

EXPLORING AFRICAN TONE SCALES

Dirk Moelants

Ghent University

Dirk.Moelants@UGent.be

Olmo Cornelis

University College Ghent

Olmo.Cornelis@hogent.be

Marc Leman

Ghent University

Marc.Leman@UGent.be

ABSTRACT

Key-finding is a central topic in Western music analysis and development of MIR tools. However, most approaches rely on the Western 12-tone scale, which is not universally used. African music does not follow a fixed tone scale. In order to classify and study African tone scales, we developed a system in which the pitch is first analyzed on a continuous scale. Peak analysis is then applied on these data to extract the actual scale used. This system has been applied to a selection of African music, it allows us to look for similarities using cross-correlation. Thus it provides an interesting tool for query-by-example and database management in collections of ethnic music which can not be simply classified according to keys. Next to this the data can be used for ethnomusicological research. The study of the intervals used in this collection, e.g., gives us evidence for Western influence, with recent recordings having a tendency to use more regular intervals.

1. INTRODUCTION

Scale recognition has a long tradition in the analysis of Western music. Already in medieval music theory, determining the mode and classifying pieces according to their mode was a central topic. Also in the music theories of the Middle-East and India, classification of music according to the scale (often connected to a certain ‘mood’) is an important topic. Not surprisingly, with the advent of computational methods, researchers started to design systems to perform the process of scale recognition automatically [1]. In recent years the focus has shifted from symbolic approaches, based on MIDI or score representations, to the analysis of musical audio files (e.g. [2 - 6]). Various systems have reached a reasonable level of success in labeling music according to the keys of the Western tonal system (cf. MIREX 2005 [7]).

Automatic analysis and classification of scales in music that is not organized according to the Western tonal system is much less developed. Some efforts that have been done to extract the scales of e.g. Australian aboriginal didjeridu music [8] or Indian classical music [9] use a reduction to Western pitch classes, thus avoiding the problems raised by irregular temperaments. Although this approach can be efficient to a certain extent, it seems limited to music with a pitch organisation that has a certain resemblance to the Western system and is problematic in terms of culture specific information. In some music the pitch-

set used is as such not very important, but rather the musical gestures associated with playing or moving from one tone to another are the most characteristic aspects. This has been used in the study of Chinese guqin music [10] and Carnatic (South-Indian classical) music [11], using prototypical gestural patterns or melodic atoms to describe the melodic content of music in which the pitch is seldom stable.

Some work has been done on the scale analysis of music of the Middle-East, more precisely on Turkish [12] and Persian [13] modes. This music is characterized by the occurrence of intervals based on (roughly) a quarter tone scale. Therefore, an analysis based on a chromatic (half-tone) division of the octave can not be used. Therefore the pitch is analysed on a more continuous scale, then transformed to pitch histograms, which can be attributed to schemata that represent the modes used in this specific repertoire.

Pitch organisation in the music of Sub-Saharan Africa does not rely on a fixed theoretical framework. Ethnomusicological research has shown that a large variety of scales is in use. Often these scales use intervals that do not conform to the European chromatic scale, e.g. the use of intervals around 240 cents in (roughly) equidistant pentatonic scales [14]. However, standardized tuning systems or culture-specific classification systems do not exist. In this paper we will propose a system to explore African scales with applications in Music Information Retrieval and ethnomusicology. First we will present the collection on which the scale-detection system is applied, as well as the test-set which will be analyzed in detail. In the next chapter we give a brief description of the pitch detection and peak extraction systems used to analyze the music and how the output can be coupled with the metadata associated with the original sound files.

2. BACKGROUND

The audio set which has been used in this research, is a selection from the audio archive of the RMCA (Royal Museum for Central Africa) in Belgium. It is one of the largest collections worldwide of music from Central Africa. The audio collection consists of about 50,000 sound recordings (with a total of 3,000 hours of music), dating from the early 20th century up to now. Aiming for durable conservation and enhanced accessibility, the audio archive has been digitized entirely. Not only the audio but also accompanying metadata and contextual information have

been digitized. A database and website were developed containing complete descriptions and fragments of the audio. The results of this project can be accessed on the website <http://music.africamuseum.be>.

