

Machine Learning Classification of Regional Swiss Yodel Styles Based on Their Melodic Attributes

Yannick Wey¹ and Cornelia Metzиг²

Abstract

A classification of wordless yodel melodies from five different regions in Switzerland was made. For our analysis, we used a total of 217 yodel tunes from five regions, which can be grouped into two larger regions, central and north-eastern Switzerland. The results show high accuracy of classification, therefore confirming the existence of regional differences in yodel melodies. The most salient features, such as rhythmic patterns or intervals, demonstrate some of the key differences in pairwise comparisons, which can be confirmed by a postanalysis survey of the relevant scores.

Keywords

Classification, folk music, machine learning, Switzerland, yodel

Submission date: 27 May 2020; Acceptance date: 4 March 2021

Introduction

Yodeling is one of the signature vocal styles in the Alpine region and beyond. In the case of Switzerland, a classification based on regional characteristics was long proposed in folkloristic literature (Fellmann, 1962; Leuthold, 1981) yet has never been investigated through music analysis. The ambiguity of the regional characteristics hypothesis gave way to concepts of yodeling as a “national” style, as listeners might not perceive any differences between the regions. In this article, we follow the approach that a classification based on melodic features extracted using the recently developed tool *MelodyFeatures* (Metzig et al., 2020, available as an R package) can be used not only for the task of classification itself, but also for an inquiry into which melodic features are particularly important for differentiation between tunes from different regions. This therefore potentially creates a method of systematically recording the style through statistics. Retrospectively, these features are then studied and interpreted based on the existing ethnomusical literature and transcription of the melodies concerned. If successful, this systematic approach can be applied to any setting in which different musical styles interact and overlap. This study therefore aims at a data-driven exploration of different yodel styles, posing two questions: (1) Is it possible to classify yodel melodies of different

regions based on melodic attributes with considerable accuracy to reject the hypothesis that these differences are purely socially constructed ideas in the minds of performers? (2) If yes, which melodic attributes are the most salient?

Yodeling is a form of singing characterized by changes between chest voice and head voice, and a vocalization using syllables without lexical meanings. While these singing techniques exist around the globe (Plantenga, 2004), the use of the term ‘yodeling’ (German: *Jodeln*) is sometimes limited to the Alpine region. Yodel melodies are predominantly orally transmitted and traditional melodies are not ascribed to a particular author. Although sometimes described as a “national” tradition (e.g., *Swiss yodeling*, *Austrian yodeling*), yodeling practices are not homogeneous within national borders, and variations within the countries mentioned are possibly larger than those between countries. The differentiation between these regional styles, which to inexpert listeners tend to sound very similar, has engaged

¹ Lucerne University of Applied Sciences and Arts, Luzern, Switzerland

² Queen Mary University of London, London, UK

Corresponding author:

Yannick Wey, Lucerne University of Applied Sciences and Arts, Zentralstrasse 18, Luzern 6003, Switzerland.

Email: yannick.vey@hslu.ch



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages

(<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

yodelers since at least the 1960s (Leuthold, 1981; Räss & Wigger, 2010; Wey, 2019). Systematic classification of melodies from different origins dates to the work of Lomax (1976). Recent studies focus on audio recordings (Li et al., 2017; Mehr et al., 2019) and analyze vocal style; others use symbolic data, which can use global features, for example, time signature or beats per minute (Velardo et al., 2016), or local features, for example, n -grams (Conklin, 2009; Müllensiefen & Frieler, 2007), which use pitch contours, among other factors. Walshaw (2018) uses hierarchical abstractions of the melody. Li et al. (2006) use symbolic data for folk song classification, but also information on chord accompaniment, and do not study discriminating features. Hillewaere et al., 2009 find that the use of event models such as n -grams outperforms global features, and, similar to our approach, use feature selection. Conklin (2013) presents a method to detect rare patterns in Basque folk music. Music genre classification is also often approached hierarchically (Arnal Barbedo & Lopes, 2006; Silla & Freitas, 2009).

The method we use, *MelodyFeatures*, differentiates from similar approaches (Cuthbert & Ariza, 2010; Eerola & Toiviainen, 2004; McKay et al., 2018; Müllensiefen, 2009) in that it only uses melodic features, and considers a much larger number of them (30,014), of which many will be zero throughout the database. Furthermore, it examines more detailed rhythm n -grams. In this way, the extracted features are particularly suited for machine learning approaches that use feature selection. The identification of informative features, which is our second research question, has been addressed in the literature through subgroup discovery (Taminau et al., 2009), distinctive pattern discovery (Conklin & Anagnostopoulou, 2011), contrast pattern mining (Neubarth & Conklin, 2016), and supervised descriptive pattern mining (Neubarth et al., 2018).

Following a comparative approach and acknowledging critical arguments that such an approach to music may often result in comparing “apples and oranges” (Nettl, 2005, p. 61), we solidify the argument by restricting samples to the small geographic areas of central and north-eastern Switzerland, both of which are home to traditional yodeling styles. We integrate our results with the available information on regional characteristics from ethnomusicological and folkloristic literature from the concerned area, as well as recent fieldwork notes. This approach is intended to meet the requirements placed on the validity of culturally relevant music analyses: “It became clear that we must also study each music in terms of the theoretical system that its own culture provides for it, whether an explicitly articulated, written system or one that must be derived from interview and analysis; and that one must study musical behavior in terms of the underlying value structure of the culture from which it comes.” (Nettl, 2005, p. 63). Criticizing the “myth of universality”, Dave (2014, p. 4) points out that “recent research on sound and affect in the human voice shows how vocal production is shaped by particular localities, cultural memories, national identities, and

histories”. In the present case, the music-related regions share many of these cultural memories and histories, not only by their proximity, but also because of the activities of the Federal Yodelling Association (*Eidgenössischer Jodlerverband*, EJV), which, since its inception in 1910, has brought together performers from all regions of Switzerland for large, triannual celebrations (*Eidgenössischer Jodlerverband*, 2010). Considering these limits and pitfalls for meaningful comparisons, the results have to be informed by, and compared with, existing theories and beliefs about stylistic differences. By comparing the data-based results with existing assumptions from the point of view of field research and exemplary transcriptions, we compare the previous findings with the new results, and evaluate the musical significance of individual characteristics. The literature on these characteristics contains treatments from the perspective of ethnomusicologists as well as expert practitioners in the field.

Yodeling Regions Within the Alpine Region

The question of regional classification has been a central problem in research on yodeling throughout time. Contrasting the idea of a “national” song, since the inception of the EJV, yodelers have spoken about yodel styles in terms of different regions rather than nations. These regions exist throughout the German-speaking Alpine territory, in Switzerland as well as in Austria, with “a bewildering variety of names and types” (Wise, 2007, p. 3). Yodeling, or in German, *Jodel*, is commonly used as an umbrella term for various traditions of vocal performance, each with its local name; there are the Bernese *Jutz*, the central Swiss *Juiz*, the Appenzell *Zäuerli* and *Rugguusseli* (Mock, 2007), the *Dudler* in and around Vienna, the *Johlar* in Vorarlberg (Fink-Mennel, 2007), among several others. In Switzerland, the term *natural yodel* refers to yodeling with only sense-neutral (“meaningless”) syllables (Wey et al., 2017).

