Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

On the Emergence of the Major-Minor System: Cluster Analysis Suggests the Late 16th Century Collapse of the Dorian and Aeolian Modes

Joshua Albrecht,[*1] David Huron[*2]

[*]*School of Music, Ohio State University, USA*
[1]albrecht.89@osu.edu, [2]huron.1@osu.edu

ABSTRACT

Stable scale-degree distributions have been observed for an idealized version of the major and minor scales. However, these scales developed out of an earlier system of modes. This paper describes a corpus study conducted on works spanning the period in which the major and minor modes were established as the dominant modes. The study involves 455 musical works by 259 composers sampled across the years 1400 to 1750. Beginning with the period 1700-1750, a series of statistical studies are carried out on the distribution of scale tones, progressively moving backward in time. The method utilizes a modified version of the Krumhansl-Schmuckler method of key determination – generalized to handle an arbitrary number of modal classifications. The results from cluster analyses on this data are consistent with the view that the modern minor mode emerged from the amalgamation of earlier Dorian and Aeolian modes, with the collapse being completed around the late sixteenth century.

I. INTRODUCTION

For many modern musicians and listeners, the major-minor system is taken for granted as the normative scales for Western music-making. However, this system did not arise out of nothing. Music historians have chronicled the emergence of this system around the sixteenth century. Prior to this, a system of modes existed.

Modal theory has a long and profoundly complex history that ranges from the ancient Greeks to modern times. The modal system of the Middle Ages has been the subject of extensive research (Apel 1958; Atkinson 1982, 1989, 1995; Auda 1979; Curtis 1992; Powers/Wiering 2001; Schlager 1985, and others).

Independent of the historical and analytic research, psychological research has also addressed the phenomenon of scales and the associated phenomenon of tonality (Krumhansl & Kessler, 1982; Krumhansl, 1990; Huron & Veltman, 2006; Butler, 1988; Butler & Brown, 1981; Lerdahl, 2001, Cuddy, 1997; Cohen, 1991). Central to this research effort has been the work of Carol Krumhansl. In 1990, Krumhansl published her seminal work *Cognitive Foundations of Musical Pitch*, in which she chronicled experimental efforts to elucidate the mental representations for "key" and "scale," and related these subjective experiences to the corresponding objective acoustical phenomena. Using a "probe tone" paradigm, Krumhansl and Kessler measured the stability of various scale tones in different contexts. After establishing a key context, Krumhansl and Kessler played one of 12 chromatic tones, and participants judged the "goodness of fit" of the probe tone with the antecedent tonal context.

The results, replicate the long-standing concept of scale-degree hierarchy described by music theorists. In the context of listening to music, the various pitches accrue certain subjective properties, such as the feeling of stability. Specifically, one pitch comes to be anchored mentally as the central or tonic pitch. The tonic tends to be stressed metrically, rhythmically, and agogically, and to begin and end melodic phrases and works. Following the tonic pitch, the mediant and dominant pitches are the next most prevalent, followed by the remaining pitches of the scale, with the pitches of the chromatic set exhibiting the least stability and occurring the least frequently. Tones are readily perceived according to their functions or positions within this diatonic set. Following up on Krumhansl's work, Huron (2006) showed how simple statistical properties of the various scale tones could be used to account for the subjective *qualia* or affective descriptions provided by listeners to characterize the different scale tones.

Krumhansl observed that there was a close similarity between the experimentally-determined "key profiles" and the distribution of scale tones in actual music. Both Youngblood (1958) and Hutchinson and Knopoff (1983) measured the distributions of pitches in samples of Schubert songs, Mozart arias, Strauss Leider, and Mendelssohn arias. The combination of these two studies represent nearly 25,000 melodies: 20,042 in the major mode and 4,810 in the minor mode. Pinkerton (1956) also calculated a distribution of 39 nursery tunes. By conducting a product-moment correlation between her key profiles and each of these collections, Krumhansl demonstrated that her profiles were highly correlated with each of the pitch distributions in these samples.

A pivotal insight in the psychological research has been the recognition of the role of pitch hierarchy and statistical learning as the foundations of the perception of tonality. Ample experimental research suggests that listeners internalize the relative frequencies of different scale tones, and use the relative prevalence of the tones and tone-successions to infer the tonal context for some musical passage. Common distributions of various scale tones can be regarded as cognitive "schemas" that are used to interpret the world of pitch tones. Moreover, these schemas appear to emerge spontaneously in the auditory system, simply through a listener's exposure to the normative pitch environment. In short, there exists a sort of cultural-cognitive loop: the culture in which a listener is immersed shapes the pitch schemas of the listener, yet these same pitch schemas in turn shape how the listener apprehends the tonal structure of the sounds they hear.

This loop implies that a scale or tonal system is likely to be highly stable. Although the system is thought to be dominated by learning-through-exposure, this learned system feeds back to the music-making in a way that resists change. Nevertheless, despite this stability, there is theoretically room for what might be called schematic "drift." Such drift is evident in certain species of songbirds. Experiments have shown that the territorial songs for some birds are learned by adolescents through exposure to the songs of adults (Tchernichovski, et al 2001). Over time, and across geographical distances, these songs diverge slightly. A similar phenomenon has been chronicled by historical linguists with regard to speech (Hock 1991). Although pronunciation is nominally dictated by one's linguistic environment, over time and with geographical separation the boundaries of categorical perception drift and diverge – ultimately resulting in different dialects and languages.

The stability of the major and minor pitch schemas is evident in the pitch distributions for music written in these modes. The distributions of pitches for passages in the major mode are remarkably similar. Although slightly less consistent, pitch distributions for passages in the minor mode are similarly remarkably alike. Using these pitch distributions, Carol Krumhansl and Mark Schmuckler devised an algorithm that was able to successfully predict both the nominal and perceived key for passages written in the major or minor modes. Aarden (2003), Huron (2006), and Temperley and Marvin (2007) questioned the value of using data generated from perceptual experiments asking for "goodness of fit" as a standard for a key profile. Since common-practice music commonly modulates, Sapp (2005) devised a hierarchical key-finding algorithm that systematically analyzes successively larger sliding windows of the composition, and analyzes the key at each one of these levels. Over the past decade, modifications to the original Krumhansl-Schmuckler algorithm by Craig Sapp, David Temperley, Bret Aarden and others have improved these predictions to better than 95% accuracy – simply by tallying the frequencies of occurrence of different scale-degrees.