For this study a selection of 901 audio files was used. In order to be sure to get a selection that consists only of music that uses a relatively fixed tone scale (and not e.g. music for percussion ensemble), we extracted music using four common types of musical instruments: musical bow ($N = 132$), zither ($N = 134$), flute ($N = 385$) and lamellophone (thumb piano) ($N = 250$). The selection was limited to music described as solo performances, mostly they contain only the sound of the instrument, in some cases the performer also sings, accompanying him/herself on the instrument.

3. ANALYSIS

3.1 Pitch detection

The pitch algorithm used in this paper has originally been designed to perform automated transcription of sung audio into a sequence of pitch classes and their duration [15]. Original goal of this tool was the development of a query-by-humming system for retrieving pieces of music from a digitized musical library. In this original system, the acoustic signal is turned into a parametric representation of the time-frequency information. A note is assigned to the segment by identifying the highest peak in the histogram of the frame-level pitch frequencies found in the segment, and by computing the average of the pitches lying in that bin. The pitch is then converted to a MIDI note rounding the computed annotation to the closest note frequency.

For the application of pitch recognition to the study of African scales, some important adaptations had to be made. First, the time segmentation, necessary to create melodies, was not of importance for building the pitch scales and was left out. Second, the quantization of the annotations into MIDI notes was unwanted, as we want to describe music that does not necessarily follows the equally-tempered scale. Therefore, the actual frequencies were used as pitch annotations. The output of this pitch algorithm consists of a list in which every line represents 10 ms, listing six potential frequencies, each with a probability. This allows extension to polyphonic textures.

In this case however, we choose to work with largely monophonic music. Therefore only the pitch with the highest probability was retained for every 10 ms, at least if this probability was higher than a minimal threshold (in this case 0.5). Then the frequencies were transformed to a

cents scale, setting C0 to zero (cents). For comparing histograms, all listed values were reduced to one octave, generating a chromavector of 1200 values, representing the scale of the piece. A typical example of the graphical representation generated by the pitch detection system is shown in Figure 1.

3.2 Peak Extraction

The pitch analysis gives us a precise representation of the pitch content used in every piece. To extract the scale, a peak analysis was performed on the histograms. As a first step, the 1200 integers are ranked by their value. Starting from the highest value, peaks are assigned. A peak is accepted if it meets all parameters. Parameters for the selection are width of the area around the peak in which another peak cannot be assigned, the size (volume) of the peak and height relative to average height. Parameters were manually optimized for this data set by trial and error. As the histograms show a large variance (small and wide peaks, high and low peaks, high and low noise levels), a mean of the best individual settings was chosen as final parameter settings (see Table 1). The analysis gives us the number of peaks, average height and a precise description of the location and size of the individual peaks (Table 2).

Parameter	Definition	Value
Places	number of lines in the input file	1200
Peakradius	width of the peaks	30
Overlap	tolerated overlap between peaks	0.25
Accept	maximal proportion: volume peak without overlap/volume peak with overlap	25
Volfact	minimal volume of a peak: volfact*(average height of histogram)*(1+2*peakradius)	1
Heightfact	minimal height of a peak: heightfact*(average height of the histogram)	1

Table 1. Parameters used in the peak detection, with the settings used in the current analysis in the Reith column.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

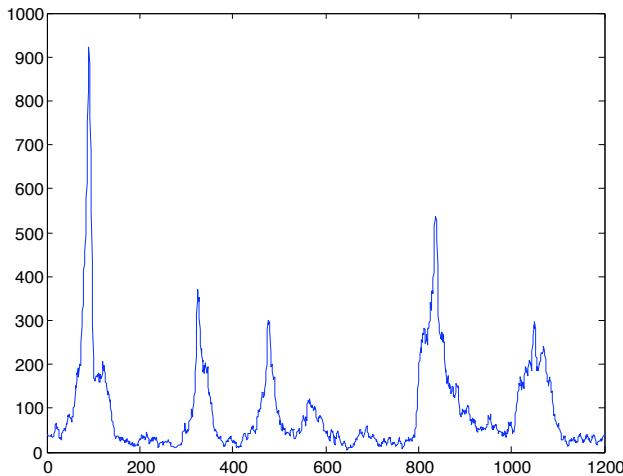


Figure 1. Example of the graphical output of the pitch analysis.