Several scholars and practitioners invested in traditional yodeling have already presented the concept of distinct regions based on their perception of aesthetic characteristics, anecdotal evidence, and narratives. As early as the mid-19th century, folklorist Heinrich Szadowsky distinguished three “basic types” of yodel singing: “Appenzell song, Bernese highlands, and Vaud song” (Szadowsky, 1864, p. 512, translated from the German by the authors). Szadowsky did not list central Switzerland, an area where yodeling culture is very active today, in this subdivision. Following the folk song collector Alfred Leonz Gassmann’s *Tonpsychologie* (Gassmann, 1936), yodel composer Robert Fellmann divided the Swiss yodeling melodies into three regions: melodies from the Central Plateau, the foothills of the Alps, and the High Alps. In the fourth edition of Fellmann’s textbook (Fellmann, 1962), a detailed appendix by the composer Max Lienert was published. Lienert distinguished between the three yodeling landscapes of Toggenburg-Appenzell, central Switzerland, and

Bern-Fribourg (Fellmann, 1962). The Obwalden yodel expert Edi Gasser follows Lienert's classification and names as regions eastern Switzerland, with the two Appenzells and Toggenburg, the Bernese highlands, and the Emmental, as well as central Switzerland, with Entlebuch and the cantons of Schwyz, Obwalden, and Nidwalden. In the 1960s, several local yodeling experts from central Switzerland and Bern led a movement toward an understanding of regional characteristics, as revealed by archival sources in the literary estate of composer Heinrich Leuthold.¹ In a letter from 20 October 1967, the then president of the EJV, Balthasar Müller, asked Leuthold for suggestions for a "natural yodel course" (Wey, 2019, p. 227). Leuthold designed a syllabus for such a course and, on 15 November 1967, introduced the study of the following regions' characteristics as the main topic of the course: eastern Switzerland, central Switzerland, and Bern (Wey, 2019, p. 227). Leuthold's emphasis makes it clear that the stylistic identities are based on regional styles and not on the ideal of a uniform, nationwide aesthetic. The borders between the yodel regions mentioned have probably been blurred during recent decades (Leuthold, 1981), but at the same time the awareness of regional styles may have strengthened their practice. The CD *Die Jodelarten der Schweiz* (Bachmann-Geiser, 2010) documents the diversity of yodeling styles in Switzerland but does not mention regional differences and characteristics in writing. In summary, the division into larger regions has remained remarkably consistent since the beginning of the 19th century. The regions of central Switzerland and north-eastern Switzerland are the most frequently distinguished. On this basis, we examine these two regions in a superordinate manner.

Features and Regional Musical Differences

The descriptions of these regions mention a number of melodic, harmonic, and aesthetic features. Some of these are of interest in comparison with the present results; some contain broader statements and cannot be confirmed or dismissed by a melodic classification. The composer and yodel enthusiast Heinrich Leuthold (1981, p. 89) lists the following attributes for the Appenzell-Toggenburg region: triads upward, an emphasis on the major sixth, and the use of the augmented fourth. In some cases, a peculiar large interval of a seventh downwards adds to the melodic particularity (Leuthold, 1981, p. 87). Differences between the so-called Rugguusseli of Appenzell Innerrhoden and the Zäuerli of Appenzell Ausserrhoden are widely discussed among their performers, yet the question of whether these differentiate at all remains open, as both neighboring regions are small and highly interconnected. Mock (2007, p. 58) provided a list of features he learned from his interviews with yodelers. Accordingly, Rugguusseli are slower and freer in meter, have a wider ambitus but fewer high notes than Zäuerli. Zäuerli, on the other hand, are vocalized, with sharper and more diverse vowels. However,

Table 1. Origin, abbreviation, and group sizes for the regional classes.

Region	Abbreviation	Group size (no. melodies)
Central Switzerland	CE	103
Nidwalden	NW	37
Obwalden	OW	66
North-eastern Switzerland	NE	114
Appenzell Innerrhoden	AI	37
Appenzell Ausserrhoden	AR	40
Toggenburg	TO	37

concrete melodic features are not named. When asked about the difference between Rugguusseli and Zäuerli, some yodelers respond that one is performed by a yodeler from Innerrhoden, the other one by a yodeler from Ausserrhoden—even if the melody is the same. Therefore, this particular regional distinction would be based on how the performers (self-)identify, not on objective melodic features. Even without measurable differences between the two yodeling styles, performers may identify some songs as "their own", because they relate to personal experiences and carry local names.

In central Switzerland, the "stereotypical form" (Leuthold, 1981, p. 95) of a harmonic progression (between the tonic and the dominant, I–V–V–I) is prevalent, and again the major sixth constitutes a salient interval. For Toggenburg, Leuthold stresses the influence of "tonguing", a technique for fast changes between syllables, which could be derived from Bavarian yodeling (Leuthold, 1981, p. 90). Edi Gasser, collector and expert on yodels of central Switzerland, states that natural yodel melodies can differ even between small local spaces, for example, neighboring valleys—mountains may form natural barriers, restricting musical exchanges between the valleys. According to Gasser (2017), Appenzell Ausserrhoden yodels are often performed with the chest voice and omit switching voice registers, while Innerrhoden yodels change to the head voice, whereas high notes and an increased agility mark those from Toggenburg. According to Gasser, the cantons of Nidwalden and Obwalden differentiate from each other by timbre, with a darker timbre in the latter case. So far, various attempts have been made to define yodeling melodies according to their regional characteristics. The motivation for this was, on one level, musicological, to provide a systematic distinction; on another level, the importance of preserving and promoting regional diversity and the traditional way of singing was emphasized.

