COMPUTATIONAL MODELLING OF INFORMATION DYNAMICS IN CHILDREN'S MELODY PERCEPTION

Lorena Mihelač

Doctoral Dissertation
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia

Supervisor: Prof. Dr. Dr. Geraint Anthony Wiggins, Vrije Universiteit Brussel, Brussels, Belgium &
School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK
Co-Supervisor: Prof. Dr. Janez Povh, University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, Slovenia &
Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

Evaluation Board:
Prof. Dr. Nada Lavrač, Chair, Jožef Stefan Institute, Ljubljana, Slovenia
Prof. Dr. Matija Marolt, Member, University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia
Prof. Dr. Frans Wiering, Member, Utrecht University, Utrecht, Netherlands

MEDNARODNA PODIPLOMSKA ŠOLA JOŽEFA STEFANA
JOŽEF STEFAN INTERNATIONAL POSTGRADUATE SCHOOL

Lorena Mihelač

COMPUTATIONAL MODELLING OF INFORMATION DYNAMICS IN CHILDREN'S MELODY PERCEPTION

Doctoral Dissertation

RAČUNALNIŠKO MODELIRANJE INFORMACIJSKE DINAMIKE V ZAZNAVANJU OTROŠKIH MELODIJ

Doktorska disertacija

Supervisor: Prof. Dr. Dr. Geraint Anthony Wiggins

Co-Supervisor: Prof. Dr. Janez Povh

Ljubljana, Slovenia, July 2023

To my loving Father & Mother.

Acknowledgments

This thesis would not have been possible without the help of a large number of people. To begin, I would like to thank my first supervisor, prof. dr. dr. Geraint A. Wiggins, for his excellent supervision over the years, for all of his time, ideas, and support in ensuring that my Ph.D. experience was rewarding and stimulating. I have gained a great deal of knowledge from our frequent and lively discussions, which have significantly contributed to the completion of this work.

In addition, I am indebted to my second supervisor, prof. dr. Janez Povh, for his encouragement and illuminating discussions, for his sincere and selfless support, and for prompt and useful advice during my research.

I am thankful to prof. dr. Nada Lavrač for supporting my studies at the Jožef Stefan Postgraduate School, for supervising me, and inspiring me to complete my studies.

I would like to take this opportunity to express my gratitude to prof. dr. Leon Stefanija for his insightful suggestions, and to all those persons who have provided me with their invaluable support and assistance.

And finally, I would like to express my gratitude to my family who endured this long process with me, always offering me support and love.

This thesis research was partially funded by the Jožef Stefan International Postgraduate School in Ljubljana and the Šolski center Novo mesto in Novo mesto.

Ljubljana, Slovenia
July 5, 2023

Lorena Mihelač

Abstract

This thesis focuses on the information dynamics of children's melody and, more specifically, on modeling children's perception of melodic surface. The topic of children's melody has not received much attention in the past, and there are currently no clear definitions of what exactly a "children's melody" is. As neither children's folk songs nor children's songs are now recognized as a unique genre, the study of children's melody is pertinent to efforts to provide new insight into both genres.

In addition, children's melody is addressed within the context of three major research objectives: (i) musical segmentation in children of different ages, (ii) (ir)regularity in children's folk songs, and (iii) (dis)similarities in the use of musical features and dimensions between and within 22 European countries. The purpose of the research presented here is to fill current gaps in all three fields.

The majority of studies on the mechanisms underlying perceptual grouping in music include adult participants, and few studies have examined how infants and children of a particular age perceive and organize musical structure. There has been little to no research on how children of different ages perceive and process music. Consequently, the first part investigates how children of varying ages perceive musical boundaries and how music is segmented in order to fully comprehend the perception of music segmentation and present a more complete picture of human segmentation.

The second part examines the (ir)regularity of children's folk songs. Numerous studies on children's folk songs focus on the musical content, the social and cultural significance of children's folk songs and their relationship to adult music, the contribution of children's folk songs to cultural preservation, and their transmission from generation to generation. (Ir)regularity and complexity of children's folk songs and their impact on children's perception of musical structure are rarely discussed.

There are currently no studies addressing the (dis)similarities between and within countries in terms of how musical dimensions and features are perceived and utilized in folk songs, children's folk songs, and children's songs. Therefore, the third part explores the (dis)similarities in musical dimensions and features across these genres in 22 European countries deemed to be geographically close or to have similar political, historical, economic, and cultural backgrounds. The decision to include not just folk songs, but also children's folk songs and children's songs as distinct genres from a given country is based on the fact that children's folk songs and children's songs are rarely used in cross-cultural studies.

A multidisciplinary approach is utilized by employing disciplines such as music theory, music psychology, computational musicology, folkloristics, and information theory. The Information Dynamics of Music (IDyOM) computational model is used for simulating and modeling the listener's perception of music and observing its structure, to uncover meaningful and unexpected findings in all three research fields.

Povzetek

Doktorska disertacija se osredotoča na informacijsko dinamiko otroške melodije, natančneje na modeliranje otrokovega dojemanja glasbene strukture melodije v otroških in otroških ljudskih pesmih, ki v preteklosti niso bile deležne velike pozornosti. Ker niti otroške niti otroške ljudske pesmi trenutno niso priznane kot edinstven in samostojen glasbeni žanr, je raziskava teh pesmi pomembna v prizadevanju ponuditi nov vpogled v oba glasbena žanra. Melodija v otroških in otroških ljudskih pesmih se v doktorski disertaciji obravnava v okviru treh večjih raziskav: (i) dojemanje in segmentacija glasbene strukture pri otrocih različnih starosti, (ii) (i)regularnost glasbene strukture v otroških ljudskih pesmih, (iii) podobnosti/razlike glede uporabe glasbenih parametrov v otroških (ljudskih) pesmih med in znotraj 22 evropskih držav. Namen teh treh raziskav je zapolnitev trenutnih vrzeli glede razumevanja omenjenih vsebin na vseh treh področjih.

Prvi del doktorske disertacije obravnava dojemanje in segmentacijo glasbe. Večina raziskav o mehanizmih, na katerih temelji segmentacija glasbe, vključuje odrasle udeležence in le nekaj raziskav je usmerjenih na to, kako dojenčki in otroci v neki določeni starosti zaznajo in organizirajo glasbeno strukturo. Ni raziskav o tem, kako otroci različnih starosti (od otroštva do adolescence) dojemajo in segmentirajo glasbo. Posledično prvi del doktorske disertacije raziskuje, kako otroci različnih starosti zaznavajo glasbene fraze v melodiji in kako jih segmentirajo, da bi bolj jasno razumeli dojemanje in segmentacije glasbe. Namen je pridobiti tudi popolnejšo sliko človeške glasbene segmentacije – od otroštva do odrasle dobe.

Drugi del doktorske disertacije obravnava (i)regularnost glasbene strukture v otroških ljudskih pesmih. Številne raziskave o otroških ljudskih pesmih se osredotočajo na njihovo glasbeno vsebino, družbeni in kulturni pomen v družbi ter na njihov prispevek pri ohranjanju ljudskega glasbenega izročila. Le redke raziskave obravnavajo (i)regularnost in kompleksnost glasbene strukture v teh pesmih ter njihov vpliv na dojemanje glasbene strukture pri otrocih.

Tretji del doktorske disertacije raziskuje podobnosti in razlike pri uporabi glasbenih parametrov v otroških, otroških ljudskih in ljudskih pesmih med 22 evropskimi državami in znotraj njih, saj trenutno ni raziskav na tem področju. Teh 22 evropskih držav je izbranih na osnovi njihove geografske bližine in na podlagi podobnega političnega, zgodovinskega, gospodarskega in kulturnega ozadja. Odločitev, da se pri iskanju teh podobnosti in razlik ne jemljejo v vzorec samo ljudske pesmi, temveč tudi otroške in otroške ljudske pesmi, temelji na dejstvu, da se slednja dva glasbena žanra ne uporabljata ali pa se zelo redko uporabljata v medkulturnih raziskavah.

V doktorski disertaciji je uporabljen multidisciplinarni pristop z uporabo disciplin, kot so glasbena teorija, glasbena psihologija, računalniška muzikologija, folkloristika in informacijska teorija. Računalniški model, Information Dynamics of Music (IDyOM) je uporabljen za računalniško simulacijo človeške percepcije in modeliranje poslušalčevega dojemanja glasbe ter opazovanje njene strukture, predvsem z namenom odkrivanja pomembnih in nepričakovanih ugotovitev na vseh treh raziskovalnih področjih.

Contents

List of Figures

List of Tables

List of Algorithms

Abbreviations

CMCS	...	Classification Model for Songs
IDyOM	...	Information Dynamics of Music
FS	...	Folk Songs
CFS	...	Children's Folk Songs
CS	...	Children's Songs
E	...	Entropy
IC	...	Information Content
e	...	Discrete Musical Event in a Melody
IR_REG	...	Algorithm for (Ir)regularity

Chapter 1

Introduction

> The voyage of discovery is not in seeking new landscapes but in having new eyes.
>
> *Marcel Proust*

1.1 Motivation

When it comes to human perception, particularly musical perception, and computer modeling, a variety of issues and debates arise, including the following: can a computer model human perception of music? If this is the case, how precise is it? How much of a similarity to a human does it exhibit? Is it advantageous? Is it possible to identify some peculiarities that cannot be resolved using more conventional methods?

The problem lies in music itself. To begin with, music exists exclusively in humans. Other species have music, but it is not processed, created, or combined in the same manner that humans do. Darwin's theory that all animals are capable of detecting and appreciating rhythm and melody due to a neurological system that is comparable in humans and animals (Darwin, 1871) has been challenged in various recent research papers (e.g., Honing et al., 2018; Patel, 2014).

Secondly, we are biological organisms, and everything our brain produces is a biological output. If we know that our brain contains around ten to the power of eleven neurons (Kandel et al., 1991, pp. 19–20), the question is if we *can* model our mind and, if so, *how* can we model something as sophisticated as *music perception*.

Finally, music is self-referential. As with any other language, music makes use of distinct symbols that are arranged hierarchically, from the most fundamental to the most sophisticated. For instance, notes (individual events) can be combined to produce motifs (many events combined), motifs can be combined to phrases, and phrases to sentences, and so on.

When we use the term "house" in a language, we may imagine what it refers to: a *house*, regardless of its shape or color, because we are merely thinking about the term "house". When a musical motif is used (which can be regarded as a word in a language), as in the first bar of Beethoven's Fifth Symphony, each of us can have our own interpretation of what the motif is providing, implying that the motif can provide anything. This is the semantic *meaning* problem in music.

While we can claim that each element of a musical piece, or the entire piece, has meaning, the extent to which it does is controversial, as there is no general agreement on the meaning of musical content. This is the musical semantics paradox: "music seems full of meaning to ordinary and extraordinary listeners, yet no community of listeners can agree [. . .] about the nature of that meaning" (Swain, 1997, p. 45).

With these three issues in mind, and returning to our initial question, "how to model human perception of music," we are posing a question that has occupied numerous scientists, some of whom are more or less interested in music, for more than two centuries. Over the years, various proposals, attempts, and techniques have been propositioned. Several of them can be considered "milestones" in terms of simulating human perception of music, including:

a) The use of musical-theory (e.g., Riemann, 1877; Schenker, 1935),

b) The use of experimental psychology (e.g., Deutsch, 1979, 1983; Lerdahl & Jackendoff, 1977; Meyer, 1956; Nan et al., 2006),

c) The use of neuropsychology (e.g., Altenmüller, 1993; E. Gordon, 1979; Hunter et al., 2010; Trehub et al., 1990), and

d) The use of computational models (e.g., Chew, 2000; Huron, 1996; Longuet-Higgins, 1962; M. T. Pearce, 2005; Temperley, 2001; Volk et al., 2008)

The problem with applying musical theory (and its rules) to a certain style of music is that it works flawlessly in one style but ultimately fails in another. This is not to argue that musical syntax (the rules governing the usage of musical dimensions/elements individually or in combination within a musical framework) should be avoided. While musical theory is necessary for modeling, it is insufficient for simulating human perception of music.

Significant findings have been made in the field of experimental psychology, such as how music is perceived/processed (e.g., pitch and time-based relations), what are the thresholds (e.g., for pitch, loudness), and so on. Typically, some type of behavior is measured, but with limitations, such as the limited number of participants who can be observed or examined, the limited number of musical examples that can be analyzed (especially if these examples are large), or the limited experimental scenarios that are available, which can be quite expensive due to the difficulty of observing multiple types of behavior simultaneously.

Musical neuropsychology has advanced our understanding of human perception significantly. This field focuses on determining which parts of the brain are active during various musical activities in order to gain a better understanding of the role of general and music-specific systems in musical perception, processing, and creativity. It has recently become more popular (as seen by a rising number of studies) because it incorporates brain imaging, which is the visualization of how the human brain functions in response to diverse musical stimuli.

The use of a computational method for the modeling of human perception is another key milestone. The development of computational models began around 1960, owing to advancements in technology and the digitization of music. The latter entails having extensive collections of digitized music at one's disposal for musicological, psychological, musical-theoretical, and other study purposes. The increasing use of computational models can also be explained by the ability to provide an objective representation of music, identify significant features within a musical structure, compare human music processing/production, and simulate human perception of music.

Several areas (tasks) in which computational models have been applied include the examination of tonality (Longuet-Higgins, 1962), melody analysis (Huron, 1996), music analysis (Smaill et al., 1993), determining stable tonal areas vs. areas of tonal modulation (Chew, 2000), exploring rhythm-meter with a rule-based approach (Temperley, 2001) and rhythm-meter with a probability-based approach (Raphael, 2001), exploring pulse patterns (Volk, 2008), segmentation research (M. T. Pearce, Müllensiefen, et al., 2010a; Tenney & Polansky, 1980), prediction of melodic expectancy and human behavior (M. T. Pearce & Wiggins, 2006a, 2012), etc.

The use of computational models continues to expand (according to research studies and the specific activities for which they are used). However, certain topics are not covered or are simply ignored when computational models are applied. These areas are discussed further in Subsections 1.3 and 1.4.

1.2 The Information Dynamics of Music (IDyOM)

When examining the musical surface with artificial intelligence, it is necessary to establish proper "communication" between music and computers, i.e., the sound must be "captured," which can be accomplished through the use of symbolic representations of notes or other musical events, or audio representations of acoustic sound waves.

In this thesis, the musical surface (musical input) is captured using the Information Dynamics of Music (IDyOM), a computational cognitive model[1] of music perception (M. T. Pearce, 2005). Because the model cannot process dynamic, timbral, or textual changes, the musical input is represented symbolically in IDyOM. IDyOM learns progressively about musical syntactic structure and its sequential regularities after being exposed to a corpus of music after importing musical examples in MIDI format.[2]

The computational perceptual model IDyOM has proven to be an accurate predictor of melodic expectancy (M. T. Pearce & Wiggins, 2006a; Sauvé & Pearce, 2019), of behaviour and neural measures (EEG) of melodic expectedness (Agres et al., 2018; M. T. Pearce, Ruiz, et al., 2010), of phrase boundaries (M. T. Pearce, Müllensiefen, et al., 2010a). It has been manifested that it provides a good quantitative model of cultural distance (M. T. Pearce, 2018).

IDyOM is based on n-gram models, which are used effectively in the biological domain, in natural language processing, statistical machine learning, artificial intelligence, from 1950 on in music research related tasks (e.g., in machine improvisation, music information retrieval, cognitive modelling). An n-gram model is a collection of sequences, s, consisting of n symbols (characters/events), each of which is associated with a frequency count. The model calculates the probability of a symbol s_n based on a history $h = s_1 \dots s_{n-1}$, $P(s_n|h)$. Where $n = 1$, a monogram model, a zeroth–order model is determining the predictions, meaning that a symbol s_n is independent from the previous context (symbols). In a bigram model, $n = 2$, the probability of a symbol depends just on the previous one, and so on.

When using fixed-order n-gram models, low orders may fail to provide a good model of the global structure on the distributions, while high orders may not capture enough of the statistical regularity in a sequence. This trade-off may be addressed by using hierarchical forms of n-gram model (e.g., G. A. Wiggins & Sanjekdar, 2019), and this is arguably a necessary feature if a model is to describe the structure of sequences that include long-term dependencies (Widmer, 2016). However, IDyOM has been shown to capture the structure of melody extremely well, suggesting that such long-term dependencies are not significant in this context.

A special case occurs when in a sequence of symbols an unseen symbol is encountered, providing an estimated probability of zero (M. T. Pearce & Wiggins, 2006b; A. G. Wiggins

[1]The word "model" has been used in this paper as a term for the IDyOM 'theory-and-system', and also for some of its components (models of data). IDyOM may be downloaded from https://github.com/mtpearce/idyom/wiki

[2]Another possibility is to import musical examples in ** kern format: more in Huron (1997)

et al., 2009). These issues are addressed in IDyOM by implementing different strategies, among other, by extending the basic n-gram modelling to a Variable-Order Markov Model (VMM) over a finite alphabet Σ, where the conditional probability distributions are combined in a way that reflects the statistics obtained from the training data (Begleiter et al., 2004). VMMs, in contrast to basic n-gram models, are able to capture contexts of different length in a single probabilistic model.

IDyOM uses a complex methodology in estimating probabilities of an event e given a history h. The central component in IDyOM is a sequence prediction model, the adaptive[3] lossless data compression algorithm PPM*, an improved version of PPM (Prediction by Partial Match). In the classic PPM algorithm, originally introduced by Cleary and Witten (1984), the maximum context length is a fixed constant k, where k denotes the number of preceding events used in the prediction task. A "suite" of fixed-order context models is used, with different values of k, from 0 up to pre-determined maximum. The learning about context-dependent conditional probability distributions is gradual (Steinruecken et al., 2015). The process of prediction starts by default with the model with the largest k, followed by orders with a smaller value of k in case an event is novel, and is terminated when all the events are encoded.

A separate probability distribution in each model is calculated from counts of all the events that have followed every subsequence of length k in that particular model. If there are models with different values of k, it means that different probability distributions are obtained from these models, which are in the end effectively combined in a single model. This is achieved with blending, using an escape method, where artificial escape symbols are put in a transition from higher to lower-order context models in case a model does not contain the input symbol (Drinic et al., 2003). In IDyOM, PPM* algorithm has been used in combination with interpolated smoothing, a technique used in generating non-zero probabilities to unseen events, as it has manifested to outperform the backoff smoothing (M. T. Pearce, 2005)

IDyOM uses statistical learning and probabilistic prediction to acquire and process the internal representations of a musical piece/style. IDyOM learns about syntactic structure simply by being exposed to it while observing and analyzing content in a corpora of musical pieces. From the perspective of the feature that is analyzed with a viewpoint or viewpoints,[4] the likelihood of a forthcoming event is determined using sequential regularities (M. T. Pearce, 2018).

IDyOM simulates a listener's expectations in music (which is based on the knowledge acquired during the entire lifetime) with a *long-term model* (LTM), which accumulates statistical information about musical structure from a large corpus. As listeners are sensitive to repeated patterns in an on-going listening experience, a second, *short-term model* (STM) is also used, in which the information about the musical structure of the current piece is learned dynamically and incrementally (M. T. Pearce, 2005, 2018). LTM and STM predictions are then combined, and it has been shown that better prediction performance is achieved by combining LTM and STM dynamically (Conklin, 1990; M. Pearce & Wiggins, 2004; M. T. Pearce, 2005; M. T. Pearce & Wiggins, 2006b).

1.3 Fundamental Question

One may infer that a children's melody is a melody written specifically for children (e.g., by adults) or by children themselves. If it is a *children's* melody, it must have all the characteristics that make it appropriate for children. The children's melody range should

[3]In the information-theoretic literature the term "adaptive" is understood as "sequential"

[4]More about viewpoints in Section 3.2.2

meet, for example, the vocal range in preschool and school-aged children, include elements of perfect chords and arpeggios, simple rhythm, stable tonality (diatonic melodies within the key centers), only occasional chromatism, small intervals (without extreme leaps between tones), simple harmonic structure (with basic chords as tonic, subdominant and dominant), binary over the ternary meter, simple form, frequent repetition of one tone, motifs, phrases, and so on (Mihelač & Panić Grazio, 2021).

Thus, in order to meet the social-emotional, cognitive-linguistic, and musical requirements of children of varying ages, the children's melody must be (to a certain extent) *simple*. This raises the fundamental question of whether or not it is appropriate to use "simple" children's melodies to answer important questions such as,

Q.1.: How is music segmented in the melodies of children of different ages?

Q.2.: How complex and (ir)regular is the musical structure of a children's melody?

Q.3.: Are children's melodies in children's songs distinct enough to be recognized as a specific genre, and can they also represent the culture of a particular country?

Herzog (1944, p. 11) writes in his paper that "many children's songs in European folk music are exceedingly simple" [. . .] (and that) "the simple songs of European children are very reminiscent also of melodic types that predominate in the styles of many primitive peoples, especially those of a very simple culture." Despite the "simplicity" of melodies found in (folk) songs in many cultures around the world, these songs are (according to a large number of scientific papers and publicly available datasets) frequently analyzed and used also in cross-cultural studies, compared to children's (folk) songs, which are rarely used or even mentioned.

The musical and lyrical simplicity of pop music is likewise evident. Some of the characteristics of these melodies include stepwise diatonic melodies, short phrases, intervals rarely exceeding a third, and simple time signatures (Warner, 2003). As with folk songs, there are many publicly accessible datasets of pop songs, and pop songs are (compared to children's songs) frequently used and analyzed from the perspective of their perception, segmentation, production, consumption, subcultural identity, impact on mood, etc. (Rojek, 2011).

In light of this, another question arises: where are the reasons to be found, if the "simplicity" of children's melodies is not an impediment to answering the previously posed questions Q.1., Q.2., and Q.3.?

First of all, "children were at the sidelines of society for centuries" [. . .] " (and) the concept of children as members of a children's culture was inconceivable" (P. Campbell, 2010, p. ix). In the late nineteenth century, it was John Dewey's progressive education movement that brought attention to children as individuals with unique abilities and requirements (Cremin, 1959). In the course of the last few decades, children have become the focus of numerous scientific disciplines, as they are increasingly recognized as individuals with distinct musical interests, identities, perception, understanding, and creating of music.

Second, Herzog's claim, that "there is no evidence to show that children's songs have ever been created by children" [. . .] "(and that) songs are imparted to small children by grown-ups or older children" (Herzog, 1944, p. 11) reflects not only the problem of the authorship of children's songs but also the unrecognized capacity of children to be creative composers of their own music, with melodies that truly represent them.

This thesis, therefore, seeks to examine and reevaluate children's melodies, which are waiting to be "(re)discovered" and recognized as a valuable source of information that has the potential to shed light on a number of significant aspects of music.

1.4 Research Aims

In music structural analysis, a musical work is decomposed (segmented) into its constituent elements. Segmentation is mostly used in a purely musical-theoretical sense to determine how a piece of music is constructed, how structural units interact, and how they are connected (Ahlbäck, 2007). Due to the ambiguity of music, determining segment boundaries may be challenging (Margulis et al., 2017). Even though two persons have similar experience of a musical work, their perceptions of the musical piece or its structural parts may be quite different.

