1^5_3\}$	1.07	$\{-1^8_4, -1^8_5, -1^8_6\}$	0.75	$\{-2^3_3, -2^4_2, -1^3_1\}$	0.33	$\{-2^8_6, -2^7_6, -1^5_5\}$	0.50	$\{0^7_5, 0^6_4, 0^5_3\}$	1.53
$\{-4^7_5, -4^6_5, -4^5_5\}$	0.31	$\{0^5_3, -1^5_3, -2^5_3\}$	0.69	$\{0^6_4, 0^5_3, -1^5_3\}$	1.00	$\{-1^7_5, -1^6_4, -1^5_3\}$	0.58	$\{-2^4_3, -2^4_2, -1^3_1\}$	0.26	$\{-2^2_2, -1^3_1, 0^1_1\}$	0.39	$\{-1^6_4, -1^5_3, 0^5_1\}$	1.09
$\{-3^4_2, -2^2_2, -1^3_1\}$	0.29	$\{-2^4_2, -1^3_1, 0^1_1\}$	0.68	$\{-2^5_3, -1^3_1, 0^1_1\}$	0.89	$\{-1^8_6, -1^8_7, -1^8_6\}$	0.54	$\{2^4_2, 3^3_1, 4^1_1\}$	0.24	$\{-1^8_6, -1^9_6, -1^8_6\}$	0.39	$\{-1^7_5, -1^6_4, -1^5_3\}$	0.98
$\{-3^4_2, -2^3_1, -2^2_2\}$	0.28	$\{-2^6_4, -2^5_3, -1^3_1\}$	0.53	$\{-2^5_3, -1^5_4, -1^5_3\}$	0.89	$\{-2^8_4, -2^8_5, -2^8_6\}$	0.54	$\{-4^5_5, -4^5_4, -4^5_3\}$	0.22	$\{0^8_4, 0^8_5, 0^8_6\}$	0.35	$\{0^6_4, 0^7_5, 0^6_4\}$	0.87
$\{-5^7_4, -4^5_3, -3^5_1\}$	0.27	$\{-1^7_5, -1^6_4, -1^5_3\}$	0.49	$\{-1^5_3, -2^5_3, -1^3_1\}$	0.65	$\{0^7_5, 0^6_4, 0^5_3\}$	0.52	$\{1^5_4, 2^4_3, 2^4_2\}$	0.22	$\{-3^8_4, -3^8_5, -3^8_6\}$	0.35	$\{3^3_1, 3^4_2, 3^5_3\}$	0.87
$\{-2^7_4, -1^5_3, 0^5_1\}$	0.26	$\{-2^5_4, -2^5_3, -1^3_1\}$	0.42	$\{0^5_4, 0^6_4, 0^5_3\}$	0.65	$\{-2^8_6, -2^8_7, -2^8_6\}$	0.46	$\{2^4_3, 2^4_2, 3^3_1\}$	0.22	$\{-1^5_3, 0^5_1, -1^8_6\}$	0.31	$\{0^5_3, -1^5_3, -2^5_3\}$	0.87
$\{-3^7_4, -2^5_3, -1^5_1\}$	0.26	$\{0^6_4, 0^7_5, 0^6_4\}$	0.41	$\{-1^5_3, -2^5_3, -1^5_3\}$	0.56	$\{-1^6_4, -1^5_3, 0^5_1\}$	0.44	$\{-4^6_5, -4^5_5, -4^5_4\}$	0.20	$\{-2^6_4, -1^5_3, 0^5_1\}$	0.31	$\{3^6_4, 3^5_3, 2^5_3\}$	0.76
$\{-5^4_3, -4^3_1, -3^1_1\}$	0.26	$\{-1^6_4, -1^7_5, -1^6_4\}$	0.38	$\{-3^7_5, -2^5_3, -1^3_1\}$	0.51	$\{-1^8_6, -1^7_5, -1^6_4\}$	0.42	$\{0^5_4, 0^6_4, 0^7_4\}$	0.20	$\{-2^8_6, -2^9_6, -2^8_6\}$	0.31	$\{-3^8_6, -3^7_5, -3^6_4\}$	0.76
$\{-2^9_5, -3^9_5, -4^9_5\}$	0.25	$\{0^6_4, 0^5_3, -1^5_3\}$	0.37	$\{0^5_3, 0^6_4, 0^5_3\}$	0.49	$\{0^5_3, 0^5_4, -1^5_4\}$	0.41	$\{0^7_4, 0^6_4, 0^5_4\}$	0.20	$\{-2^5_4, -1^5_3, 0^5_1\}$	0.31	$\{-3^9_7, -3^8_6, -3^7_5\}$	0.76
$\{-2^7_5, -2^6_5, -2^5_5\}$	0.24	$\{-1^5_3, 0^5_3, -1^5_3\}$	0.34	$\{-1^5_3, -1^3_1, 0^1_1\}$	0.47	$\{0^5_4, -1^5_4, -1^5_3\}$	0.41	$\{-2^4_4, -2^4_3, -2^4_2\}$	0.20	$\{-2^7_6, -1^5_5, 0^5_4\}$	0.31	$\{-2^5_3, -1^5_3, -1^6_4\}$	0.65