II. SOME HISTORICAL PERSPECTIVE

Past key-finding research has tended toward a synchronic perspective, focusing on a stable, unchanging representation of a major/minor system. However, a dominant perspective of historical musicologists is that the major/minor system evolved out of an earlier modal system. This modal system was originally conceived as an eight mode system in the early Medieval period. However, theorists throughout the Medieval period and into the Renaissance disagreed as to the number of modes in existence. For example, the 11th century theorist, Johannes Cotto, suggested that there were only four basic modes, and explicitly grouped together Phrygian and Hypomixolydian as being the same mode. Writing in about 1330, Jacques of Liége described how the soft hexachord (ranging from *f* to *d')* had become so commonly used as to suggest that the b^b might be considered an essential degree in the system, producing what became a new mode, known as the *cantus mollis* (Powers/Wiering 2001). In 1547, Heinrich Glarean published his famous work, the *Dodecachordon*, in which he extended the corpus of modes by 4, resulting in a total of 12 modes. However, even these 12 modes were in practice collapsed into 6 modes with those modes sharing a common final paired together. According to Glarean, each of these modes (really, pairs of modes) had its distinctive character due to the layout of its tones and semitones.

In a study of traditional Korean p'iri music, Unjung Nam (1998) analyzed the distributions of various scale tones in different works and showed that the results are consistent with a transposable tonal hierarchy. As in the case of Western music, the music exhibits stable scale-degree behavior, and these probabilistic properties can be used to infer the mode and tonic. Similarly, Kessler, Hansen and Shepard (1984) carried out a probe-tone study with Balinese participants living near the remote Gunung Agung volcano. They were able to demonstrate the existence of stable pitch hierarchies for pelog and slendro scales. The cross-cultural validity of the pitch-schema concept remains to be thoroughly investigated. Although not all cultures make use of stable pitch-related patterns, most do, and many musical cultures exhibit some sort of hierarchical organization with some pitches being more prominent than others.

Huron and Veltman (2006) applied this insight to a study of Gregorian chant. They assembled a sample of chants, and calculated the pitch distributions for each work. They found stable scale-degree distributions for the various modes. They also found that some modes, such as the Phrygian and Hypomixolydian, are statistically indistinguishable in terms of their scale-degree distributions. This latter finding is consistent with Johannes Cotto's (11th century) view that these two modes are identical. In addition, Huron and Veltman performed a cluster analysis on the scale-degree distributions of these modes and found that they loosely grouped together according to the quality of their third degree, suggesting a possible early bifurcation into proto-major and proto-minor modes.

If scales and modes evolve in a manner akin to pronunciation in historical linguistics, then it is appropriate to pursue a thoroughly diachronic study in which the distributions of scale tones are traced over a long span of music history. It is not unreasonable to suppose that the major and minor system evolved gradually over time out of the modal system. Moreover, we might well expect that the major and minor modes continued to evolve after they were established as the predominant scale system. How stable are the major and minor distributions over time? Can we pinpoint the historical moment when these distributions emerged? Is the minor scale truly one scale? Can we observe the remnants of earlier modes? In order to address such questions, we proposed to carry out a diachronic analysis spanning the period 1400 to 1750.

III. METHODOLOGICAL CONSIDERATIONS

The subject of mode is among the most venerable topics in historical music scholarship. At the same time, the subject of tonality has proven to be one of the most active areas of research in the field of music perception and cognition. In this study, we will endeavor to link together these disparate areas of scholarship and apply a cognitively inspired approach to the study of mode. More specifically, we propose to apply the principles of structural tonality to an analysis of the interrelationships among the modes of the medieval eight-fold system.

In key-finding algorithms like the Krumhansl-Schmuckler algorithm, the "ideal" distributions for the major and minor keys are pre-defined. That is, the pitch-class distribution for a musical passage is compared with pre-existing distributions for the major and minor modes. However, if one entertains the idea that the distributions for the major and minor modes have potentially changed over history, then an algorithm based on prior distributions is problematic. There is something of a chicken/egg problem in which the musical practice determines the distributions, while the distributions shape the musical practice.

Suppose that we have a musical work from the distant past and calculate the distribution of pitch classes for that work. In order to determine whether the work is in mode X or mode Y, we need pre-existing distributions for both modes. The pertinent distributions, however, would be the mode distributions as they existed at the time the musical work was created. But, to create pertinent distributions for that time period we must know how many different modes were current at the time, and how to transpose the various works so that works in the same mode would be recognized as such.

However one approaches this problem, some sort of "bootstrap" method is needed. One approach is to begin with modern times and work backwards. That is, although we have no knowledge, for example, of how many modes listeners distinguished in 1600, there is a fair amount of agreement that listeners primarily distinguish two modes today (major and minor). We might therefore apply currently existing "key profiles" to the music of an immediately preceding historical period as a way of estimating the key (mode and tonic) for the music in question. Having established the keys for each work in a sample of earlier music, we can then revise our "key profiles" (perhaps "mode profile" would be a more fitting term) for the music in question. Having established the mode for each work in a sample of earlier music, we can then revise our "mode profiles" to reflect the distributions of pitches for that period, and then apply these revised distributions to the next earlier period. In this way, one could theoretically work backwards through history, allowing whatever changes occurred to emerge.

A difficulty with this approach is that we can't be sure of the number of "modes" current at any given time. If the claims of Cotto are correct, then four modes were in use during the 11th century, while twelve modes were in use at the time Glareanus. If the claims of earlier theorists are correct, then around 900 A.D. there were eight modes in use. We cannot even be sure that two modes exist today.

Rather than essentializing these questions, our study will simply rely on a data-driven approach. An appropriate method is provided by cluster analysis. Accordingly, for each historical epoch, we may subject the pitch distributions for various works to a cluster analysis, and allow the music to suggest the number of "modes" evident at that time, based on the aggregate similarities of the zeroth-order pitch distributions.

In light of these considerations, this study will use the following method:

1) Musical works will be sampled over a series of epochs. In this study, we use 50-year epochs.
2) Beginning with the epoch with the most stable distributions of the current major/minor system, a mode-finding algorithm will be used to determine the tonic pitch for each work. The works will then be transposed to a common tonic.
3) With common tonics, the pitch distributions will be determined for each work in the epoch sample and these distributions will be subjected to cluster analysis (each of the twelve scale-degree proportions used as variables). The analysis will determine the number of modes conjectured to have existed during that epoch; for example, two clusters representing the major and minor modes might be expected to appear for music throughout the common practice period.
4) Scale-degree distributions will be determined for each cluster by amalgamating all of the works deemed to belong to their respective clusters. These distributions will then be applied to the immediately preceding epoch.
5) The process will then be repeated, moving backwards through history. Our method relies on the assumption that pitch schemas for any given epoch will be similar to the preceding 50-year epoch.