The majority of research on the mechanisms underlying perceptual grouping in music is conducted on adults. Only a few studies have examined how very young toddlers and children of particular age perceive and organize musical structure.[5] Furthermore, most of research on children is either focused on a certain age group or on specific musical dimensions and features, with little attention paid to music segmentation.

As there is currently no research on how children of various ages perceive boundaries in music or how music is segmented, the first research aim is:

"To analyze segmentation of music in children of varying ages (from preschoolers to adolescents) in order to provide a more comprehensive picture of music segmentation in humans."

In music, regularity is perceived as a highly ordered texture with prominent periodic patterns and strong neighboring relationships (Manjunath et al., 2000), in which musical ideas are organized in a way that the human mind can understand (Pole, 2014). In contrast, irregularity is experienced in an unstructured or poorly structured musical piece, where the relationship between patterns is rarely discernible and enjoyment is diminished by the enormous mental space required to process unique musical content (Kramer, 1988). The interaction between regularity and irregularity in musical structure is one of the fundamental forms of musical expression used by composers across all musical styles (perhaps with different proportions in different aesthetics).

Even though interest in children's folk songs has increased among ethnomusicologists, sociologists, educators, and folklorists since 1940, there is a lack of study on children's folk songs. The majority of research on children's folk songs is related to musical content, the social and cultural significance of children's folk songs and their relationships to adult music, the contribution of children's folk songs to cultural preservation, and the transmission of children's folk songs from generation to generation. Generally, the musical structure of children's folk songs is analyzed in terms of musical elements/dimensions, i.e., how the elements/dimensions are used (e.g., pitch span, use of meter, keys, rhythm, contour etc.).

The impact of various musical elements/dimensions on children's folk song perceptions of (ir)regularity and complexity is mostly unclear. Although cross-cultural study on melodic complexity has been undertaken, the majority of it has focused on children's lullabies (e.g., Balkwill & Thompson, 1999; Unyk et al., 1992). This may be because children's folk songs are generally considered to be regular, with a simple structure with repeated patterns, and even as a primitive layer of folk songs (Herzog, 1944). Therefore, the second research aim is:

"To examine the regularity in the musical structure of children's folk songs by observing lower and higher order musical elements."

Music is believed to be "universal." Hence, if music is universal, what distinguishes one culture's music from another, and why is music perceived as dissimilar when two or more cultures are compared, despite the fact that musical components such as tempo, pitch, and

[5] Throughout this thesis, the term "children" is used in accordance with the United Nations Convention on the Rights of the Child - UNCRC 1989 - which defines a "child" as any person under the age of 18.

rhythm are used consistently across cultures?

Numerous plausible explanations have been proposed, with biological and/or cultural explanations dominating in recent decades. Traditionally, cross-cultural research aimed at identifying (dis)similarities (either manually or computationally) has concentrated either on biological or cultural aspects. However, the study by Lumaca et al. (2018) has demonstrated, that linking biological (individual neurobiological variability and its impact, up to the population level) and cultural differences in the understanding of musical diversity could help better explain musical diversity, as music is not solely a cultural or a biological product.

Currently, studies are examining mostly the (dis)similarities between and within countries by using folk songs, in terms of, how musical dimensions and features are understood and used. As children's folk songs and children's songs are rarely employed in cross-cultural research, the third research aim is:

"To investigate the cross-cultural (dis)similarities in musical elements across three genres thought to reflect culture, namely folk songs, children's folk songs, and children's songs, in 22 European countries considered to be geographically close or with similar political, historical, economic, and cultural backgrounds."

While children's folk songs and folk songs share a number of characteristics (for example, contour, rhythm and melody patterns, and meter), the question is how comparable children's folk songs and folk songs are when adults and children share the same cultural context, and whether they can be considered distinct genres. Additionally, the fact that many children's songs contain numerous folk elements may help us understand whether this genre, together with folk songs and children's folk songs, adds to a country's recognition as a distinct culture.

1.5 Hypotheses

In the light of the content, discussed in the Subsections 1.3 and 1.4, the general hypothesis is:

It is possible to examine different aspects of music by using a dynamic computational modelling of children's melodies.

To examine the musical segmentation in children of various ages (from preschool to adolescence), and to compare musical segmentation in children of various ages with musical experts, additional hypotheses will be tested in Chapter 4.

In Chapter 4, in Section 4.2 presenting the Game Experiment, the following hypothesis will be tested:

(i) *Children are using phrases as a grouping strategy in memorizing melodies.*

In Chapter 4, in Section 4.3 presenting the Breathing Experiment, two hypotheses will be tested:

(i) *Agreement on segmentation within age groups increases with age.*

(ii) *The similarity between normative and participant's segmentation increases with age.*

To employ a computational approach to understand the segmentation of music in children of varying ages, the following two hypotheses will be tested in Chapter 5:

(i) *As age increases, higher musical features are employed for segmentation.*

(ii) *Participants with musical knowledge segment music using more complex and higher musical features.*

To investigate the regularity in children's folk songs (from the perspective of the usage of permitted rules in musical structure), two hypotheses will be tested in Chapter 6:

(i) *The frequency of repeating patterns correlates with the irregularity of musical structure in children's folk songs.*

(ii) *Repeated patterns contribute to regularity when presented at the same pitch.*

To determine which musical elements[6] differ between 22 European countries in folk songs, children's folk songs, and children's songs, two hypotheses will be tested in Chapter 7 in Section 7.3:

(i) *There are substantial variances in the use of musical features and dimensions between European countries that are regarded to share a single musical style.*

(ii) *The musical features and dimensions used in the representative music of a certain country are more similar in countries that share a similar cultural, political, historical, economic background, and are geographically close.*

To determine which musical elements differ within 22 European countries in folk songs, children's folk songs, and children's songs, one hypothesis will be tested in Chapter 7 in Section 7.4:

(i) *There are differences in the use of musical features and dimensions in genres considered to belong to the representative music of a particular country, however depending on cultural forces which homogenize and diversify these genres.*

1.6 Contributions

The contributions of this thesis are manifold:

- Three corpora have been created: (i) to investigate musical segmentation in children of various ages utilizing both a computational and experimental approach (Chapters 4 and 5), (ii) to computationally simulate the (ir)regularity of musical structure in children's folk songs (Chapter 6), and (iii) to compare and investigate the diversity of musical dimensions and features in folk songs, children's folk songs, and children's songs between and within 22 European countries using computational methods (Chapter 7). The Essen Folksong Collection (Schaffrath, 1995), and the Meertens Tune Collection (The Meertens Tune Collections, 2019) were supplemented by songs, collected by the author, from several songbooks and school textbooks featuring traditional children's (folk) songs and folk songs, courtesy of the national/school libraries from 22 European countries. All songs are freely available.[7]

[6] In this thesis, musical elements are to be considered in the same manner as musical dimensions or musical parameters. The decision to utilize "musical elements" is based on the fact that there is currently no consensus regarding which terms to employ (musical parameters, musical elements or musical dimensions).

[7] https://github.com/LMihel/LMihelac

- This thesis contributes to a deeper understanding of music segmentation in children of varying ages by conducting in-depth research in this field. The cross-sectional study of melody segmentation in children and adolescents, as well as its development from childhood to adolescence, using an experimental approach, revealed that (i) over-segmentation is present in the youngest age group (not correlated with lung capacity) and (ii) that 8–12-year-olds and 15–16-year-olds segment phrases similarly to adults, regardless of age or musical experience (Mihelač et al., 2022). Using a computational approach and comparing the computational segmentation of musical structure to the segmentation of music by children, adolescents, and musical experts, some age-related differences in segmentation were discovered while modeling human music perception.

- This thesis contributes with its experiments in the field of music segmentation to the potential repeatability of the same experiments with a larger/smaller group of children of different ages, and musical knowledge.

- An algorithm for locating "candidates" in an unknown irregular musical structure is provided, as well as a methodology for discovering (ir)regularity in musical works (Mihelač et al., 2023).

- The thesis shows that computational modelling of human perception of (ir)regularity in children's folk songs can add considerable and useful contribution to the understanding of this genre.

- A manual Classification Model (Model CMCS) for the categorization of children's songs has been developed (Mihelač & Panić Grazio, 2021). The model classifies children's songs into children's folk (traditional) songs, children's songs based on children's folk songs, and "new" children's songs, where authorship of lyrics and melody is known.

- The thesis contributes to a better understanding of the musical diversity between and within 22 European countries, using not only folk songs, but also children's folk songs and children's songs. The cross-cultural study of three genres (folk songs, children's folk songs, and children's songs) within and between 22 European countries has revealed the following: (i) significant differences in the incorporation of musical features and dimensions into the musical structure of these genres, (ii) unexpected similarities between geographically distant countries and dissimilarities between geographically close countries, as well as (iii) greater musical diversity within countries than between countries.

- The thesis demonstrates that two genres, children's folk songs and children's songs, are distinct.

1.7 Thesis Structure

The thesis is divided into eight chapters. A brief overview of each chapter is provided in the continuation.

Chapter I

Chapter 1 defines the scope of the thesis, formulates the research questions, thesis structure, and references related to the thesis's chapters.

Chapter II

Chapter 2 presents the three corpora used to investigate music segmentation in children of various ages (Corpus I), the (ir)regularity of musical structure in children's folk songs (Corpus II), and the (dis)similarities in musical features between and within 22 European countries in folk songs, children's folk songs, and children's songs (Corpus III).

Chapter III

Chapter 3 presents the musical structure in general, with an emphasis on melody and its observation using the computational model IDyOM.

Chapter IV

The segmentation of music in children of various ages (from early childhood until adolescence) is examined in Chapter 4. There are two experiments presented: Game Experiment and Breathing Experiment. The Game Experiment examines whether children and adolescents use phrases to memorize music. It tests the hypothesis that children are using phrases as a grouping strategy in memorizing melodies. The Breathing Experiment examines if breathing and phrases are coordinated while singing music without lyrics. The Breathing Experiment tests two hypotheses: (i) that agreement on segmentation within age groups increases with age, and (ii) that the similarity between normative and participant's segmentation increases with age.

Chapter V

The Breathing Experiment's results (Chapter 4) are compared to the automatic segmentation in IDyOM in Chapter 5. The musical structure is observed with different viewpoints in order to acquire new insight into how children of various ages perceive and organize music, as well as which lower and/or higher order musical features predominate in the perception and segmentation of music at a given age. Two hypotheses are tested: (i) that as age increases, higher musical features are employed for segmentation, and (ii) that participants with musical knowledge segment music using more complex and higher musical features.

Chapter VI

Chapter 6 examines the regularity of musical structure in children's folk songs using the IDyOM computational model to simulate human perception of (ir)regularity in musical structure and to find patterns that contribute to a higher/lower sense of irregularity. Two hypotheses are tested in Chapter VI: (i) that the frequency of repeating patterns correlates with the irregularity of musical structure in children's folk songs, and (ii) that repeated patterns contribute to a stronger sense of regularity when presented at the same pitch.

Chapter VII

Chapter 7 explores the (dis)similarities between and within 22 European countries using the IDyOM computational model to observe various musical dimensions and features in folk songs, children's folk songs, and children's songs. Two hypotheses are tested in examining the (dis)similarities between 22 European countries: (i) there are substantial variances in the use of musical features and dimensions between European countries that are regarded as sharing a single musical style, and (ii) that the musical features and dimensions used in the representative music of a certain country are more similar in countries that share a similar cultural, political, historical, economic background, and are geographically close.

One hypothesis is tested in examining the (dis)similarities within 22 European countries: that there are differences in the use of musical features and dimensions in genres considered to belong to the representative music of a particular country, however depending on cultural forces which homogenize and diversify these genres.

Chapter VIII

Chapter 8 outlines the main findings in relation to the research objectives and suggests possible directions for future research.

Chapter 2

Music Corpora

> The goal is to turn data into information, and information into insight.
>
> *Carly Fiorina*

To address the challenges outlined in Subsections 1.4 and 1.5, the following has been created:

- Corpus I: To use an experimental and computational approach to examine the segmentation of music in children of various ages.
- Corpus II: To study the (ir)regularity in musical structure of children's folk songs using a computational approach.
- Corpus III: To examine the (dis)similarities in musical dimensions and features between and within 22 European countries in folk songs, children's folk songs, and children's songs using a computational approach.

With the exception of folk songs and children's folk songs from "The Essen Folksong Collection," (Corpus II and III) all of the songs were initially notated as musical scores in Sibelius Notation Software and then converted to MIDI files. Due to the unique requirements of the IDyOM computational model used in this thesis, all of the songs in all of the corpora were encoded into MIDI (256 PPQN) with a piano timbre, at the same speed, and with no loudness or articulation changes before being imported into IDyOM.

Because IDyOM requires inputs to be expressed symbolically, which implies it cannot support textual modifications, only monophonic songs[1] without lyrics were used. No songs were transposed to a final tonal tone, and the original tonality was maintained in order to capture as much as possible of the tonal variance seen in European songs (e.g., Phrygian tonality in Andalusian songs, Turkish makam Hicaz and Nikriz mode, Dorian scales in Greece…) (Further refer to: Dal & Pihl, 1956; Gelbart & Rehding, 2011; Manuel, 1989). The corpora are available to the research community at https://github.com/LMihel/LMihelac and can be utilized to produce new datasets for future studies.

[1]Songs consisting of a single unaccompanied melodic line.

2.1 Corpus I

Corpus I was created to investigate the segmentation of music in children of varying ages (Chapter 4) and to compare the segmentation of songs by children, adolescents, and musical experts with the segmentation of the same songs in IDyOM (Chapter 5). There are 155 Slovenian monophonic songs, including 44 children's folk songs and 111 children's songs. The author collected children's folk songs from several songbooks and school textbooks containing traditional children's folk songs, courtesy of Slovenian national/school libraries and teachers from a number of Slovenian primary schools. Only children's folk songs were selected, either from the official Slovenian music curriculum or from songbooks used as supplementary teaching material in Slovenian kindergartens and primary schools.

Children's songs were mainly collected from various Slovenian school (educational) textbooks, which were primarily provided by school libraries and national libraries, and partially (when necessary) from the official websites of Slovenian children's music authors who were contacted for this project.

This corpus was created using current definitions and understandings of the two melody-types, namely the children's folk song melody-type and the children's song melody-type, because children's folk songs and children's songs are not yet recognized as distinct genres. In addition, the recommended criteria for recognizing songs as children's (folk) songs with a *manual* classification model, the Classification Model (Model CMCS) for the categorization of Children's Songs proposed in Mihelač and Panić Grazio (2021) were applied in the construction of this corpus.

This model CMCS takes into account (i) criteria identified as crucial in studies on song selection, (ii) the foundations of music theory, and (iii) the results of cross-cultural research on the (dis)similarity of children's (folk) songs, that can be used to define additional selection criteria. The CMCS model consists of five levels, of an initial level, and four main levels. Each level has a 'yes' or 'no' option, 'yes' option leading to another condition (or level), and a 'no' option for terminating the process, if a song does not meet the required condition(s) in a particular level.

The classification starts with the *Initial level* in which selected songs are classified by title, by applying the word 'children'.[2]

If a selected song does not meet the initial requirement, it is automatically terminated otherwise proceeded to the *First level*, in which only one condition is checked, the existence of both lyrics and melody, as a children's song is considered as a syncretic musical composition, comprising lyrics and melody, regardless of the song has an additional accompaniment or not (for example piano, organ, guitar accompaniment). If 'yes' option is chosen, the song is proceeded to the *Second level*, otherwise, it is terminated. The *Second level* has two conditions:

a) Content, suitable and understandable to children.

b) The audience consisting of children.

Even if the songs are classified as children's songs, the content of a song can be too demanding for children i.e., exceeding their understanding. The content can be also inappropriate because it contains e.g., unacceptable values, stereotypes, gender inequality, violence, etc., or does not provide a joyful, emotional experience and the possibility for self-expression (Jožef-Beg & Mihelač, 2019). In case the first condition 'a' is met, the song proceeds to the second condition 'b' which checks the audience. Although the intended audience can be

[2]In some cases, the title of a song clearly references children or their activities. When the title of a song did not indicate whether or not it is a children's melody, a unique problem arose. In this particular case, it was necessary to acquire additional information, such as the song's origin, its creation circumstances, etc.

made up of adults or children (or adults and children), in this classification model, *children* are considered as the most important audience and their acceptance of a song.

After fulfilling *both* conditions, the song proceeds to the *Third level*, otherwise, it is terminated. *Third level* deals with the authorship of lyrics and melody. At this level three possible conditions are outlined:

a) Unknown authorship of both lyrics and melody: the song is classified as a "Traditional song" (Children's folk song).

b) Unknown authorship of lyrics and known melody (or vice versa): the song is classified as a "Song based on a traditional song".

c) Known authorship of lyrics and melody: the song is classified as a "New song"

In this level, if condition 'a' is not met, the song proceeds to condition 'b' or 'c'. None of the selected songs is terminated, as this level deals only with authorship of lyrics or melody, regardless of who the author is i.e., a child or an adult (or even both, child and adult).

When a song reaches the *Fourth level*, the last and the most crucial level, the structure and features of lyrics and melody are analyzed. From the perspective of *lyrics*, analyzed are the educative contribution, pleasantness (joyfulness), understandability, stimulation to the imagination, unusual/picturesque words, syllabic text setting (lack of melisma), rhyme, assonance, alliteration, rhythm, shortness, repetition. From the perspective of *melody*, analyzed are rhythm, meter, tonality, chromatism, pitch (interval leaps), tessitura (pitch range of the song), contour, harmonic structure, basic formal units (motifs, sub-motifs, phrases, sentences/periods), song-type (e.g., one-part, simple two-part, three-part), shortness, repetition (absolute and relative).

Both lyrics and melody in this *Fourth level* have to meet the outlined criteria to proceed to two additional conditions examining the matching between melody range (tessitura) and children's vocal range: whether the song is more suitable for preschool children (condition 'a') or school-aged children (condition 'b'). The conditions 'a' and 'b' are based on the findings from studies examining the vocal range in children (Cooksey, 1992; J. Kim, 2000; Moore, 1991; Welch, 1979).

2.2 Corpus II

Corpus II was created to simulate and identify (ir)regularities in the musical structure of monophonic children's folk songs using the IDyOM computational model. The corpus consists of 736 monophonic children's folk songs from 22 European countries (see Table 2.1).

This corpus was supplemented with 44 Slovenian children's folk songs from Corpus I. The Essen Folksong Collection (Schaffrath, 1995) featured a selection of German children's folk songs (30 songs out of 124). The Meertens Tune Collection (The Meertens Tune Collections, 2019) produced twenty Dutch children's folk songs out of a total of fifty-eight.

All other songs were collected by the author, from 20 other countries from several songbooks and school textbooks featuring traditional children's folk songs, courtesy of the national/school libraries in these countries.

As with Corpus I, only children's folk songs contained in formal music curricula or in songbooks used as supplementary instructional material in kindergartens and primary schools were selected. Additionally, whenever possible, children's folk songs with a specific country of origin were included, despite the fact that the same song can frequently be found in another country, even in geographically distant European countries (for example,

Table 2.1: Countries and number of children's folk songs used in analyzing the (ir)regularity. The total is 736 songs.

Country	Num. of songs	Country	Num. of songs
Bulgaria	18	Croatia	16
Denmark	15	France	71
Germany	124	Great Britain	38
Greece	26	Hungary	27
Italy	22	Latvia	23
Netherlands	58	Norway	23
Poland	22	Portugal	27
Romania	18	Russia	21
Serbia	13	Slovenia	44
Spain	54	Sweden	29
Switzerland	23	Turkey	24

"wanderer melody" found in the Czech song *Kočka leze dírou, pes oknem*, and in the Slovenian song *Čuk se je oženil*). The CMCS model for categorizing children's songs was also applied to this Corpus.

2.3 Corpus III

Corpus III was created in order to examine, using the IDyOM computational model, the (dis)similarities in the utilization of musical dimensions and features between and within 22 European countries. Corpus III includes 2,184 songs from 22 European countries and three genres: 959 folk songs, 736 children's folk songs from Corpus II, and 489 children's songs (out of them 111 children's songs from Corpus I) (see Table 2.2).

Children's songs from European countries were selected and compiled from the songbooks of renowned children's song authors and the official websites of these European children's music authors. Additionally, the CMCS classification model was used to categorize these songs.

The compilation of 959 folk songs was assembled by combining the author's collection of 429 folk songs (using songbooks, books, and textbooks from various national and school libraries in 22 countries) with 530 folk songs from the Shaffrath Essen Collection (Schaffrath, 1995). In determining the final selection of folk songs, musical, cultural, historical, political, and other aspects of music traditions were examined where necessary.

Table 2.2: Number of folk songs (FS), children's folk songs (CFS), and children's songs from 22 European countries. The column FS (Essen) indicates the detailed number of 530 songs used from the Shaffrath Collection.

Country	FS	FS (Essen)	CFS	CS	Country	FS	FS (Essen)	CFS	CS
Bulgaria	23	0	18	18	Croatia	58	0	16	16
Denmark	22	9	15	11	France	133	131	71	22
Germany	139	139	124	21	United Kingdom	41	0	38	17
Greece	23	0	26	19	Hungary	34	34	27	16
Italy	23	7	22	12	Latvia	46	0	23	20
Netherlands	53	53	58	23	Norway	30	0	23	20
Poland	24	20	22	21	Portugal	15	0	27	15
Romania	23	23	18	18	Russia	28	28	21	21
Serbia	24	0	13	23	Slovenia	71	0	44	111
Spain	19	0	54	15	Sweden	30	11	29	17
Switzerland	75	75	23	16	Turkey	25	0	24	17

2.4 Summary

To address the research challenges outlined in Subsections 1.4 and 1.5, Music Corpora has been created: (i) Corpus I - to investigate musical segmentation in children of various ages utilizing both a computational and experimental approach (Chapter 4 and 5), (ii) Corpus II - to computationally simulate the (ir)regularity of musical structure in children's folk songs (Chapter 6), and (iii) Corpus III - to compare and investigate the diversity of musical dimensions and features in folk songs, children's folk songs, and children's songs between and within 22 European countries using computational methods (Chapter 7).

The concepts of purpose, coverage, completeness, quality, and reusability were considered in the building of corpora (Serra, 2014). The CMCS classification model and criteria were utilized in the construction of corpora containing children's folk tunes and children's songs. All songs are freely available at https://github.com/LMihel/LMihelac.

Chapter 3

The Observation of Melody

> One can imagine a melody as a distant echo of something more primal—the direct expression of emotion in the form of a raw cry.
>
> *Roger Mathew Grant*

What makes one piece of music stand out from another? Why are some musical pieces more easily recognized, recalled, or appealing to listeners than others? Which of the musical dimensions, out of all the others, retains the listener's attention the longest? "It is the melody" (Selfridge-Field, 1998), "the one quality in music which really counts" (Dunhill, 1907), according to Aristotle, "the root of our enjoyment of music [...] the source of musical creation itself" (Schoen-Nazzaro, 1978). Humans are intrinsically capable of reproducing, memorizing, and even recalling a melody among all the musical dimensions in a musical piece. "Melody is an important descriptor of music" (Marolt, 2008, p. 1619). It contains data about rhythm, (implicit) harmony, loudness, timbre, pitch, spatial position, and reverberant environment (Y. E. Kim et al., 2000).