$\{-3^5_5, -3^5_4, -3^5_3\}$	0.24	$\{-2^8_6, -2^7_5, -2^6_4\}$	0.33	$\{-2^5_3, -1^6_4, -1^5_3\}$	0.45	$\{-1^5_5, -1^5_4, -1^5_3\}$	0.41	$\{0^5_5, 1^5_4, 2^4_3\}$	0.20	$\{-5^5_4, -4^5_3, -3^5_1\}$	0.27	$\{3^5_3, 2^5_3, 1^5_3\}$	0.65
$\{-6^9_8, -6^8_8, -5^7_6\}$	0.24	$\{0^7_5, 0^6_4, 0^7_5\}$	0.31	$\{-2^6_4, -2^5_3, -1^5_3\}$	0.45	$\{-1^6_4, -1^7_5, -1^6_4\}$	0.41	$\{-4^5_3, -4^5_4, -3^4_4\}$	0.18	$\{-5^{10}_6, -5^{10}_7, -5^{10}_8\}$	0.27	$\{-1^5_3, -2^5_3, -3^5_3\}$	0.65
$\{-6^8_8, -5^7_6, -5^7_4\}$	0.24	$\{-2^5_3, -2^4_2, -1^3_1\}$	0.29	$\{-1^5_3, 0^6_4, 0^5_3\}$	0.40	$\{-1^7_5, -1^6_4, -1^7_5\}$	0.41	$\{-4^7_5, -4^6_5, -4^5_5\}$	0.18	$\{-2^7_5, -2^6_4, -1^5_3\}$	0.27	$\{-3^5_3, -3^6_4, -3^7_5\}$	0.65
$\{-5^7_6, -5^7_4, -4^6_4\}$	0.24	$\{-3^7_5, -2^5_3, -1^3_1\}$	0.29	$\{-3^5_3, -2^5_4, -2^5_3\}$	0.40	$\{-2^{10}_8, -1^8_6, -1^8_7\}$	0.39	$\{-1^5_3, 0^5_4, 0^6_4\}$	0.18	$\{-1^8_6, -2^8_6, -2^7_6\}$	0.27	$\{0^5_3, 0^6_4, 0^5_3\}$	0.54
$\{-5^9_5, -4^9_5, -3^9_5\}$	0.24	$\{-2^5_3, -2^6_4, -2^5_3\}$	0.28	$\{-2^5_4, -2^5_3, -1^5_4\}$	0.40	$\{-1^5_3, 0^5_3, -1^5_3\}$	0.39	$\{-4^5_3, -3^5_3, -3^5_4\}$	0.18	$\{-1^5_5, 0^5_4, -2^8_6\}$	0.27	$\{-2^5_3, -3^5_3, -3^6_4\}$	0.54
$\{-1^9_5, -2^9_5, -3^9_5\}$	0.23	$\{0^5_3, 0^6_4, 0^7_5\}$	0.28	$\{0^5_3, 0^5_4, 0^5_3\}$	0.40	$\{0^5_3, -1^5_3, 0^5_3\}$	0.39	$\{-5^8_7, -5^8_6, -5^8_7\}$	0.18	$\{0^5_4, -2^8_6, -4^{10}_8\}$	0.27	$\{-3^7_5, -3^6_4, -3^5_3\}$	0.54
$\{-5^4_4, -5^4_3, -4^3_1\}$	0.21	$\{0^7_5, 0^6_4, 0^5_3\}$	0.28	$\{-2^6_4, -2^5_3, -1^5_4\}$	0.40	$\{-2^5_3, -2^5_4, -2^5_3\}$	0.37	$\{-3^5_3, -3^5_4, -2^4_4\}$	0.18	$\{-2^8_6, -4^{10}_8, -3^{10}_6\}$	0.27	$\{-3^7_5, -2^5_3, -1^3_1\}$	0.54
$\{-2^9_5, -1^9_5, -2^9_5\}$	0.19	$\{-2^6_4, -2^7_5, -2^6_4\}$	0.27	$\{-2^5_3, -2^5_5, -1^5_4\}$	0.40	$\{-1^8_6, -2^{10}_8, -1^8_6\}$	0.37	$\{-3^5_4, -2^4_4, -2^4_3\}$	0.18	$\{-2^9_6, -2^8_6, -2^7_6\}$	0.27	$\{-1^5_4, -1^5_3, 0^5_1\}$	0.54
$\{-4^9_5, -3^9_5, -2^9_5\}$	0.19	$\{-3^8_6, -3^7_5, -2^5_3\}$	0.27	$\{0^6_4, 0^5_3, 0^6_4\}$	0.40	$\{0^5_3, -1^5_3, -1^5_4\}$	0.37	$\{-2^7_5, -2^6_5, -2^5_5\}$	0.18	$\{-1^7_3, 0^5_1, -1^8_5\}$	0.23	$\{-3^8_5, -3^8_7, -3^8_6\}$	0.54

