The Co-Poem Project: Didactic Resources and Pedagogical Projection for Musical Education in Primary School

El proyecto Co-Poem: recursos didácticos y proyección pedagógica para la educación musical en Educación Primaria

Javier Merchán Sánchez-Jara[a*], Sara González-Gutiérrez[b], María Navarro-Cáceres[c], Matilde María Olarte-Martínez[d], Concepción Pedrero-Muñoz[e]

[a] *Departamento de Didáctica de la Expresión Musical, Plástica y Corporal, Universidad de Salamanca, España.*
https://orcid.org/0000-0003-1828-5182   javiermerchan@usal.es

[b] *Departamento de Didáctica de la Expresión Musical, Plástica y Corporal, Universidad de Salamanca, España.*
https://orcid.org/0000-0002-5706-4705   saragnzlz@usal.es

[c] *Departamento de Informática y Automática, Universidad de Salamanca, España.*
https://orcid.org/0000-0003-2344-1012   maria90@usal.es

[d] *Departamento de Didáctica de la Expresión Musical, Plástica y Corporal, Universidad de Salamanca, España.*
https://orcid.org/0000-0001-8722-9446   mom@usal.es

[e] *Departamento de Didáctica de la Expresión Musical, Plástica y Corporal, Universidad de Salamanca, España.*
https://orcid.org/0000-0002-9151-7296   cpedrero@usal.es

ARTICLE INFO	ABSTRACT

Orally transmitted traditional music constitutes an invaluable intangible legacy for understanding the development of local cultures and observing social dynamics over time. The Co-Poem project aims to make a significant corpus of these musical cultures available to any user in the Mediterranean basin by creating the I-Folk database, which allows access and consultation through semantic technologies and artificial intelligence, to promote knowledge of this cultural reality and perpetuate its intergenerational transmission. The conceptualization of both resources is oriented to the practice of music teaching in primary education. From this point of view, the structure of metadata that musically describes each work and the implementation of access resources to the corpus under study are analyzed from the point of view of the pedagogical function. All these issues are materialized in creation of various educational resources based on the possibilities of the I-Folk database to promote meaningful and interactive learning of musical content (melody, rhythm, dance, etc.) based on the defining characteristics of this type of repertoire. Recent interventions and didactic activities in primary schools show that both the I-Folk database and the associated resources and the search and comparative analysis functionalities between repertoires from different territories constitute a tool highly appreciated by teachers and arouses great interest among students.

RESUMEN

La música tradicional transmitida oralmente constituye un legado de gran valor para comprender el desarrollo de las culturas locales y observar las dinámicas sociales a lo largo del tiempo. El proyecto Co-Poem pretende poner a disposición de cualquier usuario un corpus significativo de estas culturas musicales en el entorno de la cuenca mediterránea, mediante la creación de la base de datos I-Folk que, a través de tecnologías semánticas e inteligencia artificial, permite el acceso y consulta con el fin de promover el conocimiento de esta realidad cultural y perpetuar su transmisión intergeneracional. La conceptualización de ambos recursos está orientada a la práctica de la enseñanza de la música en la educación primaria. Desde este punto de vista pedagógico, se

(*) Corresponding Author

analiza la estructura de metadatos que describen musicalmente cada obra y la implementación de recursos de acceso al corpus objeto de estudio. Todas estas cuestiones se materializan en la creación de diversos recursos educativos basados en las posibilidades de la base de datos I-Folk, con la intención de promover el aprendizaje significativo e interactivo de contenidos musicales (melodía, ritmo, danza, etc.) a partir de las características definitorias de este tipo de repertorio. Las intervenciones y actividades didácticas recientes en las escuelas primarias muestran que tanto la base de datos I-Folk y los recursos asociados, como las funcionalidades de búsqueda y análisis comparativo entre repertorios de diferentes territorios, constituyen una herramienta muy apreciada por los docentes y que despierta un gran interés entre los alumnos.

1. Introduction

Folk music of oral tradition is a repertoire of popular roots, which is transmitted orally from generation to generation, and which is part of the daily life of societies. In addition, it represents an intangible element of its cultural heritage and binds together different social and cultural identities.

It should be recalled that since the consolidation of Ethnomusicology as an academic discipline, it has been strongly engaged in assembling and transcribing folk music of oral tradition in songbooks to ensure its preservation since the turn of the century. In this sense, the songbooks were born as a collection of traditional pieces from different periods and geographical areas and even contain songs that had already disappeared in their own rural environments where they were conceived (Navarro Cáceres, 2018). In other words, we can affirm that the songbooks act as staunch conservators of tradition. However, the conception of tradition to which they refer is a written tradition and, therefore, static. They only act as a photography of a specific time and geographical location, with little social projection, if any at all, when the music contained inside is not put into practice. Our project pursues as its ultimate goal, to value folk music of oral tradition not as a fixed repertoire, but as a living tradition that opens the doors to the creativity of children because precisely the inventive and improvisational capacity in the popular environment has been a determining factor for the increase of the popular repertoire itself (Navarro Cáceres, 2018).

In the same approach, emerging ethnomusicologists discerned the compositions' pedagogical potential. One of the most influential figures in this discipline was Zoltán Kodály, who devised a theory to improve music pedagogy. Other renowned specialists have underwritten their insights on music, ethnic culture, and musical resources. Thanks to them, it has been conceivable to bring various folk repertoires closer to mainstream musical education in this regard (Campbell, 2016).

