Exploring folk songs to educate for resilience

Juan Albacete-Maza, Antonio Fernández-Cano and Zoraida Callejas

Abstract

Purpose – Covid-19 pandemic, war, climate emergency and other recent challenges are inflicting tremendous stress to youth. However, death and tragedy are nowadays considered taboo, as there is generally no standardized nor naturalized discussion on the subject, especially with young people. The current multi-crisis scenario is intensifying the need to incorporate an education on tragedy and resilience in our learning systems. In this context, it is necessary to find suitable teaching resources for this educational challenge that are attractive, entertaining and suitable for children and youth. A resource that meets all these requirements are children’s folk songs (CFSs). Apart from the intrinsic educational potential of music, folk songs have a simplicity and musicality that make them an ideal teaching resource. Considering their oral historical transmission, their survival confirms the attraction that this type of composition causes on children. However, to consider CFSs as an adequate resource to carry out an education for death and tragedy, it is necessary to study whether they present a non-negligible proportion of tragic passages and with enough variety of themes. This paper aims to address the study of the presence of explicit tragic content in Spanish CFSs and thus could be considered a cultural resource with transformative educational potential to develop resilience capabilities on the face of tragedy.

Design/methodology/approach – An analysis of lyrics of 2,558 Spanish CFSs is presented, using a manual content analysis as well as a computerized content analysis with the aim of identifying the tragic component of these songs and, thereby, assessing their pedagogical potential as a transformative educational resource.

Findings – The results obtained show a considerable presence of death and tragedy (19.78%) and a variety of tragedy dimensions. CFSs have been transmitted orally not only as a ludic resource, but also to prepare children for life (and death). The results show the complementarity of both analyses to avoid subjectivity while considering the underlying meanings of the songs.

Originality/value – This task had previously not been approached in an automated manner in the literature, nor there had been a similar study with a sample of this magnitude. The outcomes obtained show the considerable presence of tragedy in Spanish CFSs and emphasize the interest of this currently undervalued didactic resource.

Keywords Death education, Tragic content, Death taboo, Lyric analysis, Folk songs

Paper type Research paper

1. Introduction

Music and its educational potential are an increasingly investigated research topic (Dvorak et al., 2020; Kallio, 2020; Savage, 2021). The present study aims to examine the tragic content of a very peculiar type of musical composition: children’s folk songs (CFSs).

These songs have served us since ancient times, helping our ancestors in educating subsequent generations and, in turn, they have been used by children to accompany their games and entertainment. We consider worthy of study the standardized presence of passages in these song lyrics, which refer to tragic aspects of the human condition such as death, sadness, illness, violence, heartbreak or fear. Currently, there is an educational and social taboo regarding death and all things tragic (Colomo, 2016; Macedo, 2019;
Nelson et al., 2018) wherein discussion of these topics is avoided as much as possible with children and adolescents.

However, many voices have highlighted the pressing need to overcome such taboos and to support education regarding death, tragedy and resilience (Sahin and Türk, 2021; Testoni et al., 2021b; Grigoropoulos, 2022), which offers tools and knowledge that help to normalize and admit our finite condition and the tragic nature of our existence. There is a need to normalize debate and reflection on such tragic aspects, and thereby provide a true comprehensive education to the youth, which will allow them to live balanced lives and adequately face the possible setbacks that life may bring them. In a way, the pandemic caused by Covid-19 has also come to bolster those who call for overcoming these taboos, highlighting the importance of considering tragic topics in our education and achieving a real education for life (and death), finding innovative ways to promote resilience and flexibility in the face of uncertainty (Karjalainen et al., 2022; Jasman and McIlveen, 2011).

Faced with such educational challenge, it is necessary to find suitable resources to deal with this taboo. Music can be a great ally in this regard. Specifically, CFSs present formal and musical characteristics that make them especially attractive to children and young people. If some of them have survived for centuries through oral transmission, it was because of their playful component and potential as an educational resource and cultural transmitter (Denac and Žnidaršić, 2018).