Melody is derived from the Greek term μὲλος (melos) and has both musical and metaphorical implications. It is used in the musical sense as a basic musical dimension, supplemented by descriptions of compositional or aesthetic features.

Additionally, it is a synonym for song (Beekes, 2010; Mihelač et al., 2018). Numerous and diverse definitions of melody published in the literature take various aspects of melodic characterization into account. Melody is a semi-autonomous segment of a musical piece in which the melody is recognized as the same regardless of the arrangement or context in which it is delivered (Bartlett & Dowling, 1980; Levitin, 1999; Stefani, 1987).

What is interesting is that everyone agrees that melody is a sequence of sounds and an auditory object. A sequence of more or less well-balanced and ordered sounds can be heard worldwide even in the earliest forms of melody. On the other hand, the evolution of melody demonstrates that ordered sounds and their succession in a melody were (obviously) insufficient to distinguish, attract, and pleasure listeners.

This explains the various approaches to two important aspects of melody creation: the musical scale[1] and melody design, which Dunhill defines as "the ordering of the various factors of effect in their proper places" (Dunhill, 1907, p. 102).[2] Throughout history, Western and Eastern musical traditions have revealed that these two aspects of music are

[1] Musical scale and mode will be utilized interchangeably in this thesis as a 'method,' or 'modus operandi,' i.e. the way pitches produce the melodic contour.

[2] The most important aspects of 'melody design' include duration, pitch, texture, timbre, loudness, measure, and form, despite the fact that there is no consensus on what constitutes a 'melody design'.

interpreted differently (as evidenced by the wide range of musical scales and design used to create melodies).

Melody has been studied for more than three centuries by researchers in a variety of fields (anthropology, musicology, computational musicology, socio-ethnological studies, musical theory, etc.). Melody is studied as a *subjective object* (for example, whether the melody is considered simple or complex, why listeners find some musical pieces more or less "melodic," and in which musical pieces is "melodicity" of a melody more or less present, how popular is the melody) and as a pure *technical object* (Pachet, 2009), for example, how musical features and dimensions are used in the melody.

Melody as a *technical* object can be considered as a "set of hypotheses what is heard in music,"(Cuddy et al., 1981, p. 872) that can assist in clarifying how melody is perceived and interpreted as a *subjective* object. Melody is also explored in the context of various *musical genres* throughout history (e.g., Gold et al., 2019; Montagu, 2017), as well as in Western and Eastern *musical traditions* (e.g., Bowling et al., 2012; Jouste, 2009).

Melody is examined in terms of its belonging to a *melody-type*, which Slobin defines as "a group of melodies that are related, in that they all contain similar modal procedures and characteristic rhythmic and melodic contours or patterns" (Slobin, 1996, p. 186).

Most cultures organize their music into melody types and use the melody-type as a "schema," or "model." The listener can recognize the basic outline even when a melody is improvised (van Kranenburg et al., 2012), utilizing various combinations and recombinations of traditional elements within the framework of a melody-type (Zonis, 1983, p. 274). Melody-types with features of some early layers of Gregorian chant found in Arabia (*maqām*), Syria (*ris-qole*), and Europe are examples of such a combinative approach.

It is worth noting that, despite the vast amount of literature on melodies on either a subjective or technical level (or both), and regardless of which melody-types predominate in folk music, classical music (medieval, Renaissance, Baroque, Classical, Romantic,. . .), jazz music, and popular music genres, two groups of melodies have received the least attention: children's melodies and children's folk melodies.

Neither of these melody groups is recognized as a distinct genre or as a melody-group at the present (more on this in Chapter 7). Due to the focus of this thesis on *monophonic* children's melodies and children's folk melodies from the Western musical tradition, both melody groups will be analyzed as a subjective and technical object as well as in terms of melody-types.

3.1 The Observation of Melody as a Technical and Subjective Object

Melody is considered a *horizontal* dimension of music; a "horizontal expression of a (musical) idea in a single line [. . .] incorporated in the horizontal plane as successive sounds" (Busch & Graubart, 1986, p. 10–12), as opposed to the *vertical* presentation of sound (the vertical dimension of music), in which sounds appear in the vertical plane simultaneously (e.g., chords).

Melody can be observed and evaluated both horizontally and vertically, as well as through observing and analyzing its "building blocks", where a building block is to be understood in the same manner as *feature*, "characteristic part of something [. . .] that helps to distinguish one thing from another (or one group of things from another group of things" (Huron, 2001, p. 2). It is expected that a proposed feature/building block (e.g., interval, pitch, duration, loudness, motif,. . .) is present in a melody, however, it can also be absent. In this case, we could refer to this property as "negative presence," which requires long-term listening experience or a set of established expectations, since the absence of a certain feature cannot be identified if it is not understood what is typically present (Huron, 2001).

It is necessary to perceive a melody as a "whole at some level," in which substructures and individual events are to some degree related to one another; otherwise, the substructures and events would be perceived only as a "chaotic temporal pitch structure" (Ahlbäck, 2004, p. 13). Nonetheless, an in-depth analysis of the melody's substructures (e.g., phrases, motifs) and individual events can reveal useful information about how the melody is organized (e.g., which parts/individual events of the melody are repeated frequently, how they are repeated), as well as about musical features and musical dimensions that distinguish the melody from other melodies.

When we analyze a melody, which is defined in this thesis as an entity, a sequence of events, each of which has a set of attributes that define the melodic properties of sound (Gómez et al., 2003), we have the option of analyzing it in its (i) entirety, (ii) in parts, or (iii) on its most basic level of musical detail, its surface (Jackendoff, 1987), analyzing each sound in the melody as a discrete musical event e. In this thesis, melody will be observed and analyzed as a horizontal dimension, and where necessary as vertical dimension. Features in higher and lower structural parts as phrases and motifs of the melody will be observed and analyzed as well, all the way down to the level of the discrete musical event e, which can be examined separately or in multiple ways as seen in Figure 3.1.

In Figure 3.1 there are 11 events. It is possible to observe merely *intervals*, the *distance* between two events. When we compare the distance between events $e1$ and $e2$, we can see that it is significantly larger than the distance between events $e10$ and $e11$. The melody also reveals that certain events have the same *duration* (representation of time), such as $e1$, $e2$, $e3$, $e5$, $e11$, while others, such as $e4$, $e6$, $e8$, $e10$, have different durations. Multiple features of a single event can be seen for example in event $e1$; we can notice the *duration* (quarter note), *pitch* (position of a single sound in the complete range of sound) B4, *loudness* (the intensity of auditory sensation produced) $p =$ soft , and *timbre* or color (quality of auditory sensations produced by the tone of a sound wave), to be played with piano sound. We can observe also the pitch range (*tessitura*), i.e. the general range of pitches found in this melody.

A closer examination of the same melody in Figure 3.1 indicates that the *tessitura* (pitch range) spans D4 to F#5. Drawing a line from $e1 - e11$ gives the melody its shape or *contour*. The *tonality*, or key, specifies pitch "g" as the tonal center and views all pitches as scale degrees (the position of a note in a scale) around that center (G major in this melody, marked directly after the treble key with a #). The *meter* (2/4) ensures that the events (which can have the same or different durations) are structured according to the number of beats (in this case 2 beats) and the value of the fundamental beat (in this case 4 = quarter value).

Because of the thesis's focus on monophonic melody in children's songs and children's folk songs, and to a lesser extent, folk songs, not all musical features and dimensions will be explored, due to the songs' brevity and the IDyOM computational system (see Subsection 3.2) used to analyze these three groups of songs), which does not support timbre and loudness analysis, as well as lyric analysis.

In Chapter 4, the emphasis is on the melody as a subjective object, with the musical structure (melody as technical object) being used to analyze the perception and segmentation of children's (folk) melodies in children of various ages, as opposed to Chapter 5, which examines a simulated computational perception of children's (folk) melodies as subjective and technical objects.

In Chapter 6, computational models of melody perception as a subjective object, as well as in-depth analysis of musical structure (the observation of melody as a technical object), are used to examine the (ir)regularity of musical structure in children's folk songs.

Chapter 7 gives insight into melody as a subjective and technical object in children's songs, children's folk songs, and folk songs, examining whether there are differences in how musical dimensions and features are used between and within 22 European countries, and whether distinct melody-types can be defined for both groups of melody in children's songs and children's folk songs.

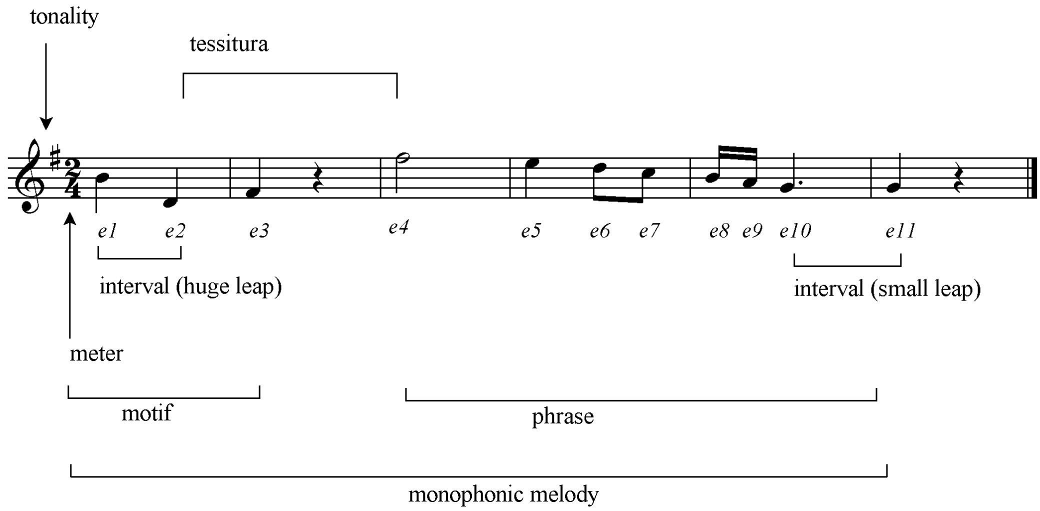


Figure 3.1: Representation of a melody.

3.2 The Observation of Melody with IDyOM

3.2.1 Entropy and Information Content

IDyOM analyzes and predicts musical events using two information-theoretic measures: entropy and information content (MacKay, 2003). Although both measurements are fundamentally similar and were used interchangeably by C. Shannon and Weaver (1949), they can be thought of as usefully different. Shannon's entropy, or simply entropy, is a measure of uncertainty that is related to the concept of entropy in physics. It is the average amount of bits necessary to transmit/represent statistical uncertainty about the character of an event chosen at random from a probability distribution. When a single event dominates a probability distribution, there is less uncertainty. When no event dominates another, as is the case with an equal or nearly uniform discrete distribution, a larger or maximum entropy is expected.

Entropy, as defined by C. E. Shannon (2001), and shown in equation 3.1, is used as baseline model for quantifying the uncertainty in the prediction of a musical event *before* it is heard:

$$H(e_i) = -\sum_{x \in X_i} p_x \log_2(p_x) \tag{3.1}$$

where X_i is the set of all possible continuations of a given musical sequence before event e_i (in our case, the alphabet of a viewpoint) and p_x is the probability of event e_x.

Hansen and Pearce (2014) provide evidence that entropy is correlated with perceived uncertainty; mathematically, it is related to the balance of likelihoods in a distribution: a uniform distribution has maximum entropy (because all events have equal probability so we can not predict them better than purely guessing) while zero (i.e., minimum) entropy entails that exactly one value of the distribution has probability 1.

The idea behind quantifying *information content* (MacKay, 2003), which is significant from the perspective of compressibility (Bell et al., 1990), is to estimate how *unexpected* an event is in the context in which it happens. Thus, information content is a measure of unexpectedness (sometimes called "surprisal") of an event which actually appears in the sequence. For a discrete event e_i, it can be calculated as shown in equation 3.2, where p_i is the probability of event e_i.

Rare events in context have a low probability, are more surprising, and have a higher information content. Frequent events in context have a high probability and low information content, and so are less surprising. If the probability of an event is 1 (when there is no surprise), the information content is zero.

$$IC(e_i) = -\log_2 p_i, \tag{3.2}$$

3.2.2 Viewpoints

When a sequence of events (notes) is given, we can define viewpoints, which are functions (variables, features) that accept initial sub-sequences of the sequence and measure a particular feature inside it, such as pitch, duration, or the relationship between two tones. Pitch (`cpitch`), and time (`dur`) are the two fundamental viewpoints from which IDyOM describes events in a sequence.

Additionally, we can compute with IDyOM so-called *derived* viewpoints (derived from basic viewpoints), *linked* viewpoints (when two or more viewpoints are linked), *threated* viewpoints (types of viewpoints defined at specific points in a piece of music, for example, the first event in each bar), and *test* viewpoints (which return a Boolean value indicating whether or not a particular condition is satisfied). An innovative alternative is to use `viewpoint selection`, a hill-climbing procedure that integrates multiple viewpoints in order to minimize the information content of a dataset.

This thesis employs a variety of viewpoints, depending on the research topic being discussed, which may include an examination of (ir)regularity, segmentation, or (dis)similarity in the way musical features and/or musical dimensions are utilized. The used viewpoints are: **cpitch, cpitch⊗dur, dur-ratio, inscale, cpint, cpint-size, cpitch-class, cpcint, cpcint-size, contour, newcontour, cpintfref, cpintfip, tessitura, cpitch⊗ioi**, cpitch⊗ioi-ratio.

Their selection is based on research showing their usefulness in investigating musical structure, simulating the listener's cognition and perception of music, as well as enculturation (e.g., Gingras et al., 2016; Mihelač & Povh, 2020b; M. T. Pearce, 2005, 2018; M. Pearce et al., 2010).

The basic viewpoint `cpitch` specifies the chromatic pitch of each event (chromatic notes are counted from middle C = 60 up and down). `Dur` is another basic viewpoint that quantifies duration in basic time units. There is currently no consensus about how listeners' pitch and temporal structure are processed, i.e. whether these dimensions are managed independently or interactively (see more in Justus & Bharucha, 2003; C. L. Krumhansl, 2000; Mihelač et al., 2023; M. T. Pearce, 2018; Volk, 2016). IDyOM supports the processing of pitch and duration independently (using two separate models) or in combination (using a single model that combines these two dimensions). `Cpitch` and `dur` are employed interactively in this thesis as a linked viewpoint `cpitch⊗dur`. The derived viewpoint `dur-ratio` represents the duration of last/duration of previous events in a sequence.

The derived viewpoint `cpitch-class` denotes a pitch class (or chroma) where note E2 is identical to notes E3, E4, and E5. The viewpoint reflects the significance in the perception of equivalence relations between two pitches, the most influential of which, according to various studies, is octave equivalence (Hoeschele et al., 2012; Patel, 2003a; Peter et al., 2008).

The derived viewpoint `cpint` denotes an interval, a relationship between two consecutive pitches. Additionally, a derived viewpoint, `cpint-size`, captures the size of intervals and can be used to explore intervals that are larger than the largest interval in the data. Intervals have been demonstrated to be significant in a variety of studies, including recognition of melodies (McDermott et al., 2008), orienting the listener in the scale, and facilitating musical tonality (Sloboda & Parker, 1982; Trehub et al., 1999). The viewpoint `cpcint` denotes the octave equivalent pitch class interval (M. T. Pearce, 2005, p. 70), and is derived from `cpint`, in the same manner as `cpitch-class` is derived from `cpitch`. `Cpcint-size` is the absolute value for cpcint.

`Tessitura` is a viewpoint derived from `cpitch`. It examines a melody's pitch range. While certain melodies have a broad range, the range of pitches is limited, and the center of that range appears to be favored (von Hippel, 2000). Thus, this viewpoint captures pitches that are extremely high or extremely low in relation to the pitch range's center.

Different intervals are used to produce the shape of a melody, the contour, which can be rising or falling. It is represented by the viewpoint `contour`. Numerous studies examining contour from a variety of perspectives, including psychological, music theoretical, and computational, confirm that contour is a vital part of musical perception (e.g., Marvin, 1991; Morris, 1993; Narmour, 1990; van Kranenburg, 2010). The viewpoint `newcontour`, derived from `cpitch`, focuses on contour changes and their relationship to the preceding contour (value 1 is given if contour is unchanged from the preceding contour, and 0 if it is different).

`Cpintfref` denotes the pitch interval between two events. This viewpoint, which has been used in numerous studies (e.g., Gingras et al., 2016; Graves & Oxenham, 2017; Marmel & Tillmann, 2008), has demonstrated that listeners' perceptions of melodic structure are influenced by the tonal hierarchical relationships within a scale, in which the tonal function (the first scale degree) is the most stable and serves as the anchor point for a key (C. L. Krumhansl, 1990). The pitch interval of an event from the first event in the piece is represented by the viewpoint `cpintfip`. This viewpoint was chosen to illustrate the impact of primacy on perceptual and structural salience (A. J. Cohen, 2000; Vos, 2000). The derived viewpoint `ioi`, which represents the Inter Onset Interval ratio, is utilized separately and in combination with `cpitch`, resulting in `cpitch`$\otimes$`ioi`, which shows to be effective in determining cultural distance (M. T. Pearce, 2018). `Inscale` reflects both "in-key" and "out-key" tones. Tonality is expressed in musical compositions through the use of tones that occur frequently together (in-key tones) rather than tones that are considered out of key (Cancino-Chacón et al., 2017).

3.2.3 Probability, Entropy and Information Content using Viewpoint `cpitch`

Figure 3.2 depicts 14 events in an excerpt from the well-known melody from "The X-Files," which were used to examine the probability, information content, and entropy of each event.

IDyOM computes the corresponding sequence of probabilities, entropy, and information content values using the viewpoint `cpitch`$\otimes$`dur`, which analyses the feature pitch *and* duration of each event in the series. As indicated in Table 3.1, the model identifies the following events as highly expected: $e10$, $e11$, and $e14$, based on their high probability

Figure 3.2: The events depicted in the "X-Files" melody (Mark Snow).

values and low information content. On the other hand, the most surprising events were discovered to be $e7$ and $e12$ (very unexpected intervals/leaps) with low probabilities and a high information content.

Table 3.1: Probability, entropy, and information content values assigned to each event in the sequence for "The X-files" using viewpoint **cpitch⊗dur**.

Event	Probability (cpitch⊗dur)	Entropy (cpitch⊗dur)	Information content (cpitch⊗dur)
e1	0.047	5.49	4.41
e2	0.018	4.57	5.81
e3	0.144	4.65	2.79
e4	0.017	2.45	5.85
e5	0.017	3.93	5.91
e6	0.098	4.33	3.36
e7	0.011	5.16	6.50
e8	0.089	5.24	3.50
e9	0.278	4.54	1.84
e10	**0.607**	2.79	**0.72**
e11	**0.525**	2.35	**0.93**
e12	0.003	2.96	8.18
e13	0.051	5.04	4.29
e14	**0.485**	3.72	**1.04**

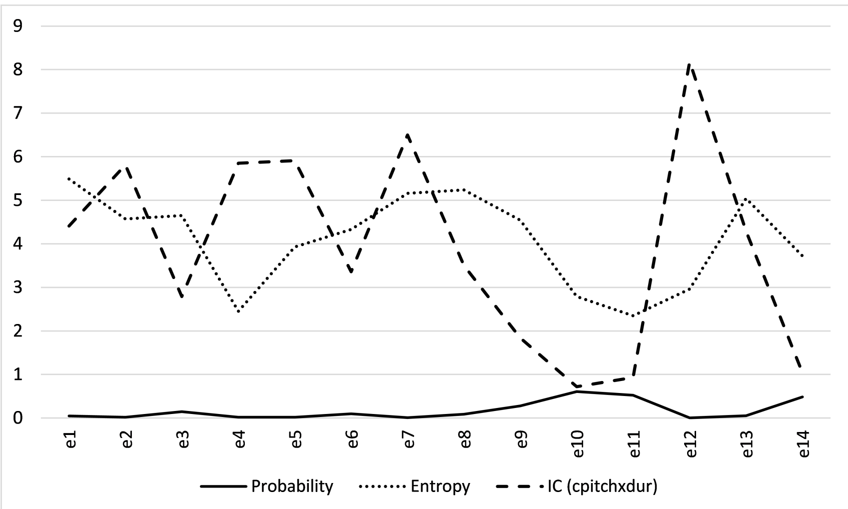


Figure 3.3: Visualizitation of probability, entropy, and information-content values for each event in "The X-files" musical excerpt obtained using the viewpoint IC_cpitch⊗dur.

3.3 Summary

This chapter presents the observation of a monophonic melody, which is described in this thesis as a sequence of discrete musical events. Melody can be examined as a subjective object (e.g., whether the melody is considered simple or complex, why listeners find some musical pieces more or less "melodic"), and technical object (e.g., how musical features and dimensions are used in the melody). In addition, melody can be examined in terms of its belonging to a melody-type (a group of melodies that are related, a certain schema or model used to organize the music).

As the topic of the thesis is *children's melody*, these approaches will be used with two genres that have received the least attention and are not currently recognized as a unique genre or melody-group: children's folk songs (CFS) and children's songs (CS). The melody of folk songs (FS) will also be studied, although to a lesser extent. The decision to use all three genres (FS, CFS, CS) is to understand how musical features and dimensions are utilized and perceived by a listener in each of these genres, and how each genre contributes to the national identity of a particular country.

It is regarded that melody is both a horizontal and vertical presentation of sound. The vertical dimension will be analyzed as implied harmony in the three previously mentioned genres. Furthermore, not all of the features of melodies will be observed and analyzed. This is due to the brevity of songs in all three genres and the computational system IDyOM, which does not support timbre and loudness analysis as well as lyric analysis.

IDyOM and viewpoints (implemented functions in IDyOM) will be used to simulate the listener's perception of melodies and to analyze monophonic songs on different hierarchical structural levels by measuring the entropy (E) and information content (IC). Viewpoints will be used depending on the research topics, which may include an examination of (ir)regularity, segmentation, or (dis)similarity in the usage of musical features and dimensions. Their selection is based on research showing their usefulness in investigating musical structure, simulating the listener's cognition and perception of music, as well as enculturation.

Chapter 4

Segmentation of Melody in Children and Adolescents

> When we are listening, the processing of music seems effortless. However, we are engaged in analyzing, segmenting and encoding a complex stream of sound.
>
> *Susan Hallam*

Segmentation (Cambouropoulos, 2006), chunking (Miller, 1956), and grouping (Lerdahl & Jackendoff, 1983b), all refer to the partitioning of musical content into sequential perceptual units during listening. Numerous studies (e.g., Knösche et al., 2005; Kragness & Trainor, 2017; Lerdahl & Jackendoff, 1983b) demonstrate the significance of segmenting auditory information for the perception and understanding of music. The segmented units may explain why some musical features or dimensions in the musical structure are perceived as perceptually more salient than others, as well as how musical knowledge influences the use of these features and dimensions in the segmentation of music (M. Phillips et al., 2020).