Materials and Method

We focus on the two overarching regions (Table 1), central Switzerland (CE) and north-eastern (NE) Switzerland, which are divided into the subregions of the cantons of Obwalden (OW) and Nidwalden (NW), and into the cantons of Appenzell Innerrhoden (AI), Appenzell

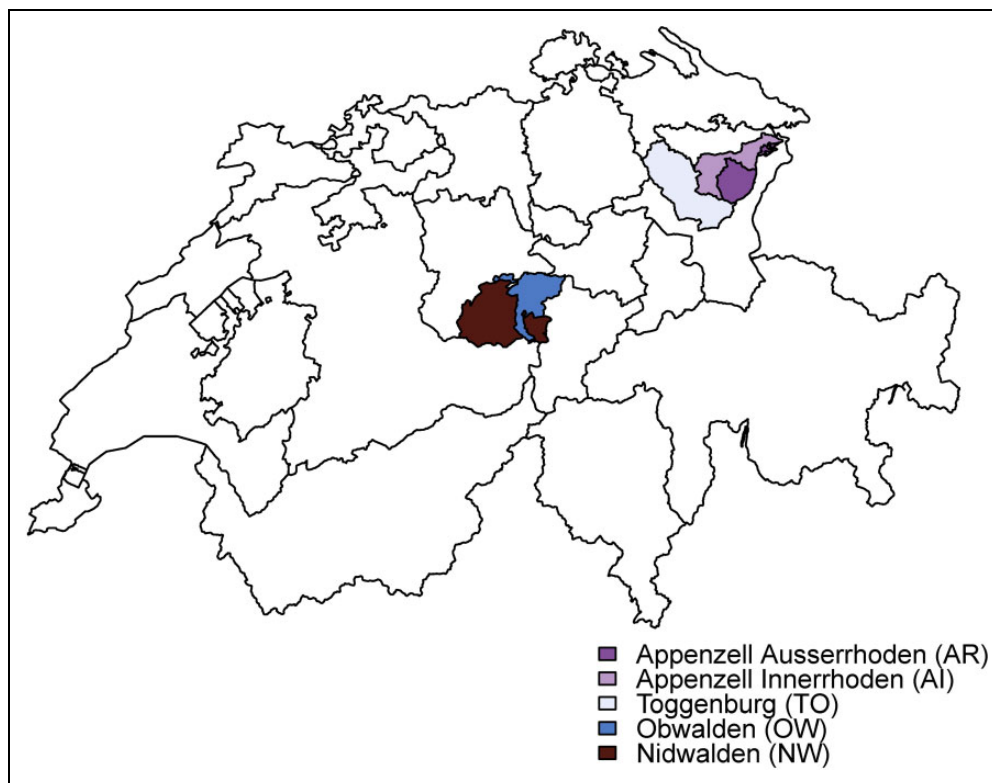


Figure 1. Map of Switzerland highlighting the geographic locations of the regional samples.

Ausserrhoden (AR) and the region of Toggenburg (TO), respectively. For our analysis we used a total of 217 yodel tunes from the five regions. Group sizes are 37 (NW), 66 (OW), 37 (AI), 40 (AR), 37 (TO) melodies.

We used transcriptions from two large databases of yodels from north-eastern and central Switzerland, from the Centre for Appenzell and Toggenburg Folk Music and the website www.naturjodler.ch. There are two ways in which yodel transcriptions ended up in the quoted archives: transcriptions were made either by collectors or by persons who left their estates to collectors. Yodelers themselves usually do not notate their melodies or claim authorship, and many transcriptions are anonymous and attributed to a local culture, tradition, or village, rather than a person. To make them available in MIDI format, we converted notation provided in PDF to MIDI files and manually corrected for errors. Those conserved in handwriting we manually copied in MuseScore. Yodeling melodies share a common musical form, consisting of two phrases of similar length. A phrase often contains eight bars. This musical form was supposedly standardized in the 19th century, when yodeling adapted, in some respects, to forms of singing taught in schools and church choirs (Wey, 2020, p. 146).

The list of regions included is by no means exhaustive; the focus on these regions can be justified through the facts laid out previously: that they harbor traditional styles, are widely regarded to differ from one another by practitioners, and provide sufficient documented material to generate

adequate samples for the present study. Figure 1 shows the geographic situation of the five regions.

We use symbolic representations of the yodel tunes in MIDI format,² that is, the differences in vocal delivery mentioned previously are not considered, as they are not notated in the MIDI files. Using the R package *MelodyFeatures*, we extracted the following features.

- Note lengths, that is, the fraction of the total melody length that is spent on a given note (of the 12 semitones with respect to the tonic note). Notation: `note_len_` followed by a number from 0 (tonic) to 11 (example: `note_len_4` is the fraction of time spent on the fourth semitone (major third) above the tonic).
- The number of occurrences of each note (i.e., the tonic and the 11 semitones above it), normalized by the total number of notes of the yodel tune. This makes tunes of varying length comparable. Notation: `note_occ_` followed by a number from 0 (tonic) to 11 (example: `note_occ_2` is the fraction of notes on the second semitone above the tonic).
- The number of occurrences of each interval. Notation: `int_occ` followed by a number from 0 (unison) to 12 (octave) (example: `int_occ2` is the fraction of minor seconds among intervals).
- Counts of n -grams (the pitch difference of n consecutive notes). We use bigrams (intervals), 3-grams (two consecutive intervals), and 4-grams, and normalize

each feature by the total number of n -grams in the tune. Notation: int (short for *intervals*), trigram, or four, followed by the number of semitones of the intervals, where “.” signifies a downward interval (example: trigram2_2 stands for a full tone upward, followed by a full tone down, for example, the notes C–D–C).

- Counts of intervals, where we consider intervals to be different features when they start on different notes in the scale (e.g., a full tone up starting on the tonic is a different feature from a full tone up starting on the first full tone). Again, these features are normalized by the total number of intervals in the tune. Notation: X, followed by the semitone with respect to the tonic on which the interval starts, followed by the interval (example: X2_1 stands for a semitone downward, starting one full tone (two semitones) above the tonic note; the underscore serves only to separate the numbers). The start note can hereby be lower than the tonic note as well, for example, X.2_1 means a semitone down, starting on a full tone below the tonic (e.g., from Bb to A if the key is C). In these features, the scale (or mode) is implicitly represented, for example a high occurrence of X0_3 (minor third upward, starting on the tonic) is an indicator for the yodel to be in a minor key.
- Counts of rhythm n -grams (3- to 6-grams). We count possible patterns of consecutive note lengths of length n , notated as multiples of the shortest note length in a given pattern. The n -grams consist of three, four, five, or six consecutive notes. They are normalized by the total number of n -grams in the tune (e.g., for rhythm 3-grams, it is the number of notes in the melody minus two). Notation: r, followed by multiples of the shortest note (example: r221123 stands for “quarter–quarter–eighth–eighth–quarter–dotted quarter note”, in an example where an eighth note constitutes the shortest note).

All these features are calculated using the R package *MelodyFeatures*, which has an example script provided.³ The rationale behind those features is that small melody parts are units that people are familiar with; they draw from that experience when composing a new melody. Implicitly, these features contain information about time signature and mode. Regarding features that refer to a tonic, *MelodyFeatures* allows the determination of the tonic note (i) from the key given in the file, (ii) from the last note in the melody, (iii) manually. All our melodies were checked, although for Alpine yodeling the convention is that the last note is the tonic note.

