Available online at ijci.wcci-international.org

International Journal of Curriculum and Instruction 17(3) (2025) 978–1001

IJCI
International Journal of Curriculum and Instruction

Turkish Music Education: An Artificial Intelligence Based Performance Analysis Design[1]

Muhammed Talha Efe[a1], Levent Değirmencioğlu[b]

[a] *Kırıkkale University, Department of Musicology, Kırıkkale 71000, Turkey*
[b] *Erciyes University, Department of Music, Kayseri, 38000, Turkey*

Abstract

The purpose of this study is to develop a software system, using the Python programming language, capable of analyzing the pitch and usul (rhythmic pattern) structures in performances of Turkish music. To achieve this goal, the research was conducted within the framework of the Design and Development Research Model (DDRM) and was completed in three stages. In the first stage, a comprehensive literature review was carried out to identify relevant publications that could provide insight into the subject; both domestic and international written sources, as well as existing applications, were examined in detail. In the second stage, based on the main objective, the needs were identified, and a software design was created to perform pitch and usul analysis on performance recordings of Turkish music. The design, developed through the iterative cycle of DDRM, analyzes frequencies using Python libraries and custom algorithms, while storing and visualizing the results through a Firebase based system. Users can track analysis results across specific time frames and access detailed visual data for each analysis, including spectral representations, error curves, and color-coded notation. In the third stage, the software design was tested through an application process based on the analysis of performance recordings. For this purpose, a study group consisting of performers and experts was formed. The performer group included two students with basic Turkish music training who could play instruments (Ud and Ney) at a beginner level, while the expert group consisted of three academics specialized in Turkish classical music and Turkish folk music. Before the application process, the students performed an etude prepared in consultation with the researcher and expert group: a 56-pitch study in the Bayati maqam and in Sofyan usul. The recorded performances were evaluated in terms of pitch and usul elements by both the expert group and the software. Subsequently, the analyses of the expert group and the software were comparatively examined, and the results were discussed. It was determined that the software could technically analyze performance recordings with a higher degree of precision than anticipated for pitch and usul elements, achieving a high level of consistency with the expert evaluations.

Keywords: Turkish music education, software, artificial intelligence, pitch analysis, usul analysis

[1] Corresponding author: Muhammed Talha EFE. ORCID ID.: *https://orcid.org/0009-0008-3136-206X*
E-mail address: talhaefe@kku.edu.tr

1. Introduction

Music is not only a universal medium for expressing emotions, thoughts, and cultural codes, but also a multidimensional form of communication in terms of education, cognitive development, and social interaction (Levitin, 2006; Sloboda, 2005). In this respect, music is not merely an aesthetic activity; it is also regarded as a data domain open to systematic analyses and modeling. In recent years, with the advancement of digital Technologies particularly artificial intelligence-based systems significant transformations have occurred in both music production and music analysis (Briot et al., 2020).

This transformation has enabled the exploration of the mathematical, statistical, and algorithmic aspects of music under the term computational musicology, making possible operations such as frequency analysis, rhythmic structure analysis, and musical pattern recognition. These advancements have contributed innovatively to various fields such as music education, music information retrieval, and digital archiving (Downie, 2003; Huron, 1999; Serra, 2011).

However, examples encountered in the literature are predominantly based on the Western music system of 12 tone equal temperament (12-TET). Due to their direct compatibility with formats such as MIDI and MusicXML, most studies have remained confined to Western music. It can be stated that the reflections of these developments are far less common in musical traditions outside Western music, one of which is Turkish music.

Turkish music, with its microtonal pitch intervals and melodic progression characteristics, has a structure far more complex and richer than that of Western music (Bozkurt, 2008b). Consequently, computational tools developed for analyzing the mathematical, statistical, and algorithmic aspects of Western music have not been effectively utilized in Turkish music analysis, and related research remains limited.

Nevertheless, recent studies show an increasing interest in computational analyses specific to Turkish music. These studies employ resources such as the SymbTr dataset (Karaosmanoğlu, 2012) and related projects (Serra, 2011), which digitally analyze Turkish maqam music, as well as n-gram models for maqam recognition, melodic segmentation, and usul detection (Ünal et al., 2014), frequency histograms (Gedik & Bozkurt, 2010), and machine learning algorithms (Abidin et al., 2018).

Based on the limited scope of research analyzing the mathematical, statistical, and algorithmic aspects of Turkish music, this study aims to develop a performance analysis software/design tailored to the structural characteristics of Turkish music for use in Turkish music education. Within this main objective, the sub goals are to ensure that the software can analyze Turkish music performances (both instrumental and vocal) in terms of pitch and rhythmic patterns, provide users with instant, weekly, and monthly feedback, and enable tracking of personal progress by recording daily performance data.

2. Method

In this study, which aims to develop a performance analysis software tailored to the structural characteristics of Turkish music, the Design and Development Research Method (DDRM) a type of qualitative research model was employed. Richey and Klein (2007, p.1) define DDRM as *"the systematic study of design, development, and evaluation processes with the aim of producing knowledge based on data systematically derived from practice, by creating new or improved models for the development of instructional and non-instructional products, tools, and their processes, in adherence to empirical research methods."*

Design based approaches involve the creation, improvement, and evaluation of a system or tool (Plomp, 2013, p.18). This method prioritizes the implementation of the developed product in a real-world context and the testing of its applicability. As it seeks to produce both theoretical and practical outcomes, DDRM is frequently preferred in research focused on developing educational software and testing its effectiveness (Akker et al., 2006, pp. 45–46).

