TOWARDS A SYSTEMATIC EXPLORATION OF MOTIF DEVELOPMENT IN ARAB-ANDALUSIAN MUSIC

Oriol Colomé i Font
UPF
oriol.colome01@estudiant.upf.edu

Satyajeet Prabhu
UPF
satyajeet.prabhu01@estudiant.upf.edu

ABSTRACT

This paper investigates motif development in Arab-Andalusian Music in the *al-Āla* tradition by studying repetition as a development technique to assess the centonization theory proposed by Amin Chachoo. We compare the significance, distribution, and repetition of critical motifs within a selected *ṭab‘ (Iraq-al-ayam)* across multiple performances and across sections of a single performance. The key motifs (*centos*) in the *ṭab‘* used in this study are taken to be the intersection between those suggested by Chachoo and the significant motifs identified by Nuttall et al. through computational analysis in their study. Our analysis reveals patterns of cento repetition within delimited distinct musical structures. These repetitions appear purposeful, suggesting intentional compositional or improvisational decisions. Our findings support the idea that cento repetition in Arab-Andalusian music is systematic rather than arbitrary, indicating a potential structural and aesthetic significance of motif development within the tradition.

1. INTRODUCTION

Computational musicology has opened doors to innovative methods for analyzing music collections, enabling researchers to delve deeper into the multifaceted nature of music and its cultural significance. [1] [2]

Now, with digital recordings and new programs, we can analyze big datasets and music performances in new ways: systematically. While such collaboration can enhance traditional musicology, we must bear in mind the subjective, cultural, and human inherent nature of the object of study. [3] [4]

1.1 Characteristics of Arab-Andalusian Music

1.1.1 History and basic theory

Arab-Andalusian music (henceforth referred to as AAM) is a music tradition born from the cultural fusion of the Muslim world and the Iberian Peninsula in Al-Andalus (Islamic Spain) from the 8th to 15th centuries.

The central musical form in Arab-Andalusian music is the "***nawba***". These suites serve as sonic journeys, encompassing instrumental and vocal compositions. Each nawba unfolds in a prescribed order dictated by the "***mīzān***" (rhythmic mode). Traditionally, a nawba is unified by a primary "***ṭab‘***" (melodic mode). Each ṭab‘, characterized by a specific diatonic scale and characteristic melodic motifs, is used to convey a certain emotional or spiritual content. [5]

1.1.2 Centonization Model

Centonization (from Latin "*cento*" meaning "patchwork") is a theory concerning the composition of melodies, or pieces, based on pre-defined melodic identities and formulas in music [6] [7].

It is an old and widely used technique [8] and similar concepts can be found in other musical traditions such as in Gregorian chant, the raga framework in Indian art music, or the pathet in Indonesian gamelan music.

1.2 Motivation

With the aid of computation, this exploratory study seeks to observe macro patterns, if any, of centos development in AAM to deepen our understanding of the musical tradition. Specifically, we study the distribution, relative importance, and repetition of centos within a single ṭab‘ using musical transcriptions. Building upon existing computational analyses, we aim to contribute to the centonization theory in AAM proposed by Amin Chachoo [9] and also to gain a deeper musicological understanding of the tradition.

2. DATASET

CompMusic project [10] is the largest source of symbolic scores in the Arab-Andalusian tradition. It consists of 158 manual transcriptions of audio recordings of performances by three different orchestras done by Amin Chachoo. The scores are stored in the MusicXML [1] format and consist of a monophonic transcription of the dominant melody. While the entire corpus, including the audio files, can be downloaded from Zenodo [2], the individual scores are also available on the open-source score repository Musescore [3] and titled in the format ’mīzān-ṭab‘’.

Our dataset is a subset limited to the ***ṭab‘ ‘irāqal-‘ajam***. We chose this particular ṭab‘ because the seven available

[1] https://www.musicxml.com/
[2] https://zenodo.org/records/1291776
[3] https://musescore.com/user/537291/sets/423121

scores for this ṭab‘ belong to a single nawba, which could help to make the findings consistent across different performances.