We then trained a random forest classifier (Breiman, 2001) to predict the origin of the tunes, with a methodology introduced in Metzsig et al. (2020). The random forest classifier attributes a value of importance to every feature used to obtain a classification result. We chose the random forest classifier because it is an ensemble method with low bias and because

we use the feature importances (i.e., the relevance a given feature had for finding the classification result) for the musical interpretation. These importances are a natural by-product of only the random forest classifier. Other classifiers were used for comparison in Metzsig et al. (2020) but showed no significant improvement. The classification was made for each pair of regions. To reduce noise, we used a statistical filter (minimum redundancy and maximum relevance, Peng et al., 2005) to reduce the number of features from about 2,700 nonzero features to 400. We then selected the 150 features with the highest importance, on which we trained the classifier again. This feature selection reduces noise and produces much better prediction accuracies. Thereby, 10% of the data were kept back as a test set, and the feature selection was performed on the remaining 90%.

The whole classification, including prefiltering, was repeated 100 times to provide robust results despite a potentially unbalanced dataset; we present here the mean of the accuracies. Repeating the filtering step for every classification run (after every time a new test set was split off) ensures that filtering is not biased by the test set. Otherwise (i.e., if feature selection were applied to the entire set before splitting the test set off), we would get higher accuracies because the test data would be overfitted. We used pairwise classification instead of 5-class classification because a crucial element of our approach is to select for informative features. In this way, every pair has been compared according to different criteria. This strongly increased the accuracy and revealed more informative features. After prefiltering, we performed the actual random forest classification. The implementation is from the R package *caret*; we used 500 trees. To reduce the variance, sampling was done with replacement, since not all predictors can be considered uncorrelated (e.g., n -grams of different lengths starting with the same intervals will be correlated). We also adjusted parameters of the random forest algorithm and found the highest accuracy for $mtry=2$ (Liaw & Wiener, 2018), which is the number of variables randomly sampled as candidates at each split. We ran the model for a fixed time and kept part of the dataset as a validation set. Despite the random forest classifier having low bias, the large number of features requires a check for overfitting; we classified the data with randomly permuted labels (before filtering and classifying the data). Averaging over 100 runs of this test gave accuracies of 0.48–0.51 (i.e., random chance), we did not constrain the minimal or maximal number of leaf nodes, to compensate for overfitting. Hyperparameter optimization was not carried out.

Results

Table 2 shows the classification accuracies and their standard deviations. The geographic regions are separable, but the classification is more accurate if one region is part of central Switzerland and the other is part of north-eastern Switzerland (accuracies between 0.73 and 0.84) than in the

Table 2. Accuracies of pairwise classifications with standard deviations (in brackets), averaged over 100 runs; feature selection was performed on the training set only.

Mean accuracy	TO	AR	AI	OW	NW
NW	0.84 (± 0.13)	0.76 (± 0.15)	0.81 (± 0.16)	0.63 (± 0.12)	—
OW	0.69 (± 0.10)	0.81 (± 0.11)	0.82 (± 0.13)	—	—
AI	0.73 (± 0.17)	0.61 (± 0.11)	—	—	—
AR	0.76 (± 0.16)	—	—	—	—
TO	—	—	—	—	—

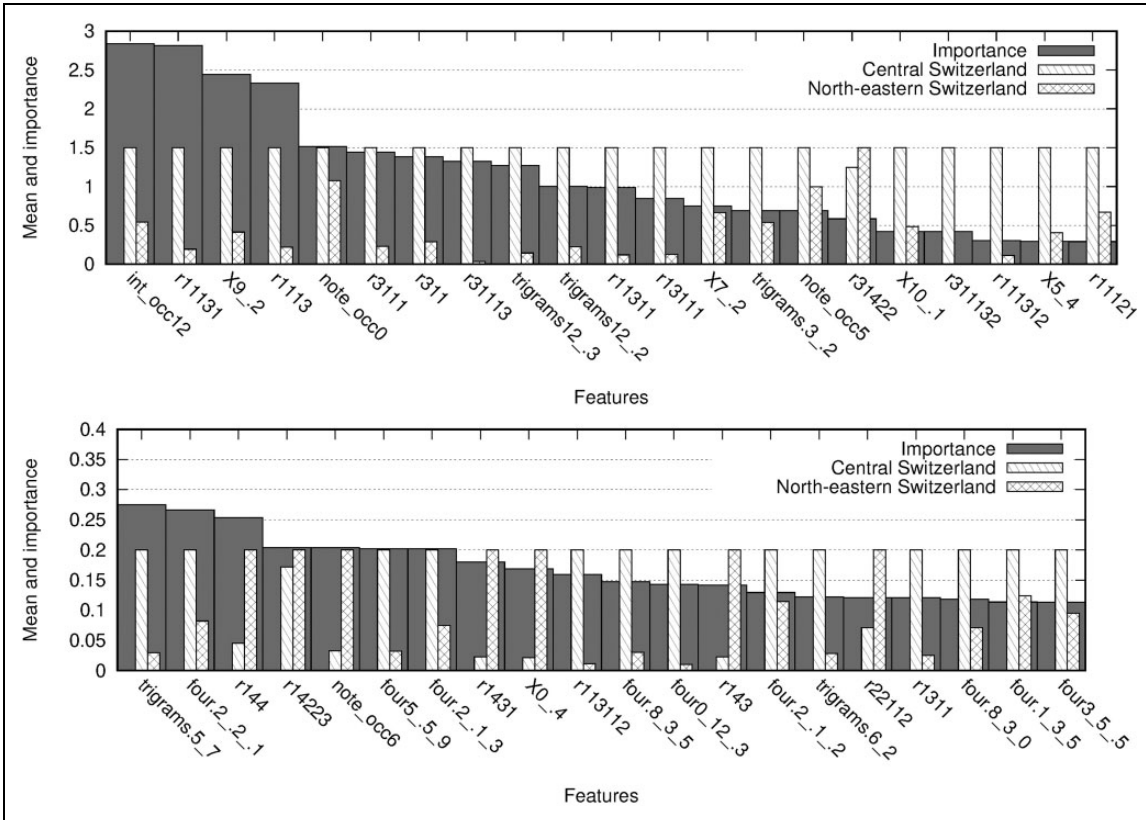


Figure 2. Features with the highest importances from the classification between central (left bars) and north-eastern (right bars) Switzerland. Wide bars: importance of the feature. Narrow bars in front: rescaled means of that feature in the respective regions (only the relative heights are of interest). Accuracy: 0.75 ± 0.10 .

cases of OW–NW (0.63) and AI–AR (0.61). To test for overfitting, we performed pairwise classification with the same method but randomized group labels, which resulted in accuracies close to 50% for equal group sizes, i.e., random chance. The accuracies were computed for the pairwise classification tasks, averaged over 10-fold cross-validation (Table 2).