Mutlu (2016, p.54), adapting from Reeves (2006), outlines the stages of DDRM as follows:

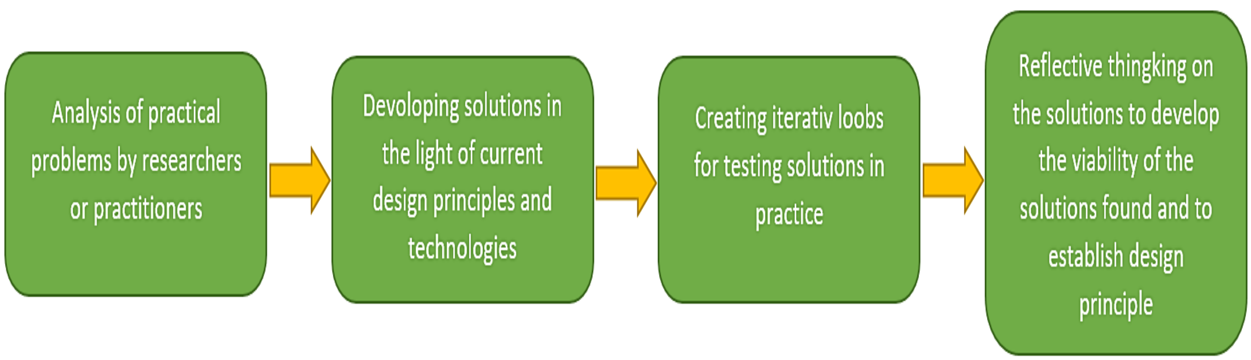

Figure 1. Stages of the Design and Development Research Model

Richey and Klein (2007, p.8) classify DDRM into two main categories: Product and Tool Development and Model Research. The Product and Tool Development processes can be defined through three stages: analysis, detailed explanation of product development processes, and evaluation of the final product (Mutlu, 2016, p.54). In this study, the Product and Tool Development type of DDRM was adopted, and the research was completed in three stages.

2.1. *Data Collection Tools*

In the first stage of the research, the functionalities that the software could provide to users in line with the main purpose of the study were determined. Within this scope, the software was designed to allow users to input musical notation, record audio in accordance with the specified usul (rhythmic pattern), and analyze the performance after execution based on parameters such as error rate, pitch deviation, and timing accuracy. Accordingly, the following data collection tools were employed in the study:

- A Frontend Interface (React) was implemented to enable students to enter musical notation and record audio through the software.
- An artificial intelligence-supported Backend (Python + Librosa + Flask) system was used to process the audio recordings made through the software, extract pitch data, and analyze the pitch-frequency alignment.
- The Firebase platform was employed to automatically store all user data, error rates, and audio recordings following each performance.
- An expert observation form was utilized, through which each expert listened to the performance and marked incorrect notes, timing deviations, and pitch inaccuracies.

The developed software analyzes audio signals with the help of Python libraries such as librosa, numpy, and flask, as well as custom algorithms, and stores the results using a Firebase-based system. The user can monitor performances over time, and for each analysis, the system produces outputs such as detailed waveforms, error curves, and color maps overlaid on the musical notation.

2.2. *Participants*

The second stage of the research constituted the testing phase. Before the testing phase, two students who had received training in Ney and Ud were selected purposively to perform the musical pieces. The primary criterion in this selection was that the students possessed sufficient performance skills and basic knowledge of Turkish music.

Subsequently, an expert group was formed to evaluate the students' performances in parallel with the software. This expert group consisted of three academics specializing in Turkish music, including both Turkish art music and Turkish folk music.

2.3. *Implementation*

In the second stage of the research, the developed software was tested with real users. At this stage, two students individually performed a pre-prepared etude in the Bayati maqam, and their performance recordings were then independently evaluated by the expert group in terms of pitch and usul elements. The experts marked incorrect pitches, timing deviations, and pitch accuracy discrepancies on the expert observation form.

In the subsequent step, the performance recordings were processed by the software, and evaluations were conducted by the system based on the same parameters. This process directly corresponds to the cycle described in the design-based research model that includes the stages of develop, implement, test and reflect (Barab and Squire, 2004, p.6).

Throughout the research process, both descriptive statistical methods (error rates, success percentages) and qualitative analyses (expert comments, comparative evaluations) were employed. In this way, not only the technical adequacy of the software but also its musical validity was assessed. Within this context, the design-based research model allows for a comprehensive examination of both technical and pedagogical dimensions (Reeves, 2006, p.56).

2.4. *Data Analysis*

The data obtained during the implementation stage were analyzed using the descriptive data analysis method. The main objective of the developed software is to detect pitch and rhythmic pattern errors in the pieces performed by the users and to present these errors in both numerical and visual formats. Therefore, the analysis process was structured on three main levels:

- Digitization of sound frequencies
- Calculation of error rates
- Generation of performance charts for the user

2.4.1. Digitization of Frequencies

The audio files processed using the Librosa library are first converted into fundamental frequency values on a frame-based basis using a customized YIN algorithm (De Cheveigné and Kawahara, 2002). For each frame, the fundamental frequency value (f_0) and a confidence coefficient are calculated, and frequencies with low confidence scores are excluded from the analysis. In this way, only sections containing musical signals are taken into consideration.

2.4.2. Pitch Frequency Error and Cent Deviation Calculation

The obtained fundamental frequency values are compared with the pitch frequencies in predefined maqam scales, and the deviation of each pitch is calculated in cents. Bozkurt (2008a) describes the calculation of cent difference using the following formula:

$$cent \ = \ 1200.log\ log\ 2\left(\frac{\int \quad actual}{\int \quad reference}\right)$$

In this context, $\int \quad actual$ refers to the actual frequency analyzed, while $\int \quad reference$ represents the theoretical frequency according to the maqam scale. These differences are directly presented to the user on the analysis screen.

2.4.3. Time Series and Progress Charts

After each analysis process, the software stores the following data in JSON format, which can be displayed to the user upon request:

- Name of the piece
- Start and end time (in seconds)
- Actual frequency value (Hz)
- Cent deviation
- Error percentage

These data are saved in a Firebase-based database to generate daily, weekly, and monthly progress curves based on the user's analysis history. The data are visualized on the React interface using Chart.js, enabling users to monitor their performance and compare it with previous analyses.

2.4.4. Classification and Labeling

The software labels the analyzed pitches as correct or incorrect. For this labeling process, a threshold value is determined parametrically based on the difficulty level selected by the user in the interface. Pitches that deviate beyond the threshold are marked as incorrect. In addition, detected pitch slides (glissando) and note onset delays are also visually presented to the user.

2.4.5. Feedback System

Immediately after the performance analysis, the data are presented to the user in two ways:

- By color-coding the notes on the notation page according to their error margin
- By generating an overall error chart in the "Ability Boards" page

The data shown on the notation page include the deviation amount of each note (in Hz), onset delay, and pitch slide amount. When the user hovers over a note with the mouse or presses and holds the note on a touchscreen, this information is displayed in a pop-up menu.

Figure 2. Display of Incorrect Notes on the Notation Panel

In the Ability Boards section, this information is provided in graphical form. The analysis chart includes the average error rate of the entire piece, the performance date, and the error amount for each note. The user can access these charts at any time.

