Harmonic syntax and high-level statistics of the songs of three early Classical composers

Wendy de Heer

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2017-167
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-167.html

December 1, 2017

Harmonic syntax and high-level statistics of the songs of three early Classical composers

by

Wendy Aileen de Heer

A thesis submitted in partial satisfaction of the
requirements for the degree of
Master of Science

in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John L. Canny, Chair
Professor Thomas L. Griffiths
Professor Keith A. Johnson

Summer 2015

Harmonic syntax and high-level statistics of the songs of three early Classical composers

Introduction

Understanding the statistical properties of music has become relevant to an increasing number of real-world applications over the last two decades, for example for electronic music distribution and storage. The fields of Music Information Retrieval (MIR) and Music Genre Recognition have correspondingly increased in popularity [1], [2]. Although a majority of research has focused on low-level, signal-processing based features of music and their corresponding statistics (e.g. [3]–[6]), it has been shown that appropriately chosen higher-level features of music, such as the melody line, can supplement and sometimes even outperform low-level features of music in classification tasks [7], [8]. Furthermore, understanding and exploring the statistics of higher-level musical features can contribute to the field of music theory [9], [10], and can potentially help with the neuroscientific and psychological investigation of the cortical processing of music [11], [12].

This paper explores the statistics of higher-order features of the musical works of three early Classical composers. We are particularly interested in chord frequency distribution [9], [13] and mapping out co-occurrences between different chords over different time-window lengths. There are several reasons that chord co-occurrences are of particular interest. First of all, context is extremely important for the musical experience: a note will have completely different meaning or generate a different emotional response depending on the notes immediately preceding it, or even preceding it on a large scale.

Secondly, in natural language processing, similar feature spaces–called word embedding spaces–have been developed and finely tuned for words [14], [15]. These spaces represent each word, or in our case chord, as a vector of frequency of co-occurrence with other words (or chords). For example, in the English language, the word *dog* often co-occurs with *cat*, or *ball*, but only rarely co-occurs with the word *sushi*. These spaces have proven to be powerful tools for analysis and dimensionality reduction, and there is evidence that they may even reflect cortical organization of information. In this paper, we create co-occurrence matrices for three composers: although they are all early classical composers, they wrote in three different times, and span two different musical subgenres (pre-tonal Renaissance music, and tonal Baroque music). A question that we therefore seek to address is: do chord co-occurrence matrices remain the same across these three composers? Or, alternatively, do they change based on musical sub-genre?

Methods

Dataset

Our dataset is a corpus of music written by three different composers from two different but closely related musical periods: Giovanni Pierluigi da Palestrina (1525-1594) from the Renaissance period, Claudio Monteverdi (1567-1643), who wrote in the period transitioning from Renaissance to Baroque, and Joann Sebastian Bach (1685-1750), from the Baroque period. 100 songs from Palestrina, 97 songs from Monteverdi, and 426 songs from Bach were analyzed. The songs are available as a part of the music21 corpus [16], in the MusicXML (.xml/.mxl) and kern (.krn) format [17].

Analysis tools

Custom code was written in python. It made heavy use of the python music analysis toolkit music21 [16] as well as the scipy [18] and numpy [19] scientific computing libraries.

Musical analysis

Transposition

We started the musical analysis by analyzing the key of each song, and then transposing all songs in major modes to C major, and all songs in minor modes to A minor. We did this for the following reason: the goal of this study is to look at higher-order statistics of music. According to musical theory, the same chord can have a different function in a musical piece, depending on the key of the song. For example, the major triad CEG, (with C as the root, E as the major third above C, and G as the perfect fifth) can have a different feel or meaning depending on whether it is the tonic chord (or I, in roman numeral notation) in the key of C major, or the dominant chord (or V, in roman numeral notation), in the key of F major. It is common for music to finish on a tonic chord, and gives listeners exposed to Western music a feeling of completion. On the other hand, finishing a song on a dominant chord, or V, can leave the listener feeling as if the musical piece is incomplete. Therefore, even though the lower-level features (in this case, the major triad CEG) are exactly the same, if this major triad is the final chord in the context of the key of C, the listener will feel like the piece is complete. If the same major triad is the final chord in the context of the key of F, the listener will expect to hear more. Composers took full advantage of these types of expectations (following them, and at times breaking them). Therefore, in order to understand the higher-level statistics of the chords, we have transposed all major mode and all minor mode pieces to the same key. We chose C major for major modes, and A minor for minor modes: both keys have no accidentals (sharps or flats) in their key signature. We analyzed the major and the minor mode songs separately.

Segmentation

Within a musical piece, the melody line (and sometimes other parts as well) will sometimes play notes that do not correspond exactly to the note from the true underlying chord. Some researchers have tried to avoid chords created with such notes. For example, Rohrmeier [9] segmented musical pieces into quarter notes, only analyzing one chord per quarter note, and choosing the least dissonant of the chords if there were several possible choices within each quarter note. However, we were interested in retaining as many chords as possible, consonant or dissonant, and we additionally did not want to impose music-theoretic notions of harmony onto our chord retention. In order to maximize the number of actual chords that we retained for analysis, and to create a data-driven analysis of the chord progressions, we used a chord segmentation heuristic similar to White's [13] *salami slice* heuristic, where we created a chord for every note transition in the musical pieces.

Most frequently occurring chords

In order to more directly compare the composers, we created a master list of the most frequently occurring chords for all composers taken together, both for minor and for major modes. To do so, we first ranked the chords according to frequency for each individual composer. We then summed the ranks of each chord over the ranking for that chord for each composer, to create the master ranking.

Chord co-occurrences

For the chord co-occurrence analysis, we subdivided each song into overlapping n-grams (sequences of n chords). The chord situated in the middle of the n-gram was the target chord: we then counted the (n-1)/2 chords before and after the target chord as co-occurring with the target chord. This was a simple binary count: we did not weight the chords according to distance from target chord.

Chord co-occurrence clustering

We performed a k-means cluster (with k = 3) analysis on the individual song major and minor mode n-gram (with n = 3, 5, 7, 9, 11, 13 and 15) chord co-occurrence matrices. This unsupervised clustering will indicate whether the co-occurrence matrices separate naturally into classes defined by composer. This analysis will also show whether there is a difference in clustering performance between different n-gram chord song co-occurrence matrices, and between major and minor mode song co-occurrence matrices.

Chord co-occurrence classification

We performed a 3-fold cross-validated decision-tree classification and a 3-fold cross-validated random-forest classification of the individual song chord co-occurrence matrices. We classified according to composer for major and minor modes, and for 3, 5, 7, 9, 11, 13 and 15-gram chord co-occurrence matrices. We further examined the classifiers' features' importance, to determine which features (or chord co-occurrences, in this case) were most important for the classification.

Results

In this paper, we analyze the statistics and harmonic syntax of three composers from three different musical periods: the Renaissance composer Palestrina, the Baroque composer Bach, and Monteverdi, who wrote in the transitionary period between the Renaissance and Baroque periods.

Chord Frequency Distribution

We first start by analyzing the frequency distributions of chords for each one of the three composers, for songs written in major keys and songs written in minor keys. Figure 1 and Figure 2 show the frequency distribution of the top major mode and top minor mode chords from the works of the Baroque composer Palestrina.

These two chord frequency distributions both follow a Zipfian (or Zeta, power-law) distribution, with the most frequent chord occurring 1511 times, but the 20th most frequent chord only appearing 98 times in the minor mode (in the major mode, these numbers are 1728 and 89). This is the case for chord frequency distributions for all three composers, for both major and minor modes (See also the Appendix for the frequency distribution of the top 20 chords for Monteverdi and Bach). For all three composers, the top 20 most frequent chords comprise a majority of all chords played (around 60%). Moreover, all three composers share many of the same top 20 chords. 31 chords total are in the top 20 most frequent chords for major mode songs for all three composers, and 29 for major mode songs.

For the major mode songs, 'I' (C major triad) is the first and 'V' is the second most frequent chord for all composers. 'vi' and 'IV' are in the top five most frequent chords, and 'I6' is in the top six. 'ii' and 'iii' are in the top ten for all three composers, as is 'IV6'. Therefore, for major mode songs, all three composers share 8 out of the top 10 chords, and of the top 6 most frequent chords, 5 are shared among the composers. For the minor mode songs, 'i' (A minor triad) is the most frequent chord for all three composers. 'v' is in the top ten for all three, as is 'III', 'iv', and 'bVI'.