3. RELATED WORK

In his foundational work, Amin Chaachoo delves into the concept of *ṭab‘ (pl. tubū)*’ in Arab-Andalusian music within the framework of modal theory [9]. Chaachoo systematically identifies characteristic motifs (centos) for each *ṭab‘*, offering a nuanced understanding of the genre’s melodic structures. Complementing this, Nuttall et al. [11] use three different computational approaches for pattern discovery, namely TF-IDF, SIA, and MGDP, to identify significant centos and cross-validate the findings with the theoretical centos proposed by Chachoo. The theoretical centos considered in the study are those with a length between 3 and 7 notes, containing more than one unique pitch and no intermediate rests. The duration and octave of the notes are disregarded. Also, only centos surpassing a minimum frequency of occurrence are considered. The significant theoretical and computationally-detected patterns for 13 of the 26 classic *ṭubū‘* are available on their project GitHub repository [4] .

4. METHODOLOGY

Our study uses this information of significant centos for *ṭab‘ ‘irāqal-‘ajam* as the basis for further computational analysis. The theoretical centos for this *ṭab‘* that meet the previously mentioned criteria are:

[‘AGB’, ‘BAG’, ‘CDE’, ‘EDC’, ‘EF#G’, ‘FED’, ‘GF#ED’]

Out of the above, the centos detected by the three computational approaches as being significant are:

- SIA = [‘BAG’, ‘EDC’, ‘FED’]
- TF-IDF = [’EF#G’, ‘BAG’, ‘GF#ED’, ‘EDC’, (GF#E)*]
- MGDP Minimal = [’EF#G’, (’GF#E’)*, (’F#ED’)*]

*addressed in section 6.3

We consider the common ground between these lists as the centos for further analysis. In other words, ‘AGB’ and ‘CDE’ do not appear significant in any of the computational searches. So out of the seven proposed theoretical centos, we only consider the following 5 centos for analysis:

‘**BAG**’, ‘**EDC**’, ‘**EF#G**’, ‘**FED**’, ‘**GF#ED**’

The python package music21 [12] is used for reading and processing the MusicXML scores. Besides the note information, staff text annotations are also extracted. The scores contain Chachoo’s manual annotations of form and structural changes, which are used for analysis.

For each score, we iterate over every measure in the score and record the presence and count of each of the centos in our list. It must be noted here that we only track the appearance of a cento within a measure and not across measure boundaries. This limits the scope to observing only the macro trends in motif development. The extracted data generates the centos distribution over the score and plots it on a graph. We overlay markers at measure positions where text annotations are found in the score. These distributions and plots are then used to infer noteworthy trends in the repetition, relative importance, and relevance of the placement of the centos in the score. Lastly, the found centos are color-coded in the original score for visual analysis. The pseudocode to retrieve the centos and its positions in the measure can be found below:

```
for i = 0 to (length(notes) - cento_length):
    if notes[i:i + cento_length] matches cento:
        melodic_contour = cento
        rhythmic_values = []
        for j = i to (i + cento_length - 1):
            rhythmic_values.append(quarterLength of note[j])
        cento_info = {
            'cento': cento,
            'position': {
                'measure': measure_number,
                'start_index': i,
                'end_index': i + cento_length - 1
            },
            'rhythmic_values': rhythmic_values,
        }
        append cento_info to found_centos_info
```

The notebook containing the code for analysis and plot generation can be found on our GitHub repository [5] . It is also accompanied by the generated plots and color-annotated scores.

5. RESULTS

Tables 1 and 2 contain the summary statistics of the centos found in each of the seven music scores for *ṭab‘ ‘irāqal-‘ajam*.

Table 1. mīzān : Bassit and Btayhi

Score	Bassit_Iraq_Ajam		Btaihi_Iraq_Ajam		Btayhi_Iraq_Ajam	
	Count	%	Count	%	Count	%
BAG	**85**	**28.7%**	**91**	**25.0%**	117	26.9%
EDC	58	19.6%	85	23.4%	79	18.2%
EF#G	75	25.3%	84	23.1%	84	19.3%
FED	42	14.2%	88	24.2%	**124**	**28.5%**
GF#ED	36	12.2%	16	4.4%	31	7.1%
	296	100.0%	364	100.0%	435	100.0%