According to the findings, listeners organize musical information hierarchically, from smaller to larger segments. This occurs mostly at the lowest level (e.g., notes, chords, motifs) and progresses to the highest level (e.g., musical section), predominantly in phrases (Deutsch, 2013; Knösche et al., 2005; C. L. Krumhansl & Jusczyk, 1990). Grouping varies among listeners and is impacted by factors such as individual differences, age, musical experience, and the musical dimensions or features (e.g., Mihelač et al., 2021; Schön & François, 2011; Tillmann et al., 2000).

Grouping in music appears to result from Gestalt-like principles (proximity; similarity; good continuation; common fate: Deliège, 1987; Lerdahl & Jackendoff, 1983b; M. T. Pearce, Müllensiefen, et al., 2010b). Two of them seem to be present early in human development: *proximity*, in which groups are formed from elements that are close together rather than those that are further apart, and *similarity*, in which configurations are built from elements that are similar (Thorpe et al., 1988; Trehub, 1987). However, little is known about the perception and processing of phrases and boundaries in melody in the 5-to 6-year-old preschool population. In addition, little is known about how children perceive musical phrases and boundaries from preschool to adolescence. In 2018 and 2020, two experiments on the segmentation of melodies by children and adolescents of various ages were conducted: the Game Experiment and the Breathing Experiment. In the first experiment, in the Game

Experiment, it was examined if phrases are used in music memory. The second experiment, the Breathing Experiment examined the relationship between breathing while singing a song and phrase perception.

This chapter is organized as follows. Section 4.1 presents the perception of phrases and phrase cues from a psychological and musical-theoretical perspective. The important findings of the Game Experiment, which examined segmentation and memorizing in children of different ages, are presented in Section 4.2. The results of the Breathing Experiment are reported in Section 4.3, and Section 4.4 provides a summary of the chapter's findings as well as suggestions for future research.

4.1 Scientific Background and Related Work

A number of studies of musical features that are significant in music in general, but especially in Western European music, have demonstrated that infants can recognize a variety of musical components (e.g., intervals, rhythm, exact pitches, contour). They are already able to analyze complex auditory data and can sequence tones, similar to adults (D. Campbell, 2002; P. Campbell, 2010; Corrigall & Schellenberg, 2015; Thorpe et al., 1988; Trehub, 1987; Volchegorskaya & Nogina, 2014).

With continued exposure to different environmental musical stimuli, preschool children can organize rhythmic information required for musical melody perception, detect melodic changes that alter not only the contour of the melody but also the interval size (Morrongiello et al., 1985), and use phrase cues as a grouping strategy in memorizing melodies (Demorest, 2000).

According to a recent study, 5– and 6–year-old children utilize statistical learning about musical structure from their environment and have already developed solid melodic expectations. These results are consistent with previous research which found that statistical aspects of auditory input are used in music and speech segmentation (François et al., 2012; Saffran et al., 1999; Schön & François, 2011).

Various studies indicate that musical experience and age improve musical content segmentation and phrase recognition (Bertrand, 1999; Eitan & Granot, 2008; Iversen et al., 2008). Surprisingly, there is little consensus that musical training has a substantial effect on segmentation accuracy. When specific musical abilities are measured, children who have been given music lessons do better than those who have not (Ireland et al., 2018); yet, in segmentation studies, small or moderate differences have been discovered between musically trained and untrained children (Bertrand, 1999; Koniari et al., 2001). Similar findings have been discovered in study on adults (Deliège, 1987; Nan et al., 2006).

On the contrary, both Schön and François (2011) and François et al. (2012) indicate improved segmentation outcomes in trained individuals in longitudinal studies. The effect of music training on speech segmentation in 8-year-old children was examined using behavioral and electrophysiological measures, as well as a test-training-retest procedure. Only the group with musical training demonstrated improved speech segmentation skills.

Language and music share a basic property: hierarchical structure (Lerdahl & Jackendoff, 1983b). A speech stream can be segmented from basic into complex units, and likewise a general audio stream (Chiappe & Schmuckler, 1997). Both entities are governed by syntax, a set of rules, and in both entities some events (units) are more expected than others (Patel, 2003b). In sung music in general, defined as "a sequence of syllables ('text') that is sung," as opposed to the more general and common "musical setting of words, as many songs contain vocables, such as 'tra-la-la', which are not words and yet are part of a song text" (Wang et al., 2010, p. 3), structural units can be found both in the lyrics and in the melody. In children's songs, in which text and melody can each have their own structure and organization of structural units, these units are usually intertwined (Jožef-Beg & Mihelač, 2019) and mostly aligned (for example, aligning stressed syllables with a strong musical beat).

It is unclear how lyrics and melody processing varies among age groups, whether as separate or combined entities, and which, if either, is more significant. According to Serafine et al. (1984) and Fedorenko et al. (2009), the processing of structural components of language and music appears to be integrated in song perception; yet, Besson et al. (1998) gives evidence for autonomous processing (Calvert & Billingsley, 1998; Lebedeva & Kuhl, 2010). Infants and preschoolers appear to pay more attention to phonemes than melody when listening to music. Morrongiello and Roes (1990) demonstrate some integration of words and melody in children as young as five, with integration increasing with age.

Many areas of the brain, some of which overlap, are involved in the processing of language and music, therefore it is natural for the perception of lyrics to influence melody and vice versa (Brown et al., 2006; Fedorenko et al., 2009). Thus, segmentation and the subsequent structural units found in lyrics and melody, if integrated, would be affected by variations in the melody or lyrics of a song (Crowder et al., 1990), i.e., when the lyrics and melody are mismatched (for example, by substituting original lyrics with meaningless words, or adding to the lyrics melody from another song).

Thus, deviations in the melody or lyrics of a song would affect segmentation and the resulting structural units when integrated (Crowder et al., 1990), i.e., when the lyrics and melody are mismatched (for example, by substituting original lyrics with meaningless words, or adding melody from another song to the lyrics). Demorest (2000) demonstrated that changing the original lyrics of a song with lyrics in a foreign language affects melody segmentation and the participants' memory of melody segmentation.

Weiss et al. (2012) report that vocal melodies are more easily memorized than instrumental melodies because vocal music induces a higher level of arousal, which (perhaps) results in enhanced processing of musical structure and enhanced memory of structure-specific features. To date, no study has demonstrated that playing a vocal song instrumentally affects segmentation accuracy in children or across age groups.

Petzold conducted a five-year longitudinal study on the development of auditory perception in the domains of melodic perception, phrase learning, melodic reproduction with varying harmonies and timbres, and rhythmic ability (Petzold, 1963; Petzold, 1966). Age plays a significant role in the development of auditory perception, which was his primary hypothesis. Auditory perception reached its peak for the majority of tasks by the age of eight, with indications that the most significant development happened between the ages of six and seven (first and second grades). The results of a study reported by Zimmerman (1971) are similar. The conclusion that can be drawn from this overview of perception study findings is that the perception of musical stimuli follows a developmental sequence.

4.2 Segmentation in Children: The Game Experiment

4.2.1 The Game Experiment (2018)

The Game Experiment conducted in 2018 partially replicates Demorest's experiment (Demorest, 2000), testing the same hypothesis: that the children are using phrases as a grouping strategy when they memorize melodies. Melody segments, representing phrases of a melody, were used in a reconstruction task that requires a participant to put the phrases in the correct sequential order, to rebuild the entire melody. As in Demorest's original experiment, the experiment in the reconstruction of melody consisted of two conditions: the first using melody fragments matching the phrases in the melody under a normative segmentation; and the second using non-normative melody segments not matching the phrase boundaries.[1] However, differently from Demorest, the number of melody segments in the two conditions differed, and followed only the Demorest's idea of putting phrase boundaries at unexpected places in the melody. This can result in the appearance of more melody segments in the second condition than in the first. The Demorest's decision to require the same number of segments in both conditions was presumably based on his success measure: the number of moves required to complete the task, which clearly requires a level playing field. However, in this study, the moves were not counted nor the number of times participants had to listen and to check the segments in order to complete the task. Measured was only the participants' success (defined as successful/not successful) in the first and second condition.

In Demorest's experiment, the usual lyrics in all the songs were removed, and substituted with lyrics in a foreign language to avoid text cues, whereas in this experiment the usual lyrics were substituted with meaningless words (e.g., "mali-bali", "mame-mimo-momu" . . .), to negate the impact of any text cues on the perception of melody, which could be present even in lyrics with unfamiliar language because of a strong bond between lyrics and melody in a song (the "integration effect": R. L. Gordon et al., 2010; Serafine et al., 1984; Serafine et al., 1986). In a period of two weeks, in March 2018, three choir sessions under the guidance of experienced music teachers were conducted in each institution (kindergarten, primary school and secondary vocational school). These sessions focused only on learning the three songs.

Hypothesis

The hypothesis is that children are using phrases as a grouping strategy in memorizing melodies.

Participants

104 out of 106 participants from two kindergartens, one primary and one secondary school took part in The Game Experiment. These institutions were selected based on their willingness to participate in the experiment and on the number of participants with formal musical experience, which was verified prior to participation of these institutions in the study. The participants were divided into three groups, depending on their age: group "early" from 5-6 years (hereafter, early), group "middle" from 8-12 years (hereafter, middle), and "adolescent", from 14-16 years (hereafter, adolescent), further dividing each group into two subgroups according to their musical knowledge:[2] subgroup "with", and subgroup "without" formal musical knowledge. In dividing children into groups, developmental milestones were taken into consideration (Shaffer & Kipp, 2009), as well as musical ability development (Berland et al., 2015; Gooding & Standley, 2011), and the specifics of the Slovenian education system.[3]

[1] In this thesis the term "normative" means the most probable interpretation according to music theory and musical experience. This terminology is used to avoid the implication that there is in general a single correct interpretation.

[2] In this thesis, formal musical knowledge is defined as knowledge obtained in formal private and state music institutions, framed and valued in curriculum. In the first group of children aged 5 to 6 years old, "musical knowledge" is defined as experience obtained through informal activities and practices (e.g., childcare musical activities, weekday concerts, music and movement activities,. . . inside and outside the family, kindergarten, and private/state music institutions.

[3] In the Slovenian school system, preschool children are enrolled (depending on the year of their birth) in kindergarten for up to six (seven) years, and then in primary schools from six (seven) to fourteen (fifteen) years, for a total of nine years divided into three three-year periods. Depending on their year of birth, children start a three- or four-year secondary school at fourteen (fifteen) years of age.

Out of 104 participants, 26 were in the group early (11 with, 15 without musical knowledge, avg. age = 5.62), 21 participants were in the group middle (6 with, 15 without musical knowledge, avg. age = 9.90), and 57 participants were in the group adolescent (28 with, 29 without musical knowledge, avg. age = 15.16).

Methodology

A binary indicator was used for the results obtained in each song for each condition (1 = solved; 0 = not solved). The results of rebuilding all the three songs in each condition and for each participant were checked by an independent musical expert.

Stimuli

Three unfamiliar children's songs were used for the Game Experiment: two Slovenian children's songs, "Mucek na dimniku" (hereafter referred to as Song 1), "Pet junakov" (hereafter referred to as Song 2), and one Slovenian children's folk song "Pozdravljena trobentica" (hereafter referred to as Song 3). The choice of these songs was determined by the features that a children's song or children's folk song should have: brevity (all of them have only 8 bars); small pitch span—within the 5th (except for Song 2 with a pitch span from D4 to C5)[4]; repeated motifs; small intervals; tessitura within the singing range of children; and major mode, which children prefer (P. Campbell, 2010; Jožef-Beg & Mihelač, 2019; Nettle, 2000; Stadler Elmer, 2015). Figure 4.1 shows three example songs and segmentations, and the number of events (eN) in each song.

The normative and non-normative phrasings (denoted by slurs in Figure 4.1) were determined by five musical experts. The musical experts were selected based on their formal musical education (of the highest level), musical experience with children's (folk) songs literature, and long-term choir leadership. Following the submission of normative and non-normative phrases in songs including meaningless words by each of these experts, a meeting with the experts was held to discuss discrepancies in phrasing. After their concordance, the final normative and non-normative phrases were obtained.

Procedure

A child-friendly on-line game was used, created especially for this experiment with two conditions (Petric, 2018). In the first condition, used were the melody segments equal to the phrases given by the normative segmentation of three children's songs. In the second condition melody segments equal to the non-normative segmentation were used more. The tasks in the first and second condition were to reconstruct the melodies in the correct order.

Before taking part in the The Game Experiment, all the participants were asked to learn three unfamiliar songs to ensure the internal validity of the song-learning task.[5] The songs were taught by teachers of the subject "Music" during three choir sessions over the course of two weeks. The teachers from all the institutions were given instructions before teaching these three songs: not to use any gestures and facial expressions that could suggest phrasing throughout the learning of the songs, and to sing each song without breathing from the beginning to the end of the song, in order not to influence the children's phrasing with their breathing in these songs. Singing each song without breathing was possible, even using a moderate tempo, because each song consists only of eight bars. Each teacher was observed in each choir session, to validate their teaching of the songs and not breathing while singing them.

[4]The American Standard Pitch Notation (ASPN) is used in this thesis.

[5]Internal validity means that the learning of songs was not biased.

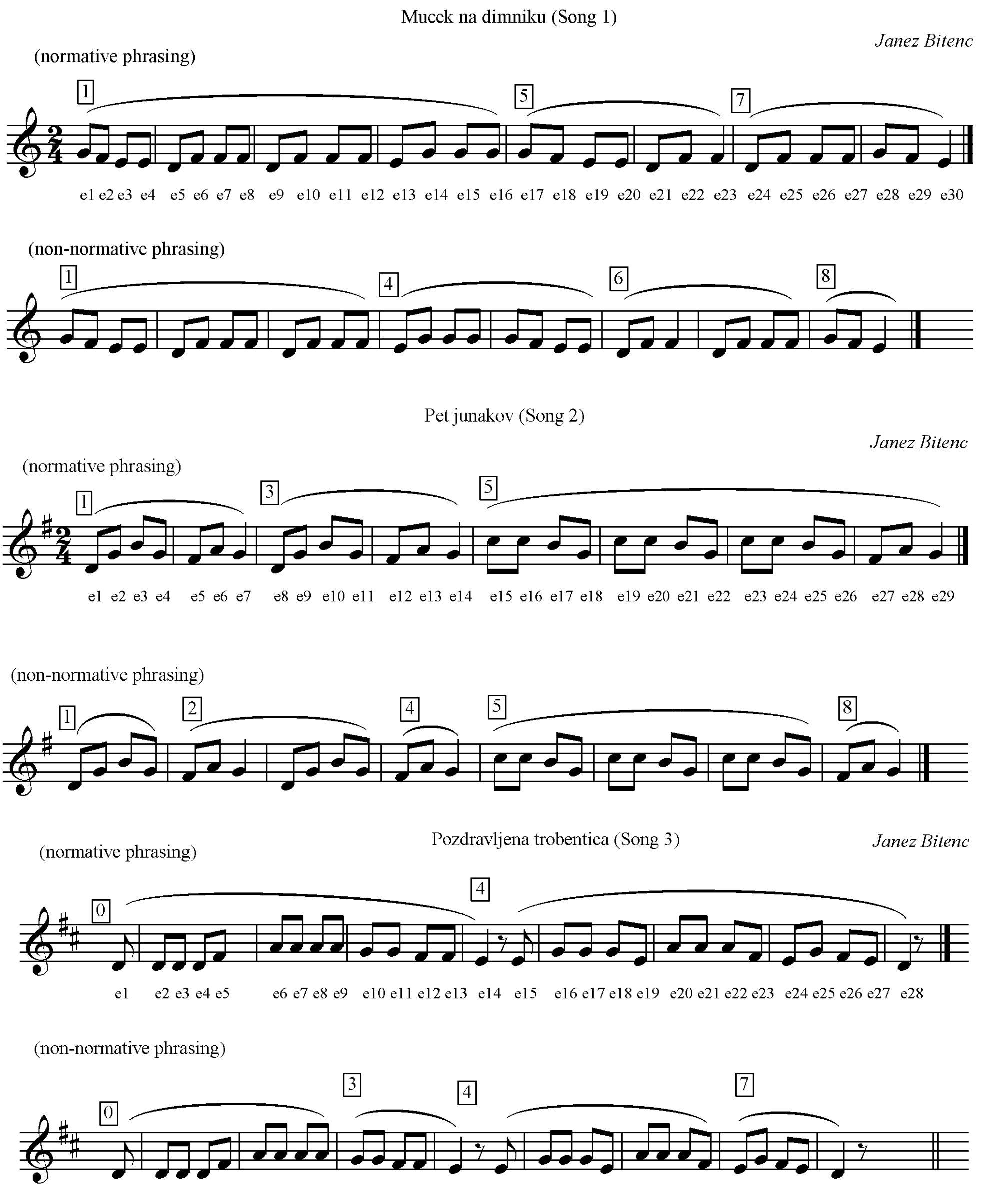


Figure 4.1: Songs used in the The Game and The Breathing experiment with phrases (normative and non-normative) provided by musical experts. Each note in each song is labeled as an event eN.

In all three songs, the usual lyrics were removed and substituted with meaningless words, as described above. The participants were informed that the Game Experiment will consist of two conditions and that in each condition the main task will be the rebuilding of the melodies. In the first condition, the normative segmentation of the three melodies was used. The participants were asked to listen to the entire song, and then to rebuild this same song with normative melody segments. The participants were told to self-pace through all the tasks, pressing the "confirm/return" key when they had finished, which put the selected melody segments together. Feedback on the accomplished task was given in the form of two different sounds, depending on a correct or incorrect result. After the first condition, there was a pause of 30 minutes, in order to avoid priming effects of songs heard in the first condition on the second condition.

In the second condition, the non-normative segmentation of the same three melodies was used. As in the first condition, the participants were asked to listen to the entire song before rebuilding it with melody segments, matching non-normative segments in the melody. No time limit was given; as in the first condition, the participants could self-pace through the tasks. By pressing the "confirm/return" key after finishing each of the tasks, the melody segments were put together, again indicating a correct or incorrect result with a sound. Figure 4.2 illustrates the presentation of Song 1, a part of the task used in the Game Experiment. The three cats (on the left side) in Condition 1 (Level 1: normative phrasing) represent the three segments, and the four cats in Condition 2 (Level 2) represent the four segments of a non-normative phrasing. By clicking on a cat, a particular segment of the Song 1 is played. After putting all the cats in their boxes, the participant may listen from the beginning to the end of all the cats and to check if the melody is rebuilt correctly. If the participant decides that one of the cats is not correct, that cat may be returned to its place, and the order of cats can be changed. The "play" button in each condition (Condition 1 and Condition 2) gives the participant the opportunity to listen to the entire song before starting with the task, the button "potrdi" (confirm) stores the rebuilt in a .txt file.

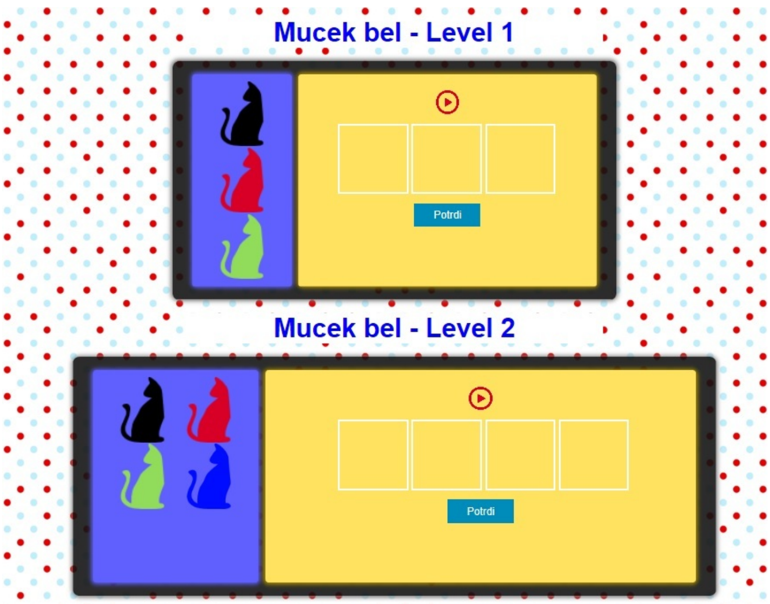


Figure 4.2: An illustration of Song 1 from the game used in the The Game Experiment.

Results

The aim of using normative and non-normative melody segments, respectively, in the two conditions of the Game Experiment was to test whether the subject's memory of a song relies only on remembering the notes of the whole song or on remembering distinct phrases. The results suggest that the reconstruction task of songs 1 and 2, but not song 3, was more difficult for all participants in the second condition than in the first condition (see Figure 4.3).

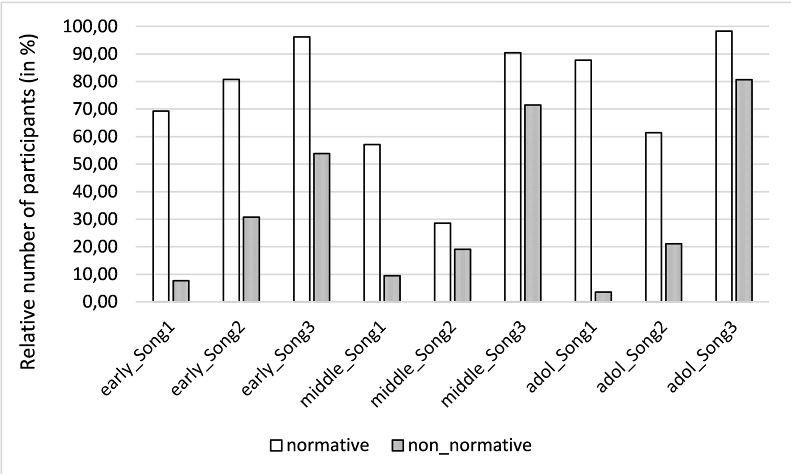


Figure 4.3: Results of the reconstruction of songs in the first and second condition in the The Game Experiment 2018.

Detailed results of the Game Experiment are presented in Table 4.1. The proportion of participants who successfully recreated the songs is higher in all the groups in the first condition with normative phrasing (proportion p1) than in the second condition with non-normative phrasing (proportion p2). The significance of differences between p1 and p2 for each group and song was tested using McNemar's test (McNemar, 1944), since comparison between proportions of two dependent samples is required. p-values of this test are reported in the last column of Table 4.1.

Table 4.1: Proportions of successfully recreated songs in the The Game Experiment (2018).

Group	**Song**	**p1**	**p2**	***p*-value**
early	Song1	.69	.08	$p < .01$
early	Song2	.81	.31	$p < .01$
early	Song3	.96	.54	$p < .01$
middle	Song1	.57	.10	$p < .01$
middle	Song2	.29	.19	$p = .68$
middle	Song3	.90	.71	$p = .22$
adolescent	Song1	.88	.04	$p < .01$
adolescent	Song2	.61	.21	$p < .01$
adolescent	Song3	.98	.80	$p < .01$

Discussion

All the participants, from all the groups, regardless of musical knowledge or age, had problems in rebuilding melodies with non-normative melody segments in the second condition. The results from the second condition in all the songs indicate that it is harder to reconstruct a song with phrases that do not match a music-theoretical normative representation of phrases and phrase boundaries. Similar results were found in Demorest's experiment.