A Spearman rank-correlation shows statistically significant correlations for the top 10 major-mode chords between the three composers: Palestrina and Monteverdi have a correlation of 0.964 (p = 7.32e-6); Monteverdi and Bach have a correlation of

Palestrina Major

rank	chord	quality	root	inv	count	frequency
0	I	major	1	0	1728	0.1478
1	V	major	5	0	1365	0.1167
2	IV	major	4	0	771	0.0659
3	vi	minor	6	0	516	0.0441
4	ii	minor	2	0	508	0.0434
5	iii	minor	3	0	326	0.0279
6	I6	major	1	1	303	0.0259
7	i54	other	1	2	280	0.0239
8	II	major	2	0	252	0.0215
9	IV6	major	4	1	201	0.0172
10	v#54	other	5	2	176	0.0151
11	v	minor	5	0	172	0.0147
12	V6	major	5	1	148	0.0127
13	ii6	minor	2	1	137	0.0117
14	I532	major	1	0	124	0.0106
15	viio#63	dimin	7	1	115	0.0098
16	vi#63	minor	6	1	112	0.0096
17	V7	major	5	0	103	0.0088
18	iv54	other	4	2	99	0.0085
19	iii#63	minor	3	1	92	0.0079
20	ii5#53	minor	2	0	89	0.0076

Figure 1: Frequency distribution of top 20 major mode chords from the works of Palestrina. Each line shows the rank, chord type (in Roman Numeral form), quality (major, minor, diminished, or other), the root number of the chord (with 1 = C, i.e. the tonic), the chord inversion, the number of times the chord appears in the corpus, and the frequency with which the chord appears in the corpus.

0.733 ($p = 0.016$); Palestrina and Bach have a correlation of 0.67 ($p = 0.033$). For the 10 top minor-mode chords, there is a statistically significant correlation between Palestrina and Monteverdi (Spearman rank-correlation: 0.794, $p = 0.006$). However, there is no significant correlation between Monteverdi and Bach (correlation: 0.491, $p = 0.150$), nor between Palestrina and Bach (correlation: 0.321, $p = 0.365$).

Chord co-occurrences

We created chord co-occurrence matrices per composer. However, for ease of comparison between composers, for each co-occurrence matrix we used the overall top 10 most common chords over all composers (see Methods for details). Each co-occurrence count is divided by the number of times that chord occurred in the musical corpus. A frequency of 1 would therefore imply that the chord occurred once per n-gram. It is possible to have frequencies above 1, especially for larger n-grams: this simply indicates that the chord in question appears more than once per n-gram.

Palestrina Minor

rank	chord	quality	root	inv	count	frequency
0	i	minor	1	0	1511	0.1351
1	v	minor	5	0	929	0.0830
2	iv	minor	4	0	609	0.0544
3	bVII	major	7	0	575	0.0514
4	III	major	3	0	567	0.0507
5	IV	major	4	0	309	0.0276
6	V	major	5	0	301	0.0269
7	bVI	major	6	0	279	0.0249
8	i54	other	1	2	249	0.0223
9	i6	minor	1	1	218	0.0195
10	iv6	minor	4	1	212	0.0189
11	I	major	1	0	191	0.0171
12	IIIb63	major	3	1	159	0.0142
13	v6	minor	5	1	157	0.0140
14	ii	minor	2	0	154	0.0138
15	iv54	other	4	2	148	0.0132
16	i532	minor	1	0	134	0.0120
17	bVIIb63	major	7	1	118	0.0105
18	II	major	2	0	102	0.0091
19	i5	minor	1	0	101	0.0090
20	i543	minor	1	0	78	0.0070

Figure 2: Frequency distribution of the top 20 minor mode chords from the works of Palestrina. As in Table 1, each line shows the rank, chord type (in Roman Numeral form), quality (major, minor, diminished, or other), the root number of the chord (with 1 = A, i.e. the tonic), the chord inversion, the number of times the chord appears in the corpus, and the frequency with which the chord appears in the corpus.

Trigram co-occurrence matrices

Major mode

The chord co-occurrence matrices for Palestrina (Figure 3, top), Monteverdi (Figure 3, middle) and Bach (Figure 3, bottom) share various similarities. There is a strong co-occurrence between I-I, V-I, V-V, II-V and V6-I for all three composers. There is a also a strong diagonal for Palestrina and Monteverdi which is barely present in Bach.

Minor mode

There is once again a strong diagonal for Palestrina and Monteverdi (Figure 4, top and middle), which is also found this time for the first 5 chords of Bach (Figure 4, bottom). Additionally, V-i, iv6-v and I-iv frequently co-occur for Palestrina, but beyond that the off-diagonal is generally weak for Palestrina. For Bach, i-i, iv6-V, and I-iv occur noticeably more frequently than other chord combinations. Also for

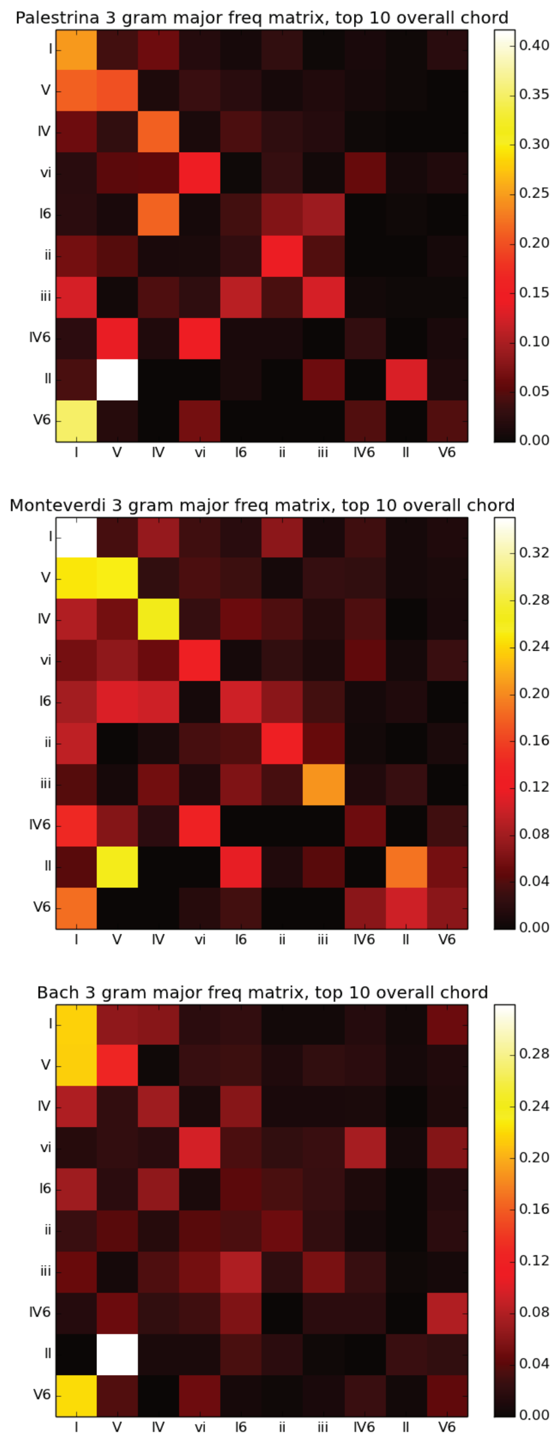

Figure 3: Major mode chord tri-gram co-occurrence matrix per composer for the top 10 chords over the whole corpus. Top: Palestrina. Middle: Monteverdi. Bottom: Bach.