Table 2. mīzān : Quddam

Score	Quddam_Iraq_Ajam(1)		Quddam_Iraq_Ajam(2)		Quddam_Iraq_Ajam		Quddam_Iraq_Ajam_Lasamir	
	Count	%	Count	%	Count	%	Count	%
BAG	37	24.2%	11	12.5%	60	22.9%	70	17.1%
EDC	38	24.8%	**46**	**52.3%**	46	17.6%	87	21.3%
EF#G	**64**	**41.8%**	26	29.5%	**94**	**35.9%**	**176**	**43.0%**
FED	11	7.2%	5	5.7%	37	14.1%	42	10.3%
GF#ED	3	2.0%	0	0.0%	25	9.5%	34	8.3%
	153	100.0%	88	100.0%	262	100.0%	409	100.0%

[4] https://github.com/centonization/centonizationtheory/tree/main/results

[5] https://github.com/satyajeetprabhu/arab-andal-motif-dev

6. DISCUSSION

6.1 Distribution and relative importance of centos across all scores

A clear hierarchy is evident in the relative importance of certain centos over others in the few scores of *ṭab‘ ‘irāqal-‘ajam* we have analyzed.

- We can see that one of the factors that might dictate the prominence of one cento over another is the mīzān. For example, the cento ’BAG’ appears to be more prominent in Bassit. Prominence is shared between ’BAG’ and ’FED’ in Btayhi. ’EDC’ and ’EF#G’ appear to be the more important centos in Quddam.
- The "EF#G" cento shows relatively consistent counts and percentages across categories, suggesting a balanced distribution.
- "BAG" and "FED" melodic identities exhibit more variability in their distributions, with larger differences in counts and percentages across categories.
- "GF#ED" has relatively lower counts than other scores across all categories, despite being considered an important motif in the ṭab‘.

6.2 Distribution and relative importance of centos across three different *mawāzīn*

We take one score from each of the three mawāzīn available for this ṭab‘ to analyze structural patterns in the appearance of a cento in the score and the positional relevance of the centos within the temporal structure of the mīzān. The plots generated for the three scores are shown in the appendix of this paper.

- In ***Bassit Iraq Ajam***, all five centos appear relevant in the piece’s second half. They all exhibit density and symmetrical patterns for sections *Rubba Laylin*, *Inshad*, *Ma Kuntu Adri*, *Nahnu Qawum Kana*, and *Marhaban*. However, for the *Tawchiya* section, the centos "EF#G" and "BAG" sections appear somewhat sporadic and irrelevant. The *Mshalia* section shows no signs of centonization. In the *Rubba Laylin* and *Ma Kuntu Adri* sections, fascinating patterns emerge, characterized by an apparent "call and response."
- In ***Btayhi Iraq Ajam***, the distribution of centos appears visually relevant across the piece in terms of density, pattern display, and inter-dependability, except for the *Tawchiya* section, where the tiny appearance of cento "EF#G" is negligible.
- The ***Quddam Iraq Ajam*** score, the motif distribution appears relevant and particularly dense in some sections, such as *Tallahi Law Jayyaruni*. The visual codependency between centos displayed in the plot is particularly evident in this score. Additionally, sections such as *Hada Elyawm*, *Atani Mina*, and the piece’s ending, *Malaktoum*, show no signs of centonization.

6.3 Overlap in centos - observations from color-annotated scores

A visual analysis of the color-annotated scores shows that some centos overlap. For example, FED (cyan) and EDC (blue) in Figure 1, as well as BAG (red) and GF#ED (orange) in Figure 2. This brings into question the notion of considering these as different centos when they appear together.

The findings of the computational approaches also confirm this. For example, TF-IDF and MGDP identify GF#E and F#ED as significant patterns, although they are a subset of the theoretical cento GF#ED. Cases such as this require further investigation into what constitutes a cento according to the practitioners of the tradition.