A plausible explanation for the better results which were obtained in Song 3, even in the second condition by using non-normative phrasing, may be that this song consists of less repeated content, and more distinguishable motifs compared to Song 1 and Song 2. Repetition is ubiquitous in music, and repeated motifs, or larger musical structures are better memorized (Hutchins & Palmer, 2008; Margulis, 2014). Furthermore, studies dealing with priming effects, and priming repetition (in general), show that repeated stimuli facilitate the processing of a structure (e.g., Cholin et al., 2004; Wheeldon & Smith, 2003). However, some studies also suggest that stimuli with low-frequency, compared to stimuli with high-frequency, receive more benefits from prior exposure (e.g., Forster & Davis, 1984; Nevers & Versace, 2003).

As can be seen from Figure 4.1, the repeated motifs were split at unusual and unexpected places in non-normative phrasings, turning them to less distinguishable stimuli (motifs) in Song 1 and 2, compared with stimuli (motifs) in Song 3, even after recreating them by using a non-normative phrasing. Several studies outline the importance of phrase boundaries in the proper segmentation of music (Deutsch, 2013; Handel, 1989; Kragness & Trainor, 2017; Lerdahl & Jackendoff, 1983b). By putting the phrase boundaries in the second condition on unusual places where no phrase boundaries were expected, problems occurred in reconstructing songs, suggesting how the participants perceive the entire song.

It seems that, while learning all three songs, the participants remembered phrases, and therefore had problems in recalling the non-normative segments while trying to rebuild the melodies in the second condition. These results are in accordance with the findings from Silva et al. (2014) and M. T. Pearce, Müllensiefen, et al. (2010a) and M. T. Pearce, Müllensiefen, et al. (2010b), outlining the importance of memory from a previous phrase in the transition from one to another phrase. These results support the hypothesis that children use phrases as a grouping strategy when memorizing melodies.

The participants were observed while solving the tasks from the second condition. They repeatedly moved the melody segments, and repeatedly listened to the melody produced, while also appearing frustrated. This may indicate that they were searching for some familiar points (boundaries) which could help them in reconstructing the song. According to the results, participants with and without musical knowledge in the three different groups perceive musical phrases and phrase boundaries in a similar way, regardless of age.

4.2.2 The Game Experiment (2020)

The same Game Experiment (2018) was conducted in 2020, from May to June, and tested the same hypothesis: that phrases are used as a grouping strategy during the memorization of melodies in children. Using the same game environment and the same three songs, two additional issues were tested, (i) how the way of learning melodies by listening only to instrumental versions of songs affects the results of the segmentation of melodies, and (ii) the effect of order, i.e., whether the randomization of songs to be recreated in tasks with normative and non-normative phrasing impacts the results of recreating the songs in both conditions. A total of 35 children and adolescents participated in the replicated Game Experiment (2020). The schools and the participants were selected, as in Experiment (2018), based on their willingness to participate in the experiment and the number of participants with formal musical experience, which was verified prior to the participation of these institutions in the study.

As in the The Game Experiment, conducted in 2018, the participants were divided into three groups, depending on their age: early childhood from 5-6 years (group early), middle childhood, from 8-12 years (group middle), and adolescent, from 14-16 years (group adolescent), further dividing each group into two subgroups according to their musical knowledge (experience): subgroup “with” and subgroup “without” formal musical knowledge.

Participants

From 35 participants, 9 participants from the group early were involved (5 with, 4 without musical knowledge, avg. age = 5.22), 13 participants from the group middle (6 with, 7 without musical knowledge, avg. age = 10.70), and 13 participants from the group adolescent (7 with, 6 without musical knowledge, avg. age = 15.38). Binary indicator (1 = solved; 0 = not solved) was used to measure success in recreating of songs in both conditions, and the same three songs with the same normative and non-normative phrasings were used.

Methodology

Procedure

The teacher, who previously sang the songs, was replaced with instrumental versions of all the three songs, created by the program Sibelius, using a piano sound. All the participants listened to the songs during three successive days in their respective institutions. Each song was played on a computer with loudspeakers and repeated three times, followed by a pause of 10 seconds before playing another song. The Game Experiment (2020) was conducted on the fourth day, after the three listening days. The order of the songs used in both conditions was randomized, and a pause of 30 min was provided after finishing the first condition of a song before starting the second condition in another song.

Binary indicator (1 = solved; 0 = not solved) was used to measure success in recreating of songs in both conditions, with the same three songs, and with the same normative and non-normative phrasings.

Comparison Between the two Game Experiments

Participants had problems in recreating the songs in the second condition, as in The Game Experiment 2018 (see Figure 4.4). Fisher exact test (Rice, 2005) was used to test the difference in the proportions of successful recreated songs in The Game Experiment 2018 (p1_Exp), and The Game Experiment 2020 (p2_Exp), for each song, each group, and each condition (altogether 18 tests). No statistically significant differences were found (see Table 4.2).

This confirms that neither the way of teaching (listening to instrumental versions), nor the order of songs in both conditions has significantly influenced the successfulness of recreating the songs.

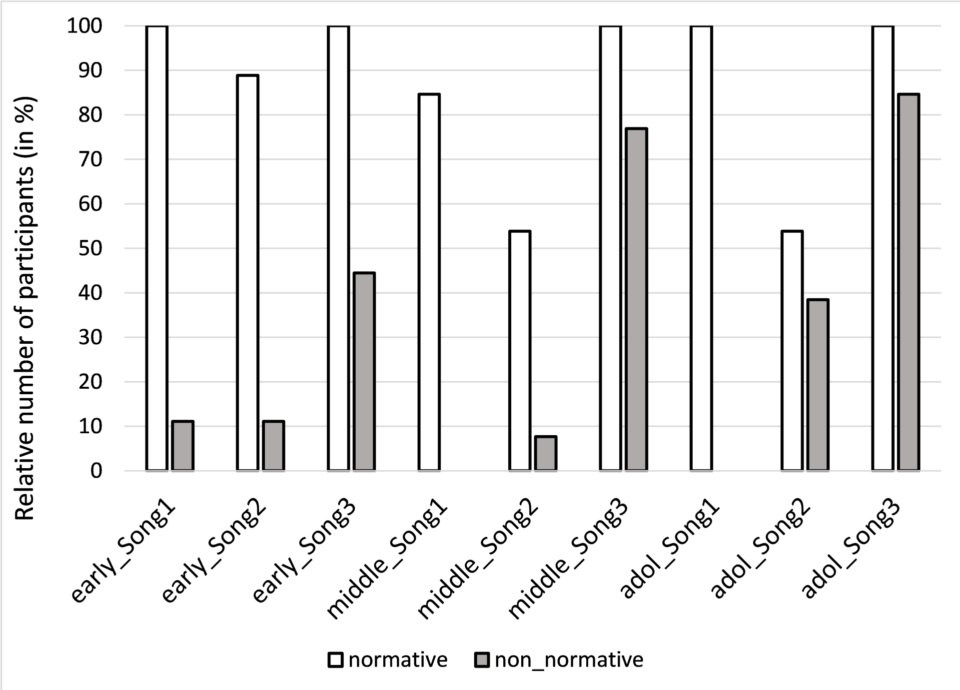


Figure 4.4: Results of the reconstruction of songs in the first and second condition in The Game Experiment 2020.

Table 4.2: Difference in the proportions of successful recreated songs with normative (norm) and non-normative (non-norm) phrasing in The Game Experiment 2018 and The Game Experiment (2020).

Group	**Song**	**Condition**	**p1_Exp**	**p2_Exp**	***p*-value**
early	Song1	norm	.69	.99	.08
early	Song1	non-norm	.08	.11	.99
early	Song2	norm	.81	.89	.99
early	Song2	non-norm	.31	.11	.39
early	Song3	norm	.96	.99	.99
early	Song3	non-norm	.54	.44	.71
middle	Song1	norm	.57	.85	.14
middle	Song1	non-norm	.09	.00	.51
middle	Song2	norm	.29	.54	.17
middle	Song2	non-norm	.19	.08	.63
middle	Song3	norm	.90	.99	.51
middle	Song3	non-norm	.71	.77	.99
adolescent	Song1	norm	.88	.99	.33
adolescent	Song1	non-norm	.04	.00	.99
adolescent	Song2	norm	.61	.54	.76
adolescent	Song2	non-norm	.21	.38	.28
adolescent	Song3	norm	.98	.99	.99
adolescent	Song3	non-norm	.80	.85	.99

4.3 Segmentation in Children: The Breathing Experiment

Two Breathing Experiments were undertaken in 2018 and 2020 to determine whether and how breathing while singing correlates with the perception of phrases in children with and without musical knowledge aged 5-6, 8-12, and 14-16 years.

The decision to conduct breathing studies was based on the fact that everything we do occurs while we are breathing and that the music we listen to (or perform) impacts our respiration. Numerous studies investigate the relation between breathing and music listening/performing/segmentation, calculating factors such as respiratory inhalation/exhalation time ratio, respiratory rate and depth, respiratory behavior and emotional responses, and changes in physiological responses as music transitions from one segment to the next (Bernardi et al., 2006; Cabredo et al., 2011; Iwanaga & Moroki, 1999; C. Krumhansl, 1997).

None of these studies investigate the relationship between breathing (respiration) and segmentation in children of various ages, and how musical structure affects breathing. The expectations were:

- The participants are using statistical information to form a model of musical syntax, and that this model should become more strongly predictive with increasing age.,
- Intra-group agreement about places of breathing increases with age.
- Agreement with normative segmentations increase with age (taking into account the possibility of ambiguity in the stimulus: M. T. Pearce, Müllensiefen, et al., 2010b).

4.3.1 The Breathing Experiment (2018)

Hypotheses

The Breathing Experiment (2018) tests two hypotheses: (i) that agreement on segmentation within age groups increases with age, and (ii) that the similarity between normative and participant's segmentation increases with age.

Participants

The same 106 children and adolescents who participated in The Game Experiment (2018) also participated in The Breathing Experiment (2018), which was divided into the same three groups as The Game Experiment (2018):

early (5-6 years), middle (8-12 years), and adolescent (14-16 years). These groups were further divided into two subgroups according to their musical experience: subgroup "with", and "without" formal musical experience. Of 104 participants, 26 participants were in the early group (11 with, 15 without musical experience, mean age = 5.62), 23 participants were in the middle group (6 with, 17 without musical experience, mean age = 9.90), and 57 participants were in the adolescent group (28 with, 29 without musical experience, mean age = 15.16). However, the number of participants' responses varied from song to song, because incomplete recorded songs, for example, where the participant could not remember the song, or sang only a part of it, were discarded.

Stimuli

The same three children's songs were used as in The Game Experiment (2018) and (2020).

Methodology

Participants learned the three songs throughout three choir sessions over the course of two weeks. Each session lasted 30 minutes. A brief rehearsal was done on the day prior to the experiment. During these sessions, the teachers were given very detailed instructions not to pay any particular attention to the boundaries in the songs, such as refraining from breathing during the singing of songs or using gestures and facial expressions that could suggest phrasing throughout the learning of the songs.

The songs were always learned in the same order, beginning with Song 1, then Song 2, and finally Song 3. The teachers sang songs with the children in an unstructured manner, allowing the children to freely breathe while singing. The same set order of songs (from Song 1 to Song 3) was adhered throughout the recording process. Each participant was separately recorded in a quiet room. Before recording a song, a trained experimenter sang the first 2–3 bars of the song to the participant to help him or her start singing. These bars were not used in any of the analyses. It was made clear to each participant that they were free to breathe whenever they felt necessary while the songs were being recorded.

Equipment

Participants' singing was recorded using a Handy Recorder H4n with a built-in stereo microphone. The researchers used Audacity[6] to visually analyze all of the captured audio data.[7]

Procedure

The locations of breathing in the visualized waveforms were compared with the participant's clear inhalation and song scores to determine where breathing occurred. Thus, both the visual depiction of Audacity and the sound of the audio recordings were used, for every single breath for every song and every child.

To analyze the breathing in detail, each event from each song and for each participant was analyzed; 30 events from Song 1, 29 events from Song 2, and 28 events from Song 3. (see Figure 4.1). Each participant's breathing locations were marked with a binary indicator (1 = breath, 0 = no breath) for each song. A comparison was made between the participants' breathing and the breathing places (normative phrasing) provided by musical experts for each song. Final breathing evaluations for each participant were made after consulting and comparing the experimenter's findings with those of an additional musical expert. This expert was chosen (as were all five musical experts) based on his formal musical training (of the highest level) and extensive choir leadership.

To measure the similarity between two different segmentations, cosine similarity was calculated (Han & Pei, 2012, p. 78) between two binary vectors whose elements corresponded to possible breathing points in each song. When the cosine value approaches 1, a stronger match between two vectors is obtained, as well as smaller angles (Han & Pei, 2012, p. 78). As the measured attributes are binary-valued, given two vectors $\vec{x}$ and $\vec{y}$, cosine similarity is interpreted in terms of shared occurrences:

$$sim(\vec{x}, \vec{y}) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}} \tag{4.1}$$

[6]https://www.audacityteam.org

[7]Visual analysis refers to the process of looking for flat lines inside a waveform, which represent the threshold (silence) and consequently indicate "not breathing."

Breathing on the very first note of each song was discarded for the same reasons as de Nooijer et al., 2008: (i) there is no information on the context before the first event in a melody, and (ii) adding breathing on the very first event (note) in a song yields to statistical anomalies. In making the analysis, only the sequential aspects of the participants' performance were considered, not the elapsed time: the focus was only on where in the sequence a participant took a breath, and not its exact timing. After consulting with an additional musical expert, a decision was reached regarding each participant's breathing in each song.

Finally, the mean similarity of each (sub)group of children to normative phrasing was estimated by averaging the computed cosine similarities across all of the children in the observed (sub)group.

Results

Song 1

93 participants out of 106 produced usable recordings of Song 1: 19 participants from the early group (13 participants with and 6 without musical knowledge, mean age = 5.68), 23 participants from the middle group (6 participants with and 17 without musical knowledge, mean age = 9.74) and 51 participants from the adolescent group (26 participants with and 25 without musical knowledge, mean age = 15.16).

Fleiss's κ (Kappa: Fleiss, 1971) was used to measure inter-participant agreement, showing moderate agreement in all groups ($\kappa = 0.59$), substantial inter-participant agreement in groups early ($\kappa = 0.64$), middle ($\kappa = 0.70$), and adolescent ($\kappa = 0.74$).

Figure 4.5 demonstrates that the group early breathed more frequently than the other two groups. In the group early, slightly different results were also discovered between children with and without musical experience: breathing is more frequent in children with musical experience. On the other hand, breathing on phrase boundaries in participants from group early with no musical experience is more comparable to breathing on phrase boundaries in participants from group middle with no musical experience.

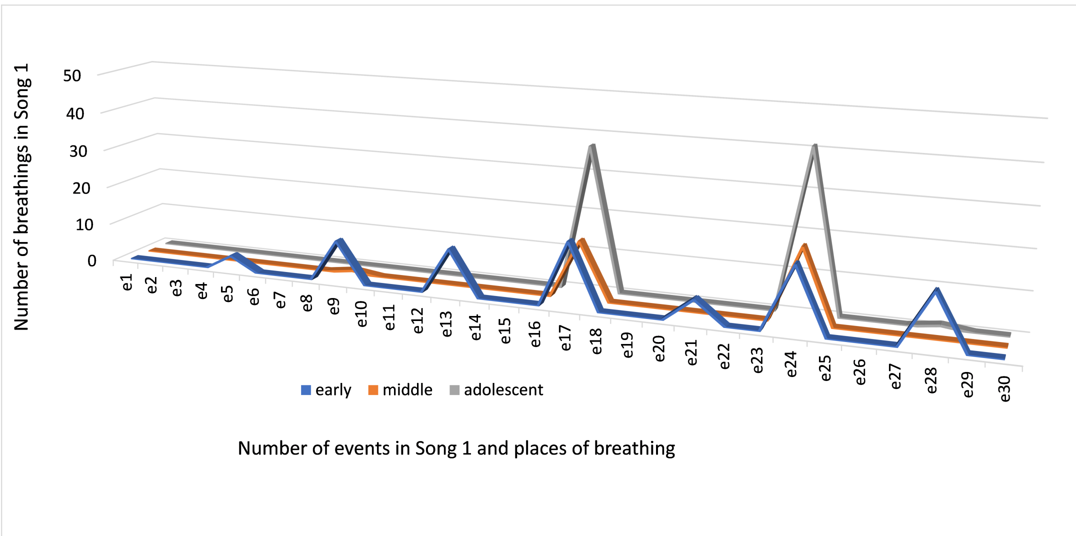


Figure 4.5: Breathing patterns in all groups in Song 1 (The Breathing Experiment 2018).

In general, it does not appear that the breathing patterns of all groups can be classified merely based on shortness of breath, which would be expected to follow a Poisson distribution. Breathing does not typically occur in the middle of a bar, but rather after a four-bar phrase at the beginning of the song (with the exception of early participants with musical knowledge), followed by two new short phrases in the fifth and seventh bars that correspond to phrases from the normative version of this song.

Table 4.3 shows the mean cosine similarity between the normative phrasing (as indicated in Figure 4.1) and breathing by all groups (depicted in Figure 4.5). There is a strong correspondence between the normative phrasing and the breathing of all groups, including the early group, despite the early group's frequent breathing.

Table 4.3: Mean cosine similarity between the normative segmentation in Song 1 and breathing of all groups.

Groups	**All**	**With**	**Without**
early	.62	.55	.80
middle	.77	.74	.83
adolescent	.78	.76	.80

Song 2

103 participants out of 106 produced usable recordings of Song 2: 25 participants from the early group (11 participants with and 14 without musical knowledge, avg.age = 5.68), 23 participants from the middle group (6 participants with and 17 without musical knowledge, avg. age = 9.74) and 54 participants from the adolescent group (26 participants with and 29 without musical knowledge, avg. age = 15.15).

Substantial agreement was found across all participants ($\kappa = 0.77$), and within early ($\kappa = 0.78$) and middle group ($\kappa = 0.76$). Almost perfect agreement was found in adolescents $\kappa =(0.81)$. Breathing patterns in all three groups are presented in Figure 4.6.

Participants with and without musical knowledge from the group early breathe frequently (every two bars or even after each bar) in this song, despite its normative phrasing consisting of two two-bar phrases in the first four bars and a four-bar phrase in the second part. There is a tendency to aggregate the last four bars as a four-bar phrase (normative phrasing), and this tendency increases from the youngest group to the oldest group, regardless of musical knowledge.

The relationship between the normative breathing in each group is shown in Figure 4.6 and Table 4.4. There is a high degree of consistency between the normative segmentation and breathing and all groups, regardless of age or musical knowledge.

Table 4.4: Mean cosine similarity between the normative phrasing in Song 2 and breathing in all groups.

Groups	**All**	**With**	**Without**
early	.88	.82	.93
middle	.85	.81	.87
adolescent	.89	.91	.86

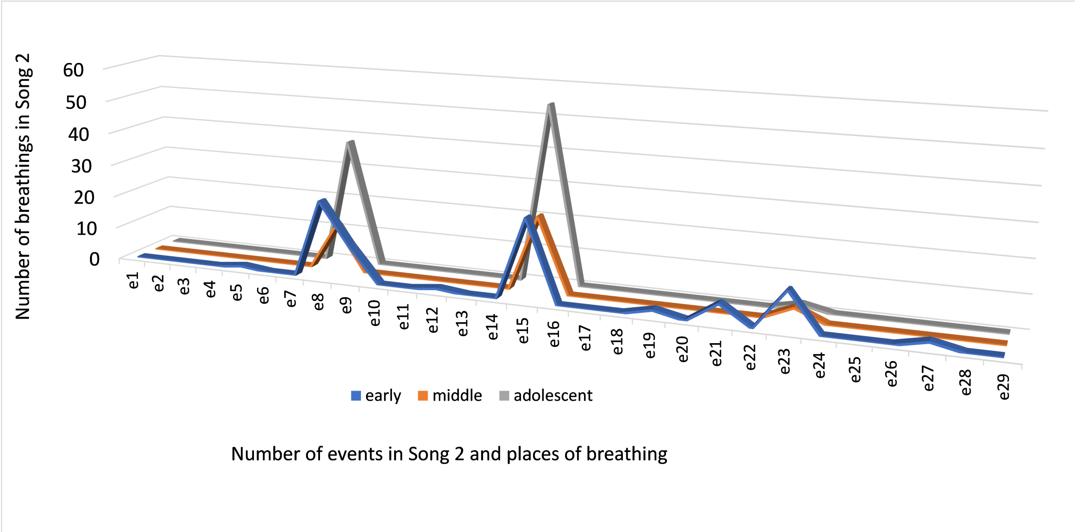


Figure 4.6: Breathing patterns in all groups in Song 2 (The Breathing Experiment 2018).

Song 3

77 participants out of 106 were involved: 9 from the group early (6 participants with and 3 without musical knowledge, avg. age = 5.56), 23 from the group middle (6 participants with and 17 without musical knowledge, avg. age = 9.74) and 45 from the group adolescent (22 participants with and 23 without musical knowledge, avg. age = 15.11). The fewest usable recordings were obtained in Song 3, possibly because the usual lyrics were not used, or because of the absence of special words with clear accents (weak-strong, strong-weak), which could help the participants recall the anacrusis at the beginning of the song.

All participants were more successful at reproducing the contour of this song than the other two, but less successful at reproducing the exact pitches and intervals. This is consistent with previous findings that contours are easier to remember than intervals under certain conditions (Dowling, 1978; Massaro et al., 1980; Müllensiefen, 2006).

Bennet (1990) demonstrates that anacruses are more easily detected when appropriate accented words are employed in conjunction with the tune. The anacrusis in this song is on a weak accent, implying that the phrase's beginning (grouping structure) is not aligned with the metrical onset (metrical structure). While it is widely accepted that grouping structure is logically independent from metric structure (Brochard et al., 2000; Liegeoise-Chauvel et al., 1998; Peretz, 1990), in practice, the two structures are frequently inextricably linked (Kragness & Trainor, 2017). This is not the case for this particular song. Additionally, the song contains multiple repeated notes with the same pitch and beat following the anacrusis that were not well learned.

Some studies indicate that smaller intervals are prevalent in children's songs (Narmour, 1990; Schellenberg et al., 2002), and that listeners have a natural tendency to anticipate smaller intervals when identifying a melody or song. Despite this, participants from all three groups found it more challenging to learn this song, which has a narrow pitch range and small intervals, than the other two.

The inter-participant agreement calculated using Fleiss's κ shows substantial agreement within all the groups ($\kappa = 0.72$). The lowest, moderate inter-participant agreement was found in group early ($\kappa = 0.44$ with 9 participants), with almost perfect agreement in middle ($\kappa = 0.96$) and in adolescent ($\kappa = 0.86$). The perception of phrases is similar in two latter groups: two four-bar phrases, the first starting with an anacrusis at the beginning

Table 4.5: Cosine similarity between the normative segmentation in Song 3 and breathing in all groups.

Groups	All	With	Without
early	.61	.66	.50
middle	.99	1	.92
adolescent	.88	.90	.86

and the second one starting in bar 5 (e16). The breathing patterns in Song 3 in all groups are shown in Figure 4.7.