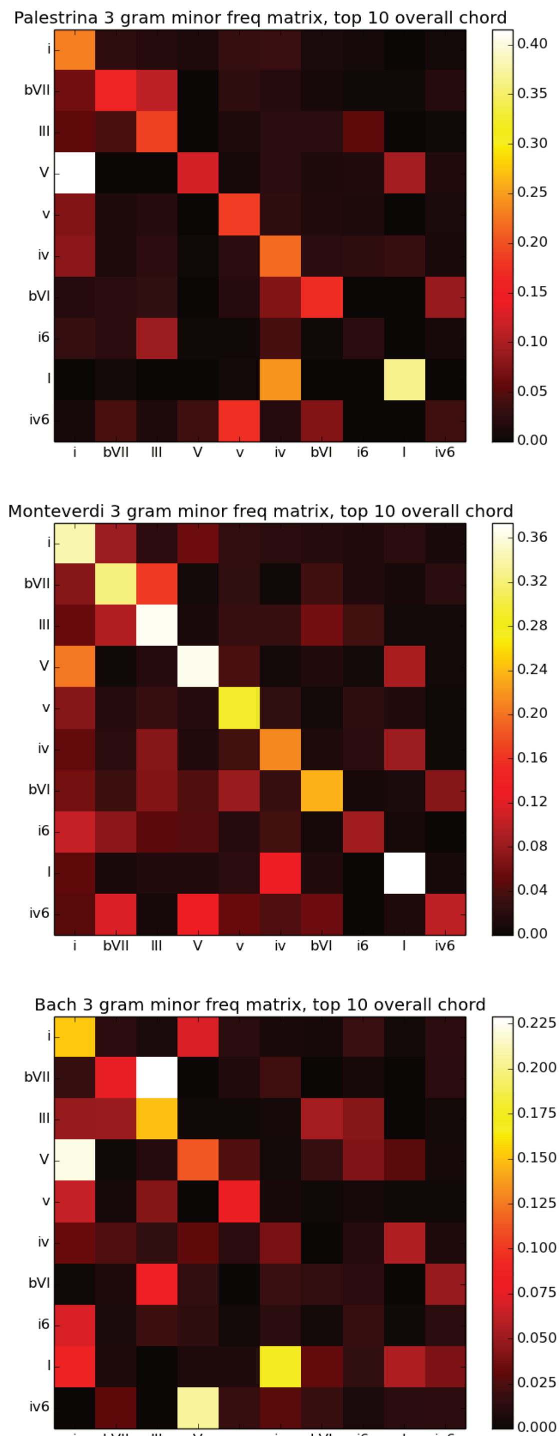

Figure 4: Minor mode chord tri-gram co-occurrence matrix per composer for the top 10 chords over the whole corpus. Top: Palestrina. Middle: Monteverdi. Bottom: Bach.

Bach, there is a trend for all chords to co-occur frequently with 'i'. This is in line with the music theoretic concept that Baroque music is the beginning of tonal music in the West, where music tends to center around the tonic ('i' or 'I') chord.

9-gram co-occurrence matrices

Major Mode

Most notable for all three composers in the 9-gram co-occurrence matrices is the dominance of co-occurrences with 'I' and with 'V' for all chords (Figure 5). The diagonal is still present for Palestrina and Monteverdi, although less noticeable than for the tri-gram co-occurrence matrices. There is no diagonal at all for the Bach pieces.

Minor Mode

The dominance of co-occurrences with 'I' for the 9-grams is still noticeable for all chords (Figure 6), but less so than for the major mode 9-grams. The diagonal is far more strongly present for Palestrina and Monteverdi in the minor than in the major mode. Additionally, there is a frequent co-occurrence between 'i' and 'iv' for all three composers.

Chord co-occurrence clustering

A k-means (k = 3) cluster analysis on the individual songs co-occurrence matrices of the top 10 chords, for all three composers, showed the following results.

Major and minor 3-grams

K-means cluster analysis for the major mode tri-gram chord co-occurrences matrices are somewhat clustered according to composer. Although there are three clusters, the majority of the songs are classified into only two categories. Palestrina and Monteverdi were clustered together (with 90% of the songs classified in Cluster 1), and Bach was clustered separately (with over 95% of Bach's songs classified in Cluster 2). Table 1 shows that of 3 different types of initialization (k-means++, random, and PCA-based), k-means initializations were most accurate. The Adjusted Rand Index (ARI, Table 1), shows an ARI of 0.794 (from a scale of 0 to 1, perfect accuracy being 1).

K-means cluster analysis for the minor mode tri-gram chord co-occurrences matrices not at all clustered according to composer. Table 2 shows that of 3 different types of initialization (k-means++, random, and PCA-based), k-means initializations were again most accurate. The Adjusted Rand Index (Table 2), shows an ARI of only 0.111.

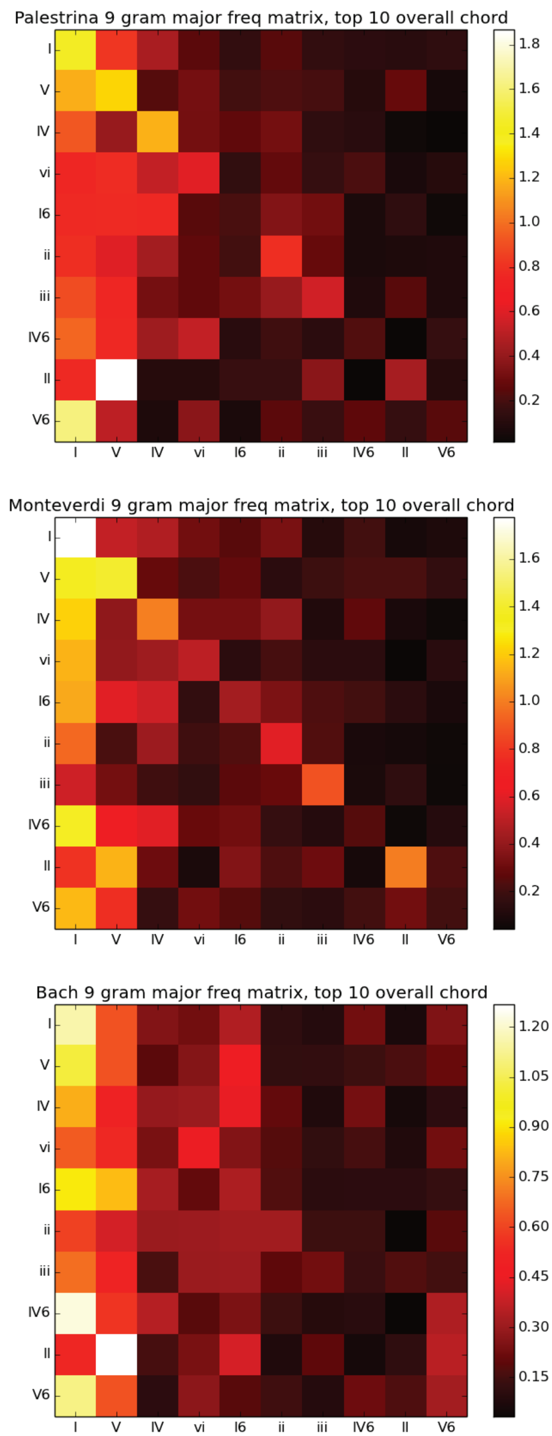

Figure 5: Major mode chord 9-gram co-occurrence matrix per composer for the top 10 chords over the whole corpus. Top: Palestrina. Middle: Monteverdi. Bottom: Bach.