The idea of a single modern "minor mode" might seem contentious. Musicians will readily distinguish harmonic and melodic minor forms, and the melodic minor itself has variant ascending and descending forms. However, cognitively inspired research does not appear to support these distinctions as basic cognitive categories. Instead, the psychological research implies that there may be a single modern "minor" schema, whose principle defining feature is the lowered mediant, with a submediant that is mostly lowered – but whose precise spelling depends on local harmonic and melodic contexts.

In order to determine the best period to begin with, a pilot study was carried out to establish which epoch in the common practice period exhibits the most stable distributions of the presumed major/minor system. Three works were sampled from each fifty year period between 1650 and 1900. The Krumhansl-Kessler key profiles were used to determine the keys of the first and last ten measures of the work. The period 1700-1750 was the only period to correctly assign the key of each piece, suggesting that this period may be especially consistent with the major/minor system. Consequently, we chose to begin our study with the 1700-1750 epoch.

IV. SAMPLING METHOD AND DATA

Selecting the musical samples for this study would ideally aim to sample music that is representative of the cognitive pitch schemas present in various historical eras. One might suppose that the appropriate sampling method would seek equal numbers of musical works composed in each of the target periods. A more subtle approach would recognize that composers live for different lengths of time, and that compositions by elderly composers may reflect older schemas and fail to be representative of the time in which they were composed.

In general, mental schemas tend to be learned in a person's formative years. Although the style of composers may change over time, much of a composer's music-related schemas will have been formed earlier in life. One might assume that this is especially true for fundamental musical elements such as scale usage or the perception of mode. That is, one might expect that most stylistic changes throughout the life of the composer would involve elements of musical language that are more ephemeral than the perception of scale or mode. For the purposes of this study, we will consider a musician's cognitive development to be "mature" at 25 years of age.

Accordingly, compositions are grouped into epochs, not by the year of a work's composition, but by the year when the work's composer turned 25 years of age. For example, Bach was born in 1685. Therefore, for this study his music is deemed to be representative of musical schemas predominant around 1710. A work written by Bach in 1740 would then be coded as belonging to schemas around 1710. That is, for the purposes of this study, all of Bach's music would be regarded as representative of 1710.

The population of interest is all music of the Western tradition that would influence listeners' and composers' mental schemas of mode. Presumably, this would include popular and folk genres, in addition to art music. We expect that it is the totality of musical exposure that would shape a listener's scale-related schemas. Unfortunately, difficulties arise in finding suitable popular and folk sources over long periods of time. Scores of popular music are commonplace only after around 1850. Although folk music materials have a very long history, notated sources similarly tend to be commonplace only after around 1850. Moreover these sources tend to originate from transcriptions of modern singers. In the case of the Western art-music tradition, however, notated sources extend back hundreds of years with no gaps or interruptions. We acknowledge that limiting our study to art music sources fails to provide a representative sample of musical exposure for past listeners. Nevertheless, practical considerations limit our sample to art-music sources.

In orchestral scores, it is common for instruments to double each other. From a sampling perspective, doubling essentially reduces data independence. Doubling is, however, much less common in music employing lighter textures. Consequently, the sample for this study will be limited to those works containing four or fewer parts. Hence, the sample includes keyboard works, string quartets and other chamber works, motets and 1-, 2,- 3-, or 4-part choral works.

Much of our sample takes advantage of a convenience sample of scores available through the Petrucci Music Library (IMSLP.org). This source includes music by 6,343 composers (accurate as of September 19, 2011). In order to increase data independence, the sample was limited to one musical work by each sampled composer for the years between 1500 and 1750. Due to the limited availability of sources before 1500, a different sampling method (to be described shortly) was used for 1400-1500. The birth dates of the composers were determined using the New Grove Dictionary of Music and Musicians. Fifty composers were selected for each 50-year epoch spanning the period 1500-1750. For the years 1400-1450, ten composers were found, and 64 scores available in the database for these composers were used. For the years 1450-1500, 151 works from 6 composers were used; this sample was augmented by a convenience sample of 112 works by Josquin already encoded.

Especially after the year 1600 or so, it became common for works to modulate to other keys within the work. Modulating to other keys typically influences the scale-degree distribution of the work. Conversely, the beginning and end of most works tend to reinforce the tonic key. It is important that the portion of the music analyzed for this study reflect just one key, so a pilot study was carried out to determine what portion of each work to sample.

In this study, the Krumhansl-Kessler key profiles were used to estimate the key (tonic and mode) of various portions of musical works. A sample of 58 piano works by Scarlatti were chosen because 1) the key of each piece was known (reflected in the title) and 2) they were written by a composer who was 25 years of age within the period with which the study would begin (1700-1750). Four portions of each work were sampled and the key was estimated using the Krumhansl-Schumkler key-finding algorithm: the first ten measures, the last ten measures, the middle ten measures, and the first and last ten measures.

The Krumhansl-Schumkler algorithm makes use of Pearson product-moment correlation to characterize the similarity of distributions. On purely mathematical grounds, there is reason to suspect that Euclidean distance provides a better way to measure the similarity of two distributions. Consequently, we also tested a Euclidean-distance version of the algorithm. The Euclidean distance method measures the proportion of each pitch class used in the work (weighted by duration) and treats each pitch class as a dimension in Euclidean space. The distance between this 12-dimensional point and points representing the key profiles for each of the 12 major and 12 minor keys are then calculated, and the key closest in Euclidean space is taken as the predicted key.

The results of this study are reported in Table 1. Of the methods in question for this sample, the Euclidean distance measure for the first and last ten measures provided the most accurate method for predicting key based on the scale-degree distributions. Therefore, we decided to sample the first and last ten measures for each selected work, and to use the Euclidean distance method for determining key.

The first and last ten measures of each sampled work were encoded in Humdrum format (Huron 1995). Specifically, the encoded material included the composer's name, the composer's birth and death dates, the title of the work, the date of the composition when available, the notated pitch (including accidentals), the durations, rests, and barlines. The data for the cluster analysis were the 12 dimensions of proportions of each of the 12 chromatic scale-degrees, transformed so that the "tonic" of each piece was encoded as the first dimension.

Table 1. The results of a pilot study to determine the best sample for a mode-finding algorithm. The first, middle, and last ten measures of 58 Scarlatti piano sonatas were sampled and the resulting pitch-class distributions were compared to 24 ideal pitch-class distributions for each mode and tonic pair. Using a Euclidean distance measure on the distributions of the first and last ten measures was most accurate, providing 98.3% accuracy.