Model validation

Statistically, Tables 1-3, together with the prior probability values shown in Fig. 7, contain all that the corpus can tell us about the characteristic features of each *song class* from the perspective of the twenty most likely n-grams. However, it remains to be clarified whether the chord progression patterns shown in Tables 1-3 can be interpreted as general characteristics for the different 'song classes' or only for the collected data sets. Following machine learning protocol, this needs to be tested by model validation experiments.

In the first validation experiment, we randomly divided the complete dataset into two subsets. For each 'song class', ¾ of the songs were randomly selected for the training of a classification algorithm, which was then applied to the remaining ¼ of the songs. This approach is called cross-validation and is commonly applied (in a variety of different versions) to evaluate the performance of machine learning models. In our case, the process was repeated one hundred times. Since the test data subsets were not used in the model training, this process actually tests the predictive power of the classifiers, in other words their generalizability to unseen data. The way we applied it in the present context was that in each of the 100 cross-validation runs and for each 'song class', we kept track of which of the songs were 'always' properly classified. Since a song might also contain patterns (n-grams) characteristic of different 'song classes', keeping only the intersection of those songs which were always classified properly (which we refer to as 'winner songs' or 'winner n-grams'), has two interesting effects.

First, it reduces the data subsets to those songs which can be considered most characteristic for the assigned 'song classes'. Because of the random selection of the subsets of songs used for the training, each of the trained classifiers for a 'song class' will capture slightly different aspects. Taking the intersections of those songs which are always classified correctly will therefore capture the most representative (pure) features of each 'song class', while those songs which contain a lot of mixed features (from different song classes) will be thrown out. In other words, this process can help to focus on the observed patterns of those songs for which we can be most confident that they represent actual features of the assigned 'song classes' and are not the results of non-specific features of individual songs or the results of mixing features from different 'song classes'.

Second, the sizes of the intersections of the successfully classified songs in all cross-validations, i.e. the sizes of the 'winner subsets', are direct measures of the numbers of songs which can be considered representative of each 'song class'. Fig. 10, which shows the confusion matrices for those 'winner songs' for n-grams with n=1-3, illustrates for example that for the *song classes* GEL, GUR, and IME, the sizes of the datasets are clearly too small to warrant meaningful interpretation. In terms of 1-grams, there is not a single song that is always classified correctly in any of these three song classes and only a few remain for the 2-grams and 3-grams for GEL and GUR. Therefore, we refrain from further interpretation of these song classes at this time.

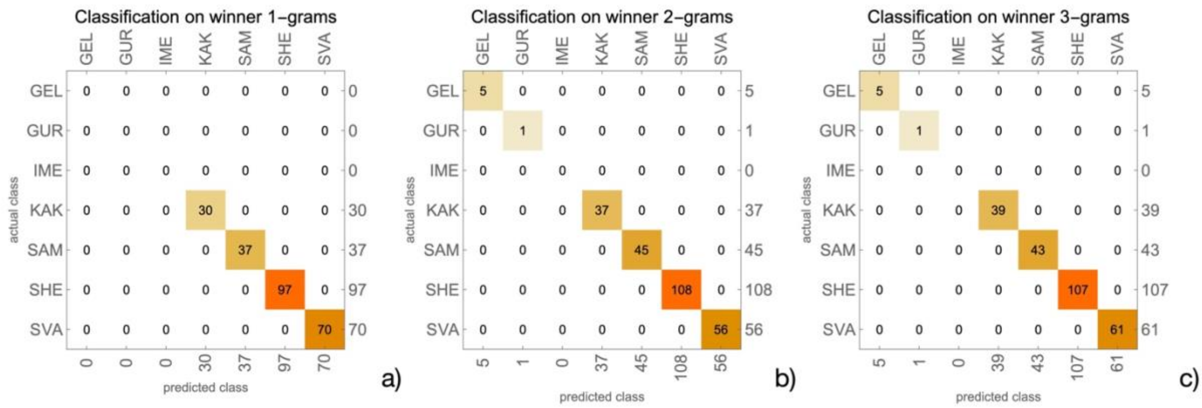

Figure 10. Confusion matrices for those songs, which are always classified correctly in all 100 cross-validation runs ('winner songs'), for n-grams with n=1-3.

The song class with the largest number of representative songs in Fig. 10 is SHE. In terms of 2-grams its size (108) is only slightly reduced with respect to the original dataset (121). Fig. 11a shows the corresponding twenty most likely 2-grams together with their likelihoods in %.

a)