Ethnomusicologists in the Iberian Peninsula have also helped to establish the inclusion of this sort of content into music education and its constant involvement in the various phases of mandatory education. In recent years, the Portuguese and Spanish authorities have become aware of the need to promote this type of popular demonstration. In the context of the depopulation of rural areas, they stand as a form of claiming social identity and a way of preserving the intangible heritage of peoples. For this reason, they have begun to show interest and support to social initiatives for the dissemination of folk music, which contributes to reduce the generation gap in its transmission chain (Manzano, 1990) and to establishing agents that stimulate a culture of great relevance in the rural world. In addition, the promotion of this repertoire provides the opportunity to awaken an interest in local, regional, and national history and aspects of traditional culture in the youngest.

However, the transfer of knowledge from these initiatives to the educational field has not been carried out with the expected success. Analyzing the situation, there are a series of factors that justify this paradox:

1) The lack of specialization of music teachers in the music of oral tradition. This fact is a consequence of the educational planning of Music studies in Higher Education since they have focused, persistently, on the teaching of Western music of a cultured tradition, traditionally known as classical music.
2) Materials are scarce, difficult to access (Manzano, 1990; Merchán Sánchez-Jara, 2016), and in most cases, they are fragmented and unfinished, making it difficult to bring didactics improvements. In this sense, this difficulty is increased by the lack of practical proposals.
3) Most professors prefer to seek multimedia content on the Web rather than in bookstores or libraries. Even when teachers find ethno music on digital repositories like the CSIC website, teachers are confused about how to introduce it into their master lectures to teach students about musical features like rhythms and melodies. In most situations, the scores lack musical analysis (Fondo de Música Tradicional, 2023).

As a result, we present a model for analyzing the folk music properties. This model relies on the Music Encoding Initiative (MEI - https://music-encoding.org/), a community-driven initiative that uses the XML standard to encode music content. Because of its semantic and flexible character, we chose MEI as our standard. This method has a significant benefit in relation with the format it employs: XML language, that is readable on virtually any device that could identify the tree structure. On the other hand, we revised several of the suggested categories to include folk music traits, as MEI is chiefly concerned with classical music, and therefore needed a readaptation towards this type of work.

The initial stage for the elaboration of our model was to locate different songbooks of the folk music tradition of the Iberian Peninsula that contained children's repertoire and lullabies. About 800 pieces were digitized, which were transcribed in XML format, later into scores, and manually analyzed by the project's musicologists, who were also in charge of developing, through the MEI standard, the template for encoding the music information.

The second stage consisted of analyzing and coding a sample of the corpus made up of a total of 350 songs (which will be increased in future stages of the research); 250 pieces of the Spanish tradition, from Asturias, Castilla y León, Catalonia, Aragon, and Andalusia, and 100 pieces of Portuguese tradition. We must emphasize that, at present, we have already begun to incorporate Italian songs with the same purpose; incorporate them into compulsory education and be able to compare repertoires, versions, and traditions.

The Co-POEM project website (http://copoem.usal.es/) contains the specific files, compilers, and other ethnomusicological data for the analyzed songbooks. Thanks to this project we show a prototype of an open-access digital database, called I-Folk: It allows specific searches for folk songs, both by searching for similarity and by musical characteristics such as genre, style, meter patterns, or geographical territories. In short, this tool aims to promote the dissemination of folk music in educational contexts.

In turn, we developed internal tools to retrieve and view all encoded content. We created a common searching engine to retrieve data on musical and ethnomusicological characteristics (author, compiler, tempo, key...) and a search for similarities to understand similarities between two songs in relation to their melody developments, their rhythmic pattern or both of them. We believe that these two methods of consultation will enable professors to retrieve materials in their institutions that are appropriate for their purposes, therefore expanding and improving their teaching.

The structure of this paper is organized as follows: the second section explains the inherent qualities of the gathered Spanish and Portuguese popular songs, as well as the MEI model adjustments applied to better incorporate them. In addition, it explains the educative potential of this initiative and how it would perform. Section 3 explains how we encoded and browsed the I-Folk database using an openly accessible web site. The fourth section delves into the creation of database navigation based on similarity inside the online site. Finally, the 5th section summarizes the findings and suggests future research directions.

2. Music Representation

Understanding the intrinsic characteristics of traditional Iberian music is a determining factor in achieving an adequate representation of said music. Consequently, this section contains both a detailed description of the repertoire and the details of adapting MEI tags to meet these needs.

2.1. Depiction of Iberian Folk Music

Following, starting from reference musicological studies (Manzano, 1990; García-Matos & Bargalló, 1982; Olmeda, 2001) and corroborated by our own experience with the analysis of songbooks, we will describe some of their most relevant characteristics and stylistic features:

- The melodic development of the works is guided by modal systems. As a result, the significance of melodic structures rises dramatically.
- The melodic gamut is narrowed; in most cases, the melody travels in a 5th or 6th interval, and in rare cases, it transcends the 8th.
- The presence of unstable or chromatic notes is quite common; notes that don't track the guidelines of academic music.

- Harmony is either non-existent or located outside of the cadential space. The prosodic rhythm is predominant in these pieces. Particularly in the dances, the steady rhythms prevail. Rhythmic patterns associated with a given genre and measure can be found throughout the heritage.
- The musical structure is usually built on chorus lyrics and stanzas being repeated.

As it has been previously commented, it is a repertoire that is disseminated orally, based on the collective memory of the individuals of their society and, therefore, plays a determining role in its preservation. However, while it is true that memory is the main pillar of its praxis and dissemination, it is also certainly limiting, nonetheless. The most common variations between similar pieces are mainly due to small memory deviations; changes produced by variation, substitution, or borrowing.