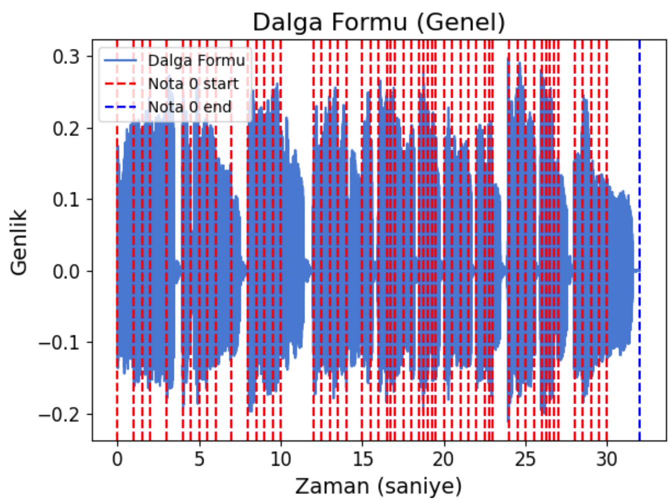

Figure 3. Overall Waveform of the Performance

In Figure 3, the audio recording of the performance is displayed as a complete waveform, and the starting point of each note is indicated with dashed red lines. This visualization allows the user to identify any interruptions in the recording or weaknesses in frequency intensity (amplitude).

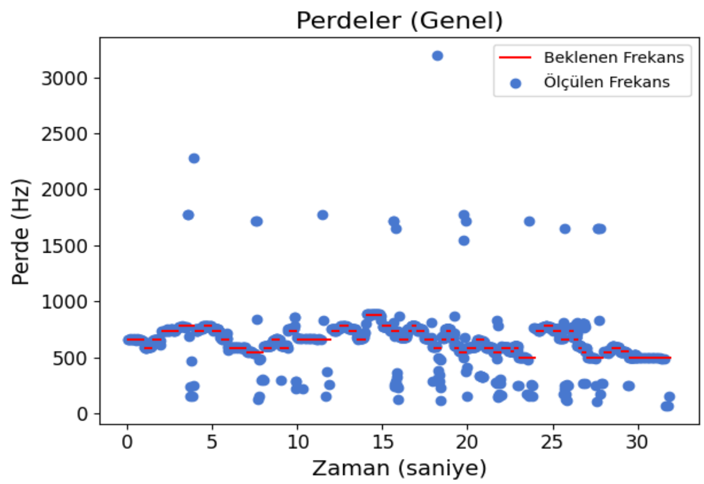

Figure 4. Comparative Analysis of Frequencies in Performances

The graph shown in Figure 4 contains two parameters: the first parameter represents the expected frequencies, that is, the frequencies of the pitches entered in the interface; the second parameter represents the measured frequencies, which are obtained by analyzing the performer's audio recording.

This graph, which presents a comprehensive comparison between the notation and the actual performance, clearly illustrates the overall correspondence between the two. Furthermore, the instantaneous deviations observed in the blue dots representing performance frequencies indicate that certain spectral leakages and harmonics were also included in the analysis.

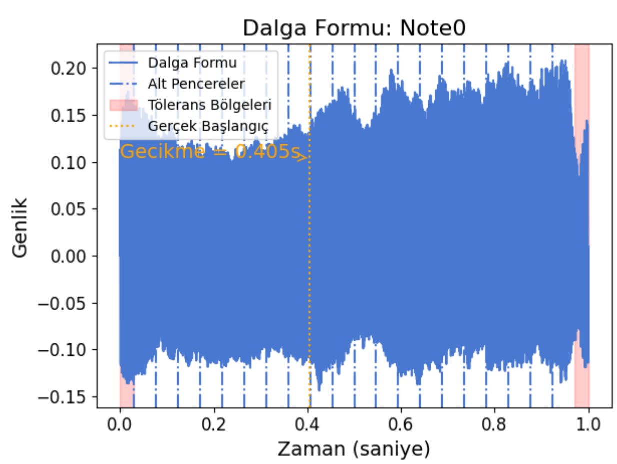

Figure 5. Frequency Graph for Individual Pitch

The frequency graph provided for each individual note includes four parameters:

- The waveform represents the frequency visualization of the current pitch
- Sub-windows indicate the segmented frames of the frequency
- Tolerance zones represent the excluded regions trimmed from the beginning and end of the pitch
- Actual onset shows the delay amount detected by the algorithm

The vertical axis represents the intensity of the frequency (amplitude), while the horizontal axis shows the duration of the pitch in seconds. The sample visual in Figure 5 corresponds to a quarter note performed at 60 bpm and divided into 20 equal frames.

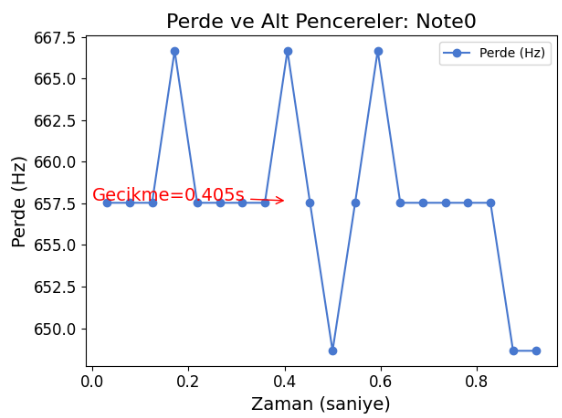

Figure 6. Pitch and Sub-windows

The graph provided for each note, exemplified in Figure 6, shows the frequency (Hz) on the vertical axis and the duration in seconds on the horizontal axis. Each point represents one of the sub-windows into which the pitch is divided. In this visual, 20 sub-windows are displayed. The parameter labeled "Delay," shown in red, represents the first moment when the performer reached the correct frequency while executing the note. In this graph, the delay was calculated as 0.405 seconds, indicating that the performer significantly missed the intended rhythm.

These data are highly valuable for both individual progress tracking and feedback mechanisms for instructors.

3. Findings

This section presents the findings obtained from the analysis of student performances. First, the findings related to the evaluations conducted by the expert group are discussed, followed by the findings derived from the analyses performed by the software, which are illustrated through visual representations.

3.1. *Interface Design and Workflow*

The design interface consists of two main pages: "Homepage" and "Hâne and Meşk Room." When the user first accesses the website, they encounter the homepage (registration page). After registering with an email address and creating a password and username, the user can access the system and begin using the application.