The classification between the overarching regions shown in Figure 2, central and north-eastern Switzerland, shows an accuracy of 0.75 ± 0.10 . Table 3 details the three most important features for any of the pairwise classifications (i.e., the column ‘Feature 1’ designates the melodic feature with the highest importance for the task of classification). In the case of the overarching regions, these are

Table 3. Most important features (1 to 3) for pairwise differentiations between regions.

Regional comparison	Feature 1	Feature 2	Feature 3
NE–CE	int_occ12	r11131	X9_2
AI–AR	note_len9	r113	int.8
AR–TO	r114	r14111	r113
AI–TO	r411	r141	X2_2
NW–OW	r22231	r11111	trigrams.9_5
NW–TO	note_occ0	note_occ4	r41111
NW–AR	int_occ12	r1131	note_occ4
NW–AI	note_len9	four.1_2_2	trigrams.2_2
OW–TO	note_occ0	note_occ5	r2221
OW–AR	int_occ12	r11111	r113
OW–AI	int_occ12	X7_2	trigrams.3_2

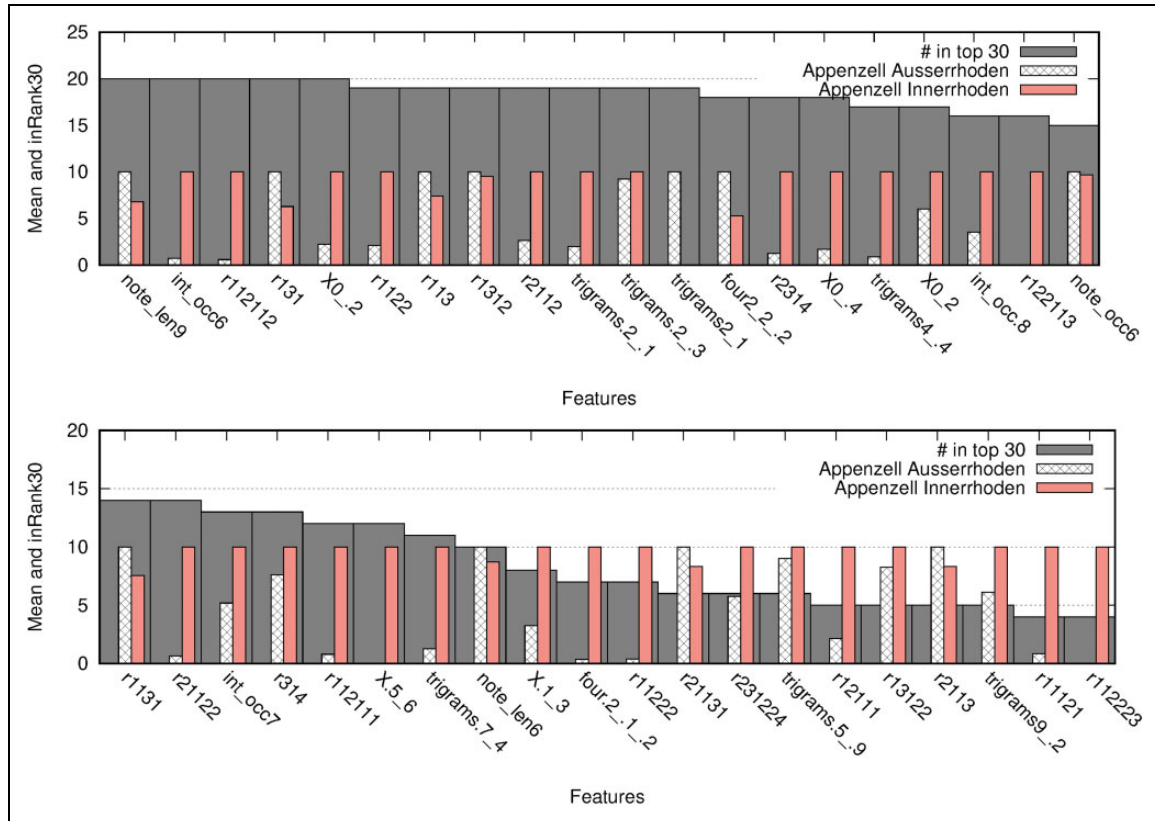


Figure 3. Features with the highest importances from the classification between Appenzell Innerrhoden (left bars) and Appenzell Ausserrhoden (right bars). We ran 20 classifications where feature selection was performed on the entire dataset and counted how often a given feature appears in the top 30 features (ranked by importance) of the random forest classification. The result represents the most informative features.

the interval of an octave (int_occ12), a rhythmic pattern involving a dotted quarter note (r11131), and a melodic motif of an upward sixth followed by a whole tone down (X9_2). All three are present in the central Swiss sample.

Figure 2 depicts the features with the highest importances from the classification between central and north-eastern Switzerland. The pairwise comparison of the subregions works in the same way as the comparison between the larger regions shown in Figure 2. Figure 3, as an example, illustrates the results of these comparisons, based on the comparison of the regions of Appenzell Innerrhoden and Appenzell Ausserrhoden. Differences between these two small regions have been discussed extensively by their performers (Mock, 2007) and their classification is therefore of particular interest. The complete set of the 11 regional pairwise comparisons is included in the supplementary dataset.

The occurrence of certain intervals of the scale is particularly decisive for a musical melody. In addition to the frequencies of the intervals, we also take into account the note lengths of each degree of the tonal scale; these allow us to make a statement about the weighting of the degrees within the scale, as longer notes take a more prominent place than shorter ones. The interval attributes (int_occ)

and note lengths (note_len) are compared in Figures 4 and 5. Figure 4 shows the differences in interval and note length for the overarching regions; Figure 5 focuses on the subregions. The large discrepancy between the degrees regarding note lengths, for example, between note_len1 (low) and note_len2 (high across all regions) stems from the fact that the melodies generally move in diatonic scales. Differences between the samples are visible but would not be decisive enough to consider a classification without considering additional categories of features.

To explore further methods to visualize the relationship between yodel tunes, we constructed a phylogenetic tree from the feature vectors; this is available in the supplementary data. For this, we used the 400 most important features (the 40 most important of each pairwise classification) and constructed a neighbor-joining tree from the R package *ape* over them (Paradis & Schliep, 2019). The tree unsurprisingly shows clusters of yodel tunes where one region dominates. Conversely, many yodels across regions also show high similarity. Since the branch length to the most common ancestor point is an indicator of the distance between feature vectors, in this method we can identify some outlier tunes. The closer to the origin the clades (i.e., hierarchical clusters) split, the bigger the difference from the other