- Changes by variation: Modifications of the melody produced by the interpreter with a generally ornamental intention.
- Changes by substitution: Modifications of musical fragments or ideas by exchange with another that is considered new or interesting.
- Changes by borrowing: Modifications of the melody, this time, produced by the assimilation of resources or styles from other musical contexts.

In this sense, through the codification of this type of repertoire and its subsequent dump in the database, we hope to favor the work of ethnomusicologists, since the comparative data of versions can contribute to uncovering the gears of creativity and spontaneous composition of the performers of folk music.

On the other hand, we may establish that, despite its closeness to the tonal system, the modal paradigm that predominates in the songbooks is attributed to the development of medieval cultures after consulting and digging into diverse research of professionals in the subject, such as that of musicologist Miguel Manzano.

For its part, the rhythmic and melodic organization of the pieces is usually determined by a text composed of eight-syllable quatrains. Several research projects (Manzano, 1993; García-Matos, 1963) reveal that folk music and cultured music present similarities between rhythmic patterns, melody lines of the songs, musical structures conditioned by the lyric rhymes or alternation of the soloist singer and the response of the listeners in small melodic motifs and repetitions. Nonetheless, other authors such as Olmeda (2001) affirm that this music's despite sharing characteristics in origin underwent very different evolutions. Along such lines, folk music has embraced different sonorities and themes as a consequence of its oral disposition; for example, this repertoire was overshadowed by flamenco "cante" and Zarzuela at the turn of the nineteenth century, and by the consolidation of the media (television and radio) in the mid-1900s.

Accordingly, the notation adopted for ethno music transcription has been inconsistent through periods and hasn't been regularized. Thus, the preservation has been a troublesome task, as only the pieces with a strong oral tradition and those that were systematically recorded at the beginning of the last century have been conserved. Unlike classical music, in which a composition is synthesized and represented in a score according to a system of rules, in folk music, each compiler developed his unique solution to handle the specifics of the music (neutral notes/microtones, ornamentation, and so on). The effort of our work consisted in making a standard interpretation of the sheet music of the songbooks to introduce these songs as a systematic process into the database.

As a result, we must not overlook the existing dissimilarities that have prompted us to alter our MEI model to better accommodate folk music. In the same way, adapting the model will make it easier for us to actively comply with the educational objectives that we have previously established:

1. Spread the importance of preserving our folk music as part of our cultural heritage in the new generations.
2. Improve the knowledge of young people about folk music in the corresponding regions.
3. Study and design teaching-learning mechanisms to incorporate the musical experience into the process of composing music from different musical models.
4. Apply new technologies to make the concept of folk music more attractive to children.
5. Train music teachers to convey the importance of folk music in our culture.
6. Analyze and design validation mechanisms for results in real educational settings.

The objectives are due to the fact that we detected an insufficient presence of folk music in the Music subject curricula in Primary Education in Spain and Portugal and the teacher training curricula in Higher Education. On the other hand, although it is true that the educational legislation indeed contemplates new technologies as a

basic competence to be introduced in the classroom, it must be understood as a cross-cutting and structuring element of the contents. The new technologies have been manufactured as an effective resource that allows complex interrelationships to be carried out in a simple way and, in turn, allow the levels of abstraction of the subject to be adapted to explain complex concepts such as, for example, the modality in music. Likewise, if we take into account the playful factor of digital resources (García-Peñalvo, 2015), they are the perfect compromise to bring a repertoire such as folk music closer to the new generations.

Once these precepts have been formulated, it is necessary to develop proposals that facilitate and promote the integration of folk music in the classroom, since, until now, the main approaches that have been made between folk music and digital technologies have revolved around the digitization and its compilation in digital format. Our initiative, taking advantage of the possibilities offered by the semantic coding of the database corpus, highlights a series of didactic activities that offer different alternatives to work on musical creation, respond to the musical stimuli, and perform (Bauer, 2020).

In short, it can be stated that the relationship between music teaching, ethnomusicology, and the introduction of digital resources is bidirectional and reciprocal. Teachers can benefit from the didactic possibilities that traditional music possesses, but ethnomusicology also finds in music education a way to spread and keep it alive, ensuring generational change in the chain of transmission.

2.2. A folk music database proposal encoded with MEI

The MEI Guidelines standard allows us to incorporate specific details about encoded folk music, musical and textual components, and knowledge implicit in the scores such as range and rhythmic pattern in the header and body of our template. On a melodic, harmonic, or rhythmic basis, the premise is to identify recurring themes among compositions from the same geographic region or even genre. It should be highlighted that the type of comparisons the system allows us to carry out opens a wide range of possibilities for future lines of research. For example, it is relevant to analyze the differences between variants of the same piece of music, as they can show, even when they were collected at the same time, important musical differences that allow certain conclusions to be drawn (Navarro Cáceres, 2018). In addition, comparing stylistic features can help to define the origin and/or approximate date of the songs.