For all these reasons, the study of the tragic components of this type of musical composition proves interesting. The standardized, recurring presence of tragic content in CFSs leads us to assume that our ancestors had a process of naturalizing the tragic aspects of the human condition, at least when educating and transmitting values to young children through these types of songs.

This paper provides a double content analysis (manual and computer) of tragic components of Spanish CFSs, and further reflection on it. We address the following research objectives:

- examine the presence of tragic content and its variety of dimensions in Spanish CFSs;
- perform the examination through manual as well as automated analysis; and
- corroborate whether the results obtained confirm a presence of tragic content that positions CFSs as a promising resource for resilience education.

The rest of the paper is organized as follows. First, the background and related work are discussed. Next, the Spanish CFSs corpus used and the analysis methods are described. Then, the results are presented and discussed. Finally, we present the conclusions reached and identify future work guidelines.

2. Related work

Nowadays, death can be considered a taboo subject in society, as there is generally no standardized reflection on the subject nor naturalized discussion about our finite condition, with conversations on the subject being mostly avoided (Caparroz et al., 2015; Colomo, 2016; Macedo, 2019). Not only is death taboo, but so are other tragic realities intrinsic to the human being: “in our society, realities inherent to the human condition such as suffering, pain, illness and death are, on the one hand, denied and hidden, when they occur on a personal and intimate level of everyday life” (Ransanz, 2015, p. 162).

Children and adolescents should be educated in tragedy, accustomed to its reflection and discussion and in possession of cognitive tools with which to adequately cope with the tragic nature of our existence. Anything else would be to ignore and sidestep an issue that should urgently be addressed to, among other purposes, avoid possible future psychological problems (Arndt et al., 2004; Wong, 2017).
Promoting a resilient attitude in children and young people in the face of the tragic aspects of the human condition is fundamental for any society. The Covid-19 pandemic has highlighted the relevance of implementing an education on tragedy and resilience, which truly prepares us for life and provides us with tools to deal with these types of negative events (Sahin and Türk, 2021; Sood and Sharma, 2020; Testoni et al., 2021a).

2.1 Educational potential of folk songs

The potential of music, and especially songs, as a didactic resource has been widely demonstrated in recent decades (Batchelor and Bintz, 2012; Debreceny, 2015; Governor et al., 2013; Klinger et al., 1998; McGregor and Mills, 2006; Napoletano, 1988; Potkay, 1982). Specifically, CFSs have enormous educational potential because of their simplicity, both musical and formal (Fernández-Poncela, 2009). In fact, the mentioned potential is proven by the very survival of this type of composition (González-Mediel, 2010). Given their simplicity, these songs were easily memorized by childhood. In addition, their cheerful music and lively rhythms allowed them to be incorporated into children’s games (Santamaría, 2006). For this reason, they have been able to survive throughout the centuries through oral transmission. If they did not possess such attractive power for childhood, they would simply have disappeared.

Although the evidence is not very extensive, there are experiences in which folk songs have been proven to be effective pedagogical resources to work with: language skills (López-Valero et al., 2009; Cerrillo, 2017; Watts and Morrissey, 2019); improvisation (Watts and Campbell, 2008); social skills (Mills, 1974; Llamas, 2005), diversity (Nompula, 2011; Campbell, 2021), cultural transmission (Minks, 2002; Arévalo Galán, 2017; Watts and Campbell, 2008; Andrews, 2021); communicating important social messages (Quan-Baffour, 2007); and, of course, musical skills (Klinger et al., 1998; Sarget, 2002; Santamaría, 2006).

We should take a cue from how our ancestors did, in fact, appreciate and fully value the potential of children’s folk music and songs. In this paper, we consider interesting the study and analysis of the tragic components of this type of musical composition and, thus, to value them as an educational resource and source of cultural information. We hypothesize that, should they address tragedy directly, the tragic content of folk songs may have great potential when it comes to educating children and youth for resilience (Higgins et al., 2020).

2.2 Analysis of songs

Songs are a cultural element of enormous significance in any society. Music and human beings have been travel companions since our most remote origins (Debreceny, 2015; McDermott and Hauser, 2005). When songs began to be accompanied by lyrics, their value as an object of study increased exponentially. Since then, songs became not only music and dance; their lyrics allowed for expressing even more the character and way of understanding the world of those who sang them (Mills, 1974).