Portion of the work	Hits for correlation	Hits for Euclidean distance
First ten measures	51 of 58 (0.88)	51 of 58 (0.88)
Middle ten measures	1 of 58 (0.017)	8 of 58 (0.14)
Last ten measures	53 of 58 (0.91)	54 of 58 (0.93)
First and last ten ms.	55 of 58 (0.95)	57 of 58 (0.98)

V. PROCEDURE

Based on the results of the two pilot studies described above, a cluster analysis was carried out on these samples of pieces from each of seven 50-year epochs, beginning with the epoch 1700-1750. Scale-degree distributions were calculated for each piece by tallying the number of notes of each scale-degree and then weighting it by the durations of the notes. In other words, total duration of each scale degree was determined rather than simply counting the number of occurrences of each scale degree.

The value for each scale degree was represented as a proportion of the total duration of all scale degrees for the purpose of comparison between works. Each of these 12 values was treated as a dimension in 12-dimensional space.

For the first epoch, 1700-1750, the 12-element vectors for each work were compared to all 24 major and minor Krumhansl-Kessler key profiles. As mentioned, the key that was closest in 12-dimensional Euclidean space was taken to be the key of the piece. The scale-degree vectors were then rotated so that the tonic pitches would coincide. The estimated tonic was compared with the last note in the lowest voice as a rudimentary means of checking the mode-finding algorithm. The last bass note of many compositions is the tonic note. However, the tonic is not the last note of all works, as in cases where an early movement in a multi-movement work ends on a half cadence. Though this measurement is therefore problematic, it provides some way of gauging how well the algorithm may be performing.

Once the tonic was assigned, the distributions of each of the 12 scale degrees were plotted to determine if there were any bimodal distributions and which scale degrees were bimodally distributed. A cluster analysis was performed to determine if there were any clusters in the scale-degree distributions for these pieces. Ward's method of hierarchical agglomerative clustering, complete linkage clustering, and centroid clustering were used, and were performed on the data set of all 50 pieces to determine whether there was convergence between methods. The number of clusters was determined by a visual inspection of the resulting dendrograms.

Once the number of clusters was established, a k-means cluster analysis was performed three times, with different starting locations randomly chosen for the k mean centroids. The result exhibiting the lowest standard squared error was chosen as the best solution. This clustering solution was then plotted using multidimensional scaling in order to provide a visual display of patterns in the clustering. Specifically, the chosen tonic pitch and mode were included in the analysis to determine if the clusters aligned with preconceived notions of major and minor modes. Averages were then calculated for each of the 12 scale degrees for each mode found. These distributions were then used as the "ideal" distributions for the epoch spanning 1650-1700; the process was repeated in this way for each 50-year epoch backward through time until the period 1400-1450.

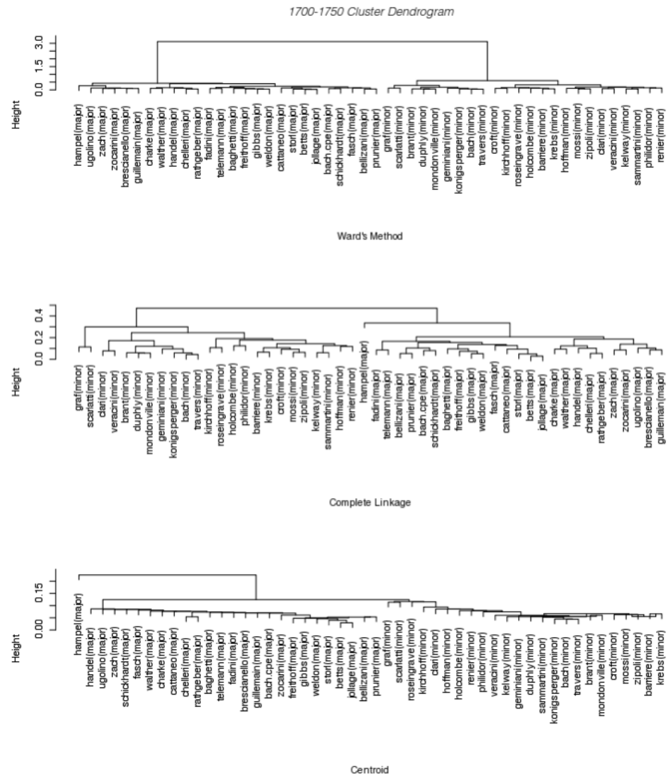
VI. RESULTS

A. 1700-1750

For the first epoch, 1700-1750, the Krumhansl-Schmuckler key-finding algorithm was applied to the sample of 50 works. The tonic was estimated for each piece and the scale-degree distributions were calculated. Of the 50 works in this period, the final bass note matched with the predicted tonic in 42 cases. Of the eight cases that did not match, the last notes of seven works ended on the dominant of the predicted tonic, consistent with the piece ending on a half cadence. The remaining piece ended on the mediant – the relative major of the minor predicted key. These results are consistent with the key-finding algorithm making an incorrect key assignment with closely related keys.

The data were then grouped using formal hierarchic cluster analysis. Ward's method, complete linkage clutering, and centroid clustering were used. The results of these methods are shown in Figure 1. The leaves of the dendrogram represent individual pieces, labeled with the composer's name and mode (as determined by the key-finding algorithm).

Figure 1. Solutions from three different cluster analysis methods for works from the 1700-1750 epoch. Regardless of the method, two clear clusters emerge from the data, a major-mode cluster and a minor-mode cluster.



The results of the three clustering methods largely agree; all three methods identify two clear clusters in the data. Moreover, single linkage clustering, group average clustering, and median clustering (not displayed here) all similarly converged on the two-cluster solution.

The presence of two clusters can be clearly seen if the data are subjected to multidimensional scaling, as illustrated in Figure 2. The graph on the right shows the eigenvalues for the first 6 dimensions. A clear "elbow" can be seen at 2 dimensions, with a stress of 0.16. Accordingly, the data were plotted in two dimensions, shown in the left graph. There are two clusters evident, with the minor cluster appearing on the right and the major cluster on the left. Notice the greater variance in the minor cluster.

With two distinct clusters, scale-degree vectors for each mode were then recalculated. These revised key profiles are displayed in Figure 3 along with the original Krumhansl-Kessler distributions. Notice that the graphs are very highly correlated, but that there are some noteworthy differences. Specifically, the differences between diatonic

Figure 2. Left figure: Multi-dimensional scaling solution for the 1700-1750 epoch. Two clusters are evident, corresponding to the major (left) and minor (right) modes. Right figure: Eigenvalues for the MDS solution. An elbow at two suggests that two dimensions are sufficient to display the clusters.