In the following lines, we detail the template of our project:

The header starts with a brief description of the file, which comprises labels that compile music belonging to different contexts, in addition to involving a large number of professionals in their preparation. Knowing this type of singularity typical of these sources is of vital importance in the decision-making of coders (Carvalho et al., 2021). In this sense, we find, for example, the same songs with significant discrepancies in songbooks with the same publisher on diverse publication dates. To handle each tome in the database, an XML identifier, incorporating this information (publisher \ publication date \ geographic area), is attached. The remaining bibliographic data of the source is mined from the work's main title, using the @role attribute to describe the functions of each authority.

```xml
<sourceDesc>
    <source xml:id="PedrellSabaté1948BCN">
        <biblStruct>
            <monogr>
                <imprint>
                    <title>Cancionero musical popular español</title>
                    <title type="subordinate"></title>
                    <title type="desc"></title>
                    <respStmt>
                        <persName role="recopilador">Felipe Pedrell Sabaté</persName>
                        <persName role="edición, introducción e índices"></persName>
                        <persName role="Biografía, bibliográfia, y discografía"></persName>
                    </respStmt>
                    <publisher>Boileau, Editorial de Música</publisher>
                    <pubPlace>Barcelona</pubPlace>
                    <date>1948</date>
                    <extent type="pages">270</extent>
                </imprint>
            </monogr>
        </biblStruct>
    </source>
</sourceDesc>
```

Musical information such as key, tempo, or musical and textual incipit is encoded with the corresponding tags through the @type attribute since it is data that does not happen in the original source and has been subsequently accomplished by the editor.

```xml
<work>
    <title type="main">2) Nana</title>
    <author type="informante"></author>
    <incip type="lyrics">
        <incipText>
            <head> A la roró mi niño</head>
        </incipText>
    </incip>
    <key mode="phrygian">G</key>
    <meter>ternary</meter>
    <tempo>Larghetto</tempo>
```

Despite what it may seem at first, unlike in classical music, including the textual incipit is not a trivial decision. The pieces of popular music of oral tradition that do not have a standardized title are named thanks to said incipit. The use of @xml: id establishes the association between the incipit and the matching role as a simulated title, and allows to identify of different versions among the analyzed corpus. In addition, a musical beginning is included, consisting of a series of measures with the first significant musical idea.

Labels related to traditional Spanish music genres and geographical origins are basic operational elements in all pieces. However, there are no specific labels to encode the delimitation of specific musical forms. Consequently, and to respond to this need, the @type attribute is used. In this case, our ultimate goal is to develop a taxonomy of shapes and styles in order to analyze the characteristics elements.

```xml
<incip type="musical">
    <score>
        <section>
            <measure copyof="#dle95"></measure>
            <measure copyof="#dle518"></measure>
        </section>
    </score>
</incip>
<langUsage>
    <language xml:lang="es"></language>
</langUsage>
<notesStmt>
    <annot>Pieza vocal tradicional con acompañamiento de piano del propio recopilador</annot>
</notesStmt>
<classification>
```

Lastly, the <editorialDecl> tag is envisioned to gather the editorial statement in case it is needed. For example, the encoder may declare that in this space it has detected a transcription error or that it specifies that the piano harmonization of a piece has been carried out by the compiler himself.

```xml
<measure n="1" xml:id="dle95">
    <staff n="1">
        <layer n="1">
            <rest xml:id="dle127" dur="8"/>
            <note xml:id="dle140" pname="a" oct="4" dur="8" stem.dir="up">
                <verse n="1">
                    <syl x="1.316">A</syl>
                </verse>
            </note>
```

It should be kept in mind that we have decided to go with the current way of grouping and notating eighth and sixteenth notes (as presented in Figure 1), which will facilitate the performing of the songs.

Figure 1. Old loose figure notation replaced by the current notation.
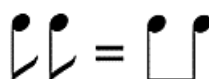
Structuring the template body also posed an organizational challenge. The necessary amendments are mostly connected to the key aspects assessed and interpreted by the coders, for example, with the rhythmic

pattern or the rest notes. The analysis and comparison of these characteristics may reveal the relationship between these variables in the songbooks and which ones are recurrently frequent, regardless of genres and territorial contexts.

```
<body>
    <mdiv>
        <score>
            <scoreDef meter.count="3" meter.unit="4" key.sig="1f">
                <staffGrp>
                    <staffDef n="1"
                        xml:id="P1"
                        label="Voice"
                        lines="5"
                        clef.line="2"
                        clef.shape="G"
                        key.sig="1f"/>
                    <staffGrp xml:id="P2" symbol="brace" label="Piano">
                        <staffDef n="2" lines="5" clef.line="2" clef.shape="G" key.sig="1f"/>
                        <staffDef n="3" lines="5" clef.line="4" clef.shape="F" key.sig="1f"/>
                    </staffGrp>
                </staffGrp>
                <ambitus>
                    <ambNote type="lowest" oct="4" pname="a"></ambNote>
                    <ambNote type="highest" oct="5" pname="f"></ambNote>
                </ambitus>
            </scoreDef>
```

The encoders further describe the information in regard to the rhythmic patterns that occur well before work's initial measure. The aim is to discover and label the rhythmic motifs that constitute each song's structure in order to spot common characteristics between genres, rhythmic-melodic elements, and/or geographical territories, as discussed previously.

```
<supplied type="rythymn pattern">
    <note dots="1" dur="4"></note>
    <note dur="8"></note>
    <note dur="4"></note>
</supplied>
```

Also, encoders mark musical phrase structures with the <supplied> tag. The beginning and ending points of each of the autonomous musical ideas are recognized in this example. Concurrently, the @type attribute is being to categorize various types of musical phrases to build the melody of the composition on a scale. In general, providing a separate capital letter to each musical notion with its own autonomy and musical semantics is the most acceptable criteria.

```
<supplied>
    <phrase n="1" startid="#d1e140" endid="#d1e1456" type="A"></phrase>
    <phrase n="2" startid="#d1e1488" endid="#d1e2893" type="B"></phrase>
</supplied>
```

The cadences are included in the endnote of each phrase with a @type attribute. The R code is used for the rests and the SR code for the semi-rests.

```
<note xml:id="d1e1456" pname="c" accid="s" oct="5" dur="4" stem.dir="down" type="R"/>
```

Finally, each of the musical notes corresponds to one of the syllables of the song lyrics. In addition, replacement texts are included in the header within <noteStmt>, along with additional comments existing in the original score that are generally related to the circumstances of the music's performance.