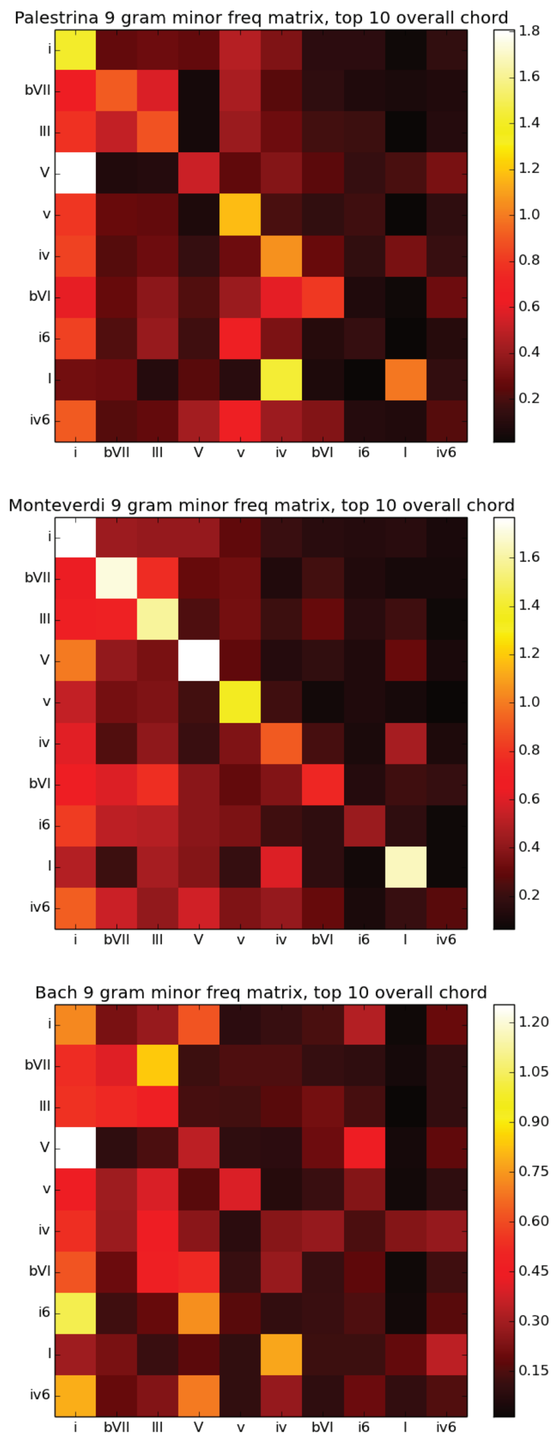

Figure 6: Minor mode chord 9-gram co-occurrence matrix per composer for the top 10 chords over the whole corpus. Top: Palestrina. Middle: Monteverdi. Bottom: Bach.

init	homo	compl	v-meas	ARI	AMI	silhouette
k-means++	0.547	0.680	0.606	0.794	0.543	0.257
random	0.583	0.616	0.599	0.783	0.579	0.212
PCA-based	0.569	0.619	0.593	0.782	0.564	0.234

Table 1: K-means cluster analysis results for major mode tri-gram chord co-occurrence matrices, for three different initialization types (k-means, random, and PCA). The tables show the homogeneity score (homo), the completeness score (compl), the V-measure score (harmonic mean between homogeneity and completeness), the Adjusted Rand Index (ARI), the Adjusted Mutual Information (AMI), and the silhouette coefficient (silhouette).

init	homo	compl	v-meas	ARI	AMI	silhouette
k-means++	0.267	0.297	0.281	0.111	0.262	0.243
random	0.260	0.292	0.275	0.103	0.255	0.240
PCA-based	0.236	0.261	0.248	0.083	0.231	0.240

Table 2: K-means cluster analysis results for minor mode tri-gram chord co-occurrence matrices, for three different initialization types (k-means, random, and PCA). The tables show the homogeneity score (homo), the completeness score (compl), the V-measure score (harmonic mean between homogeneity and completeness), the Adjusted Rand Index (ARI), the Adjusted Mutual Information (AMI), and the silhouette coefficient (silhouette).

Major and minor 9-grams

Similar to the major-mode tri-gram clustering, K-means cluster analysis for the major mode 9-gram chord co-occurrences matrices are somewhat clustered according to composer, with an ARI of 0.841 for the k-means initialization (Table 3). Of the major chord co-occurrence matrices, 99.1% of Bach was assigned to the same cluster (Cluster 0). Again, there was no differentiation between Palestrina and Monteverdi, with 83.3% of both Palestrina and Bach assigned to the same cluster (Cluster 1).

As with the minor tri-grams, the clustering did not differentiate between the three composers when classifying according to the minor chord mode co-occurrence matrices (ARI of 0.69, Table 4).

init	homo	compl	v-meas	ARI	AMI	silhouette
k-means++	0.615	0.726	0.666	0.841	0.611	0.358
random	0.544	0.665	0.599	0.788	0.540	0.356
PCA-based	0.504	0.339	0.405	0.380	0.334	0.121

Table 3: K-means cluster analysis results for major mode 9-gram chord co-occurrence matrices, for three different initialization types (k-means, random, and PCA). The tables show the homogeneity score (homo), the completeness score (compl), the V-measure score (harmonic mean between homogeneity and completeness), the Adjusted Rand Index (ARI), the Adjusted Mutual Information (AMI), and the silhouette coefficient (silhouette).

init	homo	compl	v-meas	ARI	AMI	silhouette
k-means++	0.186	0.232	0.206	0.069	0.180	0.323
random	0.173	0.224	0.196	0.064	0.168	0.378
PCA-based	0.280	0.298	0.289	0.148	0.275	0.307

Table 4: K-means cluster analysis results for minor mode 9-gram chord co-occurrence matrices, for three different initialization types (k-means, random, and PCA). The tables show the homogeneity score (homo), the completeness score (compl), the V-measure score (harmonic mean between homogeneity and completeness), the Adjusted Rand Index (ARI), the Adjusted Mutual Information (AMI), and the silhouette coefficient (silhouette).

Chord co-occurrence classification

We performed a 3-fold cross-validated decision tree classification and random forest classification of the individual song chord co-occurrence matrices. We classified according to composer for major and minor modes, and for 3, 5, 7, 9, 11, 13 and 15-gram chord co-occurrence matrices. We further examined the random forest classifiers' features' importance, to determine which features (or chord co-occurrences, in this case) were most important for the classification. Figure 5 shows an overview of both clustering (random forest and decision tree) and classification (k-means) results for major and minor mode co-occurrence matrices. This analysis confirms that there is a difference in clustering performance between major and minor mode chord co-occurrence matrices, for all n-grams analyzed, where the minor-mode chord co-occurrence matrices are not clustered according to composer, but the major-mode chord co-occurrence matrices are clustered according to composer. Although the k-means cluster analysis did not cluster the minor chord co-occurrence n-grams into clusters according to composer, both the decision-tree and the random-forest classifiers performed at well above chance levels for all minor mode chord co-occurrence n-grams (Figure 7). This shows us that there is enough information in the chord co-occurrence matrices to distinguish between composers for minor-mode songs, even though an unsupervised classifier would not naturally cluster the chord co-occurrence matrices according to composer. Major mode chord co-occurrence matrices are more accurately clustered

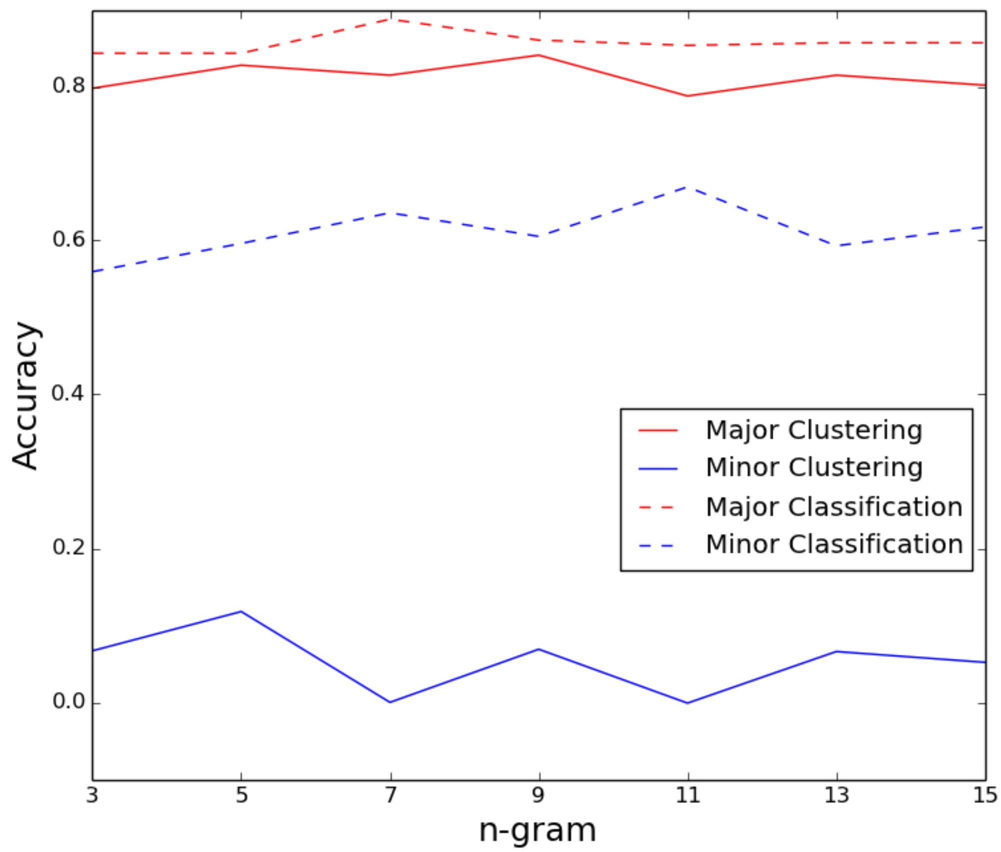

Figure 7: K-means clustering and Decision Tree and Random Forest classification performance for major and minor mode 3, 5, 7, 9, 11, 13, and 15-gram chord co-occurrence matrices.

and more accurately classified than minor-mode chord co-occurrence matrices, for all n-grams, and both random-forest and decision-tree classification always outperform k-means clustering.