The social sciences, especially ethnomusicology, have historically been interested in this type of cultural phenomenon, considering song as a fundamental source of information when analyzing and understanding any society (Giannattasio and Giuriati, 2017; Lomax, 1959; Martín-Camacho, 2003; Russo et al., 2020). Therefore, numerous researchers have been, and continue to be, interested in the analysis of song lyrics.

Ultimately, the analysis of song lyrics aims to better understand the culture to which the lyrics belong (Sotiropoulos et al., 2021). Thus, the study of different aspects and themes that characterize songs has been addressed, for example, in regard to such topics as: the affective component (Chen and Tang, 2018; Stedman and Alpher, 1987; Wu et al., 2013; Yoo et al., 2017); pedagogical potential (Kranke et al., 2016); discourse and message
(Cole, 1971; North et al., 2021); semantics and morphology (Logan et al., 2004); or topic modeling (Sotropoulos et al., 2021).

Content analysis performed by experts or evaluators is the most common methodology in the research on song lyrics. It is characterized by using different indicators (of a semantic or linguistic nature, quantifiable or not) as a reference to make logical deductions from the analyzed messages (Bardin, 1986). In recent years, fully automated or semi-supervised computerized textual analysis is also used, allowing the researcher to work with much larger song corpora and to drastically reduce the time necessary to carry out such analysis. This type of computer content analysis of song lyrics has been evolving in recent years, having been used, for instance, for explicit content localization (Rospocher, 2021); sentiment analysis (Hu et al., 2009; Jamdar et al., 2015; Ren et al., 2021; Xia et al., 2008); or identification of musical genre from lyrical analysis (Tsaptsinos, 2017).

Nevertheless, the analysis of tragic components in song lyrics is novel, given that other types of content have generally been the focus of textual analysis research, namely, gender roles (Berrocal de Luna and Pérez, 2002), explicit content (Rospocher, 2021), irony and sarcasm (Chia et al., 2021; Parameswaran et al., 2021), the affective component (Chen and Tang, 2018; Yoo et al., 2017).

In addition, there are numerous studies on the content of song lyrics, but there are very few studies that focus on this specific type of musical composition: folk songs. Much research has been done on popular songs of all musical styles (rap, pop, heavy metal, reggae, etc.); however, the analysis of the discourse and messages of folk songs is scarce, despite the significance these songs have carried throughout history in most societies.

Popular songs are by specific authors, marked by temporary success, transmitted through audio recordings and generally produced by professionals (Tagg, 1982). By contrast, folk songs are “songs which the common people have adopted as their own, regardless of origin, [which] constitute in some way or another their collective voice” (p. 22) (Roud and Bishop, 2017). Folk songs have served as entertainment to ease the burden of household chores and fieldwork, to accompany children’s games, to inform others about events in other regions and, of course, to educate children and transmit community values and traditions.

In this paper, we investigate the presence of tragedy in Spanish CFSs and analyze the type of tragic elements addressed. To do so, we use a mixed method that combines manual content analysis, to consider the nuances and subtlety of the potential tragic content observed and automated analysis, to obtain results that are not influenced by the inherent subjectivity of the manual analysis.

3. Data set and methods

Figure 1 shows the different avenues addressed for the analysis of CFSs. First, a manual content analysis was conducted on the lyrics of 2,558 CFSs, resulting in the detection of 506 CFSs with tragic content.

Manual content analysis is a good tool for song lyric analysis. However, the analysis is carried out by a researcher who is not free from bias. To alleviate this limitation, computerized analysis can be a good ally, by complementing manual analysis and being able to triangulate the veracity of the data obtained.