3. Structure and access point of the I-Folk Database from the point of view of use in educational settings

Having established the data model that supports the codification of the texts that will feed the database, we proceed to make explicit and describe the process of implementation of resources, fields, and access points that are

considered optimal for the questioning and retrieval of the information. Furthermore, the strategies followed to achieve fast access, agile information retrieval and easy navigation will also be described; all this with the main objective of reconciling completeness and accuracy in searches.

In addition, special consideration is given to the need to simplify the arrangement of elements in the interface because the group of potential users of the resource includes, among others, elementary school students (as presented in Figure 2).

Figure 2. Example of the database browser interface.
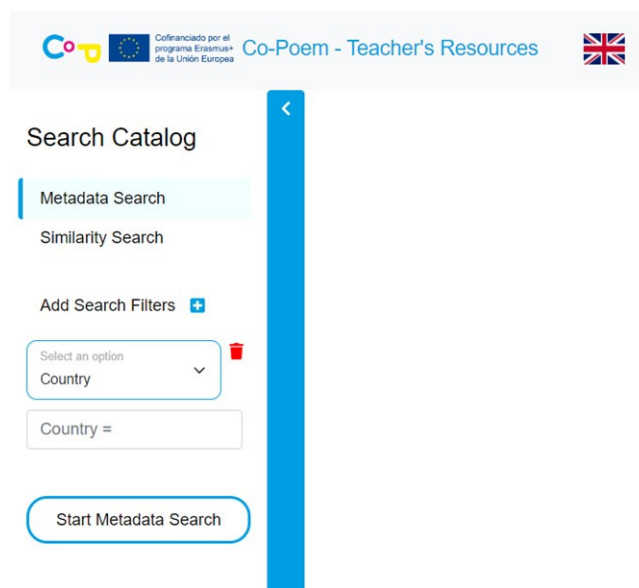
For this purpose, a web structure is designed and implemented to allow public access to the database. It establishes the necessary elements to launch advanced searches in which musical and ethnomusicological criteria such as modal scope questions, characteristic genres, rhythmic cells and motifs, melodic contours, etc. can be combined. Some of the most significant issues of the internal structure of the resource are:

1. A query interface that incorporates combinable access points to retrieve songs according to the level of coincidence concerning user-defined requirements.
2. A data visualization according to levels of similarity between repertoire items in which all songs are displayed on the screen, in groups or clusters defined by a series of specific. From a pedagogical and educational point of view, in line with the final objective of the project under the database we present is developed, which is the knowledge and learning in a meaningful way of popular music of oral tradition, the fields in the database have been conceived to be able to retrieve musical information according to access points that allow us to offer results based on essential questions concerning the music education curriculum in the environment.

On the one hand, the most essential issues related to auditory training or vocal and/or instrumental performance such as rhythmic or melodic structures and patterns. These materials are essential for the development of didactic materials of a progressive nature (by virtue of the difficulty or issues to be dealt with) which have as a substratum the components which essentially characterize the repertoire of popular music of oral tradition. The fact of being able to work on essential issues of the curriculum explicitly through a thematic repertoire of these characteristics allows, as one of the essential didactic objectives of the project, to introduce and favor the assimilation in the pupils of the most peculiar characteristics of this type of repertoire through a creative and musically significant approach.

On the other hand, the system can be questioned in terms of musically more complex or abstract issues such as the paradigmatic genres of the repertoire, on the one hand, or the melodic/harmonic characterization of the songs with issues such as tonality, modality, etc., on the other. In the first case, this aspect is essential to be able to work with students on issues such as the social projection of music and its role in the development of the daily and cultural activities of peoples and territories, limited, in this case, to issues related to children and adolescents.

The functionalities of the database about the second question, which is certainly more complex, make it possible to recover songs within the same or different harmonic/melodic scope (generally at the modal level) in conjunction with other dimensions such as genre, tempo, or rhythmic motifs. This question makes it possible not only to adapt the range or tessitura of the songs to the requirements of classroom activities but also to exemplify within the same framework the sonic characteristics of each mode or scale.

3.1. Database Structure functional projection to the educational and pedagogical environment

Retrieving information from the I-Folk dataset is carried out by means of an API using FastAPI (https://fastapi.tiangolo.com/). To query the structured data, it has been applied an interface with Javascript technology (by means of Vue 3 – https://v3.vuejs.org). Encoded information and semantic data obtained through the MEI encoded data is processed on the side of our Python application.

The bibliographic data that musically characterize each song encoded in the MEI Header section is transformed into the fields and access points that make up the database infrastructure. In order to develop the marking of each song, as a taxonomic unit of study within the different repertoires that make up the corpus of analysis, a data model has been created for the MEI Header (as detailed in point 2.2) that incorporates hierarchically (from the most general to the most particular aspects), information related to the description of sources and works, contextual, cultural or musicological characteristics, information related to the musical logic at an analytical level. Also, it has been included any other issue that may be considered relevant from the point of view of the identification, characterization or understanding of the work and the cultural context in which it is transmitted (Merchán Sánchez-Jara, 2019). On the other hand, in the <music> container, a marking process has been systematically developed with the aim of identifying all the instances and occurrences that significantly contribute to characterizing the rhythmic, melodic or harmonic aspects of the work.