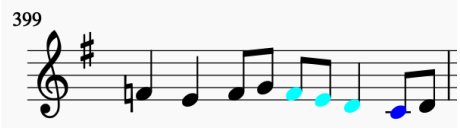


Figure 1. "FEDC" sequence due to FED (cyan) and EDC (blue) cento overlapping

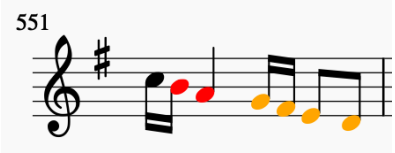


Figure 2. "BAGF#ED" sequence due to cento overlapping

6.4 Analysis for the excluded centos

The appendix also includes plots for the two theoretical centos, ‘AGB’ and ‘CDE’, which were not considered because they were not regarded as significant by the computational approaches. The plots show that although not substantial in frequency, these centos may convey crucial structural information about the form. For example, in Bassit, it appears that the cento ’AGB’ is important for the conclusion of the mīzān. This points to the limitation of using repetition as the sole metric for the importance of a cento in AAM.

6.5 General thoughts

From a high-level visual analysis of the plots and by only observing repetition, we see patterns emerging both vertically (amongst centos) and horizontally (across time and musical structure).

Given the human nature of music, we might deduce that such ordered structures observed in the plots can be attributed to deliberate melodic developments.

Horizontally, we can see that these patterns coincide with the piece’s structural sections, indicating intentional

organization and developmental processes. The plots clearly reveal that a strong structure and repeating elements help make each section feel familiar and connected.

Vertically, the observed patterns reveal an apparent interdependence and correlation between repeated centos within sections of the piece in what appears to be kind of a "call and response" phenomenon between centos.

Also, it is important to note that several sections are devoid of centos altogether and there are instances where their repetition seems somewhat random. Also, while there are observable vertical and horizontal patterns within a score, we were unable to observe any singular pattern relating centos across the board.

7. CONCLUSION AND FUTURE WORK

The study reveals a deliberate and purposeful scheme of motif repetition in Arab-Andalusian music with variations across rhythmic modes and the musical structure of a nawba in *ṭab' 'irāqal-'ajam*. It points to a possibility that motif repetition might be a tool used in this tradition for creating cohesion, guiding listeners through structural boundaries, enhancing the overall musical narrative, and partially validating [9] theory.

By utilizing systematic musicology and computational techniques to analyze motif repetition quantitatively and qualitatively, we were able to get a glimpse into implicit structural patterns that may exist in Arab-Andalusian music. These observations contribute to the centonization theory of Arab-Andalusian music and may serve as a template for similar analysis at a larger scale.

In the future, we could expand the scope to analyze other *ṭubū'* and eventually, the entire corpus. Alternately, finer statistical analysis techniques could be applied to the data generated from this study to identify concrete patterns of repetition and the role and interrelationship between the centos in a *ṭab'* in Arab-Andalusian music.

Generally speaking, a computation and data-driven approach to the study of Arab-Andalusian music could greatly contribute to the musicological study of this tradition.

8. REFERENCES

[1] G. Tzanetakis, H. Etnomüzikoloji, A. Kapur, W. A. Schloss, and M. Wright, "Computational ethnomusicology," 2007.

[2] E. Gómez, P. Herrera, and F. Gómez-Martin, "Computational ethnomusicology: perspectives and challenges," pp. 111–112, 2013.

[3] N. Cook, "Towards the compleat musicologist?" 2005.

[4] R. J. Jorna, L. 13th International Conerence on Informatics, and S. in Organisations (ICISO), *Proceedings of the Thirteenth International Conference on Informatics and Semiotics in organisation : Problems and possibillities of computational humanities*. Fryske Akademy, 2011.

[5] R. C. Repetto, N. Pretto, A. Chaachoo, B. Bozkurt, and X. Serra, "An open corpus for the computational research of arab-andalusian music," in *Proceedings of the 5th International Conference on Digital Libraries for Musicology*, ser. DLfM '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 78–86. [Online]. Available: https://doi.org/10.1145/3273024.3273025

[6] L. TREITLER, "HOMER AND GREGORY: THE TRANSMISSION OF EPIC POETRY AND PLAIN-CHANT," *The Musical Quarterly*, vol. LX, no. 3, pp. 333–372, 07 1974. [Online]. Available: https://doi.org/10.1093/mq/LX.3.333