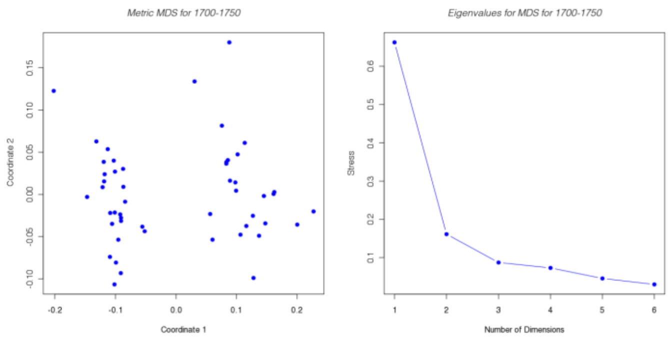
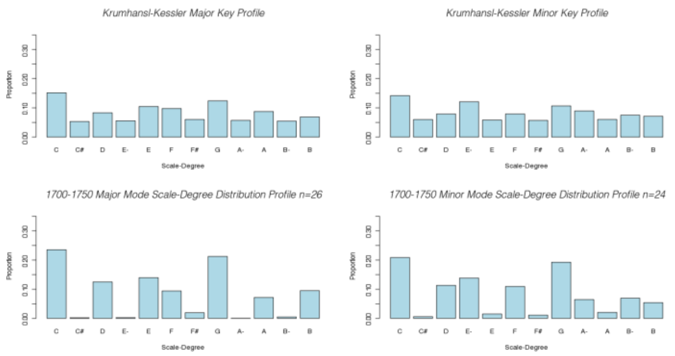


Figure 3. Comparison between Krumhansl-Kessler key profiles (higher figures) and the revised key profiles arising from the first cluster analysis for 1700-1750 (lower figures). The major mode profiles shown are on the left; minor mode profiles are shown on the right. Notice that the distinction between diatonic tones and chromatic tones is more pronounced in the new distributions.



tones and chromatic tones are more pronounced. In addition, the dominant pitch is more pronounced in the sample than in Krumhansl's experimentally-derived values. In the minor mode, the relative prevalence of the mediant and dominant have been reversed in the sample data.

B. Summary of the Remaining Epochs

Using the two revised key profiles, shown at the bottom of Figure 3, tonics were estimated for the 50 works in the preceding 1650-1700 epoch. The 50 works were subjected to the same cluster analysis, and again two clusters emerged from the cluster solution. The works within this cluster were amalgamated and the resulting distributions were used to find the tonics for the 1600-1650 epoch, and so on.

Unfortunately, space does not permit fully displaying the figures from the remaining epochs. However, a few observations are noteworthy. In the 1550-1600 epoch, the cluster consisting of minor mode works appeared to break into two sub-clusters, shown in Figure 4. In order to explore further the possible presence of a third cluster, a k-means analysis was performed with k = 3. The results are displayed in Figure 5, where factor analysis has been used to visually enhance the distinctiveness of three clusters. One distinct "major" cluster was present and two "minor" clusters were present. The scale-degree distributions for these three clusters are displayed in Figure 6. The "major mode" distribution continued to closely resemble the Krumhansl-Kessler major-mode profile (r = +0.98). The two "minor mode" distributions, however, were noticeably different from the Krumhansl-Kessler minor-mode profile (minor 1, r = +0.93; minor 2, r = +0.83): the most obvious difference is the prevalence of the minor mediant scale degree, which is much more prevalent in the "minor 1" than in "minor 2." The submediant was also noticeably different. Whereas "minor 1" had a prevalence of the minor submediant and almost no presence of the major submediant, "minor 2" had more major submediant than minor. The leading tone occurred equally infrequently in both modes, but the subtonic was much more frequent in "minor 1." These considerations are consistent with "minor 1" outlining the scale degrees of a nominal Aeolian mode while "minor 2" outlines a nominal Dorian mode.

The three distributions in Figure 6 were used to estimate the tonics of the 50 works for the 1500-1550 epoch, resulting again in three modes, roughly outlining the same mode profiles. This trend continued until the final epoch, 1400-1450, in which the major mode cluster also exhibited sub-clustering.

Figure 4. Cluster analysis of the works from 1550-1600. The minor-mode cluster on the right exhibits some sub-clustering.

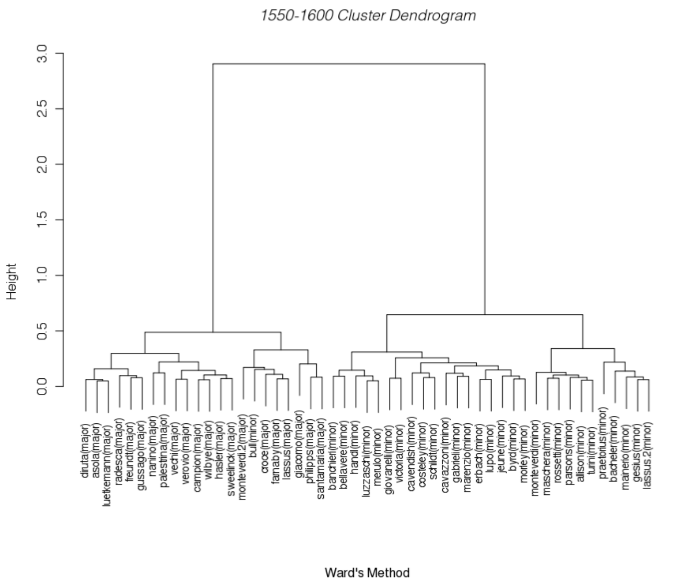


Figure 5. Factor analysis cluster plot for the 1550-1600 epoch. Three clusters are highlighted, drawn by the clustering algorithm.