More specifically, for each of the encoded songs that make up the corpus, the following files and information have been stored:

a) The digital edition of each song is encoded in the MEI standard, containing descriptive information on works, songbooks, editorial and encoding processes, as well as the marking related to the identification of musical properties and events within the musical text itself. Each of these texts (corresponding to a song as a unit of analysis) has been encoded in MEI by means of the following process: first, digital reproduction of printed sources is made through a professional scanning process. Next, the file was subjected to optical character recognition using OMR (Optical Music Recognition) technology, resulting in a file in MusicXML format. Finally, through a process based on XSLT script, it is transformed into the MEI format, which is then subjected to the semantic markup process.

b) a compressed Python pickle file, where each and every occurrence that the songs at the musical level is analyzed in the environment of a viewpoint system related to the logic and properties of the music. This multiple viewpoint system (Conklin & Witten, 1995) forms a standardized string-based representation, which develops the representation of musical content in relation to basic aspects of musical logic such as pitch and duration, by virtue of the inclusion of contextualizing information, structured at a second level, that refers, for example, to the positioning of each musical event within the song structure, the time signature, or the intervallic relationship between the different notes. These systems use domain knowledge to formulate different forms of representation to encode temporal occurrences derived from musical logic by abstracting different kinds of properties. Some of these viewpoints can be established from others, which can be defined as elementary viewpoints. These data can be used in later stages to recreate modeled musical structures that allow the automated generation of musical sequences based on a series of specific characteristics (Carvalho & Bernardes, 2020) or to establish the different levels of similarity between melodic within a song.

This functionality related to the possibility of generating new musical sequences based on questions such as the identification of characterizing elements has an extraordinary value for the achievement of one of the most complex curricular objectives from a didactic point of view: creation. The possibility of implementing functionality for guided composition based on these questions, where the student can receive suggestions proposed in an automated way by the system as a tool to guide the process, constitutes a valuable resource not only to highlight practically the most essential questions of musical creation but also that this can be contextualized from the point of view of the repertoire of popular music of oral tradition, thus reinforcing the contact and understanding of the characteristics of this repertoire.

In order to extract, structure and store semantically relevant information from the MEI files representing the encoding of each song composing the treated corpus, we analyze the content of the MEI elements and attributes that have been identified in section 2.2. Therefore, the elements and attributes are extracted, as well as the musically significant information provided in the MEI Header relevant to the characterization of most of the songs in terms of time and space, for instance, the location and geographical origin as an element of ascription to a particular corpus representative of the style of folk music of each territory. On the other hand, elements related to the information concerning the musical logic itself are also extracted, such as time signatures, the structure of phrases, semi-phrases, motifs or periods (structured according to the starting and ending points of each one of them) as well as the most common and representative rhythmic patterns of each musical genre, songbook and/ or territory. In a more detailed way, the metadata and identifying elements of the musical logic of each song that has been incorporated into the database includes the structured information:

- Song's title (if unknown, the lyrics incipit is addressed); normally, each song that is part of the popular repertoires of oral tradition is named with the factitious title by which the song is known in the environment. In cases where this is explicitly reflected in the source, it is incorporated into the MEI Header literally; in cases where this information is not present in the edited songbooks, we have chosen to identify as incipit the smallest meaningful group of words with which the lyrics of the song begin.
- Song's author (if unknown, "Anonymous" label is addressed); Although, as a general rule, popular music of oral tradition is not produced under the work-author paradigm, in certain occasions it can be attributed to a certain transmitter, arranger, or composer who exercises factually the function of author. In cases where this information is explicitly reflected in the source or is known by other means, it is coded under the <author> tag.
- Source bibliographic information, including formal title and subtitles where known; The bibliographic information on the sources that make up the corpus of analysis is incorporated in a standardized manner, as presented on the cover of the editions, in order to facilitate their precise location. Any other type of information, commentary, or clarification related to the particularities or junctures of each edition are marked through the <note> element in the MEI Header.
- Compiler and Copier of Source, if existent; An explicit distinction is made between the role of the compiler (generally the intellectual editor of the songbook) and the transcriber in cases where this information appears in the source.
- Date of Compilation; the compilation date of the songs is included according to the edition date of the songbook where they are collected, any further clarification or explanation regarding discrepancies between the compilation and edition dates are expressed through the <note> element in the MEI Header.
- Geographic Information is characteristic; The ascription of each song to the territory to which it belongs is an essential element when establishing geographic clusters that can be characterized according to a series of common characteristics at the level of genres, rhythmic and melodic structures, etc. For this purpose, the bibliographic metadata of each song is included as a source according to the songbook to which it belongs and observing the hierarchy country; autonomous community (lower autonomous division); province; locality.
- Musical genre as it stands on the sources or transcriptions; The musical genre is explicitly referenced in the bibliographic metadata incorporated in the MEI Header according to the classification that the compilers or editors of the songbooks themselves offer to classify the collected songs.
- Meter; Indication of time signature as it appears in the edition and according to the compiler's transcription criteria.
- Tempo indications when known; In this field, the tempo indications of the piece that appear in the source are explicitly marked. In cases where this is not the case, and there is no explicit indication from the source, we proceed to indicate an indicative tempo that suits the interpretation of the piece when this information is known to the musicologist in charge of the encoding.
- Key and Mode of the Song; It is indicated, under the means explicitly included for this purpose in the set of elements and attributes of the MEI standard, the aspects related to the tonality modality of each song.
- Time Signature(s); The information related to the distribution of beats in each measure (time signatures) as they appear in the sources (songbooks) is encoded, according to the transcriber's criteria. This is an element of great relevance when establishing characterizing elements of genres and repertoires.
- Pitch Range; The tonal range of each song is explicitly stated, indicating the lowest and highest notes included in the musical text. Besides making it possible to characterize the repertoire concerning the range of the melodic development, this feature is very useful in relation to its projection to the didactic

praxis, allowing the choice of the song according to the needs of each activity and the possibilities of the eventual interpreters.