[7] L. Treitler, "Centonate Chant: Centonate or Centonate," *Journal of the American Musicological Society*, vol. 28, no. 1, pp. 1–23, 04 1975. [Online]. Available: https://doi.org/10.2307/830914

[8] P. Ferretti, *Estetica Gregoriana*, ser. Da Capo Press music reprint series. Da Capo Press, 1977, no. v. 1. [Online]. Available: https://books.google.es/books?id=P-wXAQAAIAAJ

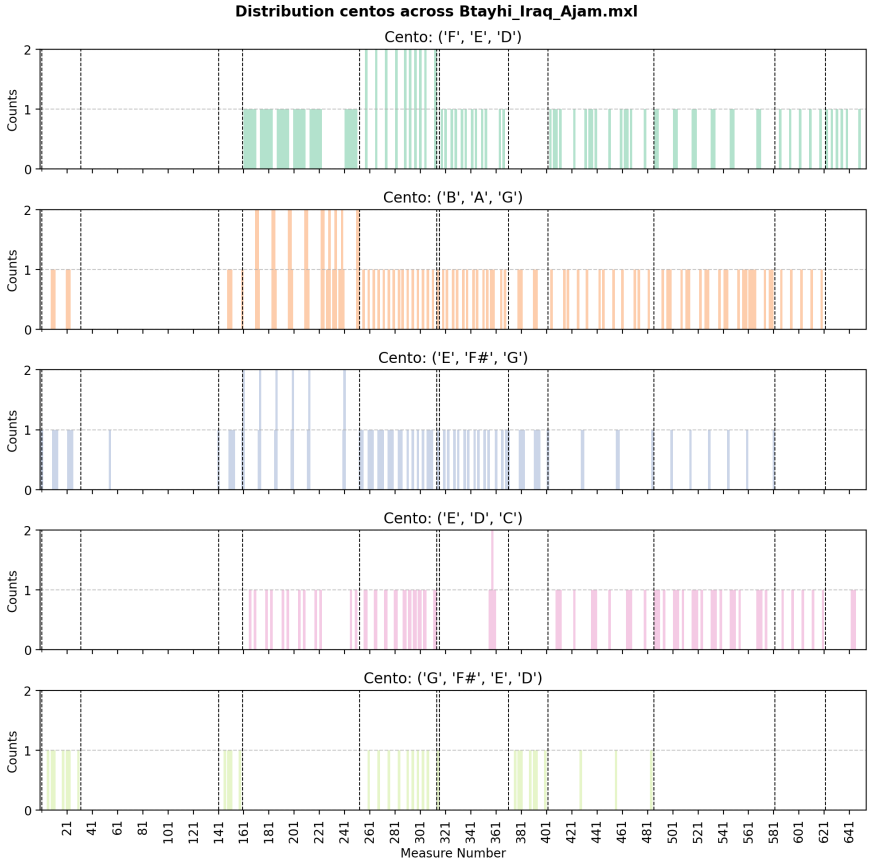
[9] A. Chaachoo, *LA MUSIQUE HISPANO-ARABE, AL-ALA*, 2011.

[10] X. Serra, "Creating research corpora for the computational study of music: The case of the compmusic project," *Proceedings of the AES International Conference*, pp. 1–9, 01 2014.

[11] T. Nuttall, M. G. Casado, A. Ferraro, D. Conklin, and R. C. Repetto, "A computational exploration of melodic patterns in arab-andalusian music," 2021. [Online]. Available: https://compmusic.upf.edu/

[12] M. S. Cuthbert and C. Ariza, "music21: A toolkit for computer-aided musicology and symbolic music data," 2010.