SHE	
$\{-1_1^3, 0_1^1\}$	1.11
$\{-2_2^2, -1_1^3\}$	0.89
$\{-4_3^5, -3_1^5\}$	0.84
$\{-4_5^9, -3_5^9\}$	0.70
$\{-2_5^9, -3_5^9\}$	0.65
$\{-2_3^5, -1_1^5\}$	0.64
$\{-3_5^9, -2_5^9\}$	0.59
$\{-1_3^5, 0_1^5\}$	0.55
$\{-5_5^9, -4_5^9\}$	0.50
$\{-3_5^6, -3_5^5\}$	0.49
$\{-3_5^9, -4_5^9\}$	0.49
$\{-4_5^8, -4_5^7\}$	0.48
$\{-4_5^9, -5_5^9\}$	0.47
$\{-2_1^3, -2_2^2\}$	0.45
$\{-5_5^9, -4_5^7\}$	0.43
$\{-4_5^7, -4_5^6\}$	0.42
$\{-6_5^9, -5_5^9\}$	0.42
$\{-4_5^6, -4_5^5\}$	0.40
$\{-2_5^7, -2_5^6\}$	0.39
$\{-6_8^9, -6_8^8\}$	0.39

b)

SHE		SHE	
$\{-1_1^3, 0_1^1\}$	1.18	$\{-1_1^3, 0_1^1\}$	1.05
$\{-2_2^2, -1_1^3\}$	0.95	$\{-2_2^2, -1_1^3\}$	0.83
$\{-4_5^9, -3_5^9\}$	0.86	$\{-4_3^5, -3_1^5\}$	0.83
$\{-4_3^5, -3_1^5\}$	0.86	$\{-2_5^9, -3_5^9\}$	0.66
$\{-3_5^9, -2_5^9\}$	0.67	$\{-2_3^5, -1_1^5\}$	0.62
$\{-2_3^5, -1_1^5\}$	0.67	$\{-4_5^9, -3_5^9\}$	0.54
$\{-2_5^9, -3_5^9\}$	0.64	$\{-1_3^5, 0_1^5\}$	0.51
$\{-5_5^9, -4_5^9\}$	0.60	$\{-3_5^9, -2_5^9\}$	0.51
$\{-1_3^5, 0_1^5\}$	0.58	$\{-3_5^6, -3_5^5\}$	0.45
$\{-3_5^6, -3_5^5\}$	0.54	$\{-3_5^9, -4_5^9\}$	0.45
$\{-5_4^8, -4_4^8\}$	0.54	$\{-4_5^8, -4_5^7\}$	0.44
$\{-6_5^9, -5_5^9\}$	0.54	$\{-4_5^7, -4_5^6\}$	0.44
$\{-3_5^9, -4_5^9\}$	0.52	$\{-2_1^3, -2_2^2\}$	0.42
$\{-4_5^9, -5_5^9\}$	0.52	$\{-4_5^9, -5_5^9\}$	0.42
$\{-4_5^8, -4_5^7\}$	0.52	$\{-2_5^9, -2_5^7\}$	0.41
$\{-5_5^9, -4_5^7\}$	0.51	$\{-5_5^9, -4_5^9\}$	0.41
$\{-5_5^9, -6_5^9\}$	0.51	$\{-2_5^7, -2_5^6\}$	0.38
$\{-2_1^3, -2_2^2\}$	0.48	$\{-4_5^6, -4_5^5\}$	0.38
$\{-4_1^3, -3_1^1\}$	0.46	$\{-6_8^9, -6_8^8\}$	0.38
$\{-4_4^8, -4_5^9\}$	0.44	$\{-5_5^9, -4_5^7\}$	0.36

c)