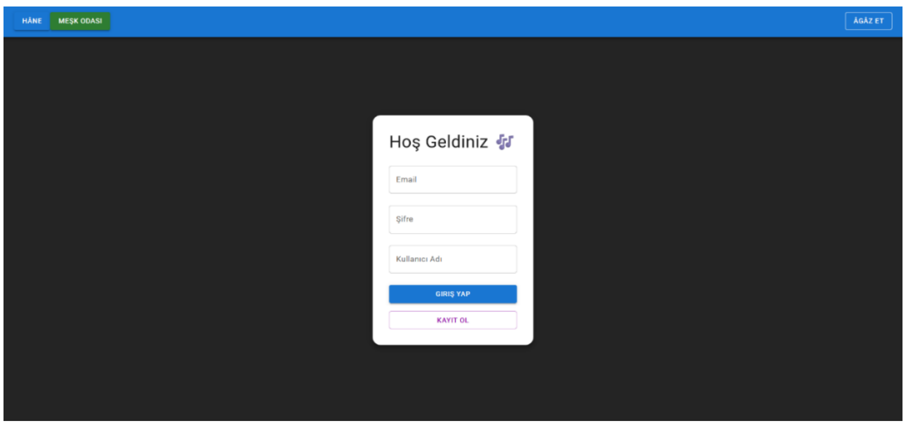

Figure 7. "Homepage"

After registration, the user is directed to the "Hâne and Meşk Room" section. This section allows the user to write musical notation, save the written score, record the performance, and conduct an analysis.

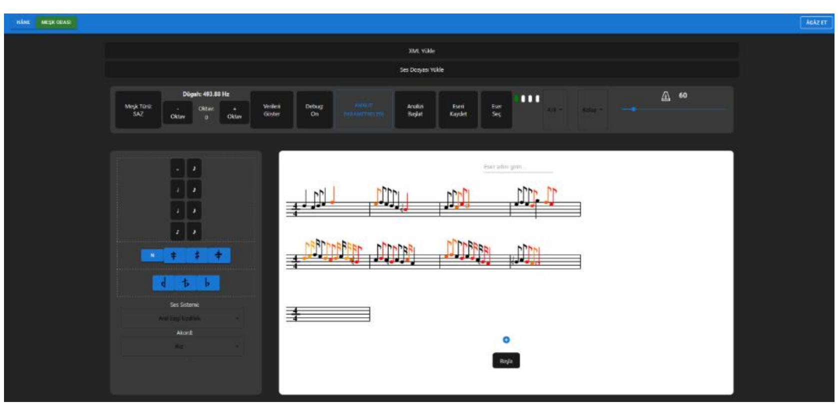

Figure 8. "Hâne and Meşk Room" Page

Errors made during the performance are displayed individually for each note in the white panel on the right side where the notation is written. The error magnitude is represented through a color gradient ranging from yellow to red according to the level of inaccuracy. Additionally, this feedback system is supported by a text indicator showing the error amount beneath each note when hovered over with the cursor.

Figure 9. Analyzed Pitches on the Staff

The analyses are stored by date. Users can compare results across different dates to directly track their progress.

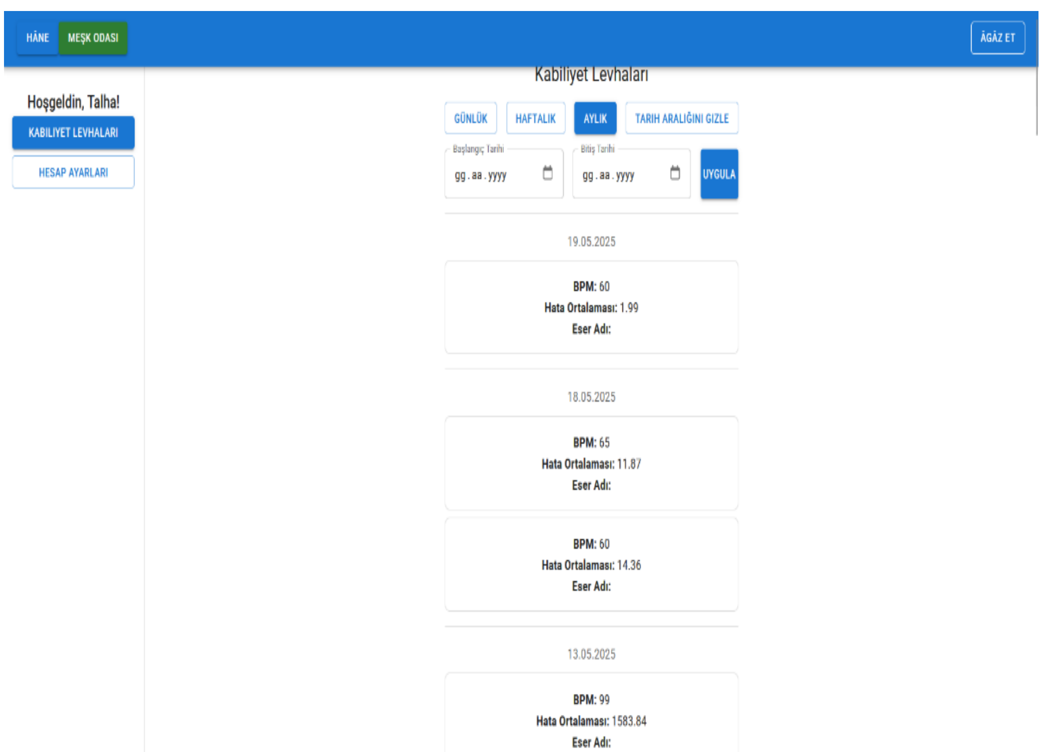

Figure 10. Analysis Results by Date

Initially, the analysis results are displayed in an overview format, and when the user clicks on the chart, a detailed view is provided. In this mode, the recorded values for each note are displayed individually.

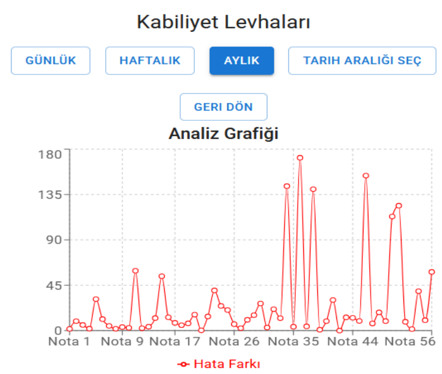

Figure 11. Detailed Analysis Chart

The following diagram presents the sequential stages of the design workflow.