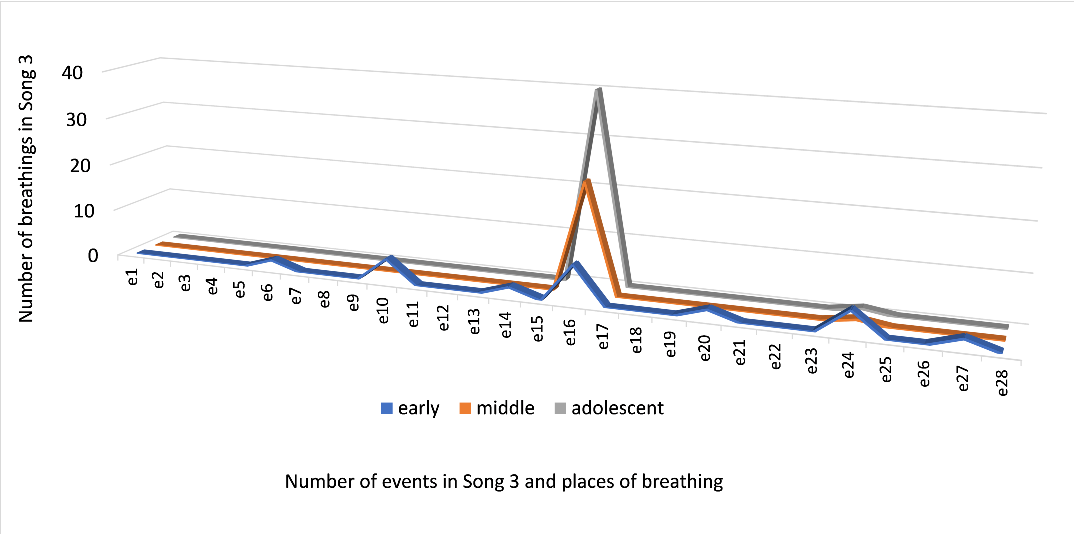


Figure 4.7: Breathing patterns in all groups in Song 3 (The Breathing Experiment 2018).

The matching between the expected breathing in bar 5 (e16), and all the groups is shown in Table 4.5. Very strong correspondence is found between the expected breathing and group middle and adolescent, moderate in the youngest group.

Discussion

Findings of The Breathing Experiment (2018) show the least inter-participant agreement in the youngest group early, very frequent breathing patterns in this group, and no statistically significant differences between participants with and without musical knowledge.

Examination of the breathing patterns in each of the three songs showed that the youngest group (group early) had the least inter-participant agreement regarding the breathing patterns. This could be owing to the unfamiliar melodies, or the absence of lyrics, harmonic accompaniment, and dynamics. Another possibility is that the three songs generated different segmentation strategies in this group of children, and that higher inter-participant agreement may be achieved by subdividing this group according to the segmentation strategies used in the three songs (M. T. Pearce, Müllensiefen, et al., 2010b).

The results of this experiment indicate (in the majority of cases) a very frequent breathing pattern, which was expected given the age group's respiratory/physiological requirements. However, it appears as though this group's breathing is unaffected by shortness of breath in any of the songs. If it were related to lung capacity, breathing could occur in the middle of bars as well as (contrary to our observations) 'rhythmically': at the beginning of each bar or every second bar, as in Song 1 and Song 2.

The following are some probable causes for frequent breathing and low inter-participant agreement on breathing (perception of phrases and phrase borders). Frequent breathing may be influenced by formal musical activities (e.g., choir singing) on the participant's understanding of the musical structure, in the sense that instructions on how to sing, where to breathe, and how frequently to breathe may be poorly understood. Another explanation could be the lack of a confident level of knowledge of basics in singing, in choir instructors themselves (K. H. Phillips & Vispoel, 1990).

According to an analysis of the implicit harmony in Song 2, frequent breathing in group early may be explained by implied chords that function as phrase boundaries, in the sense that whenever a stable chord (e.g., chord with a tonic function) is provided following chords with a high tension (e.g., dominant or subdominant), a point of relaxation/end of a phrase is achieved. This is consistent with the findings of the Lerdahl and Jackendoff (1983b)' prolongational hypothesis, as well as with the findings of other studies indicating that phrase boundaries in a musical phrase can be represented by chords that act as cues for a phrase boundary depending on their tension/relaxation (Bigand, 1993; Kragness & Trainor, 2017).

Figure 4.8 shows the stable chords in Song 2, the tonics ('T'), subdominant ('S') and dominant ('D'). These chords may be perceived as cues for phrase boundaries, less so in the first four bars of the song (where 'T' predominates) and more so in the last four bars, with a frequent changing of chords (S-T, in the last bar D-T), in which each chord marked as 'S' ('D') is deceiving and signals a (false) new "phrase." As a result, more frequent breathing (due to the 'tension/relaxation' of a chord) is observed at the beginning of each bar, particularly among participants with musical experience, and in all groups.

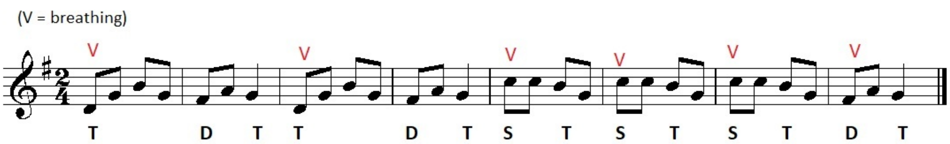


Figure 4.8: Breathing patterns in all groups in Song 3 (The Breathing Experiment 2018).

Frequent breathing in some participants, especially in participants with musical knowledge, at every bar in Song 2, suggests a high sensitivity to contour which may be used as a cue in grouping the music material. In Western musical tradition, an 'inverted U' pattern is very common for a melodic contour and several such inverted U patterns can be found in this song, in the first, second, third, fourth, and eight bar, and partially in the fifth, sixth and seventh bar (see Figure 4.9).

The frequent breathing of some participants, particularly those with musical knowledge, at each bar in Song 2 indicates a high sensitivity to contour, which might be utilized to group the musical material. An 'inverted U' pattern is a fairly prevalent melodic contour in Western musical tradition, and numerous of these inverted 'U patterns' appear in this song, in the first, second, third, fourth, and eighth bar, as well as partially in the fifth, sixth, and seventh bar (see Figure 4.9).

The song's inverted 'U patterns' (actually, each bar) are aligned with the frequent breathing of some participants from the youngest group at the beginning of each bar, using the contour as a cue for phrase boundaries (the downward slope at the end of each inverted 'U pattern'). Numerous studies indicate that very young children approximate

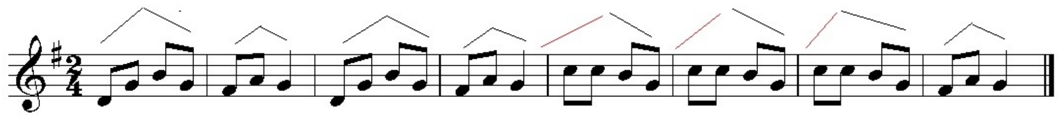


Figure 4.9: Breathing patterns in all groups in Song 3 (The Breathing Experiment 2018).

the contour of a melody at first and then gradually focus on pitches and specific intervals (Kragness & Trainor, 2017; C. L. Krumhansl, 1990; Thorpe et al., 1988) and that listeners are highly sensitive to contour in their recognition memory for melodies (Deutsch, 2013; Dowling, 1999). Due to the fact that this frequent breathing is present in only a few participants in the youngest group but not in the older groups, the results indicate that the younger participants have a high sensitivity to the contour of a recalled song. Breathing patterns in the oldest two groups (middle and adolescent) demonstrate a high degree of similarity across participants with and without musical training.

Although several studies have confirmed the effect of formal musical training on understanding musical structure (Fujioka et al., 2004; Hannon & Trainor, 2007), the obtained result contradicts the neurophysiological study of Bresson (1997), which demonstrates the importance of formal musical knowledge and shows a significant difference between those with and without musical knowledge when solving a specific and explicitly defined musical task (viz., finding the out-of-key changes in an unfamiliar melody).

None of the participants in The Breathing Experiment were informed of the experiment's purpose or where to breathe in a particular song; they were simply instructed to "sing the songs," which meant that all participants, with or without musical knowledge, were required to concentrate exclusively on singing the three learned songs without any specified or demanding tasks requiring theoretical musical knowledge. The high inter-participant agreement between participants in these two groups on their perception of phrase/phrase boundaries suggests that participants with and without musical experience employ similar grouping strategies, which is consistent with Deliège's 1987 findings.

4.3.2 The Breathing Experiment (2020)

The Breathing Experiment was conducted again in 2020. The purpose of this experiment was to obtain more information about the breathing patterns of children of various ages, particularly the youngest population, and to identify useful clues that could be used in future experimental studies.

Hypothesis

The Breathing Experiment conducted in 2020 tested the same hypotheses as the Breathing Experiment 2018: that agreement on segmentation within age groups will increase with age, and that the similarity between the normative segmentation and the participant's segmentation will increase with age.

Participants

A total of 34 children and adolescents participated in the Breathing Experiment 2020. As in The Breathing Experiment 2018, the participants were divided into three age groups, each of which was further divided according to with/without experience: early (5–6 years;

n=8, with experience n=4, mean age=5.13), middle (8–12 years; n=13, with experience n=6, mean age=10.62), and adolescent (14–16 years; n=13, with experience n=7, mean age = 15.39).

Equipment

As in The Breathing Experiment 2018, a portable Handy Recorder H4n with built-in stereo microphone was used to record the singing of participants. Breathing in all the recorded audio data was visually analyzed using Audacity.

Methodology

All three songs were recorded in their entirety as sung by all participants. In comparison to the Breathing Experiment 2018, participants were trained utilizing instrumental versions of the songs generated using Sibelius[8] and an electronic piano sound. All participants listened to the music in their respective institutions for three consecutive days (each day for 30 minutes). Each song was played three times on a computer with loudspeakers, followed by a ten-second delay before proceeding to the next song. In contrast to the Breathing Experiment 2018, the order of songs was randomly assigned during the learning sessions. A final little rehearsal occurred the day before the recording. The order of songs was randomly assigned to each participant throughout the recording process.

To denote breathing, a binary indicator was employed (1 = breathing; 0 = not breathing). Breathing locations were checked (as in the Breathing Experiment 2018) using Audacity, marked also within the scores, and checked by an additional music expert. For the same reason as in the Breathing Experiment 2018, breathing on the very first note of each song was discarded.

Results

Song 1

Of 34 participants, 8 early participants (4 with experience, mean age = 5.13), 13 middle participants (6 with experience, mean age = 10.62), and 13 adolescent participants (7 with experience, mean age = 15.39) made usable recordings. The inter-participant agreement calculated using Fleiss's κ shows substantial agreement between all the groups ($\kappa = 0.63$), and in groups early ($\kappa = 0.74$) and adolescent ($\kappa = 0.78$). Moderate agreement was found in group middle ($\kappa = .49$).

Breathing patterns in all groups are shown in Figure 4.10. The cosine similarity between the normative segmentation and participants is shown in detail in Table 4.6.

Table 4.6: Mean cosine similarity between the normative segmentation in Song 1 and breathing in all groups.

Groups	All	With	Without
early	.83	.86	.68
middle	.58	.59	.57
adolescent	.82	.71	.95

[8] https://www.avid.com/sibelius

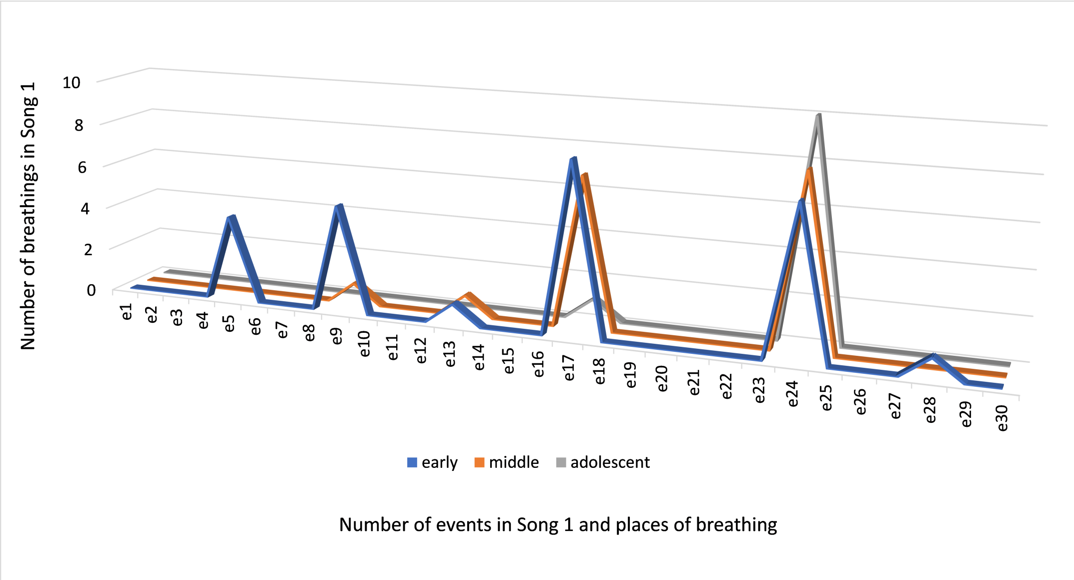


Figure 4.10: Breathing patterns in all groups in Song 1 (The Breathing Experiment 2020).

Song 2

Song 2 was also recorded by the same participants as Song 1. The inter-participant agreement calculated using Fleiss's κ shows almost perfect agreement between all the groups ($\kappa = 0.86$), and in groups early ($\kappa = .84$) and adolescent ($\kappa = 0.83$), substantial agreement in group middle ($\kappa = 0.75$).

Breathing patterns in all groups are shown in Figure 4.11. The cosine similarity between the normative segmentation and breathing patterns in all groups in Song 2 are shown in detail in Table 4.7.

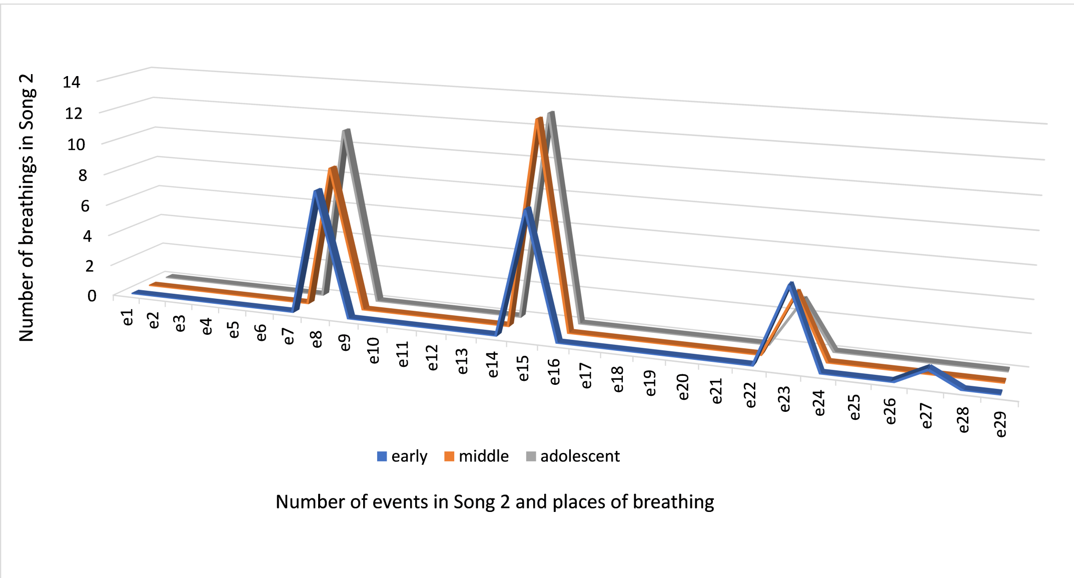


Figure 4.11: Breathing patterns in all groups in Song 2 (The Breathing Experiment 2020).

Table 4.7: Mean cosine similarity between the normative segmentation in Song 2 and breathing in all groups.

Groups	All	With	Without
early	.87	.85	.90
middle	.85	.86	.87
adolescent	.91	.92	.91

Song 3

The same participants from the middle and adolescent group were involved in the recording of Song 3. From the early group, only 6 out from 8 participants produced usable recordings (mean age = 5.17). The inter-participant agreement calculated using Fleiss's κ shows almost perfect agreement between all the groups ($\kappa = 0.87$), substantial intra-group agreement in group early ($\kappa = 0.70$), and almost perfect agreement in groups middle ($\kappa = 1$) and adolescent ($\kappa = 0.93$).

Breathing patterns in all groups are shown in Figure 4.12. Cosine similarity between the normative segmentation and breathing in all groups is shown in Table 4.8.

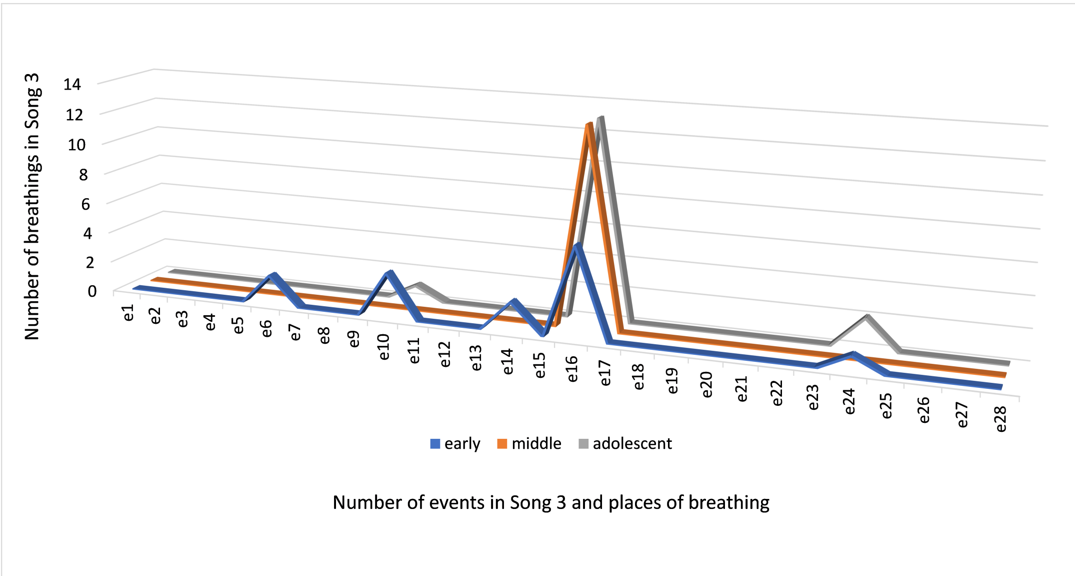


Figure 4.12: Breathing patterns in all groups in Song 3 (The Breathing Experiment 2020).

Table 4.8: Mean cosine similarity between the normative segmentation in Song 3 and breathing in all groups.

Groups	All	With	Without
early	.83	.80	.86
middle	1	1	1
adolescent	.98	1	.95

Discussion

The Breathing Experiment (2020) was conducted to understand more about the breathing patterns of children of varying ages, especially the youngest population, and to identify relevant clues for future experimental studies. Compared to The Breathing Experiment (2018), a smaller population was included in this study. When children and adolescents' breathing was compared to normative phrasing, comparable segmentation results as in The Breathing Experiment (2018) were seen, regardless of whether instrumental versions of songs were used, or the songs were randomly ordered during training. Except for group middle, the results indicate that agreement within age groups and matching between normative and participant segmentations increase with age.

Comparison Between the two Breathing Experiments

The Breathing Experiment (2020) yielded comparable results to the Breathing Experiment(2018): segmentation similar to normative phrasing was observed in older groups, and Song 2 had the best segmentation regardless of musical knowledge. Using instrumental versions rather than vocal versions to learn songs and recording songs in a random order did not provide statistically significant differences.

The Breathing Experiment (2020) revealed segmentation closer to normative segmentation in all three songs and all three groups, including the youngest. A plausible explanation could be that there were more musically experienced participants in The Breathing Experiment (2020) than in the Breathing Experiment (2018), which would also account for the younger group's better results in intra-group agreement, although intra-group agreement results in the Breathing Experiment 2020 also increase with age. As was the case in the Breathing Experiment (2018), the best segmentation was found in Song 2. The fact that all participants were successful in segmenting this song points to the musical structure's (clearly) unambiguous segmentation cues.

The fact that older groups (middle and adolescent) and normative segmentation matched better in Song 1 and Song 2 in both breathing experiments suggests that older participants are using, in addition to low-order musical features (e.g., pitch, rhythm), high-order musical features (e.g., interval, implied harmony), or a combination of low- and high-order musical features to locate phrase cues. This finding is consistent with numerous studies (Cohrdes et al., 2016; Cohrdes et al., 2014; Lamont, 1998; Trehub & Hannon, 2006) demonstrating that children become more sensitive to higher-order (interval content) musical features with age than to lower-order (surface) musical features (e.g., contour), and that cognitive processing of musical features progresses from basic to combined (Cohrdes et al., 2016; Trehub & Hannon, 2006).

Both Breathing Experiments (2018) and (2020) demonstrate that the matching between participants' segmentation and the normative segmentation increases with age, which is consistent with the findings from (Cohrdes et al., 2014; Iversen et al., 2008). As was the case in the previous studies by Koniari et al. (2001) and Melen and Wachsmann (2013), no significant differences in segmentation results were found between participants with and without musical knowledge (or experience in the youngest group), indicating that musical training had no discernible effect on song segmentation.

The results of The Breathing Experiment (2020) indicate that using instrumental versions of the identical songs, as well as the randomization of the sequence of the songs, had no effect on how the three songs' phrases were perceived. Due to the fact that the lyrics were omitted from both versions used to teach the songs (sung version in The Breathing Experiment 2018 and the instrumental version in The Breathing Experiment 2020), all participants were relying on music to locate phrase cues.

4.4 Summary

Two experiments, The Game Experiment and The Breathing Experiment described in this chapter examined the perception of melody in a specific population, aged 5–6 years, and two older groups (middle, aged 8–12 years, and adolescents, aged 14–16 years), both with and without musical knowledge. The experiments were designed to address a need in the literature, as there are currently no studies that explore how these age groups perceive musical phrases.

The Game Experiment, conducted in 2018 and 2020, examined whether a subject's memory of a song is dependent solely on recalling the entire song's notes or on recalling particular phrases. The Breathing Experiment, conducted in 2018 and 2020, examined two hypotheses: (i) that agreement on segmentation within age groups grows with age, and (ii) that similarity between the normative and participant segmentation also increases with age. Hypotheses were validated by the results in both experiments.

By using normative and non-normative melody segments in recreating songs, singing as a teaching method, and no randomized order of songs in The Game Experiment (2018), participants were more effective in rebuilding melodies with normative phrasing, regardless of age or musical knowledge. The significance of prior phrase memory in the transition from one phrase to the next has been proven, providing support for the hypothesis that children and adolescents use phrases to recall melodies.