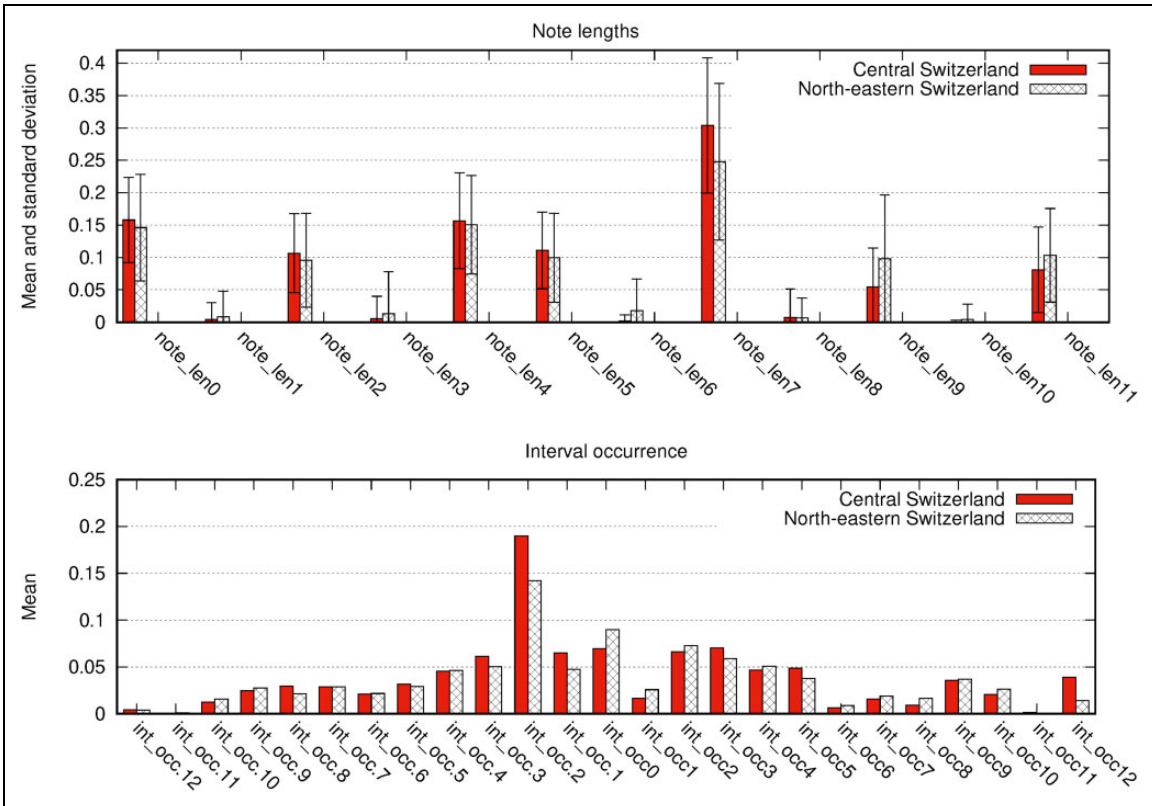


Figure 4. Note lengths and interval occurrences for the regions of central Switzerland (left bars) and north-eastern Switzerland (right bars).

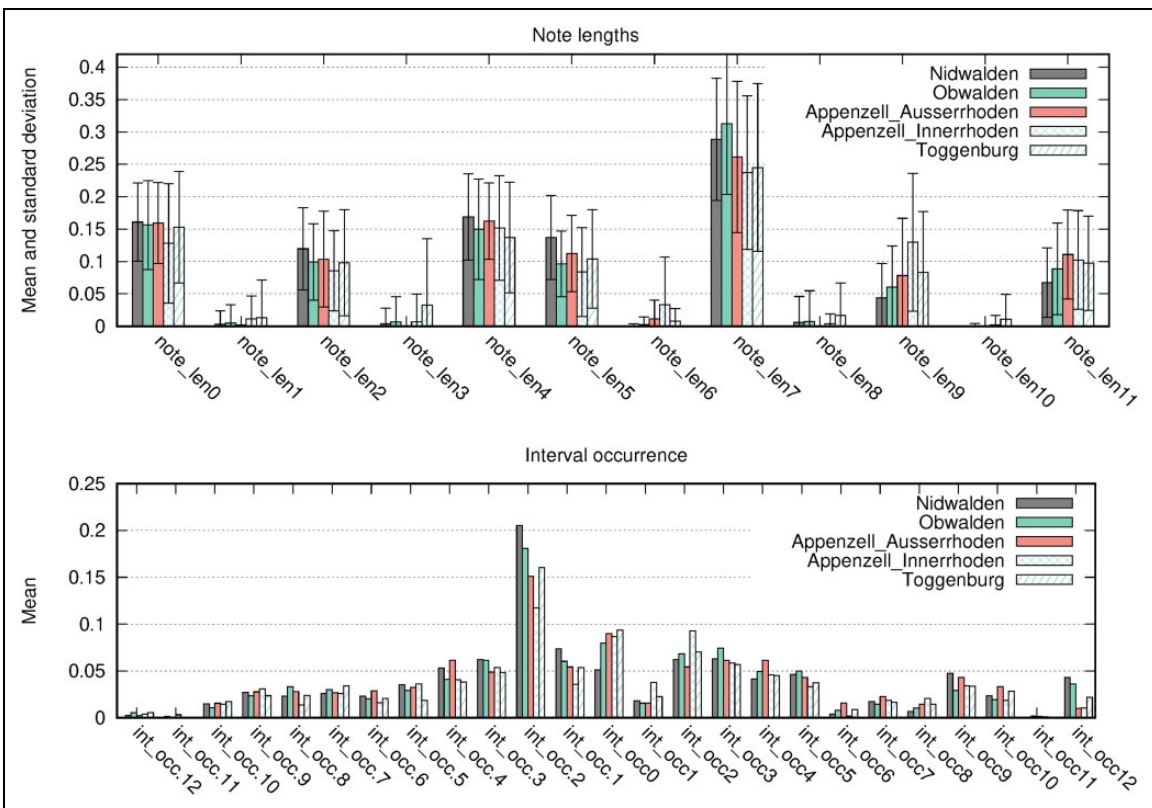


Figure 5. Note lengths and interval occurrences for the five subregions.

yodel tunes. The 400 used features were selected by aggregating the pairwise most important features to achieve some level of precision of the within-group similarity. The supplementary dataset contains a PDF of Figure 7, which allows for a precise reading of specific tune labels.

Discussion

The accuracy of the pairwise classification of yodels by region supports the narrative that yodel styles are footed in regions; however, we found that yodel tunes from the regions Nidwalden and Obwalden are less separable and can be predicted with lower accuracy, as well as those from Appenzell Innerrhoden and Appenzell Ausserrhoden. An explanation for this discrepancy is provided by the geographic proximity of these two pairs of regions and their high level of music-cultural exchanges. The defining melodic differences between any two regional samples are demonstrated through pairwise comparison and the interpretation of the features with the highest random forest importances. The most salient features demonstrate some of the key differences in pairwise comparisons, for example, the frequent use of dotted rhythms or the prevalence of an upward augmented fourth, confirmed by a postanalysis survey of the relevant scores.