Figure 12. Design Workflow Diagram

3.2. *Findings on Ney and Ud Performances*

During the implementation phase of the study, a 56-note etude in the Bayati maqam was performed separately by two students who played the Ney and Ud at a beginner level. Audio recordings of these performances were collected and evaluated by three domain experts, who analyzed the accuracy or inaccuracy of pitches and usul.

3.2.1. Findings from Expert Group and Software Analyses of the Ney Performance

The student performed the Bayati etude on a Kız Ney, encountering the piece for the first time. The performance recording was captured in a single take. The Bayati etude consists of a total of 56 pitches (notes). The expert group listened to the performance recording and conducted accuracy analyses based on pitch and usul. The markings made by the experts on the notation in a digital environment while listening to the audio recordings are presented in the figures below.

3.2.1.1. Results of the Analysis Conducted by Expert 1

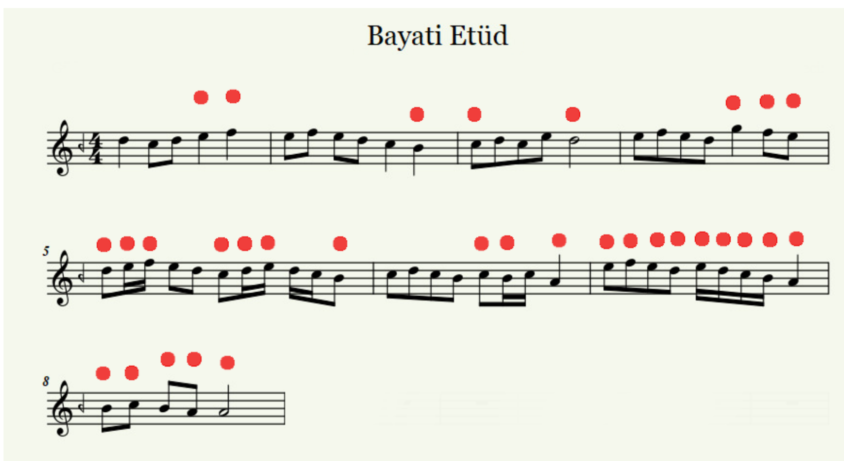

Figure 13. Visual Representation of the Analysis Conducted by Expert 1

The figure displays the annotations made by Expert 1 on the Bayati etude. The red markings indicate that the corresponding pitch was either played incorrectly or not performed with a duration consistent with the usul. According to Expert 1's findings, during the performance of the Bayati etude, the following pitches were played incorrectly or outside the appropriate duration: 4, 5, 11, 12, 16, 21, 22, 23, 24, 25, 26, 29, 30, 31, 34, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, and 56. Based on this assessment, 32 out of the 56 pitches in the etude were performed incorrectly.

3.2.1.2. Results of the Analysis Conducted by Expert 2

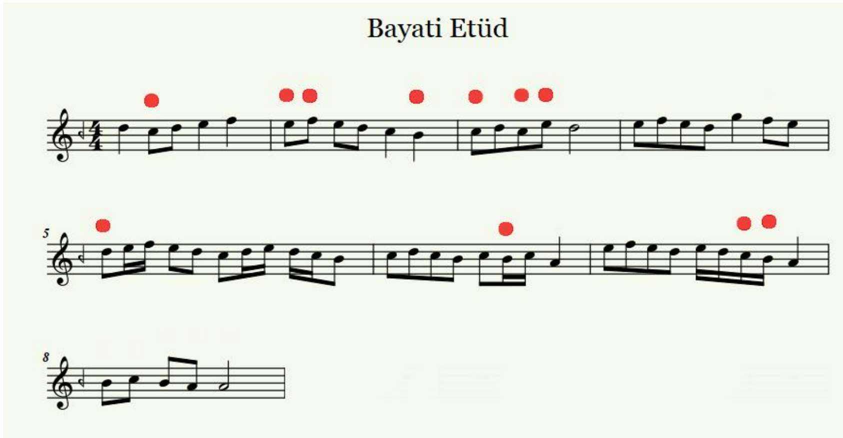

Figure 14. Visual Representation of the Analysis Conducted by Expert 2

The figure displays the annotations made by Expert 2 on the Bayati etude. The red markings indicate that the corresponding pitch was either played incorrectly or not performed with a duration consistent with the usul. According to Expert 2's findings, during the performance of the Bayati etude, the following pitches were played incorrectly or outside the appropriate duration: 2, 6, 7, 11, 12, 14, 15, 24, 40, 49, and 50. Based on this assessment, 11 out of the 56 pitches in the etude were performed incorrectly.

3.2.1.3. Results of the Analysis Conducted by Expert 3

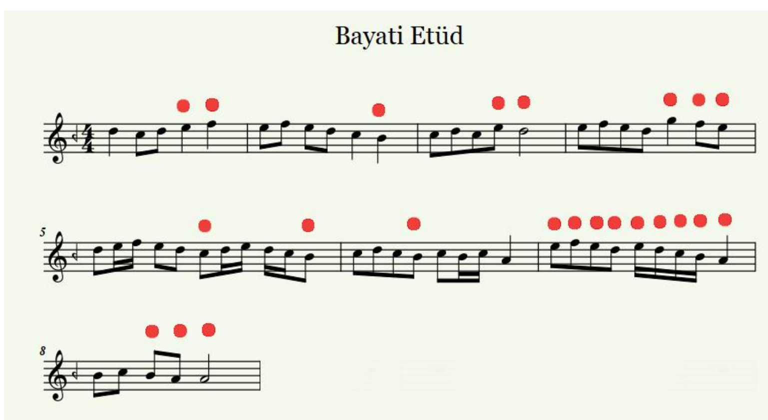

Figure 15. Visual Representation of the Analysis Conducted by Expert 3

The figure displays the annotations made by Expert 3 on the Bayati etude. The red markings indicate that the corresponding pitch was either played incorrectly or not performed with a duration consistent with the usul. According to Expert 3's findings, during the performance of the Bayati etude, the following pitches were played incorrectly or outside the appropriate duration: 4, 5, 11, 15, 16, 21, 22, 23, 29, 34, 38, 43, 44, 45, 46, 47, 48, 49, 50, 51, 54, 55, and 56. Based on this assessment, 23 out of the 56 pitches in the etude were performed incorrectly.