- Rhythmic Pattern; independently for each song in the corpus and according to the academic judgment of the musicologist in charge of its codification, the rhythmic motifs (reduced to their smallest expression) that in some ways are representative or characterize the song are identified and marked. As has been pointed out in previous sections, the marking of these rhythmic patterns is done by explicitly indicating the figures that compose them as an abstraction outside the musical text itself.
- Phrase types; In order to characterize the songs under study at the structural level, we proceed to identify and mark their melodic development according to their division into musical phrases (taxonomic units with melodic meaning) and semi-phrases (subdivisions and fragments that make up the identified phrases). For this purpose, the system conceptualizes these units as the section comprised between a starting point and an end point concretized in certain notes identified with the attribute @xml:id. In the same way, these notes that represent the end are qualitatively identified as cadential elements, distinguishing between cadential or semi-cadential.

To retrieve multiple viewpoint systems from MEI dataset, the use of Music21 (http://web.mit.edu/music21), a Python toolkit developed to support the computation of structured musical data in the musicological representation and analysis environment is considered. Multiple viewpoint systems have been developed in order to abstract music surface information into a complete set of music surface descriptions. In order to achieve this, properties are extracted that allow us to characterize the song at the musical level from the information encoded in MEI. Then, we go through the primary MEI data and collect the information of all the events and occurrences proper to the musical discourse, in order to form a sort of dictionary of points of view. Each and every one of these events of the musical discourse is indexed according to a standardized value related to its starting time, nesting these values hierarchically in the dictionary of points of view.

3.2. Retrieving information from the Dataset

Two protocols have been developed for querying the database, which are briefly described below:

1) an interface with the basic fields and access points obtained from the metadata, in which the user can select in a combined way which characteristics and criteria musically represent the song or songs to be retrieved, including fields such as genre, mode, rhythmic cells structuring the song, melodic structure and melodic range, etc. and 2) an interface for browsing and querying the database based on similarity, in which the songs that make up the database are shown grouped graphically, under the criterion that the physical distance between them corresponds to the degree of concordance in relation to the essence of their musical characteristics. The first approach to querying the database corresponds to the structure and logic common to information systems of this type, leaving the combination of fields and access points for meaningful and relevant retrieval of the desired songs to the user's discretion and level of expertise. For the less expert user in the use of these search systems, a series of functionalities are implemented in parallel that allows an intuitive search through the graphical representation of the data collected in the platform through a two-dimensional space where the level of similarity between the items of the musical corpus is represented. The level of user expertise knowledge is not fundamental to this approach, as they can freely navigate the whole information as a set of "spatial" infrastructure. In addition, a set of different visualizations is offered, related to a series of criteria that can be defined by the user, such as the geographical environment to which the song is attached, its melodic contour, etc.

3.2.1. Querying metadata information

The metadata query interface includes a composite query system that allows the user to specify multiple simultaneous criteria (Carvalho et al., 2021). Each field or access point effectively constitutes an application filter to the database, and for each of these fields or criteria, different values can be applied from an a priori unlimited set of each of them.

The query results returned by the system are presented by default sorted alphabetically in successive pages showing a maximum of ten items each. As an abstraction of the content, and in an identifying way, the name or

title of the song, the internal code used for indexing, an audio preview playable in MIDI format, and a link with the information collected about the song.

4. Likeness analysis in repertoires of popular music of oral tradition

The design of the web platform that hosts and allows querying the I-Folk database incorporates an iterative approach regarding the identification and visualization of the similarity of the songs that compose it. To optimize its level of operability, different sets of views have been analyzed and tested as inputs to calculate the pairwise similarity ratio.

Concerning the I-Folk database, as detailed in section 2, this type of repertoire is characterized by a narrow melodic range, and most of its most characteristic motives and rhythmic patterns are directly related to specific genres or time signatures that have been perpetuated and consolidated along the history.

5. Conclusions and future work

This article presents particularities of the codification of Iberian folk music and proposes a series of contributions in this area. In the first place, we present a standardized model, based on the MEI, which reflects in detail the musical content of this type of sheet music. In addition, on the dump of this information, we have created a new database, which we call I-Folk, with tools that allow to retrieve different data based on the query made by the users.

Our future research will look into a more systematic evaluation of similarity metrics within the I-Folk database to see how beneficial the system is in providing assistance in the learning of the repertoire and its qualities. In this regard, it is important to point out that our database facilitates the work of teachers in music classrooms, because even if they are not familiar with this repertoire, our website provides them with a multitude of analyzed scores accompanied by pedagogical guidelines. Allowing the teachers to have the analyzed scores helps searching for specific musical elements that might be useful to work with the students. For example, when a teacher needs to work with binary time signatures, dozens of songs that present this type of time signature will appear on the web. Ultimately, all these efforts are focused on the promotion of Spanish and Portuguese folk music and, consequently, on the proposed application of this music as a pedagogical resource.