In central Switzerland, the octave occurs frequently and constitutes the most important attribute in contrast to the interval's absence in the north-eastern samples. The octave also accounts for the differentiation between the OW sub-region sample and both Appenzell samples (see Table 3). The styles from the two Appenzell cantons (AR and AI), often perceived as identical by listeners, were classified at an accuracy of 0.61. The most salient attributes can be reconstructed from the survey of the notation in the sample. For AI, it is the occurrence of relatively long notes sung a major sixth above the tonic (*note_len9*), and a dotted rhythm (*r131*); for AR, it is the use of the interval of a minor sixth (*int_occ8*). However, we were not able to anticipate these key features before the experiment based on the literature: the stylistic differences between these regions mentioned in the literature do not go into the details of melodic progression and the rhythmic structure.

The Toggenburg region sample distinguishes itself from the bordering Appenzell regions by an assortment of rhythmic motifs, generally involving a repetition of eighth notes (see Table 3). This analysis can be corroborated with data from ongoing fieldwork in these regions, as interviewees across the regional boundaries concur in their statements that the Toggenburg style can be recognized by its relatively fast-paced rhythms. Specific interval patterns (full tone down that ends on the tonic *X2_.2*) are to be investigated more closely. Although the Toggenburg region neighbors the two Appenzell regions, the melodies (TO–AI and TO–AR) were distinguished with relatively high accuracy. This could be for geographic reasons: while the borderland between AI and AR is flat, large parts of their border to Toggenburg is covered by mountains, the

Alpstein massif, which historically probably led to fewer music-cultural exchange.

As in the case of the two Appenzell regions, the Nidwalden and Obwalden samples overlap sufficiently that the accuracy of classification is lower than between samples from geographically distant regions. Both regions share an overarching yodeling tradition, and the intersection of both yodel regions is not surprising, based on the literature referred to previously. However, outstanding attributes are the rhythmic patterns, *r22231* and *r11111*, which occur more frequently in Nidwalden (see Table 3; Figure 6). The distinction between the sub-regions of central Switzerland and north-eastern Switzerland is strongly influenced by the increased occurrence of the octave as an interval in yodeling melodies from the regions of Obwalden and Nidwalden (see Table 3). To follow up the discovered features for the differentiation of regional styles, we checked whether they can be identified in the underlying musical scores. In each case, we were able to retrace the salient features based on sample notation. Figure 6 lists examples for the most important feature in each comparison between two subregions. The examples are drawn from the notation provided in the supplementary data.

Conclusion

The two questions posed in this study, as stated in the introduction, can be answered using the methods described. A classification of yodel styles based on regional provenience was implemented and traced through the observation of the most important features, based on melody alone. The present analysis is limited to melodic features and therefore omits such aspects as tempo and vocalization, which could be more important to performers and to the overall cognitive perception of stylistic differences. Through our analysis of 4-grams and long rhythm *n*-grams of up to six consecutive notes, we considered longer-range correlations than the cited literature that pursues an *n*-gram approach for prediction of origin (Conklin, 2009; Cuthbert & Ariza, 2010; Eerola & Toivianen, 2004; Hillewaere et al., 2009; McKay et al., 2018; Müllensiefen, 2009; Müllensiefen & Frieler, 2007). Compared with the literature, our method relies heavily on feature selection particularly suited for the discovery of musical differences in given regions, since very specific features will be selected. The length of (rhythm) *n*-grams used is longer, and their number much greater than in the literature on pattern discovery. For this reason, we performed robustness tests.

The accuracy of classification based exclusively on melodic information signifies the importance of melodic features and solidifies the argument that regional differences are inscribed in the music and not only constructed based on individual perception. We interpret the results of the classification as supportive of the described classifications in 20th- and 21st-century folkloristic literature on the topic of yodels. This represents an important step in our understanding of vocal music tradition in the Alpine region, as these stylistic descriptions have so far only been based








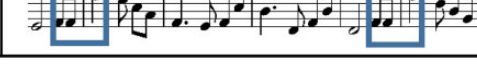
Comparison	Feature (highest importance)	Musical example	key	Time signature
<u>AI</u> – AR NW – <u>AI</u>	note_len9		C	6/4, 4/4
AR – <u>TO</u>	r114		Bb	3/4
<u>AI</u> – <u>TO</u>	r411		Bb	3/4
<u>NW</u> – OW	r22231		Bb	2/4
<u>NW</u> – TO	note_occ0		Bb	4/4, 3/4
<u>NW</u> – AR	int_occ12		Bb	3/4
<u>OW</u> – TO	note_occ0		Eb	3/4
<u>OW</u> – AR <u>OW</u> – AI	int_occ12		Bb	3/4

Figure 6. Musical examples for the most important feature in each comparison between two subregions. The underlined region is where the feature is prominent.

on anecdotal evidence. Regarding the study of music more widely, this novel method could inspire new research questions in musicology; for instance, tunes of unknown origin could be attributed to a likely geographic origin, for example to identify the closest region, using a hierarchical approach (first classify it into the region, and then within that region into the canton) Results could be presented to yodel practitioners for further comments. The method allows emerging new genres of music to be distinguished through the classification of their melodic features and could stimulate a discussion about reciprocal influences and formative characteristics.

Action Editor

David Meredith, Aalborg University, Department of Architecture, Design and Media Technology.

Data Availability Statement

The supplementary dataset is available on <https://zenodo.org/record/3859093>.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Peer review

One anonymous reviewer.
Reinier de Valk, Moodagent A/S.

ORCID iD

Yannick Wey  <https://orcid.org/0000-0002-2416-1285>

Notes

1. Leuthold's literary estate is located in the Nidwalden State Archive, Stans (index: StANW P 137/1–P 137/18).
2. Scores were imported using <https://musescore.com/import> and then manually checked for errors and duplications.
3. www.github.com/cmetzig/MelodyFeatures

References

- Arnal Barbedo, J. G., & Lopes, A. (2006). Automatic genre classification of musical signals. *EURASIP Journal on Advances in Signal Processing*, 2007(1), 064960.
- Bachmann-Geiser, B. (2010). *Die JodelARTen der Schweiz*. CD. Zytglogge.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