SHE		SHE		SHE		SHE	
$\{-1_1^3, 0_1^1\}$	1.20	$\{-4_3^5, -3_1^5\}$	1.18	$\{-1_1^3, 0_1^1\}$	1.34	$\{-1_1^3, 0_1^1\}$	1.05
$\{-2_2^2, -1_1^3\}$	0.88	$\{-1_1^3, 0_1^1\}$	0.93	$\{-2_2^2, -1_1^3\}$	1.06	$\{-4_5^9, -3_5^9\}$	0.86
$\{-4_3^5, -3_1^5\}$	0.81	$\{-2_2^2, -1_1^3\}$	0.85	$\{-2_3^5, -1_1^5\}$	0.99	$\{-2_2^2, -1_1^3\}$	0.80
$\{-2_3^5, -1_1^5\}$	0.71	$\{-4_5^8, -4_5^7\}$	0.71	$\{-2_5^9, -3_5^9\}$	0.88	$\{-4_3^5, -3_1^5\}$	0.80
$\{-4_5^9, -3_5^9\}$	0.68	$\{-3_5^6, -3_5^5\}$	0.60	$\{-3_5^9, -2_5^9\}$	0.74	$\{-2_5^9, -3_5^9\}$	0.72
$\{-3_5^6, -3_5^5\}$	0.65	$\{-4_5^9, -3_5^9\}$	0.60	$\{-1_3^5, 0_1^5\}$	0.67	$\{-3_5^9, -2_5^9\}$	0.64
$\{-2_5^9, -3_5^9\}$	0.58	$\{-3_1^3, -2_1^1\}$	0.58	$\{-4_5^9, -3_5^9\}$	0.63	$\{-5_4^8, -4_4^8\}$	0.61
$\{-3_5^9, -2_5^9\}$	0.58	$\{-4_5^7, -4_5^8\}$	0.55	$\{-3_5^7, -3_5^6\}$	0.60	$\{-5_5^9, -4_5^9\}$	0.61
$\{-2_5^9, -2_5^7\}$	0.58	$\{-5_4^8, -4_4^8\}$	0.55	$\{-2_1^3, -2_2^2\}$	0.56	$\{-4_5^9, -5_5^9\}$	0.61
$\{-3_5^9, -3_5^7\}$	0.58	$\{-2_5^7, -2_5^6\}$	0.55	$\{-3_4^7, -2_3^5\}$	0.56	$\{-3_5^9, -4_5^9\}$	0.61
$\{-1_3^5, 0_1^5\}$	0.52	$\{-6_5^9, -5_5^9\}$	0.52	$\{-5_5^9, -4_5^7\}$	0.53	$\{-2_3^5, -1_1^5\}$	0.50
$\{-3_5^5, -3_4^5\}$	0.49	$\{-4_5^9, -5_5^9\}$	0.52	$\{-4_5^7, -4_5^6\}$	0.49	$\{-3_4^8, -2_4^8\}$	0.50
$\{-3_4^5, -3_3^5\}$	0.49	$\{-5_5^9, -4_5^9\}$	0.52	$\{-4_3^5, -3_1^5\}$	0.49	$\{-5_5^9, -6_5^9\}$	0.50
$\{-3_5^9, -4_5^9\}$	0.45	$\{-1_3^5, 0_1^5\}$	0.52	$\{-4_5^6, -4_5^7\}$	0.46	$\{-1_3^5, 0_1^5\}$	0.50
$\{-5_5^9, -4_5^7\}$	0.45	$\{-4_5^9, -4_5^8\}$	0.49	$\{-4_5^6, -4_5^5\}$	0.46	$\{-4_5^7, -4_5^6\}$	0.44
$\{-6_8^9, -6_8^8\}$	0.42	$\{-5_3^4, -4_1^3\}$	0.49	$\{-3_2^4, -2_1^3\}$	0.46	$\{-2_1^3, -2_2^2\}$	0.44
$\{-3_2^4, -2_2^2\}$	0.42	$\{-4_1^3, -3_1^1\}$	0.49	$\{-3_5^6, -3_5^5\}$	0.46	$\{-4_4^8, -3_4^8\}$	0.44
$\{-4_5^8, -4_5^7\}$	0.42	$\{-4_5^9, -4_5^7\}$	0.47	$\{-2_5^7, -2_5^6\}$	0.46	$\{-2_4^8, -2_5^9\}$	0.41
$\{-4_4^8, -3_4^8\}$	0.39	$\{-5_5^9, -4_5^7\}$	0.47	$\{-2_4^7, -1_3^5\}$	0.46	$\{-4_5^6, -4_5^5\}$	0.41
$\{-5_5^9, -4_5^9\}$	0.39	$\{-6_8^9, -6_8^8\}$	0.47	$\{-5_5^9, -4_5^9\}$	0.46	$\{-4_5^8, -4_5^7\}$	0.39

Figure 11. Twenty most likely 2-grams for the 'song class' SHE and the dataset built from the 'winner songs' of the cross-validation experiment described above, for the 2-grams split up into different equally sized subsets. Fig. 11a-c show the results for the original dataset, the original dataset split up into two equally subsets, and 4 equally sized subsets, respectively.

Finally, to test the extent to which the likelihoods shown in Fig. 11a can be trusted as expressing the harmonic structure of the song class SHE, we conducted another validation experiment. For this purpose, we split the original data set into two and four equally sized sub-data sets, respectively, to check whether the essential properties of the n-gram distributions are preserved in this process. If they are so preserved, we see this as an indication that they are indeed properties of the song class because they are present in each of the subdivisions. To visualize the n-grams that remain as the most frequently occurring ones during splitting, we have chosen a simple color code. The fields coded green in Fig. 11b mark n-grams contained in both halves of the original dataset. The fields coded red and yellow in Fig. 11c are the n-grams that are retained in all 4 or 3 of the four 1/4 subsets, respectively. What we can see in Fig. 11 is that the majority of the chord progressions in Fig. 11a appear also in both halves of Fig. 11b, each containing 54 songs. However, when split into four sub-datasets, each still containing 27 songs, the representational power of the individual sub-datasets decreases significantly. Only the 3-4 most frequent chord sequences still represent the entire data set, as can be seen from the red color coding.

The conclusion we draw from this experiment is that for a corpus size of approximately 50 songs, which would correspond to the two subsets shown in Fig. 11b, we can expect the 10 to 15 most frequently observed chord progression patterns (n-grams) to reflect the syntactic structure of their corresponding 'song class', but that this is not guaranteed for smaller corpus sizes. In our case, this leaves only the song classes SHE and SVA available for interpretation. Their twenty most likely n-grams for n =1-3 is shown in Fig. 12.