Peaks (cents)	volume	height	left side	% peak height	right side	% peak height
91	20441	922	61	0,15	121	0,21
837	17953	538	807	0,47	867	0,29
325	9603	371	295	0,13	355	0,29
476	8305	301	446	0,17	506	0,16
1050	12313	296	1020	0,57	1080	0,52

Table 2. Example of the output of the peak analysis for the piece shown in Figure 1, showing the pitches, together with information on the size of the peak.

3.3 Metadata

All the meta-data that were originally associated with the collection were digitized. Thus we get a large number of data fields from different categories: identification (number/id, original carrier, reproduction rights, collector, date of recording, duration), geographic information (country, province, region, people, language), and musical content (function, participants, instrumentation). Unfortunately, not for every recording all fields are available and often these data cannot be traced, as a large part of the collection is made up of unique historical sources.

The results of the pitch and scale analysis can be coupled with existing meta-data such as instrumentation, geographical information or date of recording. This can give us valuable information on the use of certain scales, such as geographical spread or evolution through time. As the current selection of pieces is relatively small, we used broad categories for the geographical origin (West-Africa, Southern Africa,...) and the recording time (before 1960, between 1960 and 1975, after 1975). An example of such a coupling is given in Figure 2. It gives the amount of peaks for each piece for each of the three time periods. This shows that in recent recordings, hexatonic and heptatonic scale become relatively more important while the importance of pentatonic and tetratonic scales diminishes.

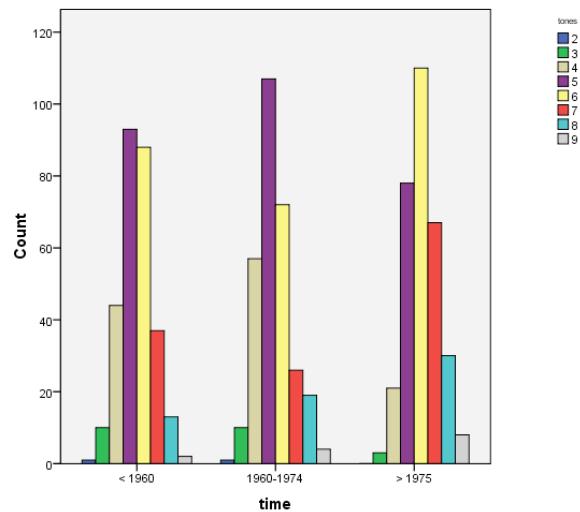


Figure 2. Bar chart representing the amount of peaks (2-9), for each the three categories of recording time: before 1960 ($n = 288$), between 1960 and 1975 ($n = 296$) and after 1975 ($n = 317$).

4. APPLICATIONS

The analyses made can be applied in different areas. First we will show the application of the pitch detection for data-mining applications, using cross-correlation of the pitch profiles to look for similarities. Next we will show an application of the techniques for ethnomusicological research, studying the intervals used in African scales.

4.1 Correlation analysis

The chromavectors given by the pitch analysis can be cross-correlated with each other in order to search for similar scales. As African music doesn't have a standardized tuning pitch, we need to allow a shift of pitch. Therefore, the cross-correlation technique uses every permutation of the original chromavector and returns the highest correlation from a list of 1200 correlations together with amount of cents it had to be shifted (Figure 4). Thus this method can be used for a query-by-example in which the output is a list of pieces with similar scales. This application allows to retrieve a song from a database without knowing any concrete fact about it, which is an important element for the usability of a search engine in a database of largely unknown music.

Next to this, the method can be useful for database management. The technique allows to check whether some songs are already present in their archive (so called double listed items), looking for perfect correlations without pitch shift. It can also help to establish groups of pieces with a similar origin, detecting possible links between recordings from different origins (cf. Figure 3). This could eventually lead to determination of missing meta-data.

Although the results of this analysis are promising, still some optimizations have to be done. Thus e.g. noisy pitch profiles with broad peaks, indicating less stable pitches

(e.g. from singing) (cf. Figure 3) are more likely to generate high correlations compared to pieces with very clearly defined pitches. Similarly the larger the number of peaks the more difficult it gets to obtain high correlations. Some mechanisms to deal with these differences should still be developed.