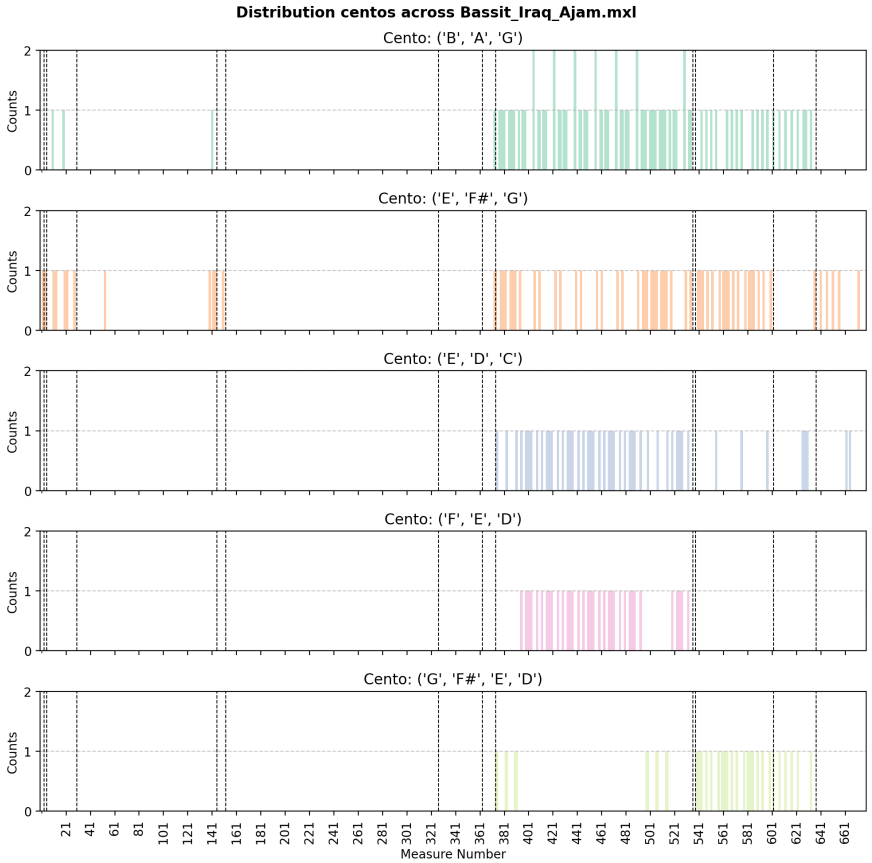
9. APPENDIX




```
Cento Counts:
(F, E, D)         124
(B, A, G)         117
(E, F#, G)         84
(E, D, C)          79
(G, F#, E, D)      31
Name: cento, dtype: int64
```

Legend of vertical markers

Measure	Text Expression
0	Mshalia
31	Tawchiya
140	Mshalia
159	Ayna elwafa
252	Ahlan Bikum
313	accel.
315	Ahda lana
370	Mshalia
401	Ya qawmi hajarni
485	Ya qalbi man yuridek
581	Ya warda zawan
621	Qadru al-Kalam

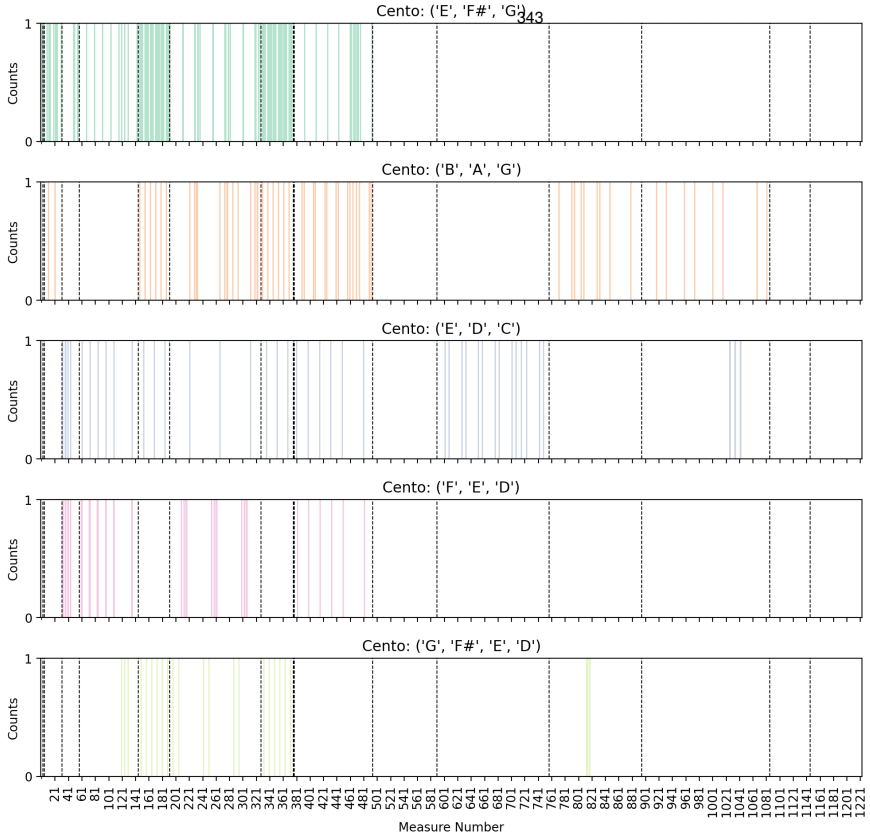



```
Cento Counts:
(B, A, G)         85
(E, F#, G)        75
(E, D, C)         58
(F, E, D)         42
(G, F#, E, D)     36
Name: cento, dtype: int64
```