Furthermore, we are evaluating the possibility of including some of the few available recordings of the works that make up the analyzed repertoire. Generally, it is music recorded during the fieldwork of ethnomusicologists, with which this musical corpus has been transcribed. It must not be forgotten that folk music has been transmitted from generation to generation thanks to its inherent orality. That is, imitation and repetition have acted as the natural learning mechanisms of this type of traditional repertoire. Thus, with the future incorporation of the recordings, the auditory memory of the students will be promoted, and their comprehensive understanding of the pieces will be enhanced. Most of these songs present simple and repetitive rhythmic-melodic patterns, which favour auditive education, creativity, and musical improvisation.

From this point of view, we are creating a space pointed at children to advance Iberian society music through intuitively recreations and accomplish, employing a perky viewpoint, transmit to the most youthful the significance of protecting our cultural legacy. Through this platform, we have proposed a few activities that will enhance this objective; "Compose your Melody", "Dance with the music", "Make a Rhythm" and "Sing with the music". There will be different grades of difficulty that varies on the user. Since the inventive and spontaneous capacity in the prevalent culture has been a determining factor for the rise of the popular repertoire itself, our project's overall aim is to value popular music of oral tradition not as a fixed repertoire, but as a living tradition that opens the doors to children's creative expression.

In short, we can affirm that the environment of our project allows us to observe the construction of a data model expressly aimed at enhancing the didactic and pedagogical aspects of this type of documentary corpus, as well as highlighting both the difficulties and the opportunities offered by the new technologies in the field of musical education in the early stages. Although the relationship between music teaching and digital technologies has been very fruitful and close in recent decades, and the presence of ICT in this field is widespread in any context and educational environment, the application of the developments of the Digital Humanities offer a significant advance and a new horizon of perspectives for the field of Music Education with the clear and inalienable objective of making this a more creative practice that encourages a deep understanding of this immaterial and idiosyncratic artistic expression.

Acknowledgment

References

Bauer, W. I. (2020). *Music learning today: Digital pedagogy for creating, performing, and responding to music* (2nd Ed.). Oxford University Press. https://doi.org/10.1093/oso/9780197503706.001.0001

Campbell, P. S. (2016). World music pedagogy. In C. R. Abril, & B. M. Gault (Eds.), *Teaching general music: Approaches, issues, and viewpoints* (pp. 89-111). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199328093.003.0006

Carvalho, N., & Bernardes, G. (2020). Towards balanced tunes: A review of symbolic music representations and their hierarchical modeling. In *Proceedings of the 11th International Conference on Computational Creativity (ICCC'20)* (pp. 236-242). CEUR Workshop Proceedings.

Carvalho, N., González-Gutiérrez, S., Merchán Sánchez-Jara, J., Bernardes, G., & Navarro-Caceres, M. (2021). Encoding, Analysing and Modeling I-Folk: A New Database of Iberian Folk Music. In *Proceedings of the DLfM '21: 8th International Conference on Digital Libraries for Musicology* (pp. 75-83). ACM. https://doi.org/10.1145/3469013.3469023

Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, *24*(1), 51-73. https://doi.org/10.1080/09298219508570672

Fondo de Música Tradicional. (2023). A Spanish Collection of Traditional Music Heritage | Fondo de Música Tradicional. IMF-CSIC. https://musicatradicional.eu/

García-Matos, M. (1963). Pervivencia en la tradición actual de canciones populares recogidas en el siglo XVI por Salinas en su tratado "De musica libri septem". *Anuario Musical*, *18*, 67.

García-Matos, M., & Bargalló, J. C. I. (1982). *Cancionero popular de la provincia de Cáceres. Lírica popular de la Alta Extremadura. Vol. II*. Consejo Superior de Investigaciones Científicas.

García-Peñalvo, F. J. (2015). Mapa de tendencias en Innovación Educativa. *Education in the Knowledge Society*, *16*(4), 6-23. https://doi.org/10.14201/eks2015164623

Manzano, M. (1990). El folklore musical en España, hoy. *Boletín Informativo de la Fundación Juan March*, *204*, 3-18.

Manzano, M. (1993). Una cuestión sobre música popular de tradición oral en La ciudad de Dios: las canciones populares y la tonalidad medieval. In *Música en el Monasterio del Escorial: Actas del Simposium (1, 4-IX-1992)* (pp. 731-749). Real Centro Universitario Escorial-María Cristina.

Merchán Sánchez-Jara, J. (2016). Lectura Musical en el ámbito digital; aplicaciones para tablets. *Education in the Knowledge Society*, *17*(1), 109-128. https://doi.org/10.14201/eks2016171109128

Merchán Sánchez-Jara, J. (2019). Music Encoding Initiative (MEI), un estándar para la edición, codificación y la descripción de documentos musicales en el entorno digital: características, prestaciones y ámbitos de aplicación. *Education in the Knowledge Society*, *20*(13). https://doi.org/10.14201/eks2019_20_a9

Navarro Cáceres, M. (2018). *Etno-music: plataforma para la generación colaborativa basada en el análisis musical de la música popular modal en Castilla y León*. [Tesis doctoral, Universidad de Salamanca]. Repositorio Documental de la Universidad de Salamanca (GREDOS). https://gredos.usal.es/handle/10366/139856

Olmeda, F. (2001). *Folk-lore de Castilla o Cancionero popular de Burgos*. Editorial MAXTOR.