Feature-importance analysis for the best-performing classifier (Random Forest classifier, Figure 8) shows that different features are most important for classifying different size n-gram chord co-occurrence matrices, and this difference is more pronounced for minor than for major chord co-occurrence matrices.

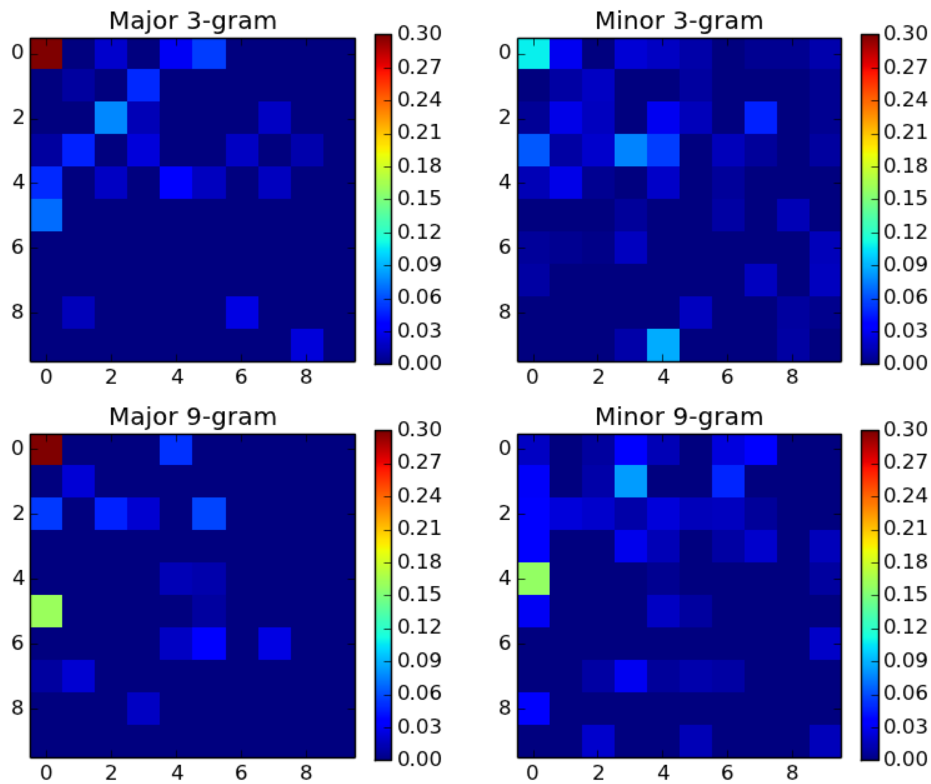

Figure 8: Feature importance for Random Forest Classifier, classifying chord co-occurrence matrices according to composer for the top 10 chords in major and in minor-mode songs. Top left: major mode tri-gram feature importance. Top right: minor tri-gram feature importance. Bottom left: Major 9-gram feature importance. Bottom right: Minor 9-gram feature importance.

Conclusion

In this paper, we performed a statistical analysis of three composers from the Baroque and the Renaissance musical traditions. Our goal was threefold: to contribute to the empirical knowledge-base of harmonic syntax; to help create a base that could be used as a feature space for future analysis (such as neuroscience), and to contribute to the understanding of the evolution of harmonic syntactical statistics.

We found that the chord frequency distributions are all Zipfian. We also found an overlap in the frequency distributions of the most common chords for all three composers, with greater overlap in major than in minor modes. In particular, 'I', 'V', 'IV' and 'vi' are in the top 5 most frequent chords for the major-mode songs for all three composers. There is a statistically significant correlation between the top 10 major-mode chords for all three composers. However, there is only a statistically significant correlation between the top 10 minor-mode chords for the two earliest composers, Palestrina and Monteverdi.

The chord co-occurrence matrices differ depending on n-gram window size, and on composer. Monteverdi and Palestrina maintain a stronger diagonal than Bach. Likewise smaller n-gram windows also tend to have a stronger diagonal. The larger n-gram windows favor 'I' or 'i', as does the Baroque composer (Bach). This is in line with the music theoretic concept that Baroque music is the beginning of tonal music in the West, where music tends to center around the tonic ('i' or 'I') chord.

Songs were equally accurately clustered according to composer for n-grams ranging from 3 to 15 chord co-occurrence matrices. However, while major-mode chord co-occurrence matrices were relatively well clustered (with Palestrina and Monteverdi falling within the same cluster, and Bach clustered separately), minor-mode chord co-occurrence matrices were not at all clustered according to composer (with 0-1.1% Adjusted Rand Index). This difference between major and minor-mode analysis was also present in classification analysis. A random forest classifier classified the songs over 90% correctly according to composer for the major-mode songs. While the random forest classifier for the minor mode songs performed far better than the clustering, it still perfomed worse than the major-mode classifier, with its best classification only reaching 68% correct.

De Clercq & Temperly [20] analyzed harmony of a rock corpus, creating a set of 20 songs per decade from Rolling Stone's list of the 500 greatest rock songs, ranging from the 1950s to the 1990s. They found that 'I' was the most common chord, followed by 'IV'. Interestingly, it therefore seems that the most frequently used chords do not seem to have changed in the past 500 years. However, the difference between the pre-tonal (Renaissance) composer chord co-occurrence matrices and the post-tonal (Baroque) composer chord co-occurrence matrices in our study points to the necessity of creating individualized co-occurrence matrices per musical genre, and possibly even per musical sub-genre. Previous research supports this finding, both in classical and in more modern music. White [10] found that harmonic progressions can be defined by time-period in classical music. Again, the De Clercq & Temperley [20] study revealed changes over time in harmonic composition. So it seems that while chord frequencies may remain very stable, chord transition practices can change quite rapidly, and add the color that helps to define a specific musical time-period.

Implementation and future directions

It would be very interesting to expand this analysis to a much wider variety of composers and musical genres; the application of this research is limited if it only remains in the early Classical musical period.

We intend on implementing a web-based application that allows users to analyze and create their own musical feature space, based on .xml or .mxl musical files. This will allow researchers who want to use chord co-occurrences as a feature set to easily translate individual chords into their representation in frequencies of common chords, and provide musical context to their analyses. We are particularly excited about its potential applications in neuroscience research: this chord co-occurrence feature set could create a high-level model of cortical responses to music, similar to high-level models that have been created and have worked remarkably well for speech.

Bibliography

[1] B. L. Sturm, "A survey of evaluation in music genre recognition," in *Adaptive Multimedia Retrieval: Semantics, Context, and Adaptation*, Springer, 2014, pp. 29–66.

[2] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *Speech and Audio Processing, IEEE transactions on*, vol. 10, no. 5, pp. 293–302, 2002.

[3] M. Muller, D. P. Ellis, A. Klapuri, and G. Richard, "Signal processing for music analysis," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 6, pp. 1088–1110, 2011.

[4] J. George and L. Shamir, "Unsupervised analysis of similarities between musicians and musical genres using spectrograms," *Artificial Intelligence Research*, vol. 4, no. 2, p61, 2015.

[5] Y. Panagakis, C. Kotropoulos, and G. R. Arce, "Music genre classification via sparse representations of auditory temporal modulations," in *Signal Processing Conference, 2009 17th European*, IEEE, 2009, pp. 1–5.

[6] Y. Panagakis, C. L. Kotropoulos, and G. R. Arce, "Music genre classification via joint sparse low-rank representation of audio features," *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, vol. 22, no. 12, pp. 1905–1917, 2014.

[7] J. Salamon, B. Rocha, and E. Gómez, "Musical genre classification using melody features extracted from polyphonic music signals," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, IEEE, 2012, pp. 81–84.

[8] C. McKay and I. Fujinaga, "Automatic genre classification using large high-level musical feature sets.," in *ISMIR*, Citeseer, vol. 2004, 2004, pp. 525–530.