- Conklin, D. (2009). *Melody classification using patterns*. Second International Workshop on Machine Learning and Music, Bled, Slovenia, pp. 37–41.
- Conklin, D. (2013). Multiple viewpoint systems for music classification. *Journal of New Music Research*, 42(1), 19–26.
- Conklin, D., & Anagnostopoulou, C. (2011). Comparative pattern analysis of Cretan folk songs. *Journal of New Music Research*, 40(2), 119–125.
- Cuthbert, M. S., & Ariza, C. (2010). Music21: A toolkit for computer-aided musicology and symbolic music data. *International Society for Music Information Retrieval*, 11, 637–642.
- Dave, N. (2014). Music and the myth of universality: Sounding human rights and capabilities. *Journal of Human Rights Practise*, 7(1), 1–17.
- Eerola, T., & Toivainen, P. (2004). *MIDI toolbox: MATLAB tools for music research*. University of Jyväskylä, Department of Music.
- Eidgenössischer Jodlerverband. (2010). *Lebendiges Schweizer Brauchtum 1910–2010*. Eidgenössischer Jodlerverband.
- Fellmann, R. (1962). *Schulungsgrundlage für Jodlerinnen und Jodler*. Eidgenössischer Jodlerverband.
- Fink-Mennel, E. (2007). *Johlar und Juz. Registerwechselnder Gesang im Bregenzerwald* (mit Tonbeispielen 1937–1997). Neugebauer.
- Gasser, E. (2017). *Naturjodel-Regionen der Schweiz*. Unpublished Manuscript. Giswil.
- Gassmann, A. L. (1936). *Zur Tonpsychologie des Schweizer Volksliedes*. Hug.
- Hillewaere, R., Manderick, B., & Conklin, D. (2009). Global feature versus event models for folk song classification. In *International Society for Music Information Retrieval Conference* (pp. 729–733). Kobe, Japan.
- Leuthold, H. (1981). *Der Naturjodel in der Schweiz*. Robert Fellmann-Liederverlag.
- Li, J., Ding, J., & Yang, X. (2017). The regional style classification of Chinese folk songs based on GMM-CRF model. In *Proceedings of the 9th International Conference on Computer and Automation Engineering* (pp. 66–72). New York, NY, USA.
- Li, X., Ji, G., & Bilmes, J. A. (2006). A factored language model of quantized pitch and duration. In *International Computer Music Conference* (pp. 556–563). New Orleans, LA, USA.
- Liaw, A., & Wiener, M. (2018). *Package 'randomForest'*. University of California.
- Lomax, A. (1976). *Cantometrics. An approach to the anthropology of music*. University of California Extension Media Center.
- McKay, C., Cumming, J., & Fujinaga, I. (2018). JSYMBOLIC 2.2: Extracting features from symbolic music for use in musicological and MIR research. In *International Society for Music Information Retrieval Conference* (pp. 348–354).
- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A., Hopkins, J., Howard, R., Hartshorne, J., Jennings, M., Simson, J., Bainbridge, C., Pinker, S., O'Donnell, T., Krasnow, M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, 366(6468), eaax0868. <https://doi.org/10.1126/science.aax0868>
- Metzigg, C., Gould, M., Noronha, R., Abbey, R., Sandler, M., & Colijn, C. (2020). Classification of origin with feature selection and network construction for folk tunes. *Pattern Recognition Letters*, 133, 356–364. <https://doi.org/10.1016/j.patrec.2020.03.023>
- Mock, B. (2007). *Rugguusseli. Zur Tradierung der Naturjodelkunst in Appenzell Innerrhoden* (Doctoral dissertation). <https://d-nb.info/98472589x/34>
- Müllensiefen, D. (2009). *Fantastic: Feature analysis technology accessing statistics (in a Corpus)* (Technical Report v1). Goldsmiths, University of London, pp. 140–144.
- Müllensiefen, D., & Frieler, K. (2007). Modelling experts' notions of melodic similarity. *Musicae Scientiae*, 11(1_suppl), 183–210.
- Nettl, B. (2005). *The study of ethnomusicology. Thirty-one issues and concepts*. University of Illinois Press.
- Neubarth, K., & Conklin, D. (2016). Contrast pattern mining in folk music analysis. In D. Meredith (Ed.), *Computational music analysis* (pp. 393–424). Springer.
- Neubarth, K., Shanahan, D., & Conklin, D. (2018). Supervised descriptive pattern discovery in native American music. *Journal of New Music Research*, 47(1), 1–16.
- Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526–528.
- Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8), 1226–1238.
- Plantenga, B. (2004). *Yodel-ay-ee-oooo: The secret history of yodeling around the world*. Routledge.
- Räss, N., & Wigger, F. (2010). *Jodel – Theorie & praxis*. Mülirad.
- Silla, C. N., & Freitas, A. A. (2009, October). Novel top-down approaches for hierarchical classification and their application to automatic music genre classification. In *2009 IEEE International Conference on Systems, Man and Cybernetics* (pp. 3499–3504). San Antonio, TX, USA.
- Silla, C. N., & Freitas, A. A. (2011). A survey of hierarchical classification across different application domains. *Data Mining and Knowledge Discovery*, 22(1–2), 31–72.
- Szadrowsky, H. (1864). Nationaler Gesang bei den Alpenbewohnern. *Jahrbuch des Schweizer Alpenclub*, 1, 275–352.
- Taminau, J., Hillewaere, R., Meganck, S., Conklin, D., Nowé, A., & Manderick, B. (2009, September). *Descriptive subgroup mining of folk music*. MML 2009: International Workshop on Machine Learning and Music, Bled, Slovenia, 1–6.
- Velardo, V., Vallati, M., & Jan, S. (2016). Symbolic melodic similarity: State of the art and future challenges. *Computer Music Journal*, 40(2), 70–83.
- Walshaw, C. (2018). A visual exploration of melodic relationships within traditional music collections. In *2018 22nd International Conference Information Visualisation IV* (pp. 478–483). Fisciano, Italy.

- Wey, Y. (2019). *Transkription wortloser Gesänge. Technik und Rückwirkungen der Verschriftlichung des Jodelns und verwandter Gesänge im deutschsprachigen Alpenraum*. Innsbruck University Press.
- Wey, Y. (2020). Transformations of tonality: A longitudinal study of yodeling in the Muotatal Valley, Central Switzerland. *Analytical Approaches to World Music*, 8(1), 144–163.
- Wey, Y., Kammermann, A., & Ammann, R. (2017). Naturjodel und Naturtonreihe – eine gemeinsame Musikästhetik des Alphorns und des Jodels? *GVS / CH-EM Bulletin 2017*. <http://doi.org/10.5281/zenodo.1244141>.
- Wise, T. (2007). Yodel species: A typology of falsetto effects in popular music vocal styles. *Radical Musicology*, 2. <http://www.radical-musicology.org.uk/2007/Wise.htm>.

Appendix A

Table A1. Accuracies of 10-fold cross-validation if feature selection is done before splitting the test set off. They are higher than the accuracies in Table 2, where the features are selected only for the training set.

Mean accuracy	TO	AR	AI	OW	NW
NW	0.94 (± 0.08)	0.88 (± 0.12)	0.91 (± 0.12)	0.79 (± 0.10)	—
OW	0.85 (± 0.19)	0.90 (± 0.09)	0.91 (± 0.08)	—	—
AI	0.84 (± 0.11)	0.78 (± 0.13)	—	—	—
AR	0.89 (± 0.11)	—	—	—	—
TO	—	—	—	—	—