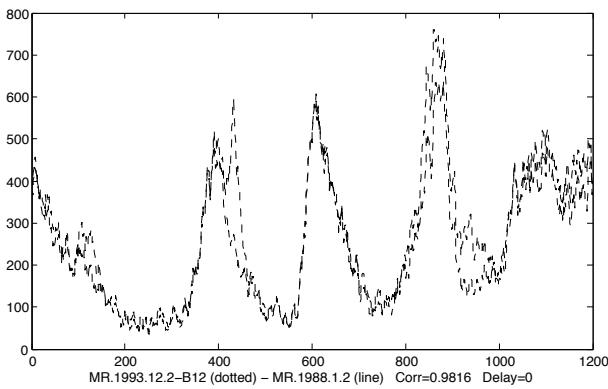


Figure 3. Graphical representation of a query-by-example, in this case the correlation is very high ($r = .98$) and no shift in pitch is necessary, which could point to a similar origin, despite the different sources.

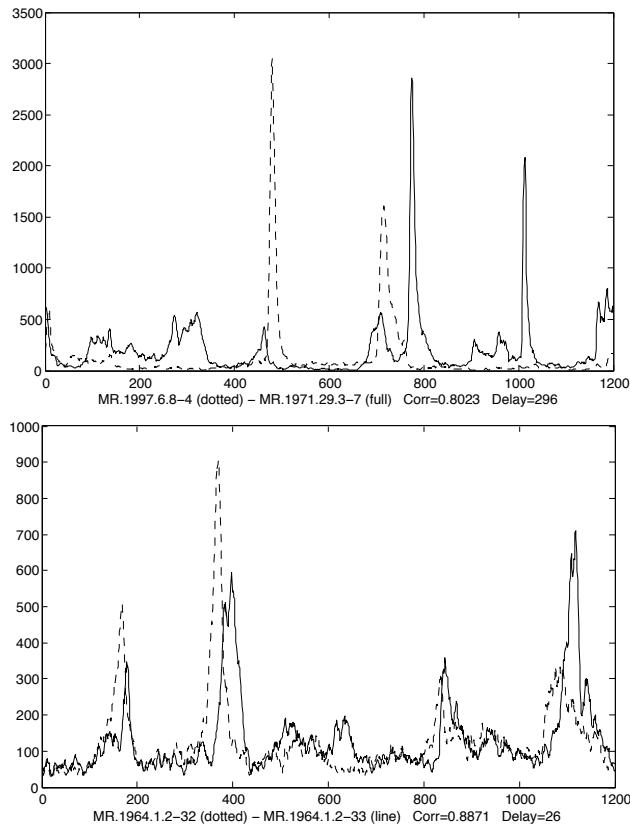


Figure 4. Two examples of a cross-correlation analysis, where the optimal correlation is found through a pitch shift. In the upper example a relatively large shift of 296 cents (about a minor third) reveals the highest similarity ($r = .80$), while in the lower example only a small shift of 26 is necessary to obtain the maximum ($r = .89$).

4.2 Interval analysis

In the analysis of 20th century Western classical music, so-called ‘interval vectors’ are used to express the intervallic content of a pitch-class set [16]. Using a Western chromatic scale, interval vectors are limited to an array of six numbers, expressing the amount of occurrences of each possible pitch interval (from a minor second to a tritone). With the variety of intervals found in African scales, this reduction to six numbers is not possible. Nevertheless, creating a global view on the intervals that constitute the scales can give us interesting insights in the pitch structure of the music. Are there for example any specific intervals that occur often, can we see regional differences or is there an evolution through time.

For this analysis the scales obtained from the peak analysis are transformed to an array of all possible intervals that can be built with this scale. As we work with scales reduced to one octave, the distinction between rising and falling intervals can not be made. Therefore the maximum interval size is set at 600 cents (a tritone or half an octave). For the analysis presented here, the intervals were grouped in bins of 5 cents, which gives us interval vectors of 120 elements.