SHE

1-gram		2-gram		3-gram	
$\{-4_5^9\}$	3.19	$\{-1_1^3, 0_1^1\}$	1.11	$\{-2_2^2, -1_1^3, 0_1^1\}$	0.90
$\{-3_5^9\}$	2.90	$\{-2_2^2, -1_1^3\}$	0.89	$\{-2_1^3, -2_2^2, -1_1^3\}$	0.42
$\{-4_5^7\}$	2.68	$\{-4_3^5, -3_1^5\}$	0.84	$\{-4_5^7, -4_5^6, -4_5^5\}$	0.32
$\{-5_5^9\}$	2.29	$\{-4_5^9, -3_5^9\}$	0.70	$\{-3_2^4, -2_2^2, -1_1^3\}$	0.30
$\{-4_3^5\}$	2.24	$\{-2_5^9, -3_5^9\}$	0.65	$\{-3_2^4, -2_1^3, -2_2^2\}$	0.29
$\{-2_5^9\}$	2.07	$\{-2_3^5, -1_1^5\}$	0.64	$\{-5_4^7, -4_3^5, -3_1^5\}$	0.29
$\{-3_5^7\}$	1.65	$\{-3_5^9, -2_5^9\}$	0.59	$\{-2_4^7, -1_3^5, 0_1^5\}$	0.27
$\{-4_4^8\}$	1.65	$\{-1_3^5, 0_1^5\}$	0.55	$\{-5_3^4, -4_1^3, -3_1^1\}$	0.27
$\{-3_5^5\}$	1.55	$\{-5_5^9, -4_5^9\}$	0.50	$\{-2_5^9, -3_5^9, -4_5^9\}$	0.26
$\{-5_5^7\}$	1.48	$\{-3_5^6, -3_5^5\}$	0.49	$\{-3_5^5, -3_4^5, -3_3^5\}$	0.26
$\{-4_6^8\}$	1.44	$\{-3_5^9, -4_5^9\}$	0.49	$\{-2_5^7, -2_5^6, -2_5^5\}$	0.25
$\{-4_5^8\}$	1.42	$\{-4_5^8, -4_5^7\}$	0.48	$\{-6_8^9, -6_8^8, -5_6^7\}$	0.25
$\{-3_4^8\}$	1.39	$\{-4_5^9, -5_5^9\}$	0.47	$\{-6_8^8, -5_6^7, -5_4^7\}$	0.25
$\{-3_3^5\}$	1.24	$\{-2_1^3, -2_2^2\}$	0.45	$\{-5_6^7, -5_4^7, -4_4^6\}$	0.25
$\{-6_5^9\}$	1.23	$\{-5_5^9, -4_5^7\}$	0.43	$\{-1_5^9, -2_5^9, -3_5^9\}$	0.24
$\{-4_5^6\}$	1.21	$\{-4_5^7, -4_5^6\}$	0.42	$\{-5_5^9, -4_5^9, -3_5^9\}$	0.24
$\{-2_5^7\}$	1.19	$\{-6_5^9, -5_5^9\}$	0.42	$\{-3_4^7, -2_3^5, -1_1^5\}$	0.24
$\{-1_1^3\}$	1.18	$\{-4_5^6, -4_5^5\}$	0.40	$\{-5_4^4, -5_3^4, -4_1^3\}$	0.22
$\{-4_5^5\}$	1.17	$\{-2_5^7, -2_5^6\}$	0.39	$\{-2_5^9, -1_5^9, -2_5^9\}$	0.20
$\{-3_5^8\}$	1.15	$\{-6_8^9, -6_8^8\}$	0.39	$\{-4_5^9, -3_5^9, -2_5^9\}$	0.20

a)