3.2.1.4. Findings from the Analysis Conducted by the Software

Figure 16. Visual Representation of the Analysis Conducted by the Software

The figure illustrates the annotations generated by the software within its own interface for the Ney performance of the Bayati etude. The different colors represent the error rates in the articulation of pitches and rhythmic patterns. The error magnitude is expressed through a color gradient ranging from yellow to red, based on the level of inaccuracy. Red markings indicate that the corresponding pitch was either played completely incorrectly or not performed with a duration consistent with the usul. According to the software's findings, during the performance of the Bayati etude, the following pitches were played incorrectly: 5, 6, 11, 14, 15, 16, 20, 22, 23, 24, 25, 28, 29, 30, 32, 33, 34, 36, 38, 41, 43, 44, 46, 48, 50, 51, 54, 55, and 56. Based on this assessment, 29 out of the 56 pitches in the etude were performed incorrectly.

3.2.1.5. Overlaps and Divergences in Error Detection Between the Expert Group and the Software

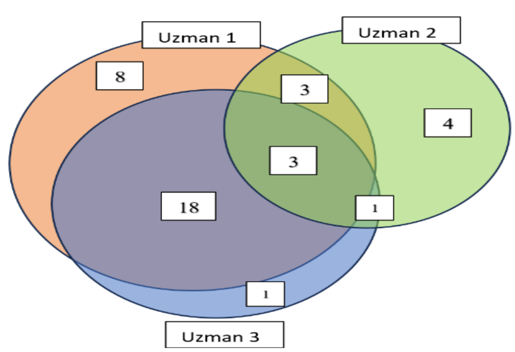

Figure 17. Venn Diagram Showing Overlaps and Divergences in Error Detection

The Venn diagram presents the overlapping and divergent results based on the analyses conducted by the expert group and the software. According to the diagram, subjective differences were observed among the experts' evaluations, which are believed to be related to variations in pitch perception stemming from their specific areas of expertise (Turkish classical music and Turkish folk music). On the other hand, the diagram indicates that the expert group and the software agreed on the incorrectness of a total of 18 pitches. This ratio demonstrates the precision of the software in error detection. The overlap ratios are provided in the table below.

Table 1. Overlap Ratios Between Expert Evaluations and the Software

Expert	Number of Notes Matching Software	Overlap Ratio (%)
Expert 1	20	60.6
Expert 2	6	54.5
Expert 3	19	79.1

An examination of the table shows that Expert 3's analyses exhibited the highest level of agreement with the software analyses at 79.1%. This is followed by Expert 1 with 60.6% and Expert 2 with 54.5%. These ratios indicate that there are individual differences in the evaluation criteria, which are thought to stem from variations in domain expertise.

In the graphical analyses, the most frequently marked incorrect pitches were identified, and recurring patterns of these pitches across all four sources (three experts and the software) were analyzed. Among the most frequently marked pitches were 5, 11, 23, 44, 50, 54, 55, and 56. These findings are considered important for understanding the pitches or rhythmic patterns that posed the greatest performance challenges.

3.2.2. Findings from Expert Group and Software Analyses of the Ud Performance

The student performed the Bayati etude on an Ud for the first time using the Bolahenk (Yerinden) tuning. The performance recording was captured in a single take. The Bayati etude consists of a total of 56 pitches (notes). The expert group listened to the performance recording and conducted accuracy analyses based on pitch and *usul*. The annotations made by the experts on the notation in a digital environment while listening to the audio recordings are presented in the figures below.

3.2.2.1. Results of the Analysis Conducted by Expert 1

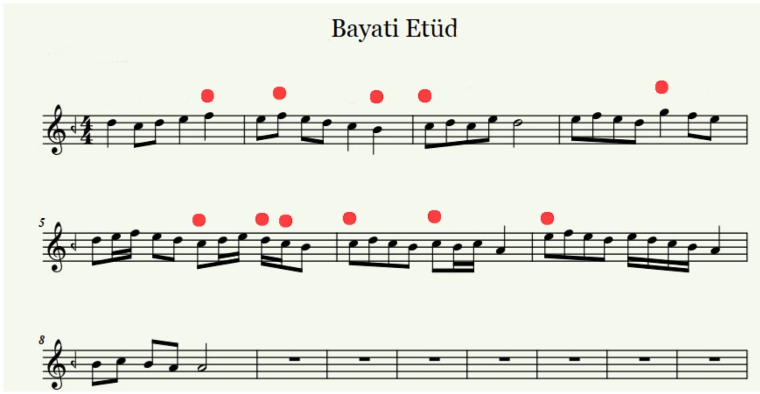

Figure 18. Visual Representation of the Analysis Conducted by Expert 1

The figure displays the annotations made by Expert 1 on the Bayati etude. The red markings indicate that the corresponding pitch was either played incorrectly or not performed with a duration consistent with the usul. According to Expert 1's findings, during the performance of the Bayati etude, the following pitches were played incorrectly or outside the appropriate duration: 4, 5, 6, 7, 8, 10, 11, 12, 14, 15, 17, 18, 21, 25, 26, 33, 34, 37, 38, 39, and 43. Based on this assessment, 21 out of the 56 pitches in the etude were performed incorrectly.

3.2.2.2. Results of the Analysis Conducted by Expert 2

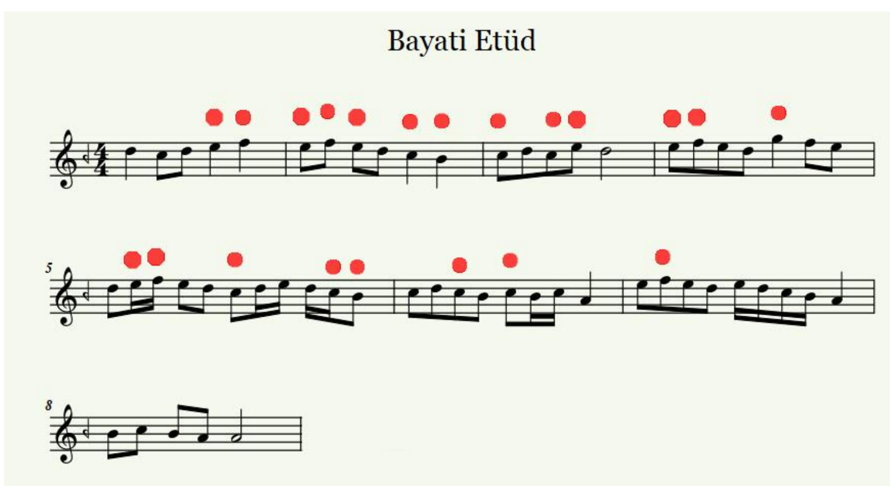

Figure 19. Visual Representation of the Analysis Conducted by Expert 2

The figure displays the annotations made by Expert 2 on the Bayati etude. The red markings indicate that the corresponding pitch was either played incorrectly or not performed with a duration consistent with the usul. According to Expert 2's findings, during the performance of the Bayati etude, the following pitches were played incorrectly or outside the appropriate duration: 5, 7, 11, 12, 21, 29, 32, 33, 35, 39, and 43. Based on this assessment, 11 out of the 56 pitches in the etude were performed incorrectly.