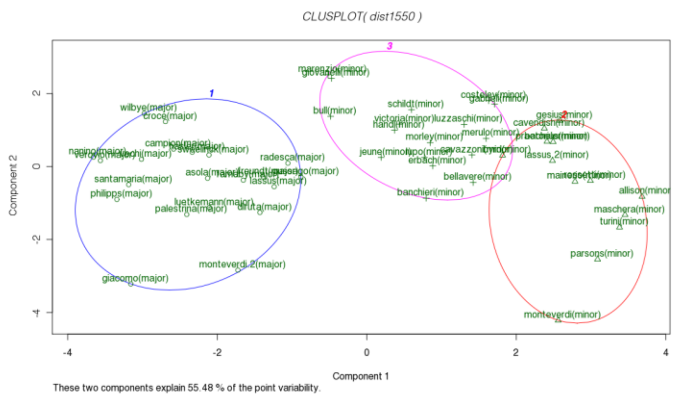
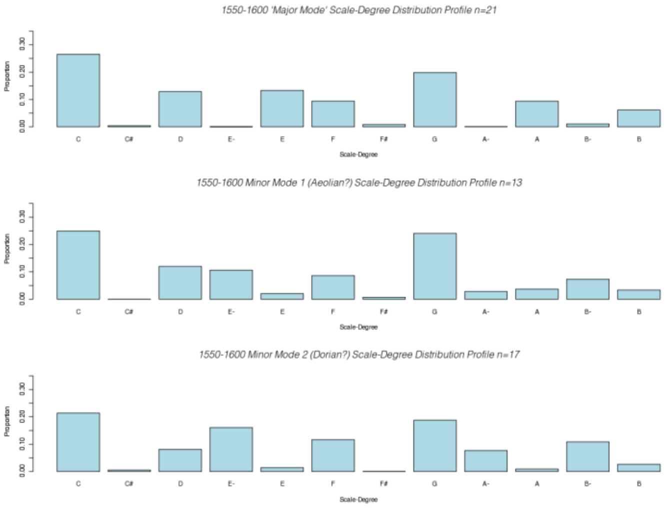


Figure 6. Scale-degree distributions for three modes in the 1550-1600 epoch. The minor scale exhibits two variants. The center distribution is suggestive of a nominal Aeolian mode, with more A than Ab and more Bb than B, whereas the bottom distribution is suggestive of a nominal Dorian mode, with more Ab and Bb than their natural counterparts.



A full-page time-line showing the scale-degree distribution clusters by epoch is shown in Figure 7. *R* values indicate Perason's correlation coefficient – a measure of mathematical similarity, where 1.0 represents a perfect match. Dashed lines show correlations between initial and final epochs. Minor 1 bears similarity to a nominal "Aeolian" mode, whereas minor 2 bears similarity to a nominal "Dorian" mode.

VII. CONCLUSION

The emergence of the major/minor system is one of the seminal historical events in Western music. The major/minor system likely arose through a complicated process involving historical, stylistic, perceptual, cognitive, and other factors. One should not be sanguine about the complexity of this historical development.

The method used in this study might be dismissed by some scholars as ahistorical and mechanical. We do not presume that the method used in this study is the only way to address the history of the major/minor system. Instead, we claim that any comprehensive story about the emergence of the major/minor system will need to draw on converging evidence from several independent sources employing contrasting methods. Psychological research on the phenomenon of tonality will surely be part of this story.

The method used here rests on foundations that, while not universally accepted in the field of music psychology, nevertheless represents the prevailing majority view concerning the role of statistical learning in the perception of tonal schemas such as evident in Western tonality. As noted earlier, the seminal work of Carol Krumhansl (1990) implicated mere exposure to some stable pitch distribution as the bootstrap process for perceiving the pitch relations referred to as key. Over the past several decades, a wealth of experimental research has reinforced the importance of mere exposure to a pitch environment as critical to tonality perception (e.g., Aarden, 2003; Huron, 2006; Krumhansl & Kessler, 1982; Nam, 1998; Saffran, et al., 1996, 1999). This research suggests that, in order to understand how pitches are perceived in some culture, one needs to attend to the frequency of occurrence of the different pitches.

In our study, one might treat each 50-year epoch as though it were a separate musical "culture." However, pitch schemas are both shaped by the music-making, and the music-making is also shaped by the prevailing pitch schema. Accordingly, we have used one historical period as a bootstrap for helping to understand a neighboring period. Since we have little idea about the prevailing pitch schemas in the 15th century, our method began with modern perceptual data and worked backwards.

VIII. CAVEATS

In any research project of this size, a large number of assumptions are made, and it is appropriate to assemble together and make clear all of the known assumptions and caveats.

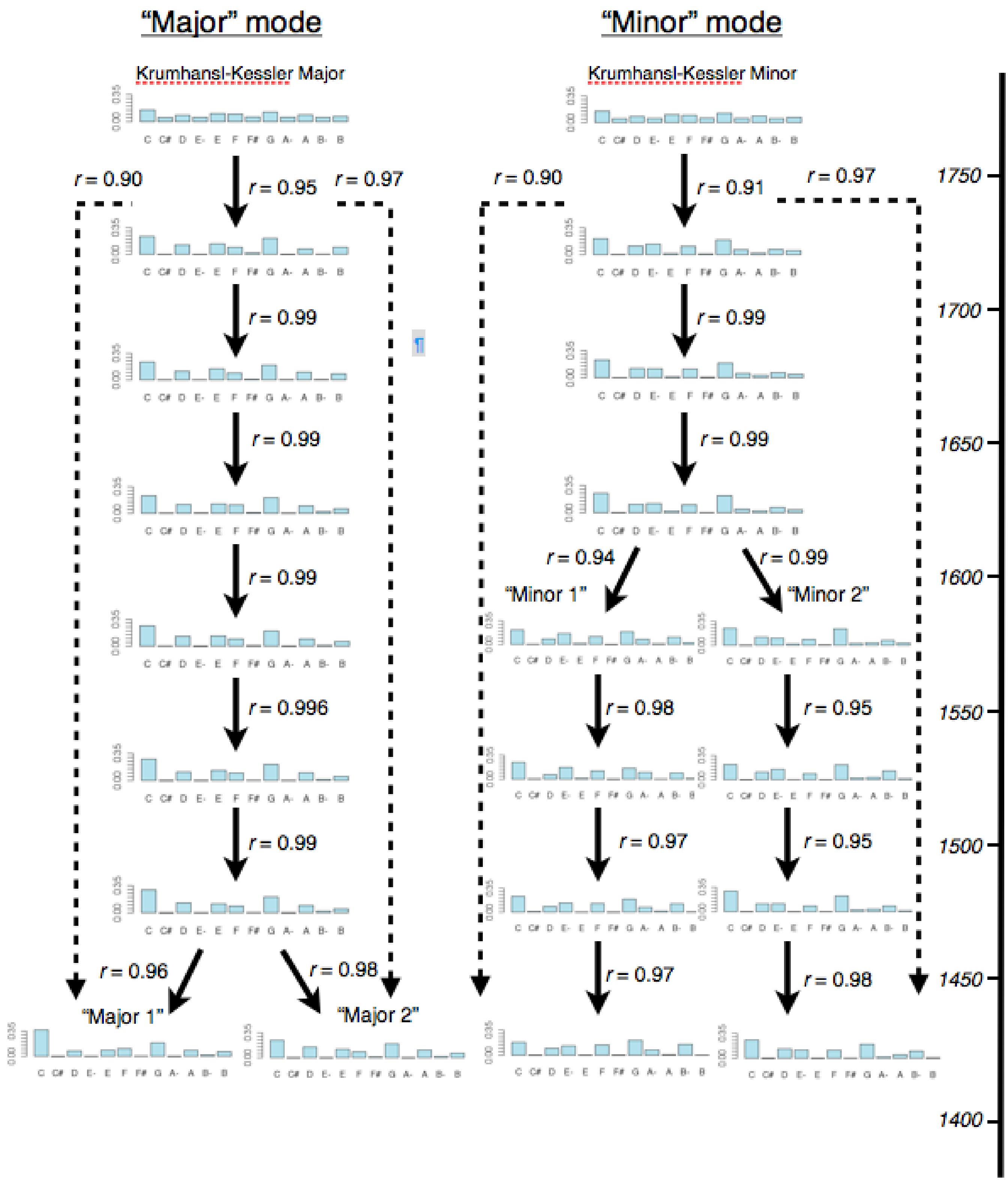
In carving up the "epochs," we chose artificial boundaries. We assumed that 50-year epochs might provide sufficient resolution for deciphering historical changes in tonality. We assumed that our selected musical samples are truly representative of their respective periods. In addition, we assumed that there is little change in how composers use mode over the course of their careers, and that their mode-related practices may be viewed as "stable" around the age of 25 years.