SVA

1-gram		2-gram		3-gram	
$\{-1_3^5\}$	16.30	$\{0_3^5, -1_3^5\}$	4.53	$\{0_3^5, -1_3^5, -2_3^5\}$	2.20
$\{0_3^5\}$	9.76	$\{-1_3^5, -2_3^5\}$	4.40	$\{-2_3^5, -1_4^5, -1_3^5\}$	1.14
$\{-2_3^5\}$	9.04	$\{-1_4^5, -1_3^5\}$	3.37	$\{-1_3^5, -2_3^5, -1_1^3\}$	0.98
$\{0_1^1\}$	5.42	$\{-1_1^3, 0_1^1\}$	2.43	$\{0_4^5, 0_3^5, -1_3^5\}$	0.91
$\{-1_4^5\}$	5.18	$\{0_4^5, 0_3^5\}$	2.18	$\{0_4^6, 0_3^5, -1_3^5\}$	0.91
$\{-1_1^3\}$	4.97	$\{0_4^6, 0_3^5\}$	1.85	$\{0_4^5, 0_4^6, 0_3^5\}$	0.87
$\{0_4^5\}$	4.73	$\{-2_3^5, -1_1^3\}$	1.56	$\{-1_3^5, -1_1^3, 0_1^1\}$	0.79
$\{0_4^6\}$	3.70	$\{-1_3^5, -1_4^5\}$	1.44	$\{-1_4^5, -1_4^6, -1_3^5\}$	0.64
$\{0_1^3\}$	3.25	$\{-2_3^5, -1_3^5\}$	1.44	$\{-2_5^5, -1_4^5, -1_3^5\}$	0.61
$\{-1_4^6\}$	2.68	$\{-1_4^6, -1_3^5\}$	1.44	$\{-1_3^5, -1_4^5, -1_3^5\}$	0.57
$\{0_1^4\}$	2.65	$\{-2_3^5, -1_4^5\}$	1.40	$\{-1_3^5, -2_3^5, -1_3^5\}$	0.57
$\{-1_3^3\}$	2.41	$\{-1_3^5, -1_1^3\}$	1.32	$\{-1_4^5, -1_3^5, -2_3^5\}$	0.57
$\{-2_5^5\}$	2.26	$\{0_1^1, 0_3^5\}$	1.15	$\{0_1^4, 0_1^3, 0_1^4\}$	0.57
$\{-2_4^6\}$	1.96	$\{0_1^4, 0_1^3\}$	1.15	$\{-2_3^5, -1_1^3, 0_1^1\}$	0.53
$\{-1_1^5\}$	1.72	$\{-1_3^5, 0_1^1\}$	1.11	$\{-1_4^5, -1_3^5, -1_1^3\}$	0.53
$\{1_1^3\}$	1.63	$\{0_4^5, 0_4^6\}$	1.03	$\{0_1^3, 0_1^4, 0_1^3\}$	0.53
$\{-3_5^7\}$	1.51	$\{-2_4^6, -2_3^5\}$	0.99	$\{-1_3^5, -1_2^4, -1_1^3\}$	0.49
$\{-2_5^7\}$	1.27	$\{0_1^3, 0_1^4\}$	0.95	$\{-2_3^5, -2_5^5, -1_4^5\}$	0.49
$\{0_1^5\}$	1.23	$\{-1_3^3, 0_1^1\}$	0.95	$\{0_3^5, 0_4^6, 0_3^5\}$	0.42
$\{1_1^1\}$	1.20	$\{-2_3^5, -2_5^5\}$	0.95	$\{0_3^5, 0_4^5, 0_3^5\}$	0.42

b)

Figure 12. Twenty most likely n-grams for n =1-3 for the 'winner songs' of song classes SHE (a) and SVA (b), together with the corresponding likelihood values in %.

Even at first glance, Fig. 12a and 12b show strong differences in the chord inventories but also the chord progression structures between the two subsets for SHE and SVA. However, we refrain from a more detailed musicological interpretation of the observed patterns in the probability distributions for the individual n-grams and leave this to our follow-up study.

Discussion

The present study demonstrates how the components of a Markov-model-based classification algorithm can be used as a key component of a workflow to analyze the syntactical harmonic structure of traditional Georgian music from digital scores. Our analysis represents a delayed follow-up study to the work of Arom and Vallejo (2008, 2010). Their key questions are still open, and a specific answer is yet beyond the scope of the present work. However, by showing that – for the extraction of chord sequences - the differences between the notation of a traditional Georgian song in Western notation and a more appropriate heptatonic tuning system can be accounted for, we believe that we have demonstrated that in principle all the required information can be extracted from transcribed traditional Georgian music by computational analysis. We are aware that the size of the data set used in our study, which is already hugely enlarged in comparison to the original dataset of (Arom and Vallejo, 2008; 2010), needs to be further increased in order to establish stronger confidence in the generalizability of the observed features and we are also aware that trying to understand the syntax of the songs needs more than n-gram frequencies. However, since the entire procedure used in this paper was implemented as a scalable workflow in Mathematica (Wolfram Research, 2020), it will be straightforward to extend the analysis to arbitrarily large data sets once these become available in digital form.

REFERENCES

Akhobadze, Vladimer. (1957). *Collection of Georgian (Svan) Folk Songs*. Tbilisi: Shroma da Teknika.

Arom, Simha; Vallejo, Pollo. (2008). "Towards a Theory of the Chord Syntax of Georgian Polyphony." [The Fourth International Symposium on Traditional Polyphony] Eds. Rusudan Tsurtsumia and Joseph Jordania: pp. 321–335. Tbilisi: International Research Center for Traditional Polyphony of Tbilisi State Conservatoire (in Georgian and English).

Arom, Simha; Vallejo, Pollo. (2010). "Outline of a Syntax of Chords in Some Songs from Samegrelo." [The Fifth International Symposium on Traditional Polyphony] Eds. Rusudan Tsurtsumia and Joseph Jordania: pp. 266–277. Tbilisi: International Research Center for Traditional Polyphony of Tbilisi State Conservatoire (in Georgian and English).

Bernard, Etienne. (2021). *Introduction to Machine Learning*. Champaign, Illinois: Wolfram Media, Inc.

Center of Church Chants of the Patriarchate. (2006a). *Georgian Church Chant, Gelati School, The Hymns of Immovable Celebrations and Twelve Celebrations of Our Lord*. Vol. Volume 2. Tbilisi.

Center of Church Chants of the Patriarchate. (2006b). *Georgian Church Chant, Gelati School, Vespers-Matins-Liturgy*. Vol. Volume 1. Tbilisi.