It is unclear how motifs with a high frequency of repetition vs distinguishing motifs that are not repeated have an effect on the memorization of tones/phrases. By dividing motifs at unique and unexpected locations in non-normative phrasings, it was anticipated that this would influence the perception of learned phrases in normative phrasings. However, the splitting of motifs in Song 3, which has no repeating motifs, was likewise performed in highly improbable locations. Unexpectedly, the finest results in replicating this song were achieved by all groups, irrespective of age or musical experience. Future research comparing motifs that are repeated frequently with those that are repeated infrequently could shed some light here.

The Game Experiment (2020) utilizing only instrumental versions of the same songs and testing the effect of the order, i.e., whether the randomization of songs to be recreated in tasks with normative and non-normative phrasing impacts the results of recreating the songs, produced no statistically significant differences. As in the 2018 Game Experiment, all participants had difficulty recreating songs with non-normative phrasing, proving that neither the teaching method (listening to sung or instrumental versions) nor the order of songs in both versions (normative and non-normative) affects the success of recreating the songs.

The Breathing Experiment (2018) revealed that the absence of meaningful lyrics pushes linguistic components into the background and that the music and its phrases and phrase cues become the dominant focus, since phrase cues in the lyrics cannot be used. Involving children of varying ages in two distinct experiments yielded fruitful results, revealing possible causes for the observed over-segmentation, particularly among the youngest. Future research could investigate why the Breathing Experiment's issues (implicit harmony and sensitivity to contour on phrasing) appear to impair the perception of phrases and phrase boundaries in the youngest group.

The Breathing Experiment (2020) investigated two additional issues: (i) how the method of learning affects the perception of phrases; and (ii) whether the random order of songs influences the results of song segmentation. As with the Game Experiment, there were no statistically significant differences between the 2018 and 2020 Breathing Experiments.

In the two older groups (middle and adolescent), there were no statistically significant differences between participants with and without musical knowledge; however, there were modest differences in the youngest group. Consequently, The Breathing Experiment could be conducted on a larger sample of the youngest children, allowing for a more thorough evaluation of the differences in phrase perception between children with and without musical expertise. In addition, it would be intriguing to investigate whether general participation in choral activities influences the melody phrasing of children of this age.

The perception of phrases and phrase boundaries by the youngest target population members and across age groups has not been compared to a computer model or automatic segmentation. By employing a computational model, it may be possible to gain a deeper understanding of the given data by analyzing specific musical components that have been shown to be significant in children's segmentation. A computational model may provide further answers to some of the study's unanswered questions, including: Is oversegmentation connected to the fact that various age groups interpret the same sound pattern (or song) differently?

According to the study by Berland et al. (2015), age affects hearing, demonstrating a change from non-logical thinking based on perceptual representations (in 6-year-old children) to logical reasoning (in older children) as a result of development, which could be a plausible explanation for why differences in listening between age groups were observed in the Breathing Experiment.

Deutsch (2013, p. 183) argues that the auditory system arranges elements in a sound pattern according to certain qualities (rules), such as frequency, loudness, timing, and timbre. However, the conditions under which these attributes are used as grouping criteria remain unknown. Examining the differences in conditions between two older groups of children (with a similar and adult-like segmentation) and the youngest group may help to further clarify the concept of over-segmentation. In music perception, a two-way process is used: grouping elements using various attributes and assigning them values; and synthesis, in which the values of various attributes are combined, which can also be done incorrectly (Deutsch et al., 2007). The latter, element synthesis, raises a last critical question: Is the over-segmentation of the youngest group the result of 'incorrect' music synthesis? These highly intriguing issues will be the focus of a future study.

Chapter 5

Computational Segmentation of Melody

> If we can give people tools to understand music better, we will also provide access to a lot of new music. It will benefit not only the individual, but the music itself – and the diversity of the entire ecosystem of music.
>
> *Olivier Lartillot*

A musical piece is decomposed (segmented) into its constituent elements in music structure analysis. Segmentation's primary function[1] from a purely musical theoretical standpoint, is to ascertain how a piece of music is constructed, how its structural elements interact, and how they are connected (Ahlbäck, 2007).

In a cognitive approach, segmentation is focused on identifying segment boundaries, which can be difficult given the ambiguous nature of music (Margulis et al., 2017). Even if two listeners interpret a musical composition similarly, there will always be discrepancies in their perception of identical structural elements, as it is exceedingly difficult for a performer to express explicitly how the musical structure is intended to be perceived. Additionally, musical units themselves can be ambiguous, as they can be transformed by changes in rhythm, meter, tempo, and instrumentation (Despić, 2007).

When computational models are used for segmentation, the tasks vary according to the domains to which they are applied, for example, music information retrieval, audio engineering, generative arts, computational musicology, music cognition, and music psychology (López & Volk, 2012). This chapter compares the automatic melody segmentation performed by the IDyOM model (M. T. Pearce, 2005) with the segmentation of melodies performed by children of varying ages and musical experts in The Breathing Experiment conducted in 2018 (more in: 4.3.1) in order to determine which lower and/or higher order musical features are used at different ages and how they vary depending on the musical structure.

There are three reasons for this. To begin with, there are currently no studies comparing computational models to segmentation of music in children of various ages and musical experts. Secondly, preliminary results from a study comparing children's, adolescent's, and musical expert's segmentation to IDyOM's automatic segmentation suggest that IDyOM

[1] Grouping, or segmentation, is distinct from streaming in that it refers to the grouping of sequential contiguous musical elements: While streaming associates a succession of elements with a common source, grouping or segmentation divides that sequence into contiguous groups.

may be capable of replicating human perception of music and its segmentation across ages (Mihelač et al., 2021). Thus, broader use of a computer model in the segmentation process would lead to a better understanding of how children of various ages perceive phrases and phrase boundaries. Thirdly, the objective is to compare the human segmentation process to several computational music representations in order to identify the most effective computational models for the segmentation task.

The following summarizes the chapter's organization. Section 5.1 discusses psychological, musical-theoretical, and computational methods to music segmentation. Section 5.2 presents the computational model IDyOM and the peak-picking algorithm used for segmentation. The methodology, results and discussion are presented in section 5.3. The concluding remarks and future work are summarized in section 5.4.

5.1 Scientific Background and Related Work

Segmentation studies suggest that listeners to music divide and group music unconsciously into separate parts (chunks: Gobet et al., 2001). This grouping of musical structure seem to be done in a more or less similar way (Koelsch et al., 2013; Lerdahl & Jackendoff, 1983b; Nan et al., 2006), regardless of musical background knowledge (Deliège, 1987; Hartmann et al., 2016). Findings from different studies outline that the grouping of the musical content occurs hierarchically, from the lowest level (e.g., notes, chords, motifs) to the highest level (e.g., musical section), mostly somewhere in between these levels, in parts defined as *phrases* (Chiappe & Schmuckler, 1997; Deutsch, 2013; Kragness & Trainor, 2017).

The review of musicological-theoretical, musicological-psychological, and computational literature reveals that the term "phrase" is frequently defined as a musical unit with certain structural (general) and special characteristics, typically with a strong (definable) beginning and an (expected) end, which may or may not have distinct (identifiable) "strengths". According to these "strengths," the beginnings, or the ends of a phrase, as well as the expectations of the phrase ends, are perceived as (more or less) "strong" or "weak". A strong, definable beginning to a phrase may result in a weak phrase ending (and vice versa), or a weak ending to a phrase may result in the necessity of continuation, which does not mean the phrase is not perceived. If the components of a phrase result in a number of weaker closes (endings), the requirement for a strong closure may prompt the creation of one (Spiro, 2007).

Musical features play a significant role in the perception of phrases, according to studies (e.g., Deliège, 1987; Palmer & Krumhansl, 1987; Schaefer et al., 2004; Sloboda & Gregory, 1980; Temperley, 2001). Several of these musical features are thought to contribute, in some extent, to the establishment of boundaries, to the points at which one phrase finishes and another begins. Gaps (a large pitch interval), rests, long notes, changes in timbre, dynamics, note repetitions, articulation (the manner in which a single note is sounded/performed), and the contour of the melodic line are just a few of the most frequently detected phrase cues. These features can be employed alone (interdependent) or in combination (totally interwoven) in the segmentation of music (Neumann & Stevens, 1993, p. 259).

According to Bamberger (1982), there are two distinct ways of processing music: the 'know-how' approach, which is characterized by an intuitive figural approach and is dominant in pre-operational thinking (e.g., in preschool young children), and the 'know-that' approach, which is characterized by conceptual thinking and prior experience (e.g., in middle childhood), in which a (musical) principle or concept is abstracted and transferred to a new musical experience. Bamberger's findings are consistent with those of Christ (1983), who discovered that the cognitive ability to apply many musical features and dimensions simultaneously in the perception of music is more likely to arise throughout children's operational stage.

It was demonstrated in the study by Iversen et al. (2008) that perceived auditory grouping is connected with experience (which is associated with age). By the age of 4–5 years, children can detect simple pitch (tonal) or rhythm patterns (C. Stevens & Gallagher, 2004), however these two fundamental dimensions develop gradually in children with age and experience (Cox, 1970; Drake, 1993; Petzold, 1966). Similar findings were reported in the study by Lamont (1998), indicating that children's sensitivity to higher-order musical features (e.g., interval content) increases with age, as opposed to lower-order (surface) features (e.g., contour). The abilities to process music units develop with age, beginning with very brief melodic or rhythmic units (Cohrdes et al., 2014), as does the cognitive processing of musical elements, progressing from basic to combined (Cohrdes et al., 2016; Trehub & Hannon, 2006).

There are currently no studies addressing how and which musical features are employed independently or in combination in the segmentation of music across ages in children using a computational approach. After the development of powerful and improved algorithms for automatic segmentation, human segmentation of music has been compared to several computational models (e.g., de Nooijer et al., 2008; Melucci & Orio, 2002; M. T. Pearce & Wiggins, 2006a; Potter et al., 2007; Thom et al., 2002); however, only the study from Cambouropoulos (2006) compared children to automatic computational segmentation. The computational model was validated (among other) against the empirical data obtained in Koniari et al.'s 2001 segmentation experiment. 41 children listened to the rondo from Anton Diabelli's Sonatina No. 2 in C Major and highlighted the boundaries with the space bar on a keyboard. This data was compared to Cambouropoulos' computational model using the same musical sample (rondo from the Diabelli Sonatina). The computational model identified 14 peaks that corresponded to the children's segmentations.

In a recent (preliminary) study using the computational model IDyOM, it was discovered that lower and higher-order musical features are utilized differently across the ages, and that these differences may be captured by the computational simulation of human perception (Mihelač et al., 2021).

5.2 Peak Picking Algorithm and IDyOM

Peak picking refers to the detection of signal peaks. The fundamental concept is to locate local maxima in any digitized signal and to use statistical approaches to eliminate false peaks. Methods for audio analysis and synthesis based on sinusoidal modeling commonly employ peak picking techniques.

In IDyOM, the melody is segmented using a peak picking algorithm that uses the boundary strength profile to determine the locations of boundaries.

the parameter k is allowed to vary depending on the nature of the boundary strength profile.

This assumes that segmentation boundaries (the peaks) are to be found in places where information-theoretic measure (information content) has a higher numerical value than in the preceding or following locations, keeping those, which are k times the standard deviation greater than the mean boundary strength (S_n), linearly weighted (w), from the beginning of the melody to the preceding event. The parameter k may vary based on the characteristics of the boundary strength profile. (M. T. Pearce, Müllensiefen, et al., 2010b):

$$S_n > k\sqrt{\frac{\sum_{i=1}^{n-1}(w_i S_i - \bar{S}_{w,1\ldots n-1})^2}{\sum_{i=1}^{n-1} w_i}} + \frac{\sum_{i=1}^{n-1} w_i S_i}{\sum_{i=1}^{n-1} w_i} \tag{5.1}$$

As illustrated in Figure 5.1, the information content and entropy decrease in event $e14$, at the conclusion of a phrase from event $e8$-$e14$, whereas they increase in event $e15$, at the start of a new phrase. Thus, a significant rise in the information content and entropy at a local level (a 'peak') indicates the start of a new segment. Following the conclusion of this new phrase (from event $e15$ to event $e29$), both the information content and entropy of the last event $e29$ decrease.

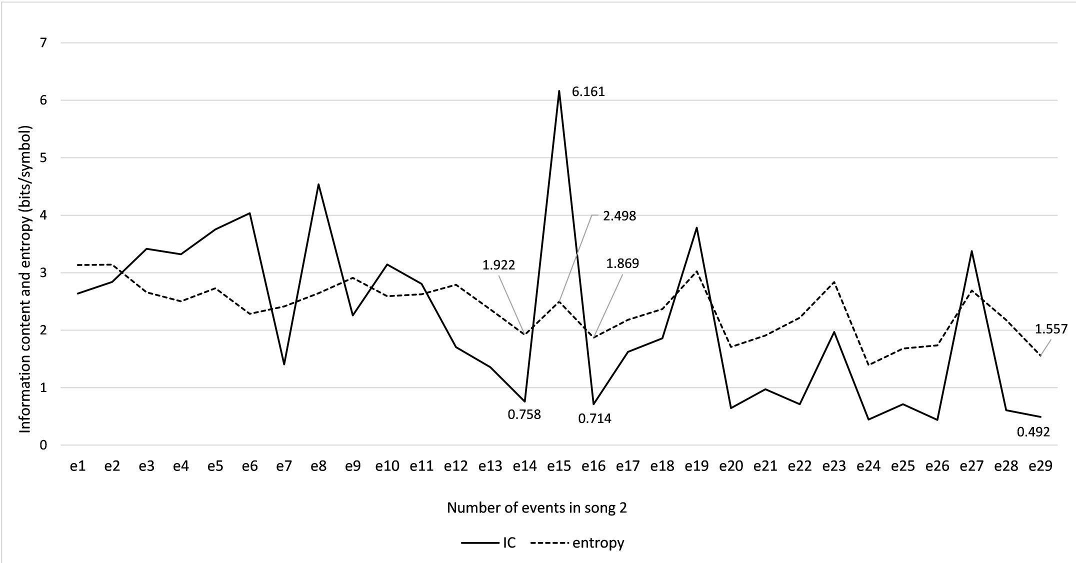


Figure 5.1: Information content and entropy in Song 2. Depicted are the information content and entropy values at the end of a phrase in $e7$, $e14$, and $e29$, and the beginning of new phrases in $e8$ and $e15$

.

5.3 Comparison of Human and Computational Segmentation

Hypotheses

As a broader musical listening experience is accomplished with the increase of age, it is expected that the listening experience will affect the way how the monophonic musical structure is segmented, and that in this segmentation mostly higher order and complex (combined higher and lower) musical features are used. It is also expected that participants with broader musical listening experience and musical knowledge will utilize higher and complex musical features than participants without musical knowledge and a limited musical listening experience.

Two hypotheses are tested: (i) that as age increases, higher musical features are employed for segmentation, and (ii) that participants with musical knowledge segment music using more complex and higher musical features.

Methodology

Stimuli

The normative phrasings of the same three songs (see Figure 4.1 in Chapter 4) were used to compare human and computational segmentation, and Corpus I (see Subsection 2.1), comprising of 155 children's folk songs and children's songs from Slovenia was used for training the IDyOM computational model. All the songs, notated in Sibelius, were encoded into MIDI (256 PPQN) with piano timbre, at the same speed, and without any changes in loudness or articulation. As it is not possible to embed any explicit information about phrases in MIDI as it is possible in e.g., **kern files, where phrases can be marked between curly braces (for detail see Huron, 1997), none of these MIDI files carry any explicit information about the phrases, which means that IDyOM has neither been optimized to predict boundaries nor given access to any boundary information.

Participants

The segmentation of three songs by three groups of children and adolescents, five musical experts were examined (see Table 5.1). Children and adolescents were further divided into subgroups according to their musical knowledge, with (w) or without (wh).

Table 5.1: Groups and numbers of children, adolescents, and musical experts involved in each song.

Group		Age	Number of participants		
			Song 1	Song 2	Song3
EARLY	all	4–6	19	25	9
	with		13	11	6
	without		6	14	3
MIDDLE	all	8–12	23	23	23
	with		6	6	6
	with		17	17	17
ADOLESCENT	all	14–16	51	55	45
	with		26	26	22
	without		25	29	23
EXPERTS	all	31–44	5	5	5

Procedure

The breathing in three chosen songs was collected during The Breathing Experiment conducted in 2018 (more in Subsection 4.3), and normative phrasings from musical experts were obtained by annotating phrases in the musical score. Using Audacity[2] and waveforms of the children's and adolescents' recorded breathing, further analysis of boundary occurrences was conducted.

Phrasings in selected songs were generated by training the model using the Corpus I. To determine the model's generalizability and to avoid overfitting, a cross validation technique was utilized (Berrar, 2019), as well as a value of $k = 10$, which is less biased than smaller values of k and has less variance than larger values of k (Rodriguez et al., 2010). When observing the musical structure with viewpoints, the peak-picking algorithm was utilized to segment it. Viewpoints for modeling the pitch structure are predominant, as pitch is the most complex dimension in children's (folk) songs (Mihelač et al., 2023), but tonality and duration were also considered.

Basic viewpoints (`cpitch, dur`), derived viewpoints (`dur_ratio, cpint, cpint_size, io, contour, cpitch_class, cpintfip, cpintfref, inscale`), and other combinations of the chosen viewpoints were employed. Additionally, 12 viewpoints were used for viewpoint selection: `cpitch, cpitch_class, tessitura, cpint, cpint_size, cpcint_size, cpcint, contour, newcontour, cpintfip, cpintfref, inscale`.

After preliminary testing 30 models using basic, derived, and linked viewpoints, 14 segmentation models out of 30 were chosen to determine if lower- and/or higher-order musical

[2]https://www.audacityteam.org/

features are dominating in music processing and are utilized to segment musical structure at a specific age. Other 16 segmentation models were excluded either because of their (i) to high information content,[3] (ii) because they did not contribute to a better understanding of the utilization of lower/higher order musical features in the segmentation of musical structure, or (iii) because they did not provide any information about boundaries.

Due to its position between a preceding note at the end of a phrase and a succeeding note at the beginning of a phrase, a boundary between phrases could be considered a “problem for a pair of notes”. As illustrated in Figure 5.2, phrase comes to an end on the seventh note, at event $e7$, and a new phrase begins on the eighth note, at event $e8$. The waveform illustrates the distinction between these two events in greater detail, with the long flat line representing breathing at this exact place. As IDyOM functions at the musical surface level (at the note level), the boundary is defined as the first note of a new phrase, which is event $e8$.

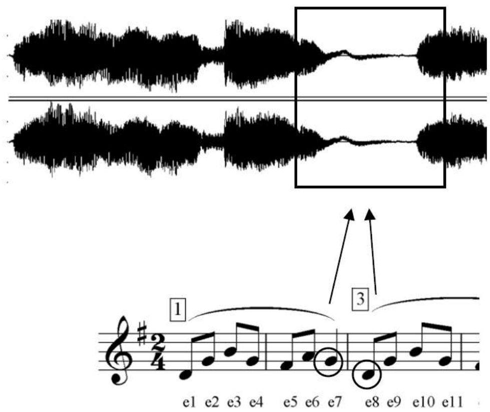


Figure 5.2: Boundary.

A binary indicator was used to indicate a boundary (1 = boundary; 0 = no boundary) for each event in a song, and for each participant separately. As each event (note) in the melody can be interpreted and assigned a value of 1 or 0, the boundary occurrences in each melody were independently compared for each participant with the IDyOM boundary occurrences for each event. Boundaries on the first note of each melody were discarded for the same reason as in the de Nooijer et al. (2008) study: there is no information about the context prior to the first event, and including a boundary on the first event would induce anomalies into the statistical analysis.

To analyze the match between each participant and IDyOM, cosine metric was used to determine the similarity of two binary vectors. The greater the cosine similarity, the closer the segmentations match one another. Finally, by averaging the computed cosine similarities across all of the observed (sub)groups, the mean similarity of each observed (sub)group of children to IDyOM’s phrasing was calculated.

Results

In the initial comparison of human and computational segmentation, 30 segmentation models were utilized, and 18 models were found to be useful in the next stage of analysis. Finally, 14 out of 18 models were used for displaying the matching between groups,

[3]Lower information content of a segmentation model indicates a higher probability of being more similar to the human segmentation of musical structure.

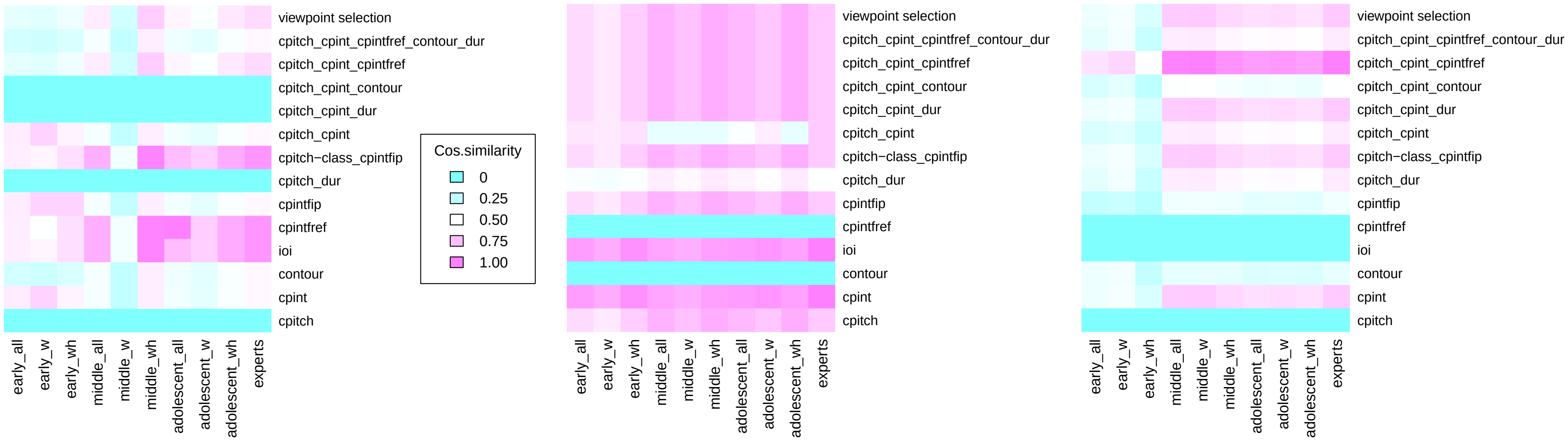
subgroups, experts, and IDyOM, as well as for identifying which musical dimensions and higher/lower order musical features are used across the ages, following a thorough comparison of human and computational segmentation and analysis of information content. Table 5.2 summarizes the information content of each employed model, which is averaged over the full dataset and for each song separately.