[9] M. Rohrmeier and I. Cross, "Statistical properties of tonal harmony in bach's chorales," 2008.

[10] C. W. White, "Changing styles, changing corpora, changing tonal models," *Music Perception: An Interdisciplinary Journal*, vol. 31, no. 3, pp. 244–253, 2014.

[11] M. Pearce and M. Rohrmeier, "Music cognition and the cognitive sciences," *Topics in cognitive science*, vol. 4, no. 4, pp. 468–484, 2012.

[12] V. Tsatsishvili, F. Cong, T. Ristaniemi, P. Toiviainen, V. Alluri, E. Brattico, *et al.*, "Generation of stimulus features for analysis of fmri during natural auditory experiences," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, IEEE, 2014, pp. 2490–2494.

[13] C. W. White, "A corpus-sensitive algorithm for automated tonal analysis," in *Mathematics and Computation in Music*, Springer, 2015, pp. 115–121.

[14] P. D. Turney, P. Pantel, *et al.*, "From frequency to meaning: vector space models of semantics," *Journal of artificial intelligence research*, vol. 37, no. 1, pp. 141–188, 2010.

[15] K. Lund and C. Burgess, "Producing high-dimensional semantic spaces from lexical co-occurrence," *Behavior Research Methods, Instruments, & Computers*, vol. 28, no. 2, pp. 203–208, 1996.

[16] M. S. Cuthbert and C. Ariza, "Music21: a toolkit for computer-aided musicology and symbolic music data," 2010.

[17] ——, (08/2015). Music21 corpus, [Online]. Available: `http://web.mit.edu/music21/doc/systemReference/referenceCorpus.html#referencecorpus`.

[18] E. Jones, T. Oliphant, and P. Peterson, "{scipy}: open source scientific tools for {python}," 2014.

[19] T. E. Oliphant, *A guide to NumPy*. Trelgol Publishing USA, 2006, vol. 1.

[20] T. De Clercq and D. Temperley, "A corpus analysis of rock harmony," *Popular Music*, vol. 30, no. 01, pp. 47–70, 2011.

Appendix

Monteverdi Major

rank	chord	quality	root	inv	count	frequency
0	I	major	1	0	655	0.1435
1	V	major	5	0	302	0.0662
2	IV	major	4	0	228	0.0499
3	ii	minor	2	0	201	0.0440
4	vi	minor	6	0	170	0.0372
5	I6	major	1	1	147	0.0322
6	iii	minor	3	0	145	0.0318
7	v	minor	5	0	142	0.0311
8	i	minor	1	0	108	0.0237
9	IV6	major	4	1	92	0.0202
10	II	major	2	0	69	0.0151
11	ii6	minor	2	1	66	0.0145
12	I64	major	1	2	63	0.0138
13	vii	minor	7	0	59	0.0129
14	iv	minor	4	0	59	0.0129
15	V6	major	5	1	59	0.0129
16	i5	minor	1	0	50	0.0110
17	i54	other	1	2	49	0.0107
18	v5	minor	5	0	44	0.0096
19	viio#63	dimin	7	1	44	0.0096
20	V7	major	5	0	40	0.0088

Figure 9: Frequency distribution of top 20 major mode chords from the works of Monteverdi. Each line shows the rank, chord type (in Roman Numeral form), quality (major, minor, diminished, or other), the root number of the chord (with 1 = C, i.e. the tonic), the chord inversion, the number of times the chord appears in the corpus, and the frequency with which the chord appears in the corpus.

Monteverdi Minor

rank	chord	quality	root	inv	count	frequency
0	i	minor	1	0	2247	0.1108
1	bVII	major	7	0	1383	0.0682
2	III	major	3	0	1380	0.0680
3	v	minor	5	0	1298	0.0640
4	V	major	5	0	1066	0.0525
5	iv	minor	4	0	766	0.0378
6	I	major	1	0	699	0.0345
7	IV	major	4	0	550	0.0271
8	bVI	major	6	0	485	0.0239
9	ii	minor	2	0	469	0.0231
10	i6	minor	1	1	364	0.0179
11	bvii	minor	7	0	293	0.0144
12	iii	minor	3	0	264	0.0130
13	i5	minor	1	0	258	0.0127
14	v6	minor	5	1	246	0.0121
15	iv6	minor	4	1	241	0.0119
16	II	major	2	0	237	0.0117
17	IIIb63	major	3	1	231	0.0114
18	i64	minor	1	2	221	0.0109
19	bVIIb63	major	7	1	189	0.0093
20	v7	minor	5	0	167	0.0082

Figure 10: Frequency distribution of the top 20 minor mode chords from the works of Monteverdi. Each line shows the rank, chord type (in Roman Numeral form), quality (major, minor, diminished, or other), the root number of the chord (with 1 = A, i.e. the tonic), the chord inversion, the number of times the chord appears in the corpus, and the frequency with which the chord appears in the corpus.

Bach Major

rank	chord	quality	root	inv	count	frequency
0	I	major	1	0	3315	0.1345
1	V	major	5	0	2167	0.0880
2	I6	major	1	1	1157	0.0470
3	vi	minor	6	0	1003	0.0407
4	IV	major	4	0	816	0.0331
5	V7	major	5	0	727	0.0295
6	V6	major	5	1	704	0.0286
7	ii	minor	2	0	563	0.0229
8	IV6	major	4	1	543	0.0220
9	iii	minor	3	0	391	0.0159
10	viio#63	dimin	7	1	341	0.0138
11	II	major	2	0	329	0.0134
12	i54	other	1	2	316	0.0128
13	V42	major	5	3	287	0.0116
14	ii65#3	minor	2	1	284	0.0115
15	I42	major	1	3	266	0.0108
16	I64	major	1	2	266	0.0108
17	V65	major	5	1	241	0.0098
18	ii6	minor	2	1	210	0.0085
19	II75#3	major	2	0	196	0.0080
20	IV7	major	4	0	193	0.0078

Figure 11: Frequency distribution of top 20 major mode chords from the works of Bach. Each line shows the rank, chord type (in Roman Numeral form), quality (major, minor, diminished, or other), the root number of the chord (with 1 = C, i.e. the tonic), the chord inversion, the number of times the chord appears in the corpus, and the frequency with which the chord appears in the corpus.

Bach Minor

rank	chord	quality	root	inv	count	frequency
0	i	minor	1	0	1993	0.1078
1	V	major	5	0	1134	0.0613
2	III	major	3	0	1009	0.0546
3	bVII	major	7	0	751	0.0406
4	i6	minor	1	1	601	0.0325
5	V7	major	5	0	448	0.0242
6	iv6	minor	4	1	428	0.0231
7	bVI	major	6	0	418	0.0226
8	iv	minor	4	0	390	0.0211
9	v	minor	5	0	381	0.0206
10	IIIb63	major	3	1	348	0.0188
11	bVIIb63	major	7	1	316	0.0171
12	V6	major	5	1	276	0.0149
13	I	major	1	0	273	0.0148
14	iio6	dimin	2	1	270	0.0146
15	iio65	dimin	2	1	255	0.0138
16	i54	other	1	2	212	0.0115
17	i64	minor	1	2	201	0.0109
18	viio#63	dimin	7	1	191	0.0103
19	V65	major	5	1	190	0.0103
20	IV	major	4	0	189	0.0102

Figure 12: Frequency distribution of the top 20 minor mode chords from the works of Bach. Each line shows the rank, chord type (in Roman Numeral form), quality (major, minor, diminished, or other), the root number of the chord (with 1 = A, i.e. the tonic), the chord inversion, the number of times the chord appears in the corpus, and the frequency with which the chord appears in the corpus.