Center of Church Chants of the Patriarchate. (2008). *Georgian Church Chant, Eastern School (Karbelashvili Mode), Vespers-Matins-Liturgy*. Vol. Volume 3. 2nd edition. Tbilisi.

Chokhonelidze, Evsevi. (2003). *Georgian Folk Music: Samegrelo*. Tbilisi: International

Center for Georgian Folk Song, Tbilisi State Conservatoire International Research Center for Traditional Polyphony.

Folklore state Centre of Georgia, Georgian Chanting Foundation. (2018a). *Georgian Chant Anthology, The Chanting Tradition of Western Georgia, The Divine Liturgies of Saint John Chrysostom, Saint Basil the Great, and the Presanctified Gifts Paschal Hymns Festal Hymns of the Liturgy Festal Communion Verses*. Vol. Volume 5. Tbilisi.

Folklore state Centre of Georgia, Georgian Chanting Foundation. (2018b). *Georgian Chant Anthology, The Chanting Traditions of Eastern Georgia, Chants for the Liturgy of St. John Chrysostom (Part I)*. Vol. Volume 2. Tbilisi.

Folklore state Centre of Georgia, Georgian Chanting Foundation. (2018c). *Georgian Chant Anthology, The Chanting Traditions of Eastern Georgia, Chants for the Liturgy of St. John Chrysostom (Part II)*. Vol. Volume 3. Tbilisi.

Folklore state Centre of Georgia, Georgian Chanting Foundation. (2020a). *Georgian Chant Anthology, Chants of the Triodion and Pentecostarion, Chant for Matins*. Vol. Volume 16. Tbilisi.

Folklore state Centre of Georgia, Georgian Chanting Foundation. (2020b). *Georgian Chant Anthology, Chants of the Triodion and Pentecostarion, Chant for Matins*. Vol. Volume 15. Tbilisi.

Folklore state Centre of Georgia, Georgian Chanting Foundation. (2020c). *Georgian Chant Anthology, Feast Day Troparia, Chants of the Triodion and Pentecostarion, Supplicatory Paracleses for the Most Holy Theotokos and Our Sweet Lord Jesus*. Vol. Volume 13. Tbilisi.

Folklore state Centre of Georgia, Georgian Chanting Foundation. (2020d). *Georgian Chant Anthology, The Chanting Tradition of Western Georgia, Chants of Vespers*. Vol. Volume 12. Tbilisi.

Jordania, Joseph. (2004). *99 Georgian Songs: A Collection of Traditional Folk, Church and Urban Songs from Georgia*. (Ed.) J. Mills. Wales-UK: Black Mountain Press.

Jordania, Joseph. (2022). "Continuing Discussions on Scale Systems in Georgian Traditional Music." *Anzor Erkomaishvili and Contemporary Trends in the Performance and study of Georgian traditional and sacred music*, Eds, J. Jordania and R. Tsutstumia: pp. 116-

146. Cambridge, UK: Cambridge Scholars Publishing.

Patarava, Dimitri. 2003. *Georgian Church and Salkhino Hymns, Guruli Mode*. Tbilisi.

Scherbaum, Frank, Nana Mzhavanadze, Simha Arom, Sebastian Rosenzweig, and Meinard Müller. (2020). *Tonal Organization of the Erkomaishvili Dataset: Pitches, Scales, Melodies and Harmonies*. Potsdam: Universitätsverlag Potsdam.

Scherbaum, Frank, Nana Mzhavanadze, Sebastian Rosenzweig, and Meinard Müller. (2022). "Tuning Systems of Traditional Georgian Singing Determined From a New Corpus of Field Recordings." *Musicologist* 6(2):142–68. doi: 10.33906/musicologist.1068947.

Shugliashvili, Davit. (2014). *Georgian Church Hymns, Shemokmedi School*. Tbilisi: Georgian Chanting Foundation & Tbilisi State Conservatory.

Tarkhnishvili, Maro. (2008). *Monument Protection and Sports of Georgia*. Tbilisi: The Folklore State Centre of Georgia, Ministry of Culture.

Tbilisi State Conservatoire. (2005). *Georgian Folk Musical Creativity*. Tbilisi.

Tsereteli, Zaal; Levan Veshapidze. (2014). "On the Georgian Traditional Scale." [The Seventh International Symposium on Traditional Polyphony] Eds. Rusudan Tsurtsumia and Joseph Jordania: pp. 288–295. Tbilisi: International Research Center for Traditional Polyphony of Tbilisi State Conservatoire (in Georgian and English).

Tsereteli, Zaal; Levan Veshapidze. (2015). "Video of the Presentation 'The Empirical Research of a Georgian Sound Scale.'" *2015 IAML/IMS Congress*. New York City, USA.

Veshapidze, Levan. (2006). *Gurian Folk Songs*. Tbilisi.

Veshapidze, Levan; Sopho Kotrikadze. (2016). *Folk Songs of Ajara*. Tbilisi: Ministry of Education, Culture and Sport of Ajara Autonomous Republic, Folklore State Centre of Georgia.

Wolfram Research, Inc. (2020). *Mathematica*. Version 12. Champaign, Illinois: Wolfram Research, Inc.