A computerized content analysis was performed using a balanced number of tragic and non-tragic songs (the 506 tragic CFSs detected in the manual analysis and 506 without tragic content, 1,012 CFSs in total). With this aim, a computer content analysis software has been designed that can identify the presence or absence of a tragic component in the lyrics of CFSs based on glossaries of tragic vocabulary as will be explained in more detail in the following sections.
3.1 Description of the corpus of Spanish children’s folk songs studied

A sample of 2,558 CFSs was compiled, consisting of Spanish CFSs available in several prestigious sources: 137 CFSs by Gil (1964); 112 CFSs by Hidalgo (1969); four CFSs by Martínez-Blanco (2011); 423 CFSs by Córdova y Oña (1947); and 1,882 CFSs from Instituto Milà i Fontanals (2021), belonging to the Consejo Superior de Investigaciones Científicas: CSIC (Spanish National Research Council). These are songbooks collected during the 20th century, most of them contain handwritten documents with lyrics and scores transcribed live by ethnomusicologists while subjects from different regions of Spain sing traditional songs. Only the Fondo Música de Tradicional (Traditional Music Holdings) of the CSIC provides a digital scanned version of these documents, in image format.

For this study, the 2,558 songs of the mentioned songbooks have been analyzed, labeling them as tragic or non-tragic following a manual analysis of content, as previously described. From them, 1,012 songs were selected, 506 tragic and 506 non-tragic whose lyrics have been manually transcribed in plain text format from the scanned handwritten lyrics provided in the digital version of the songbook. The corpus contains the name and transcribed lyric of each song in plain text format, with a minimum number of 6 words/37 characters, and a maximum of 617 words/3,350 characters. The average number of words is 55.46, and the average number of characters is 291.86.

It should be noted that other related studies on Spanish CFSs (Banderas, 2018; Berrocal de Luna and Pérez, 2002; Fernández-Poncela, 2005, 2013; Malvido, 2008; Nogueira, 2018) do not analyze more than 50 CFSs each, which underlines the robustness of the results presented in this study, given the considerable increase in sample size.
3.2 Manual content analysis

A manual analysis of the content evident in this corpus was performed, where CFSs were considered to have tragic content if their lyrics presented at least one tragic element, understood as a lyrical passage in which themes of a tragic nature are mentioned. Consequently, the CFSs were labeled as tragic or non-tragic, identifying a total of 506 Spanish CFSs with tragic content. For the CFSs identified as tragic, up to six categories of tragic content could be inferred from the content analysis, namely: death, violence, fear, sadness, heartbreak and illness.

The data obtained from the manual content analysis was collected in terms of the number of tragic elements identified in the tragic CFSs and their corresponding categorization. Each tragic CFS was transcribed into a text file in which each tragic element observed was identified, categorized and commented.

3.3 Computer content analysis

State-of-the-art approaches using deep learning and other machine learning techniques are based on large databases of popular songs (not folk songs) that are easily accessible to researchers and allow for training of classifiers with an adequate number of samples. In the case of CFSs, the available songs come from ethnomusicological studies in which lyrics are usually manually transcribed while an individual interprets them, which makes them a very valuable resource and results in a limited volume of available lyrics.

As previously mentioned, although the corpus of songs compiled and transcribed in this study is much larger than those available in the state of the art, the number of songs used does not allow us to satisfactorily perform an automatic classification based on deep learning.

Moreover, there are sentiment analysis resources in Spanish for processing based on the detection of polarity (positive or negative), and yet, the distinction we are pursuing (tragic vs non-tragic) does not correspond to the categorization of positive vs negative provided by the sentiment analysis approach. The categories with which we intended to work are very specific, and none of the reviewed alternatives provides distinctions of this nature.

To address this challenge, we have opted for the creation of our own resource with a glossary-based approach, with which satisfactory results have been obtained in other types of studies, such as the analysis of emotion conveyed by song lyrics (Jamdar et al., 2015), for the classification of song genres (Ying et al., 2015) or even for the calculation of an author’s lexical novelty (Ellis et al., 2015). In the literature reviewed, there is no reference to any glossary of tragic elements, and therefore, it was necessary to develop a glossary of vocabulary related to each of the categories (death, violence, fear, sadness, heartbreak and illness), which we make available to the scientific community for further research along the path we have undertaken with this work.