Legend of vertical markers

Measure	Text Expression
2	Mshalia
4	Mshalia
29	Tawchiya
144	Mshalia
151	Tawshiya 2
326	accel.
362	accel.
373	Rubba laylin
535	INSHAD
537	Ma kuntu adri
601	Nahnu qawmun lana
636	Marhaban

Distribution centos across Quddam_Iraq_Ajam.mxl



Cento Counts:

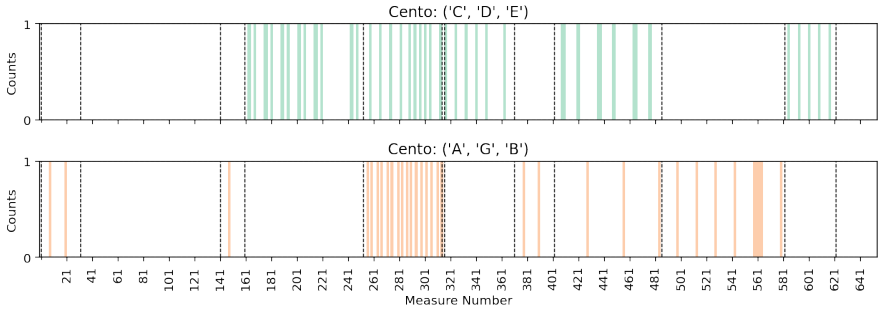
(E, F#, G)	94
(B, A, G)	60
(E, D, C)	46
(F, E, D)	37
(G, F#, E, D)	25

Name: cento, dtype: int64

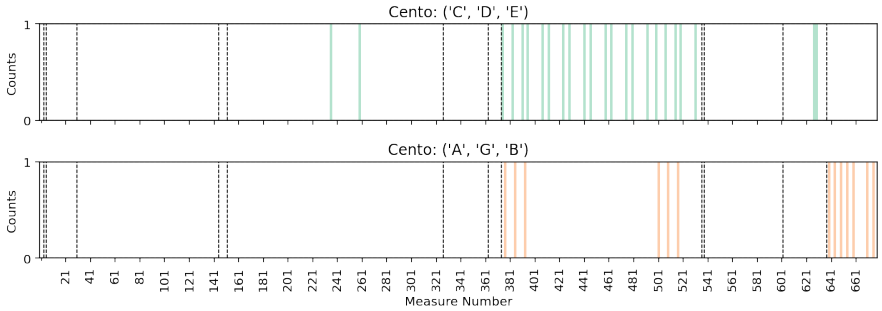
Legend of vertical markers

Measure	Text Expression
2	Mshalia
4	Mshalia
30	Tawshiyya Quddam
56	Indama ji`tu liddiyyar
144	Tallahi law jayyaruni
191	Antum sururi
327	Antum hayati
375	INSHAD
376	Man la raqad
493	Hada elyawm
589	Ida dahik
756	Ana kulli milkum
894	Aya munadi
1085	Atani mina
1145	Malaktoum

Excluded centos in Btayhi_Iraq_Ajam.mxl



Excluded centos in Bassit_Iraq_Ajam.mxl



Excluded centos in Quddam_Iraq_Ajam.mxl