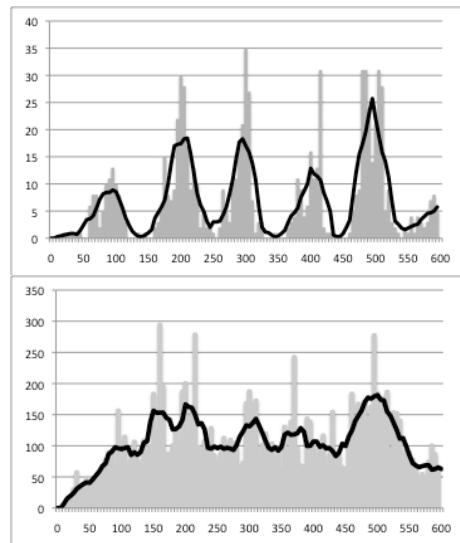


Figure 5. Comparison of all the pitch intervals found in the scale analysis from (above) the 42 pieces from the J.S. Bach’s six cello suite and (below) our collection of 901 pieces of African music.

First we can make a global analysis of the intervals. In figure 5, a comparison is made between our complete collection of 901 pieces and a small sample of Western tonal music (Johann Sebastian Bach’s six cello suites, played by Mstislav Rostropovich, a collection of 42 movements). In the interval analysis of the Western music we clearly see peaks corresponding to the standard intervals of 100 cents. For the African music the situation is much less clear. One similarity could be the importance of the 500 cents intervals (corresponding to the pure fourth/fifth), but the other

peaks are much less well-defined and in some cases sharp peaks appear at odd intervals (e.g. 160, 370 cents).

Now we can also couple the interval vectors with the meta-data. As an example we can look if we can find some differences in interval content between the three time periods. The analysis of the meta-data already revealed that tone scales with larger number of pitches became more important in recent recordings (cf. *supra*). Do we also find an influence on the pitch content? All three interval profiles are very irregular and show peaks at unexpected places, as seen in the global analysis. An interesting evolution is seen if we look at the relative share of the intervals corresponding to the Western equally-tempered scales. Counting the relative share of the 5 relevant intervals by taking the two bins around the correct interval (e.g. 95-105 cents for the minor second), we see that the share of these intervals almost doubles in the recent recordings (Table 3). Only for the minor third we don't see an increase, and the change is especially remarkable for the major seconds (also containing the minor sixths) and the pure fourths/fifths. A detailed view on the area in which pure fourths and fifths are found reveals an interesting evolution (Figure 6). The main peak seems to shift from 530 cents in the earliest recordings to 515 cents in the middle period to end up at 500 cents in the most recent recordings. This possibly also indicates a gradual evolution to a Western pure-fifth based tuning.

Interval	< 1960	1960-1975	> 1975
min. 2nd	1,46	1,87	2,20
maj. 2nd	1,57	1,71	5,20
min. 3rd	2,26	3,39	2,84
maj. 3rd	1,25	1,28	2,78
4th/5th	2,55	2,56	5,31
sum	9,10	10,81	18,33

Table 3. Relative share (in %) of pitches in an area of 10 cents around the Western equally-tempered intervals, for three recording time periods.

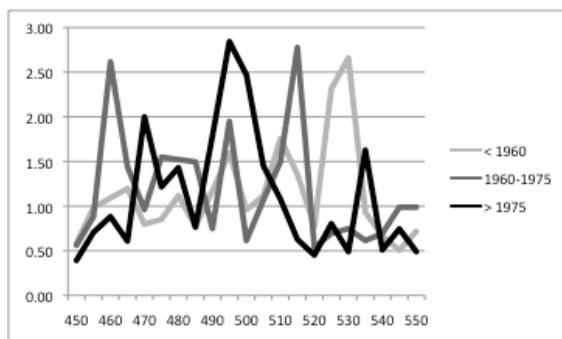


Figure 5. Relative share (in %) of pitches in bins of 5 cents between 450 and 550 cents for three recording time periods.

A detailed analysis of regional differences goes beyond the scope of this paper. Yet, we can see some interesting elements in relation to the global analysis of intervals. We see for example that the peak at 160 cents is present in every region. This shows that it is not a feature of a particular culture, but a 'pan-African' characteristic. Further ethnomusicological work is necessary to find a possible explanation for the importance of this interval. Interestingly similar interval sizes are found in the music of the Middle-East, where they are classified as 'neutral seconds' (neither minor nor major, but in between). The pitch system there however is organized according to completely different principles, so it is not clear if a direct link can be established.