The result of the process was a glossary of 760 tragic terms. To give the process more flexibility, a stemming of the words in the glossaries was performed. Stemming consists of reducing the inflection of words so that a group of words (e.g. words with different numbers or genders) can be matched to the same stem that represents them.

To automatically classify the CFSs into tragic or non-tragic, a Python script was programmed so that for each CFS in the corpus (stored as txt files), the following steps are implemented:

- **Stemming**: the stem of all the words of the lyrics of the CFS is obtained.
- **Counting**: the occurrences of the stems corresponding to words in the glossary of each category are counted.
- **Statistical calculation**: the percentage of words in each category and in all tragic categories is calculated, considering the number of occurrences of words in each category and in all tragic categories, with respect to the total number of words (total length of the lyrics).

- **Binary classification**: a song is considered tragic when it contains words from one or more of the glossaries and non-tragic when it does not.

- **Detailed classification**: in the case a CFS is considered tragic, the predominant category is indicated.

4. Results

The following sections present the results corresponding to the manual content analysis of the 2,558 Spanish CFSs analyzed as well as the results resulting from the computer content analysis of the 506 CFSs identified as tragic by the manual content analysis, and another 506 non-tragic CFSs.

4.1 **Manual content analysis**

There were 506 songs in which at least a tragic element was identified, which represents 19.78% of the total number of CFSs analyzed. In total, 1,257 tragic elements were identified, resulting in an average of 2.48 elements per song. The maximum number of tragic elements identified in a CFS is 16 elements in the one entitled *Grandes guerras se publican* (Great wars are announced), followed by *Carmela se paseaba* (Carmela strolled around) and *En el pueblo de Sevilla* (In a village in Seville), with 14 tragic elements identified, and *Don Pedro/La muerte oculta* (Don Pedro/The hidden death), with 13.

All the tragic elements identified were categorized into violence, fear, sadness, illness, death or heartbreak. **Table 1** shows the total number of tragic elements identified by category.

Thus, death is the category with the highest number of tragic elements identified, followed by violence, sadness, heartbreak, illness and fear. The weight of each category with respect to the total number of tragic elements identified is shown in **Figure 2**.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Number of tragic elements identified per category in the manual content analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death</td>
<td>Violence</td>
</tr>
<tr>
<td>347</td>
<td>325</td>
</tr>
</tbody>
</table>

**Source**: Authors’ own work

**Figure 2**  
Percentage of tragic elements in each category with respect to the total number of tragic elements identified in the manual content analysis

**Source**: Authors’ own work
In addition, Table 2 shows the number of CFSs in which at least one tragic element was identified belonging to the category indicated in each column.

In this case, the category for which at least one tragic element has been identified more times is violence, followed by death, sadness, heartbreak, illness and fear. Figure 3 further illustrates this showing the number of CFSs with at least a tragic element in each category with respect to the total number of CFSs identified as tragic (506). As can be observed, for the most relevant category, violence, 43.87% CFSs have at least one tragic element related to violence, followed by death with 41.50%.

4.2 Computer content analysis

As described before, for the computer content analysis, we used the 506 CFSs identified as tragic by the manual analysis and 506 identified as non-tragic. Table 3 presents the results achieved by the computer content analysis in a confusion matrix.

As can be observed, the computer content analysis identified as tragic 363 out of the 506 tragic CFSs, thus producing 143 false negatives. In addition, it identified 417 of the 506 non-tragic CFSs, producing 89 false positives.

The precision of the computer content analysis (the percentage of songs tagged as tragic/non-tragic that were really tragic/non-tragic) is 80.30% for the tragic category and 74.46% for the non-tragic, while the recall (the percentage of the songs in the specific category that it is able to find) for tragic is 71.73% and for non-tragic is 82.41%. This means that when the software identifies a CFS as tragic, it is really tragic (according to the manual analysis) 80.30% of times. However, it is only able to identify 71.73% of the tragic CFSs.
Nevertheless, the overall concordance between the manual and computer content analysis (accuracy) is 77.07%.