3.2.2.3. Results of the Analysis Conducted by Expert 3

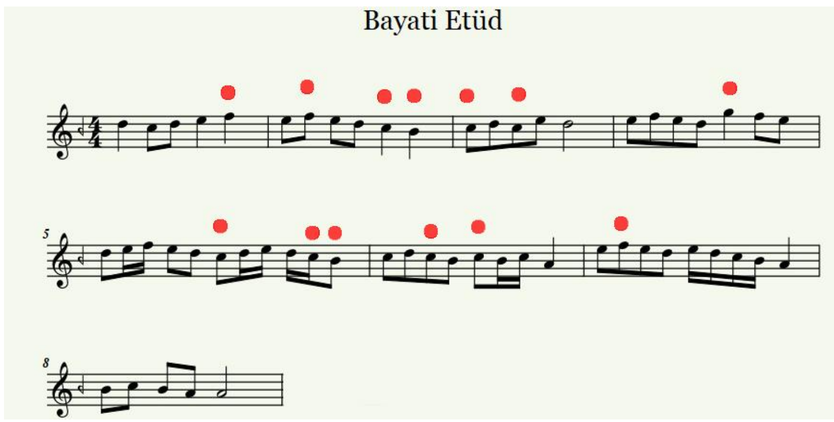

Figure 20. Visual Representation of the Analysis Conducted by Expert 3

The figure displays the annotations made by Expert 3 on the Bayati etude. The red markings indicate that the corresponding pitch was either played incorrectly or not performed with a duration consistent with the usul. According to Expert 3's findings, during the performance of the Bayati etude, the following pitches were played incorrectly or outside the appropriate duration: 5, 7, 10, 11, 12, 14, 21, 29, 33, 34, 37, 39, and 44. Based on this assessment, 13 out of the 56 pitches in the etude were performed incorrectly.

3.2.2.4. Findings from the Analysis Conducted by the Software

Figure 21. Visual Representation of the Analysis Conducted by the Software

The figure illustrates the annotations generated by the software within its own interface for the Ud performance of the Bayati etude. The different colors represent the error rates in the articulation of pitches and rhythmic patterns. The error magnitude is expressed through a color gradient ranging from yellow to red, based on the level of inaccuracy. Red markings indicate that the corresponding pitch was either played completely incorrectly or not performed with a duration consistent with the usul. According to the software's findings, during the performance of the Bayati etude, the following pitches were played incorrectly: 5, 6, 7, 9, 10, 12, 14, 16, 18, 21, 25, 26, 29, 30, 31, 32, 33, 34, 37, 40, 41, 44, 46, 47, 48, 49, and 50. Based on this assessment, 27 out of the 56 pitches in the etude were performed incorrectly.

3.2.2.5. Overlaps and Divergences in Error Detection Between the Expert Group and the Software

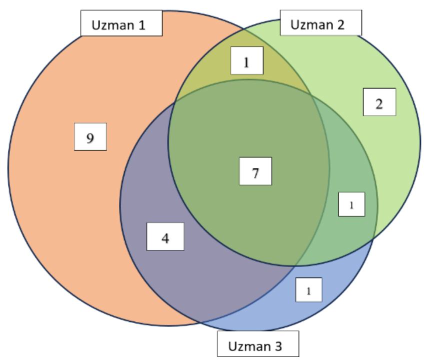

Figure 22. Venn Diagram Showing Overlaps and Divergences in Error Detection

The Venn diagram illustrates the overlapping and divergent results based on the analyses conducted by the expert group and the software. According to the diagram, subjective differences were observed among the experts' evaluations, which are thought to be related to variations in pitch perception stemming from their specific areas of expertise (Turkish classical music and Turkish folk music). On the other hand, the diagram indicates that the expert group and the software agreed on the incorrectness of a total of seven pitches. This ratio demonstrates the precision of the software in error detection. The overlap ratios are provided in the table below.

Table 2. Overlap Ratios Between Expert Evaluations and the Software

Expert	Number of Notes Matching Software	Overlap Ratio (%)
Expert1	19	59.4
Expert2	6	54.5
Expert3	18	78.3

An examination of the table shows that Expert 3's analyses exhibited the highest level of agreement with the software analyses at 78.3%. This is followed by Expert 1 with 59.4% and Expert 2 with 54.5%. These ratios indicate that there are individual differences in the evaluation criteria, which are thought to stem from variations in domain expertise.

In the graphical analyses, the most frequently marked incorrect pitches were identified, and recurring patterns of these pitches across all four sources (three experts and the software) were analyzed. Among the most frequently marked pitches were 5, 7, 11, 12, 21, 29, 33, 39, and 43. These findings are considered significant for understanding the pitches or rhythmic patterns that posed the greatest performance challenges.

4. Conclusion, Discussion, and Recommendations

In this study, which aimed to develop a specialized artificial intelligence-supported software using the Python programming language to analyze pitch and rhythmic patterns in Turkish music performances, the following results were obtained:

- The software successfully identified general tendencies in both instrument groups, such as pitch and rhythmic inaccuracies.
- During the implementation phase, the software accurately detected the correctness or incorrectness of specific pitches (microtonal) in the melodic movement of the Bayati maqam.
- The software successfully identified the correctness or incorrectness of rhythmic patterns (*usul*) within the melodic structure of the Bayati maqam selected for the implementation phase.
- In the analysis of Ney performances, the software produced results that closely matched the expert evaluations and detected a higher number of errors compared to the experts. This demonstrates the sensitivity of the program in detecting pitch and rhythmic inaccuracies.
- In Ney performance, the total number of errors was higher than in Ud performance, and discrepancies were observed among experts in terms of error detection. This is thought to result from the technical challenges associated with the absence of fixed pitches in the Ney and from deficiencies in student performance.
- In the analysis of Ud performances, the software produced results that closely matched the expert evaluations and also detected a greater number of errors than the experts, again indicating the software's precision in identifying pitch and rhythmic inaccuracies.