Most importantly, we have assumed that the research on statistical learning is indeed applicable to the case of tonality perception, and that the same cognitive principles observed today apply to earlier eras. This assumption is concretely expressed in our reliance on the Krumhansl-Schmuckler method of key-finding. We also assumed that the beginnings and endings musical passages provide the best key-defining material. We began with the assumption that, for modern listeners, there are only two pitch-related schemas: major and minor. For example, we assumed that modern listeners do not hold more than one pitch schema for the minor mode.

In the case of the clustering method, we acknowledge that clustering analysis provides no clear criterion for determining the number of salient clusters, and that what constitutes a salient cluster is a matter of researcher interpretation.

Apart from these assumptions, a major issue is the problem of tonic finding. Since the method of mode determination relies on the distributions of scale tones, failing to accurately identify the tonic will add noise to the amalgamated mode distributions. Moreover, due to the iterative nature of the method, this noise can be amplified over time. For epochs that conform well to the major-minor system, the key-finding algorithm appears to be very effective. (Recall, for example, that the algorithm correctly identified the keys for 57 of 58 Scarlatti sonatas in which the key information was explicitly given by the composer in the titles.) However, if a mode doesn't occur very often in a musical sample, then it is more likely to be "shoe-horned" into some already existing mode distribution. That is, the procedure gains robustness at the expense of ignoring modes that are represented in the sample by only a small number of works. Specifically, the mode-finding algorithm tends to transpose away less common modes by rotating them out of existence. Mis-classified modes then influence the distributions in which they are included so that these new modes are more difficult to

Figure 7. The full results from this study, shown as a timeline. *R* values indicate Perason's correlation coefficient – a measure of mathematical similarity, where 1.0 represents a perfect match. Dashed lines show correlations between initial and final epochs. Minor 1 bears similarity to a nominal "Aeolian" mode, whereas minor 2 bears similarity to a nominal "Dorian" mode.



find when they are more prevalent. This is especially problematic for tetrachordally-similar modes, such as Ionian/Lydian, Aeolian/Phrygian, and Dorian/Aeolian.

IX. DISCUSSION

These caveats notwithstanding, the method used in this study offers a number of tentative conclusions. The main results of this study are summarized graphically in Figure 7. As can be seen, the scale degree distributions are relatively stable over time. For example, the major mode distribution for the epoch 1650-1700 is statistically identical to the major mode distribution for 1700-1750 (with a correlation of +0.99). At the same time, we see evidence of slowly changing distributions over time. For example, the major 1 distribution from 1400-1450 correlates well (but imperfectly) with the major distribution from 1700-1750 (with a correlation of +0.90). In short, the method traces some gradual evolution in these scale-degree distributions, suggesting a slowly changing musical practice, and also implying parallel changes in listening-related tonal schemas.

Perhaps the most salient result from the study is the suggestion that the modern minor schema arose from an amalgamation of Dorian-like and Aeolian-like modes. This amalgamation is hardly surprising to historians, but it is reassuring that a data-driven empirical method converges with this common view. The emergence and disappearance of perceptual categories is well known in cognitive psychology. Perhaps the best documented examples are found in the field of historical linguistics. Linguists have shown that it is commonplace for two previously distinguished phonemes to lose their distinctiveness in some language and, over time, to join to form a single phoneme category. For example, all English speakers used to distinguish between the "ah" sounds found in the words "caught" and "cotton." This distinction is clearly retained by speakers in New York and New England, but many speakers of American English no longer distinguish between the two pronunciations. Linguists call this process "merger" (e.g., Katamba, 1989). Perceptual merger is most likely to happen when the two categories are especially similar, and so easily confused. It may be that the similarity between pre-existing Dorian-like and Aeolian-like modes ultimately failed to maintain their distinctiveness and so merged into a single schema (see also Huron & Veltman, 2006).

Another result from our study is the observation that the leading tone increasingly displaces the subtonic over the history of the minor mode. This change is reminiscent of a similar historical transformation documented by Nazir Jairazbhoy regarding rāg Tori in Indian music. In his classic book on Indian rāgs, Jairazbhoy traced the history of rāg Tori as it evolved from a Phrygian-like scale to its modern form (1971: 98-99). The rāg was transformed in four steps, with each scale modification providing addition opportunities for semitone movement to a structural tone. In effect, Tori evolved to increase the expressive capacity of the rāg for what might be dubbed “yearning semitone” relations. It is possible that a similar phenomenon occurred in the Western minor scale, although the speed of this historical transformation is considerably slower than occurred for rāg Tori.