All composers, Major tri-gram frequency co-occurrence matrices

Palestrina

	I	V	IV	vi	ii	iii	I6	i54	II	IV6
I	0.2490	0.0401	0.0643	0.0195	0.0289	0.0024	0.0083	0.0437	0.0047	0.0112
V	0.2122	0.2018	0.0141	0.0357	0.0082	0.0171	0.0223	0.0208	0.0045	0.0089
IV	0.0636	0.0286	0.2117	0.0130	0.0273	0.0182	0.0442	0.0195	0.0000	0.0026
vi	0.0213	0.0523	0.0543	0.1492	0.0310	0.0058	0.0019	0.0504	0.0097	0.0601
ii	0.0681	0.0501	0.0100	0.0120	0.1403	0.0481	0.0301	0.0000	0.0000	0.0000
iii	0.1258	0.0061	0.0460	0.0276	0.0429	0.1258	0.1104	0.0000	0.0031	0.0061
I6	0.0231	0.0099	0.2145	0.0066	0.0759	0.0924	0.0396	0.0099	0.0033	0.0000
i54	0.0000	0.7750	0.0000	0.0000	0.0107	0.0036	0.0000	0.0071	0.0000	0.0000
II	0.0437	0.4167	0.0000	0.0000	0.0000	0.0635	0.0119	0.0000	0.1270	0.0000
IV6	0.0249	0.1393	0.0149	0.1443	0.0100	0.0000	0.0100	0.0995	0.0000	0.0299

Monteverdi

	I	V	IV	ii	vi	I6	iii	v	i	IV6
I	0.3510	0.0343	0.0733	0.0686	0.0328	0.0187	0.0094	0.0062	0.0094	0.0328
V	0.2492	0.2525	0.0236	0.0067	0.0370	0.0303	0.0269	0.0034	0.0034	0.0236
IV	0.0881	0.0573	0.2643	0.0396	0.0264	0.0529	0.0176	0.0176	0.0088	0.0396
ii	0.0950	0.0000	0.0100	0.1300	0.0350	0.0400	0.0500	0.0150	0.0050	0.0050
vi	0.0588	0.0706	0.0529	0.0235	0.1294	0.0059	0.0118	0.0412	0.0118	0.0471
I6	0.0822	0.1096	0.1027	0.0685	0.0068	0.1027	0.0342	0.0000	0.0137	0.0068
iii	0.0414	0.0069	0.0552	0.0345	0.0138	0.0621	0.2069	0.0207	0.0000	0.0138
v	0.0567	0.0071	0.0284	0.0213	0.0284	0.0213	0.0284	0.1631	0.0071	0.0355
i	0.0769	0.0385	0.0000	0.0192	0.0192	0.0096	0.0096	0.0000	0.2212	0.0096
IV6	0.1413	0.0652	0.0217	0.0000	0.1304	0.0000	0.0000	0.0326	0.0109	0.0543

Bach

	I	V	I6	vi	IV	V7	V6	ii	IV6	iii
I	0.2161	0.0642	0.0234	0.0175	0.0601	0.0034	0.0481	0.0045	0.0141	0.0041
V	0.2153	0.1264	0.0282	0.0255	0.0028	0.1685	0.0130	0.0120	0.0176	0.0218
I6	0.0717	0.0182	0.0406	0.0095	0.0648	0.0009	0.0138	0.0328	0.0104	0.0259
vi	0.0152	0.0233	0.0344	0.0952	0.0172	0.0051	0.0588	0.0223	0.0750	0.0274
IV	0.0790	0.0235	0.0605	0.0099	0.0716	0.0099	0.0111	0.0086	0.0099	0.0086
V7	0.6286	0.0344	0.0028	0.0990	0.0000	0.0344	0.0000	0.0000	0.0028	0.0014
V6	0.2217	0.0372	0.0072	0.0486	0.0000	0.0215	0.0429	0.0029	0.0272	0.0086
ii	0.0267	0.0391	0.0338	0.0391	0.0160	0.0231	0.0178	0.0498	0.0053	0.0231
IV6	0.0147	0.0479	0.0571	0.0295	0.0221	0.0129	0.0810	0.0000	0.0184	0.0184
iii	0.0463	0.0051	0.0797	0.0514	0.0360	0.0051	0.0051	0.0206	0.0257	0.0540

Figure 13: Tri-gram frequency co-occurrence matrices for the top 10 major mode chords, for the three different composers

All composers, Major 9gram frequency co-occurrence matrices

Palestrina

	I	V	IV	vi	I6	ii	iii	IV6	II	V6
I	1.4527	0.8210	0.4575	0.2480	0.1454	0.2456	0.1509	0.1271	0.1124	0.1326
V	1.1768	1.2867	0.2337	0.3213	0.1952	0.2198	0.1991	0.1053	0.2790	0.0561
IV	0.9324	0.4122	1.1838	0.3176	0.2622	0.3162	0.1378	0.1149	0.0351	0.0149
vi	0.7579	0.7837	0.5218	0.6091	0.1409	0.2738	0.1567	0.2123	0.0615	0.1032
I6	0.7718	0.7752	0.7651	0.2450	0.2081	0.3490	0.3054	0.0671	0.1309	0.0336
ii	0.7963	0.6070	0.4486	0.2654	0.1934	0.7922	0.2778	0.0720	0.0782	0.0823
iii	0.8927	0.7571	0.3218	0.2650	0.3123	0.4101	0.5647	0.0820	0.2429	0.0852
IV6	0.9751	0.7612	0.4279	0.5224	0.1095	0.1841	0.1194	0.2289	0.0199	0.1592
II	0.7686	1.8678	0.1074	0.1074	0.1653	0.1653	0.3760	0.0165	0.4545	0.1074
V6	1.6181	0.5139	0.0764	0.3750	0.0694	0.2500	0.1736	0.2569	0.1528	0.2431

Monteverdi

	I	V	IV	vi	I6	ii	iii	IV6	II	V6
I	1.7765	0.5214	0.4754	0.3170	0.2536	0.3360	0.1252	0.2044	0.0761	0.0998
V	1.3191	1.4043	0.2908	0.2305	0.2801	0.1383	0.1915	0.2234	0.2199	0.1702
IV	1.2168	0.3938	1.0088	0.3274	0.3274	0.3982	0.1106	0.2743	0.0885	0.0531
vi	1.1429	0.3988	0.4345	0.5060	0.1369	0.2143	0.1369	0.1369	0.0417	0.1310
I6	1.1189	0.6014	0.5455	0.1678	0.4476	0.3427	0.2378	0.2098	0.1399	0.0839
ii	0.9492	0.2183	0.4264	0.1980	0.2386	0.5990	0.2437	0.0863	0.0711	0.0558
iii	0.5448	0.3241	0.2000	0.1586	0.2621	0.2966	0.8828	0.0897	0.1517	0.0552
IV6	1.3152	0.6739	0.5978	0.2935	0.3152	0.1739	0.1196	0.2500	0.0543	0.1196
II	0.7812	1.1406	0.3125	0.0938	0.3594	0.2344	0.3125	0.0781	1.0000	0.2344
V6	1.1579	0.7719	0.1754	0.3158	0.2456	0.1579	0.1404	0.2105	0.3158	0.2105

Bach

	I	V	IV	vi	I6	ii	iii	IV6	II	V6
I	1.1594	0.6331	0.2547	0.2277	0.3421	0.1129	0.0885	0.2288	0.0647	0.2532
V	1.0248	0.6311	0.1909	0.2637	0.4716	0.1151	0.1223	0.1388	0.1631	0.2141
IV	0.8092	0.5046	0.2889	0.2980	0.4431	0.2052	0.0784	0.2366	0.0601	0.1085
vi	0.6525	0.5320	0.2431	0.4478	0.2559	0.1812	0.1151	0.1567	0.0853	0.2271
I6	0.9129	0.8294	0.3240	0.2024	0.3330	0.1760	0.1034	0.1053	0.1071	0.1261
ii	0.5946	0.4036	0.3009	0.3063	0.3153	0.3135	0.1423	0.1405	0.0360	0.1856
iii	0.6873	0.5175	0.1590	0.3019	0.3046	0.1941	0.2264	0.1375	0.1725	0.1509
IV6	1.2218	0.5679	0.3537	0.1836	0.2486	0.1434	0.0918	0.0994	0.0325	0.3365
II	0.5292	1.2677	0.1538	0.2400	0.4062	0.0800	0.1938	0.0523	0.1138	0.3569
V6	1.1115	0.6351	0.1053	0.2779	0.1832	0.1466	0.0901	0.2229	0.1695	0.3206

Figure 14: 9-gram frequency co-occurrence matrices for the top 10 major mode chords, for the three different composers