5. DISCUSSION AND CONCLUSIONS

We proposed a number of methods to deal with non-standardized tone-scales, as they are found in African music. Avoiding working with *a priori* determined categories (such as the pitches of the chromatic scale), allows a representation and analysis of a large variety of tone scales. This was illustrated by a sample of solo-music on four different instrument types taken from the archive of the Belgian Royal Museum of Central-Africa. The results are promising, both for data-mining application and as a starting point for ethnomusicological research. Before we can expand these techniques to the whole database, several problems still have to be solved. Some important obstacles are the presence of unaccompanied vocal music, which usually has a fluctuating pitch. This makes it very hard to extract the exact scale automatically, without applying a kind of pitch correction first. Also there is a problem with the use of percussion. The presence of percussive sounds tends to obscure the actual pitch scale used and to generate one large peak associated with the pitch of the percussion instrument. Therefore a system to suppress these percussive sounds should also be developed.

Using this relatively small sample of 901 pieces, we could already develop some methods for ethnomusicological research, creating a more elaborate view on scales and temperaments in African music in an automated way. A global comparison between the intervals found in Western and in African scales, shows that African music does not conform to the fixed chromatic scale nor has another fixed scale, however in recent recordings there seems to be a tendency to the use of more elaborate, equally-tempered scales. Further research has to be done in these historical aspects as well as on the geographical aspects of African tone scales. These techniques lead to usable applications for query-by-example, database management and classification.

6. REFERENCES

- [1] H. C. Longuet-Higgins & M. J. Steedman: "On interpreting Bach," *Machine Intelligence*, vol. 6, pp. 221–241, 1971.
- [2] M. Leman: *Music and Schema Theory*, Berlin, Springer Verlag, 1995.
- [3] S. Pauws: "Extracting the Key from Music," in W. Verhaegh, E. Aarts & J. Korst (eds.) *Intelligent Algorithms in Ambient and Biomedical Computing*, pp. 119–132, 2006.
- [4] Ö. Izmirli: "Localized Key Finding from Audio Using Nonnegative Matrix Factorization for Segmentation," in *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR2007)*, 2007.
- [5] C. Chuan & E. Chew: "Audio key finding: considerations in system design and case studies on Chopin's 24 preludes," *EURASIP Journal on Applied Signal Processing*, Vol. 2007, No. 1, 15 pp., 2007.
- [6] E. Gomez: *Tonal Description of Music Audio Signals*, Ph.D. Thesis, Universitat Pompeu Fabra, Barcelona, 2006.
- [7] J. S. Downie, K. West, A. Ehmann & E. Vincent: "The 2005 Music Information Retrieval Evaluation eXchange (MIREX 2005): Preliminary Overview", in *Proceedings of the Sixth International Conference on Music Information Retrieval (ISMIR 2005)*, pp. 320-323, 2005.
- [8] A. Nesbit, L. Hollenberg & A. Senyard: "Towards Automatic Transcription of Australian Aboriginal Music," in *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004)*, pp. 326-330, 2004.
- [9] P. Chordia, M. Godfrey & A. Rae: "Extending Content-Based Recommendation: the Case of Indian Classical Music," *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR2008)*, pp. 571-576, 2008.
- [10] H. Li & M. Leman: "A Gesture-baseed Typology of Sliding-tones in Guqin Music," *Journal of New Music Research*, Vol. 36, pp. 61-82, 2007.
- [11] A. Krishnaswamy: "Multi-dimensional Musical Atoms in South-Indian Classical Music," *Proceedings of the 8th International Conference on Music Perception and Cognition (ICMPC8)*, 2004.
- [12] B. Bozkurt: "An Automatic Pitch Analysis Method for Turkish Maqam Music," *Journal of New Music Research*, Vol. 37, pp. 1-13, 2008.
- [13] P. Heydarian, L. Jones & A. Seago: "The Analysis and Determination of the Tuning System in Audio Musical Signals," *Paper presented at the 123rd convention of the Audio Engineering Society*, 5 pp., 2007.
- [14] G. Kubik: *Theory of African Music*, Willemshaven, F. Noetzel, 1994.
- [15] L.P. Clarisse, J.P. Martens, M. Lesaffre, B. De Baets, H. De Meyer & M. Leman: "An Auditory Model Based Transcriber of Singing Sequences," *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR2002)*, pp. 116-123, 2002.
- [16] A. Forte: *The Structure of Atonal Music*, Yale University Press, 1973.