In general, we can observe fewer non-scale tones in earlier epochs. For example, in 1500-1550 the proportion of scale tones in the major mode is 97.7%; whereas, in 1700-1750 the comparable proportion for major is 97.2%. This is consistent with both informal observations and formal studies, such as the study of the history of chromaticism by Perttu (2007). In addition, we see a greater proportion of the use of the tonic in earlier epochs. For example, the tonic pitch appears 27.7% of the time in the major mode in 1500-1550, but just 23.5% in 1700-1750. As in the case of the minor mode, the leading tone increases in usage for the major mode, from 5.3% in 1500-1550 to 9.4% in 1700-1750. As speculated above, this might possibly be symptomatic of an attraction toward increased “yearning semitone” relationships.

With regard to future studies, the method employed here could be extended in a variety of ways. A larger sample of music would likely result in greater accuracy and increased validity. The sample might be extended further back in time, and could also be projected forward beyond 1750 into more recent times. Apart from tracing changes associated with atonality and various forms of extended tonality, the same method could be applied to changing scale forms used in popular music and jazz.

In order to achieve greater reliability in tonic-finding, one might employ an iterative approach: going backward in time, then using the resulting mode distributions to go forward in time – and repeating this back-and-forth approach until some stable scenario emerges. At the risk of introducing anachronistic preconceptions, another approach might rely on the judgments of modern theorists to assign a tonic to each extracted excerpt. One might also expand the study to examine 2- or 3-note transitional probabilities, rather than relying exclusively on zeroth-order distributions. With a larger sample of musical works, one might be able to compare trends in different countries or cultural regions, or compare possible differences among individual composers. With a sufficiently large musical sample, one might even be able to identify the geographical locus for the origin of the major/minor system, or to even trace its geographical spread over time. With a sufficient volume of sampled musical materials, one might also examine shorter epochs (e.g. 20-year periods), and these might provide greater temporal resolution chronicling the historical changes.

REFERENCES

Apel, Willi. *Gregorian Chant.* Bloomington: IndianaUniversity Press, 1958.

Atkinson, Charles M. “The Parapteres: Nothi or Not?” *Musical Quarterly*, Vol. 68, (1982): 32-59.

Atkinson, Charles M. “On the interpretation of ‘modi, quos abusive tonos dicimus.’” In: *Hermeneutics and Medieval Culture.* Albany: State University of New York, (1989): 147-161.

Atkinson, Charles M. “Modus.” In *Handwörterbuch der Musikalischen Terminologie.* Wiesbaden: F. Steiner, 1995.

Auda, Antoine. *Les modes et les tons de la musique et specialement de la musique medievale.* Hilde- sheim: Olms, 1979.

Butler, David. “Tonal bootstrapping: Re-thinking the intervallic rivalry model.” In *Music, Mind, and Science,* edited by Suk Won Yi, 7-12. Seoul: Seoul National University Press, 1999.

Butler, David and Brown, Helen. "Diatonic trichords as minimal tonal cue-cells." *In Theory Only,* Vol. 5 (1981): 39-55.

Cohen, A. J. "Tonality and perception: Musical scales primed by excerpts from *The Well-Tempered Clavier* of J. S. Bach." *Psychological Research/Psychologische Forschung* 53 (4) (1991): 305-314.

Cuddy, Lola. "Tonal relations." In I. *Perception and Cognition of Music,* edited by Irene Deliège and Juslin Sloboda, 329-352. Hove, East Sussex: Psychology Press, 1997.

Curtis, Lola. "Mode." In *Companion to Medieval and Renaissance Music*, 255-264. London: Dent, 1992.

Everitt, Brian S. *Cluster Analysis*. New York, NY: Halsted Press, 1993.

Hock, Hans Heinrich. *Principles of Historical Linguistics.* New York, NY: Walter de Gruyter & Co, 1991.

Huron, David. *The Humdrum Toolkit: Reference Manual*. Menlo Park, CA: Center for Computer Assisted Research in the Humanities, 1995.

Huron, David, and Joshua Veltman. "A Cognitive Approach to Medieval Mode: Evidence for an Historical Antecedent to the Major/Minor System." *Empirical Musicology Review,* **1**, No. 1 (2006): 33-55.

Hutchinson, William, and Knopoff, Leon. "The significance of the acoustic component of consonance in Western music." *Journal of Musicological Research*, 3 (1979): 5-22.

Jairazbhoy, Nazir A. *Rāgs of North Indian Music: Their Structure and Evolution*. London: Faber and Faber, 1971.

Katamba, Francis. *An Introduction to Phonology.* London: Longman, 1989.

Kessler, Edward J., Christa Hansen and Roger Shepard. "Tonal schemata in the perception of music in Bali and the West." *Music Perception*, 2 (1984): 131-165.

Krumhansl, Carol. *Cognitive Foundations of Musical Pitch*. New York, NY: Oxford UP, 1990.

Krumhansl, Carol. & Kessler, Edward J. "Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys." *Psychological Review,* Vol. 89 (1982): 334-368.

Kruskal, Joseph B., and Myron Wish. *Multidimensional Scaling (Quantitative Applications in the Social Sciences)*. Beverley Hills, CA: Sage Publications, 1981.

Lerdahl, Fred. *Tonal Pitch Space*. New York: Oxford University Press, 2001.

Nam, Unjun. "Pitch distribution in Korean court music: evidence consistent with tonal hierarchies." *Music Perception*. Vol. 16, No. 2 (1998): 243-247.

Perttu, Daniel. "A quantitative study of chromaticism: Changes observed in historical eras and individual composers." *Empirical Musicology Review*. Vol. 2, No. 2 (2007): 47-54.

Pinkerton, Richard C. "Information theory and melody." *Scientific American*, **194** (1956): 77-86.

Powers, Harold S. / Wiering, Frans. "Mode: II. Medieval modal theory." In *The New Grove Dictionary of Music and Musicians*, 2nd ed, edited by Stanley Sadie New York: Grove, 2001.

Sapp, Craig. "Visual Hierarchical Key Analysis." *Computers in Entertainment*, **3**, No. 4 (Oct 2005).

Schlager, Karlheinz. "Regionaltradition und Modalität in der liturgischen Monodie des Mittelalters." *Studia musicologica Academiae Scientarum Hungaricae*, Hungary, Vol. 27, Nos. 1-4, (1985): 117-122.

Temperley, David and Marvin, Elizabeth W. "Pitch-Class Distribution and the Identification of Key." *Music Perception: An Interdisciplinary Journal*, **25**, No. 3 (2007): 193-212.

Tchernichovski, Ofer, Mitra, Partha P., Lints, Thierry, Nottebohm, Fernando. "Dynamics of the vocal imitation process: how a zebra finch learns its song." *Science,* 291 (5513) (2001): *2564-2569.*

Youngblood, Joseph E. (1958). "Style as information." *Journal of Music Theory*, **2** (1958): 24-35.