All composers, Minor tri-gram frequency co-occurrence matrices

Palestrina

	i	bVII	III	V	v	iv	bVI	i6	I	iv6
i	0.2314	0.0262	0.0201	0.0134	0.0322	0.0355	0.0114	0.0067	0.0000	0.0054
bVII	0.0661	0.1617	0.1113	0.0000	0.0261	0.0191	0.0087	0.0035	0.0035	0.0209
III	0.0564	0.0423	0.1887	0.0000	0.0141	0.0212	0.0212	0.0564	0.0000	0.0018
V	0.4152	0.0000	0.0000	0.1246	0.0069	0.0242	0.0138	0.0173	0.0969	0.0173
v	0.0755	0.0142	0.0186	0.0000	0.1860	0.0274	0.0175	0.0142	0.0000	0.0120
iv	0.0796	0.0133	0.0249	0.0017	0.0232	0.2189	0.0232	0.0265	0.0332	0.0100
bVI	0.0179	0.0215	0.0287	0.0000	0.0179	0.0753	0.1720	0.0000	0.0000	0.0860
i6	0.0321	0.0229	0.0917	0.0046	0.0046	0.0413	0.0046	0.0229	0.0000	0.0092
I	0.0000	0.0064	0.0000	0.0000	0.0064	0.2436	0.0000	0.0000	0.3654	0.0000
iv6	0.0094	0.0425	0.0142	0.0377	0.1745	0.0189	0.0755	0.0000	0.0000	0.0377

Monteverdi

	i	bVII	III	V	v	iv	bVI	i6	I	iv6
i	0.3418	0.0823	0.0231	0.0579	0.0267	0.0199	0.0176	0.0154	0.0190	0.0095
bVII	0.0701	0.3225	0.1634	0.0058	0.0253	0.0036	0.0347	0.0152	0.0080	0.0195
III	0.0552	0.0936	0.3665	0.0102	0.0298	0.0283	0.0588	0.0356	0.0051	0.0058
V	0.2038	0.0038	0.0171	0.3657	0.0381	0.0048	0.0152	0.0048	0.0895	0.0057
v	0.0709	0.0164	0.0296	0.0171	0.2931	0.0257	0.0055	0.0234	0.0148	0.0039
iv	0.0516	0.0198	0.0701	0.0146	0.0357	0.2130	0.0119	0.0198	0.0820	0.0040
bVI	0.0599	0.0331	0.0682	0.0434	0.0806	0.0289	0.2355	0.0083	0.0103	0.0702
i6	0.1044	0.0742	0.0467	0.0440	0.0165	0.0357	0.0082	0.0852	0.0082	0.0000
I	0.0503	0.0091	0.0152	0.0137	0.0213	0.1341	0.0137	0.0000	0.3735	0.0076

Bach

	i	bVII	III	V	v	iv	bVI	i6	I	iv6
i	0.1512	0.0142	0.0071	0.0699	0.0120	0.0060	0.0049	0.0197	0.0033	0.0136
bVII	0.0173	0.0759	0.2290	0.0000	0.0093	0.0213	0.0000	0.0053	0.0000	0.0133
III	0.0485	0.0495	0.1486	0.0020	0.0020	0.0040	0.0536	0.0435	0.0000	0.0030
V	0.2222	0.0009	0.0101	0.1129	0.0266	0.0028	0.0174	0.0404	0.0285	0.0037
v	0.0638	0.0053	0.0426	0.0000	0.0771	0.0053	0.0027	0.0053	0.0027	0.0027
iv	0.0335	0.0258	0.0155	0.0309	0.0129	0.0387	0.0000	0.0103	0.0567	0.0077
bVI	0.0024	0.0072	0.0837	0.0167	0.0000	0.0191	0.0167	0.0120	0.0000	0.0478
i6	0.0701	0.0067	0.0200	0.0150	0.0033	0.0117	0.0033	0.0184	0.0017	0.0134
I	0.0880	0.0080	0.0000	0.0080	0.0080	0.1760	0.0320	0.0160	0.0560	0.0400
iv6	0.0000	0.0304	0.0000	0.2056	0.0187	0.0280	0.0187	0.0070	0.0117	0.0140

Figure 15: 3-gram frequency co-occurrence matrices for the top 10 minor mode chords, for the three different composers

All composers, Minor 9gram frequency co-occurrence matrices

Palestrina

	i	bVII	III	V	v	iv	bVI	i6	I	iv6
i	1.4157	0.2604	0.2921	0.2617	0.4793	0.3315	0.1209	0.1222	0.0276	0.1360
bVII	0.6567	0.9155	0.5687	0.0440	0.4472	0.2324	0.1285	0.0792	0.0651	0.0898
III	0.7718	0.5098	0.8841	0.0463	0.4011	0.2923	0.1872	0.1729	0.0196	0.0945
V	1.8083	0.0902	0.0940	0.5301	0.2519	0.3496	0.2406	0.1504	0.1992	0.3195
v	0.7975	0.2785	0.2617	0.0761	1.1767	0.1991	0.1342	0.1779	0.0134	0.1298
iv	0.8348	0.2285	0.2900	0.1529	0.2917	1.0650	0.2759	0.1336	0.3216	0.1564
bVI	0.6036	0.2691	0.3709	0.2182	0.4036	0.6000	0.8036	0.0800	0.0327	0.2873
i6	0.8287	0.2222	0.3981	0.1806	0.6759	0.3287	0.1019	0.1481	0.0139	0.1019
I	0.3000	0.2917	0.0917	0.2333	0.1083	1.4417	0.0750	0.0167	0.9917	0.1417
iv6	0.9052	0.2275	0.2607	0.4313	0.6730	0.4076	0.3365	0.0948	0.0853	0.2275

Monteverdi

	i	bVII	III	V	v	iv	bVI	i6	I	iv6
i	1.7673	0.4431	0.4191	0.4176	0.2968	0.2081	0.1538	0.1383	0.1553	0.1067
bVII	0.6420	1.7042	0.7738	0.3053	0.3360	0.1332	0.2269	0.1332	0.1018	0.0996
III	0.6913	0.7075	1.5924	0.2566	0.3409	0.2119	0.3050	0.1569	0.2185	0.0777
V	0.9990	0.4122	0.3521	1.7653	0.2969	0.1414	0.1805	0.1274	0.3109	0.1063
v	0.5352	0.3498	0.3700	0.2300	1.3806	0.2227	0.0899	0.1263	0.0988	0.0648
iv	0.6135	0.2641	0.4085	0.2050	0.3700	0.9106	0.2352	0.1142	0.4677	0.1238
bVI	0.6858	0.6030	0.7792	0.3949	0.3036	0.3737	0.7495	0.1401	0.2229	0.1975
i6	0.8260	0.5221	0.5083	0.3923	0.3619	0.2210	0.1768	0.4337	0.1713	0.0746
I	0.5016	0.2168	0.4693	0.3835	0.1974	0.5987	0.1715	0.0890	1.6537	0.0744
iv6	0.9237	0.5551	0.4110	0.5678	0.3686	0.4195	0.3093	0.1144	0.1992	0.2754

Bach

	i	bVII	III	V	v	iv	bVI	i6	I	iv6
i	0.7211	0.2257	0.2795	0.6221	0.0891	0.1140	0.1424	0.3304	0.0243	0.1950
bVII	0.5294	0.4090	0.8427	0.1245	0.1518	0.1505	0.1081	0.0958	0.0342	0.1040
III	0.5423	0.5168	0.4740	0.1356	0.1325	0.1692	0.2171	0.1366	0.0143	0.1050
V	1.2545	0.0954	0.1463	0.3478	0.0933	0.0891	0.2078	0.4475	0.0339	0.1835
v	0.4454	0.2923	0.3934	0.1667	0.4016	0.0765	0.1202	0.2459	0.0301	0.0956
iv	0.5341	0.2834	0.4387	0.2561	0.0845	0.2507	0.2752	0.1471	0.2452	0.2752
bVI	0.6221	0.2005	0.4704	0.5141	0.1131	0.2776	0.1157	0.1774	0.0283	0.1260
i6	1.0333	0.1259	0.1907	0.7333	0.1667	0.1019	0.1204	0.1537	0.0333	0.1667
I	0.2946	0.2232	0.1161	0.1696	0.0982	0.7768	0.1250	0.1250	0.1875	0.3482
iv6	0.7908	0.1898	0.2409	0.6959	0.0973	0.2725	0.0949	0.2019	0.1022	0.1582

Figure 16: 9-gram frequency co-occurrence matrices for the top 10 minor mode chords, for the three different composers