- The total number of errors in Ud performance was lower than in Ney performance, and agreement among experts in error detection was relatively higher.
- The data indicate that the software's analytical capability may vary depending on the sound characteristics of different instruments and that some instruments may yield higher accuracy rates in error detection.
- The analysis of expert evaluations revealed that interpretive differences in musical expression could directly influence error detection. The fact that the same pitch was evaluated as correct by one expert and incorrect by another demonstrates the impact of individual interpretative perceptions on the assessment process in Turkish music.
- Based on these interpretive differences, it was concluded that the software system should be developed to detect not only technical accuracy but also interpretive variations in performance.

In conclusion, the developed software provides a structure capable of technically analyzing beginner-level student performances in terms of pitch and rhythmic accuracy. However, the system still requires enhancements to better capture interpretive differences, to be trained with a larger dataset of expert evaluations, and to be tested across various maqams and a wider range of musical pieces.

The Python-based audio analysis software developed in this study presents an original approach aimed at identifying microtonal structures and rhythmic inaccuracies specific to Turkish music. The findings show that while the system aligns strongly with expert evaluations in terms of technical accuracy, it does not fully account for interpretive performance variations. This highlights the necessity of incorporating performance style (*tavır*) and interpretive elements into the software to improve its analytical capabilities.

In the comparative analysis of two instruments, Ney performances showed both a higher error rate and greater divergence among expert evaluations. This can be attributed to the significant influence of interpretive control on the Ney, where sound production relies heavily on breath, diaphragm, and oral technique (Bozkurt et al., 2014). Conversely, Ud performances exhibited higher levels of agreement between software and expert evaluations, indicating that interpretive variability played a less prominent role. This suggests that the software may perform better when analyzing instruments with relatively stable pitch structures.

The highest overlap ratio between the software and Expert 3 further reveals that evaluation criteria vary among experts. This indicates that factors such as individual listening habits, professional experience, pedagogical approach, and musical style significantly shape assessment outcomes. These differences are directly related to the nature of Turkish music, which allows for multiple interpretive approaches rather than a single standard of correctness (B. Çini, 2024).

Additionally, the software provides the ability to track student performance graphically, not only through frequency analysis but also by analyzing temporal characteristics (note duration and timing) and ornamentation elements (glissando, vibrato, etc.). This feature supports individual music education by offering measurable feedback and progress tracking, while also providing instructors with detailed observation and evaluation tools.

However, the software was tested only with beginner-level performances, indicating that its effectiveness in complex compositions and advanced interpretations remains unverified. The system should be expanded by incorporating more expert evaluations and recordings from different instruments. Based on the results of this study, the following recommendations are proposed:

- Tests for Advanced Performances: The software should also be tested with intermediate and advanced-level performances to assess its limits in more complex musical contexts.
- Development of Ornamentation and Style Detection Algorithms: The system should be capable of detecting ornamentation elements (glissando, vibrato, trill, etc.) and interpretive styles essential in Turkish music. Machine learning-based models should be integrated for this purpose.
- Integration of an Educational Module: The pedagogical functionality of the software should be enhanced by adding a module that provides hints for incorrectly performed passages and supports repetitive practice.
- Instrument-Specific Algorithms: Parametric filters and trained models tailored to the acoustic characteristics of instruments such as Ud, Ney, and violin should be developed to improve analytical accuracy.
- Improved Firebase Integration: A dashboard capable of generating statistical analyses of performance records and providing progress charts over time should be implemented to enhance performance tracking features.

5. Acknowledgement

This article has been generated from the first author's Master's Thesis titled "Analysis of Pitch and Rhythmic Structures in Turkish Music Using the Python Programming Language", which was supervised by the second author.

References

Abidin, D., Özacar, T., & Öztürk, Ö. (2018). Usıng Classıfıcatıon Algorıthms for Turkısh Musıc Makam Recognıtıon. Selcuk University Journal of Engineering, Science and Technology, 6(3), 377-393. https://doi.org/10.15317/Scitech.2018.139

Akker, J. V. den, Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). Educational Design Research. Routledge.

Barab, S. A., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. Journal of the Learning Sciences, 13(1), 1-14. https://doi.org/10.1207/s15327809jls1301-1

Bozkurt, B. (2008a). An Automatic Pitch Analysis Method for Turkish Maqam Music. Journal of New Music Research, 37(1), 1-13. https://doi.org/10.1080/09298210802479252

Bozkurt, B. (2008b). An experiment-based pitch interval analysis of Turkish maqam music. Journal of New Music Research, 37(1), 1-13.

Briot, J.-P., Hadjeres, G., & Pachet, F.-D. (2020). Deep learning techniques for music generation. Springer.

De Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. The Journal of the Acoustical Society of America, 111(4), 1917-1930. https://doi.org/10.1121/1.1458024

Downie, J. S. (2003). Music information retrieval. Annual Review of Information Science and Technology, 37(1), 295-340.

Gedik, A. C., & Bozkurt, B. (2010). Pitch-frequency histogram based music information retrieval for Turkish music. Signal Processing, 90(4), 1049-1063.

Huron, D. (1999). Music in the era of cognitive science. Musicae Scientiae, 3(1_suppl), 5-36.

Karaosmanoğlu, M. K. (2012). A symbolic music representation for Turkish makam music. Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR).

Levitin, D. J. (2006). This Is Your Brain on Music: The Science of a Human Obsession. Dutton.

Mutlu, N. (2016). Qualitative Research Methods for Production Based Studies in Education: Anı Publishing.

Plomp, T. (2013). Educational design research: An introduction. İçinde T. Plomp & N. Nieveen (Ed.), Educational Design Research. SLO.

Reeves, T. C. (2006). Design research from a technology perspective. İçinde J. V. den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Ed.), Educational Design Research (ss. 52-66). Routledge.

Richey, R. C., & Klein, J. D. (2007). Design and Development Research: Methods, Strategies, and Issues. Routledge. https://doi.org/10.4324/9781410614701

Serra, X. (2011). A multicultural approach in music information research. Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR).

Sloboda, J. A. (2005). Exploring the Musical Mind: Cognition, Emotion, Ability, Function. Oxford University Press.

Ünal, E., Bozkurt, B., & Karaosmanoğlu, M. K. (2014). A Hierarchical Approach to Makam Classification of Turkish Makam Music, Using Symbolic Data. Journal of New Music Research, 43(1), 132-146. https://doi.org/10.1080/09298215.2013.870211