As previously described, this is, to the best of our knowledge, the first attempt to process the tragic content in songs, so the comparison using other classification approaches as a baseline would be inevitably biased. As a proof-of-concept, we considered the most similar task in the literature: sentiment analysis, which aim is not to differentiate tragic vs non-tragic, but negative vs positive or neutral content. For an initial comparison, we have run state-of-the-art sentiment analysis algorithms with our corpus. The best results were attained with a model of convolutional neural networks trained over 800,000 online reviews in Spanish. The accuracy obtained was 58.99% (while our approach reached 77.07% accuracy), and the recall of the tragic category was 21.74% (while we obtained 71.73%).

5. Discussion of results

As for the results of the manual content analysis, the percentage of Spanish CFSs with tragic content (19.78%) is significant. The presence of tragic elements in this type of Folk compositions intended for children is frequent. There are even cases in which tragedy is the thematic center of the composition and everything revolves around it. It is not uncommon to find CFSs that narrate, with total naturalness, episodes such as murders, betrayals, agonies, wars, suicide or stories of love/heartbreak with a fatal outcome.

When analyzing the tragic components, the categories of death and violence are the most salient. Death is the most frequent, whereas violence is the category than appears in a higher number of different songs. It should be noted that it is very common to find elements that belong simultaneously to both categories, because, for example, the narration of murders is very common. On the other hand, Fear is the category with the most discrete weight in this type of compositions, although its presence is very frequent in compositions aimed at warning children of the dangers of everyday life, especially in lullabies using scary characters.

To complement the manual analysis and account for the effect of subjectivity, we performed a computer content analysis to check whether it could also find such a prominent presence of tragic elements in CFSs. The agreement between both analyses was high, which corroborates the findings of the manual analysis. However, the number of false negatives (tragic songs tagged as non-tragic) obtained by the computer content analysis is considerable (143, 71.73%). A more detailed analysis of the false negatives has shown that most are caused by passages in which the tragic content is expressed in deeper or underlying meanings in phrases and literary resources of greater complexity, such as metaphors, use of historical characters or colloquialisms. The fact that computer content analysis of lyrics of this type of composition is not sensitive to tragic elements presented in the forms mentioned highlights the need to use a semi-supervised approach (wherein human annotators work in collaboration with software) in the analysis of CFSs.

These results show that the current tendency to ignore and hide death and tragedy from new generations is not so strongly evident in Spanish CFSs. It is surprising to see the naturalness with which lyrical passages speak plainly about death, violence, sadness, illness, fear or heartbreak. Nevertheless, academic literature on the subject is very scarce. The presence of tragic content in a large percentage of CFSs clashes with the current conception of songs aimed at children.

6. Conclusions

A double content analysis (manual and computerized) has been presented to study the tragic content of CFSs, which has been used to analyze a corpus of 2,558 Spanish CFSs. The results show a considerable presence of tragedy: 19.78% of the CFSs analyzed
present tragic elements, with a variety of dimensions including death, violence, heartbreak, sadness, fear and illness.

Our ancestors did not hesitate to transmit this tragic and at the same time indissoluble part of our existence to the youngest, in a natural way and far from any kind of taboo. Tragedy was more than present in everyday life in Spain in past centuries and, for this reason, it appears without complex in this type of folk compositions; together with countless CFSs in which joy, humor and beauty monopolize all the attention. Reflecting, in short, life itself.

Having justified the relevance of a proper education on tragedy with the main objective of overcoming the taboo that characterizes death and the tragic elements of life, especially in school environments, educational interventions on these topics are presented as more than necessary to strengthen the resilience of recipients and foster a certain normality and naturalness. The formal and musical characteristics of CFSs, together with their historical survival, highlight their widely demonstrated educational potential. Our study now confirms that they also contain a considerable amount and variety of tragic themes in their lyrics.

Therefore, CFSs can be an ideal didactic resource to carry out an appropriate and necessary education for death, tragedy and resilience. As future work, we will address pedagogical interventions for resilience through music and songs based on the results presented.

References


**Corresponding author**

Zoraida Callejas can be contacted at: zoraida@ugr.es