Table 5.2: Information content obtained for each used model, averaged on the entire dataset, and on each of the chosen songs separately.

viewpoint	all songs	song 1	song 2	song 3
cpitch	2.339	1.917	2.383	2.229
cpint	2.528	2.107	3.009	2.626
contour	4.003	3.673	4.718	3.582
ioi	1.062	0.576	0.628	0.688
cpintfref	3.216	2.686	3.157	3.306
cpintfip	2.412	1.975	2.436	2.498
cpitch⊗dur	2.340	1.917	2.383	2.229
cpitch-class⊗cpintfip	2.370	1.946	2.495	2.472
cpitch⊗cpint	2.272	1.784	2.561	2.306
cpitch⊗cpint⊗dur	2.211	1.859	2.535	2.225
cpitch⊗cpint⊗contour	2.371	2.029	2.806	2.347
cpitch⊗cpint⊗cpintfref	2.205	1.723	2.472	2.222
cpitch⊗cpint⊗cpintfref⊗contour⊗dur	2.209	1.857	2.566	2.191
viewpoint selection	2.058	1.704	2.155	2.006

The cosine similarity for each group and IDyOM in each song is shown in Table 5.3, where "0" indicate no cosine similarity between IDyOM and groups, subgroups, and experts. Figure 5.3 is showing heatmaps from Song 1, Song 2, and Song 3. In Song 1 we can see an increase in matching across the ages in linked viewpoints, and especially in Song 3, indicating that multiple psychological representations of pitch (e.g., pitch height, pitch interval, contour), pitch combined with duration, and higher-order musical features are used in older groups. Intriguingly, Song 2 does not demonstrate an increase in age-based matching between groups and IDyOM as Song 1 and Song 3 do when viewpoints are linked.

As an additional example, the matching between all groups and IDyOM utilizing `viewpoint selection` is displayed in Figure 5.4. The chosen model created using this procedure has an average IC of 2.058 bits/symbol and is primarily composed of higher-order musical features: ((`cpint cpint-size)(cpitch-class cpintfip)(cpint cpintfref) (cpitch cpintfref`)). Only the greatest peaks indicate a phrase boundary in the depiction of the IDyOM's segmentation; consequently, boundaries are located at $e6$ and $e24$ in Song 1, $e15$ in Song 2, and $e16$ and $e25$ in Song 3.



(a) Song 1. (b) Song 2. (c) Song 3.

Figure 5.3: Heatmaps showing cosine similarity in Song 1, Song 2 and Song 3. We can see increased matching between older groups and IDyOM in Song 1 and Song 2 when linked viewpoints are used, while this is not the case in Song 2.

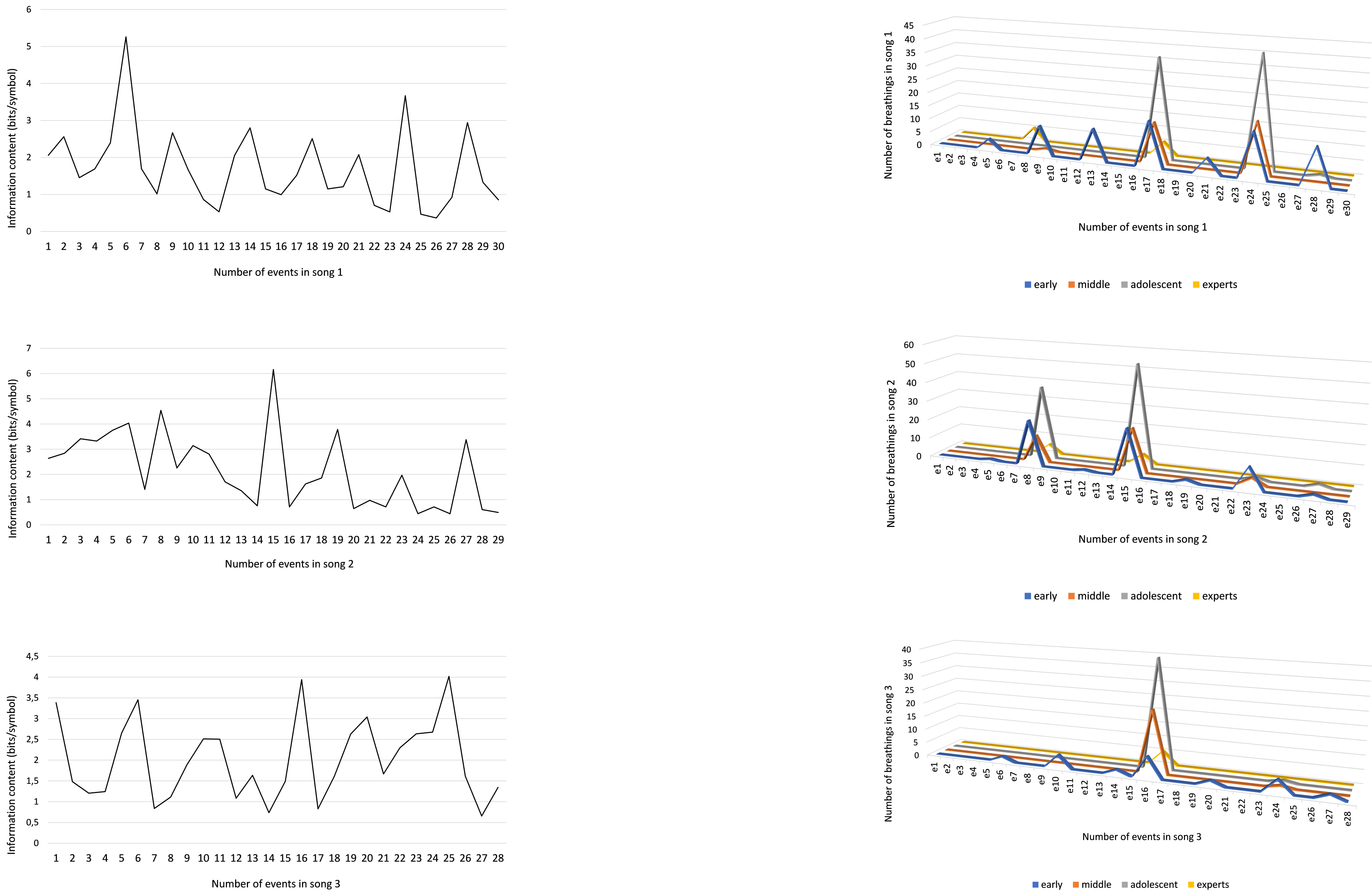
Table 5.3: Mean cosine similarity between segmentations of IDyOM and subgroup of children and experts.

Viewpoint	Song	early_all	early_w	early_wh	middle_all	middle_w	middle_wh	adolescent_all	adolescent_w	adolescent_wh	experts
cpitch	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpint	1	0.442	0.525	0.422	0.363	0.204	0.440	0.347	0.310	0.371	0.408
contour	1	0.253	0.241	0.279	0.363	0.204	0.440	0.340	0.307	0.371	0.408
ioi	1	0.439	0.417	0.483	0.629	0.353	0.761	0.589	0.532	0.643	0.707
cpintfref	1	0.439	0.387	0.483	0.629	0.354	0.761	0.777	0.537	0.643	0.707
cpintfip	1	0.442	0.521	0.422	0.363	0.204	0.440	0.347	0.310	0.371	0.408
cpitch⊗dur	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpitch-class⊗cpintfip	1	0.439	0.417	0.483	0.629	0.353	0.761	0.589	0.532	0.643	0.707
cpitch⊗cpint	1	0.442	0.525	0.422	0.363	0.204	0.440	0.347	0.310	0.371	0.408
cpitch⊗cpint⊗dur	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpitch⊗cpint⊗contour	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpitch⊗cpint⊗cpintfref	1	0.310	0.295	0.342	0.445	0.250	0.538	0.416	0.376	0.455	0.500
cpitch⊗cpint⊗cpintfref⊗contour⊗dur	1	0.253	0.241	0.279	0.363	0.204	0.439	0.340	0.307	0.371	0.408
viewpoint selection	1	0.310	0.295	0.342	0.445	0.250	0.538	0.416	0.376	0.455	0.500
cpitch	2	0.643	0.583	0.691	0.799	0.740	0.820	0.772	0.720	0.817	0.707
cpint	2	0.882	0.824	0.927	0.853	0.811	0.869	0.886	0.909	0.862	1.000
contour	2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ioi	2	0.882	0.824	0.927	0.853	0.811	0.869	0.886	0.909	0.862	1.000
cpintfref	2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpintfip	2	0.643	0.583	0.691	0.799	0.740	0.820	0.770	0.720	0.817	0.707
cpitch⊗dur	2	0.478	0.465	0.489	0.565	0.523	0.580	0.546	0.507	0.577	0.500
cpitch-class⊗cpintfip	2	0.643	0.583	0.691	0.799	0.740	0.820	0.772	0.720	0.817	0.707
cpitch⊗cpint	2	0.596	0.583	0.620	0.408	0.407	0.408	0.481	0.566	0.403	0.707
cpitch⊗cpint⊗dur	2	0.643	0.583	0.691	0.799	0.740	0.820	0.772	0.720	0.817	0.707
cpitch⊗cpint⊗contour	2	0.643	0.583	0.691	0.799	0.740	0.820	0.772	0.720	0.817	0.707
cpitch⊗cpint⊗cpintfref	2	0.643	0.583	0.691	0.799	0.740	0.820	0.772	0.720	0.817	0.707
cpitch⊗cpint⊗cpintfref⊗contour⊗dur	2	0.643	0.583	0.691	0.799	0.740	0.820	0.772	0.720	0.817	0.707
viewpoint selection	2	0.643	0.583	0.691	0.799	0.740	0.820	0.772	0.720	0.817	0.707
cpitch	3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpint	3	0.431	0.464	0.354	0.698	0.707	0.653	0.624	0.633	0.615	0.707
contour	3	0.430	0.454	0.272	0.403	0.408	0.401	0.360	0.366	0.355	0.408
ioi	3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpintfref	3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
cpintfip	3	0.273	0.294	0.224	0.442	0.447	0.439	0.395	0.401	0.383	0.447
cpitch⊗dur	3	0.401	0.455	0.289	0.570	0.577	0.533	0.510	0.517	0.502	0.577
cpitch-class⊗cpintfip	3	0.431	0.464	0.354	0.698	0.707	0.653	0.624	0.633	0.606	0.707
cpitch⊗cpint	3	0.352	0.379	0.289	0.570	0.577	0.533	0.509	0.517	0.497	0.577
cpitch⊗cpint⊗dur	3	0.431	0.464	0.354	0.698	0.707	0.653	0.608	0.633	0.610	0.707

Continuation of Table 5.3

Viewpoint	Song	early_all	early_w	early_wh	middle_all	middle_w	middle_wh	adolescent_all	adolescent_w	adolescent_wh	experts
cpitch⊗cpint⊗contour	3	0.347	0.394	0.250	0.494	0.500	0.462	0.441	0.448	0.429	0.500
cpitch⊗cpint⊗cpintfref	3	0.610	0.657	0.500	0.987	1.000	0.924	0.882	0.896	0.870	1.000
cpitch⊗cpint⊗cpintfref⊗contour⊗dur	3	0.401	0.455	0.289	0.570	0.577	0.533	0.509	0.517	0.502	0.577
viewpoint selection	3	0.431	0.464	0.354	0.698	0.707	0.653	0.624	0.633	0.606	0.707

Figure 5.4: Matching between all the groups and IDyOM when using `viewpoint selection` in the segmentation task.



Discussion

The purpose of comparing the IDyOM computational segmentation of musical structure to the segmentation of music by children, adolescents, and musical experts was to determine whether IDyOM depicts age-related differences in segmentation while modeling human music perception. Corpus I (155 Slovenian children's (folk) songs) was used to train the computational model and afterward applied to the three songs, for the same reason that they were used in the Game and Breathing Experiments (see Chapter 4): their distinct structure, as two of them (Song 1 and Song 3) lack specific segmentation indications, while the third (Song 2) has an extremely clear melodic structure and unambiguous segmentation clues.

In order to simulate human segmentation of musical content, IDyOM was used to generate 30 models with separate and linked viewpoints, tested in the initial stage, however, 18 models were shown to be useful in the human and computational comparison. The final decision was to use only 14 models out of 18. The excluded four models had either too high information content or did not contribute to a better understanding of how lower- and higher-order musical features are used in the segmentation of music in various age groups.

During the segmentation task, differences in the efficacy, processing, and usage of lower- and higher-order musical features were detected (e.g., `cpintfref`), as represented by the similarity (matching) between human and automatic segmentation. The fact that increased matching was found between IDyOM and older groups when using linked viewpoints (Table 5.3), indicates (i) that older groups are using more multiple representations of pitch and time in the segmentation task than younger groups (Deutsch, 1982; C. L. Krumhansl, 1990; Levitin & Tirovolas, 2009). This is in accordance with the study by Costa-Giomi (2003), in which was found that older children have fewer limitations on their memories and more ways to direct their attention to the pertinent cues of the stimuli, which makes older children better able to process information.

There was considerable matching between all the groups and IDyOM in Song 2 when `cpint` was used in all three songs, and moderate matching when `contour` and higher-order musical features such as interval (`cpint` and tonality `cpintfref`) were used for song segmentation, even in the youngest group, in the linked viewpoints. The findings indicate that, regardless of age or experience, the use of musical features in music segmentation is dependent upon (among other factors) the musical structure, which may or may not give unambiguous clues for musical content segmentation.

In comparison to Song 2, Songs 1 and 3 include small intervals (see Figure 4.1), only one rhythmic clue ($e23$) in bar five of the first song and an eight pause in bar four of the third song), and, particularly in Song 1, a non-distinctive contour. When the information stored in the musical content is insufficiently relevant, higher-order musical features are processed and used in the segmentation task, rather than just a few lower-order musical features, however not sequentially from lower to higher order (Umemoto, 1990), but separately or in combination (depending on the musical structure) and are more effective in older age groups. When viewpoints are linked in various combinations, ranging from two to five, increasing matching outcomes across the ages are noticeable (see Table 5.3 and Figure 5.4).

In summarizing the findings, comparing human and machine segmentation in this chapter, the first hypothesis that higher-order musical features are used as participants get older was only partially supported. First, only the results of three songs were examined. Further research is needed to identify why and how lower- and higher-order musical features are used in song segmentation at different ages, as well as a comparison of automatic and human segmentation utilizing a larger sample size.

Second, it is unclear whether the over-segmentation in the youngest group early (Figure 5.4) is related to unfamiliarity with songs, resulting in a difference in the use of lower/higher-order musical elements in song segmentation, as this data was not available for study. Adults' perceptions of familiar melodies, according to Morrongiello et al. (1985), are based on interval information rather than absolute pitch information, whereas novel melodies are based on less exact contour information. It is possible that lower- and higher order musical features are used differently in familiar and unfamiliar songs, resulting in an over-segmentation in the youngest group.

The second hypothesis that participants with musical knowledge segment music using more complex and higher musical features was partially supported as well. According to Bigand (2003), participants without musical knowledge process music using the same principles as those with musical knowledge. Saari et al. (2018), using six musical features representing low-level (timbre) and high-level (rhythm and tonality) aspects of music perception, demonstrated that musical training had the greatest effect on the processing of high-level musical features. As all participants without musical knowledge performed better in Songs 1 and 2, while participants with musical knowledge performed better in Song 3 (Table 5.3), it is unclear whether musical knowledge or the musical structure itself influences the use of lower- and higher-order musical features regardless of age.

5.4 Summary

In this chapter, the automatic segmentation from IDyOM was compared to human segmentation across ages. Differences in the perception of musical phrases and phrase boundaries from the perspective of the use of lower and higher-order musical features in the segmentation of musical content were discovered using IDyOM for the simulation of human segmentation and using different viewpoints for the observation of the musical structure. The IDyOM's capacity to capture the human's perception of musical segments at different ages was demonstrated (to some extent) by increased matching from the youngest group (early) to the oldest (musical experts) in the similarity measurement between human and computational segmentation.

The findings suggest that each musical feature can play a varied role in phrase identification when used alone or in combination with other musical features, depending on a variety of factors such as experience, age, and feature combination in a specific part of the song. More research is needed to identify what factors influence how certain events are used in the segmentation of music in different age groups, and why some aspects of the musical structure are perceived to be more essential than others. The impact of song familiarity on the perception of phrases and phrase boundaries, independent of age or experience, is of special interest, as is the analysis of the (potential) impact of song familiarity on over-segmentation.

As it was discovered in The Breathing Experiment (2018) (Chapter 4), that anacrusis has an influence on the memory and perception of phrases and phrase boundaries in Song 3, it is intended to pay special attention to all Corpus I songs that contain anacrusis a future study.

Although oversegmentation was detected in the youngest group early in this song, using IDyOM and various viewpoints either separately or in combination (the `viewpoint selection` being an exception), numerous plausible boundaries were identified that correspond to the segmentation in this group. Given that phrase structure is considered to be hierarchical, the most confusing segmentation occurs in the smallest units, which can be subject to not just individual interpretations (Friberg & Battel, 2002), but also to extremely contradicting perceptual cues present in the musical structure (Spiro, 2007).

Chapter 6

Computational Detection of (Ir)regularity in Children's Folk Songs

> The beauty of irregularity–which in its true form is actually liberated from both regularity and irregularity–the asymmetric principle contains the seed of the highest form of beauty known to man.
>
> *Sōetsu Yanagi*

Music may be perceived as "difficult," "complex," or even "incomprehensible by the ear." Many listeners will reject such music and seek out alternatives, unless they are explicitly interested in the reasons for a particular musical piece's perception. Behind each piece of music is a story about the creation of a musical structure, which encompasses various musical elements and dimensions, and a musical syntax, which is a more or less formal description of the rules that define the permissible structure, i.e., how the constituent parts of a piece may be formed and combined over time (Rohrmeier & Pearce, 2018).

Syntactic features include, but are not limited to, the hierarchical structure generated through the combination of perceptually discrete linked elements over a range of timescales. Another syntactic feature is the sequential relationship between musical elements, as the "syntactic functional-psychological qualities" of certain musical elements are dependent upon their relationship with others (Bigand et al., 2014, p. 2).

Numerous studies (e.g., Bigand et al., 2014; Bigand & Poulin-Charronat, 2006; Koelsch & Jentschke, 2008; Mihelač et al., 2018; Mihelač & Povh, 2020a) emphasize the significance of musical syntax and its contribution to the concept of structure that is *irregular*, or adheres to syntactic rules, as opposed to structure that is *irregular*, or nonconforming. These studies, in general, follow a traditional empirical approach (including listeners in experiments and music rating). In addition, these studies investigate the relationship between musical syntax and deviation in a musical structure, which occurs when structure deviates from syntax, and the listener's perception of the (ir)regularity of musical structure, as well as their acceptance, enjoyment, and comprehension of musical works and/or musical genres.

Taking a more traditional approach can be costly and scientifically demanding, as it is difficult to eliminate subjectivity in evaluation tasks, both among musical experts and among untrained listeners. Additionally, it is difficult to replicate the findings, as the response to a task might vary significantly between hearings, even when the same participants are used in the replication (Mihelač & Povh, 2020b).

In recent decades, computational approaches to music analysis have enabled a more objective examination of music (Potter et al., 2007), allowing for the identification of significant features (such as semiotic structure: G. A. Wiggins, 2010) in a musical structure, in comparison to human music processing/production. These computational methods are now capable of simulating the human perception of music to a certain extent, particularly when based on an underlying cognitive theory.

The fact that computational models can be efficiently used in the simulation of the human perception of music was the motivation for a recent study by Mihelač and Povh (2020a). Human experts involved in the detection of (ir)regularity in the musical structure in a previous study (Mihelač & Povh, 2020b), and the evaluation of (ir)regularity of musical excerpts by listeners were replaced by simulating their responses with the computational model, IDyOM (M. T. Pearce, 2005). The artificial model of the perception of (ir)regularity obtained by Mihelač and Povh (2020a) is shown to accord with human perception of irregularity in the previous study (Mihelač & Povh, 2020b), suggesting that expert-based detection of (ir)regularity in musical structure can usefully be replaced by a suitable computational model.

In this chapter, a similar approach to the (ir)regularity in musical structure is used, comparing (ir)regularity in children's folk songs (Corpus 2.2) with the IDyOM computational model, by observing the information content and entropy of musical structure with then viewpoints (`IC_cpitch, E_cpitch, IC_cpint, E_cpint, IC_cpintfref, E_cpintfref, IC_cpitch⊗dur,E_cpitch⊗dur,IC_contour, E_contour`),[1] and with the algorithm Ir_Reg, which classifies melodies according to regularity of the musical structure. The decision to use this corpus is to show that children's folk songs, often presumed to be simple and regular in structure (Herzog, 1947; Ling, 1997; Nettl, 1983; Pond, 1981; Romet, 1980), can also be complex, and considered as an alone standing genre.

Following is the chapter's structure: In Section 6.1, research on (ir)regularity is provided. The horizontal and vertical approaches to monophonic melody, as well as the algorithm Ir_Reg for the detection of (ir)regular songs, are explained in Section 6.2. Methodology and findings are detailed in Section 6.3. The most significant findings and future efforts are outlined in the concluding Section 6.4.

6.1 Scientific Background and Related Work

Regularity exists in both natural and man-made objects, including biology, physics, engineering, architecture, and art, and is crucial to human life. The detection of recurring structures (patterns)[2] is important since it determines our ability to recognize and understand the world (Pauly et al., 2008). Thus, identifying patterns that are repeated and form a regular structure might help in comprehending and analyzing structural irregularities caused by certain criteria (e.g., unexpected use of chords in the harmonic progression and its impact on listener enjoyment of a musical piece). Regularity is a term that refers to a class of configurations that an observer is capable to use or identify when they occur (Feldman, 1997, p. 3,):

$$X_R \subset X \tag{6.1}$$

where R (regularity) is a logical predicate defined on X, which holds on some subset X_R.

[1] 'E' denotes entropy, and 'IC' information content.

[2] In this thesis, pattern is defined as the succession of event features.

Music is an art form composed of "humanly organized sounds" (Godt, 2005, p. 84). Thus, sounds are not randomly distributed, but are organized in a specific order: discrete units (sounds) into smaller units, and smaller units into larger units, all the way up to the level of a musical piece, establishing a hierarchical structure that is a cornerstone of music theory and cognition (Levitin, 2000). A set of permitted and rigid rules, referred to as *musical syntax*, dictates how this structure must be arranged, encompassing a broad variety of fundamental musical elements (Berezovsky, 2019).

According to studies (e.g., D. Cohen, 2003; D. Cohen & Katz, 2013; Marcus, 2003), musical syntax varies not only between cultures (in non-Western and Western tonal music traditions), but also within Western tonal music styles (e.g., the Middle Ages, the Renaissance, the period of tonal music from approximately 1600 to 1918, and the Modern period after 1918), due to the evolution of syntactic rules in music over time. (Klein & Jacobsen, 2014; Meyer, 1989; Vuvan & Hughes, 2019). Changes within these time periods are seen as compositional choices (see Meyer, 1989, for details), rather than as changes in musical language.

Numerous empirical studies indicate that musical syntax is cognitively represented (e.g., J. J. Bharucha & Stoeckig, 1986; Tillmann et al., 2000) and that observable neural correlates exist (e.g., Janata et al., 2002; Koelsch & Siebel, 2005). Manipulating any of the fundamental musical elements inside a structure has an effect on the reported *feeling* of regularity (Mihelač & Povh, 2020a; Mihelač & Povh, 2020b; Pole, 2014; Rohrmeier, 2011; Rohrmeier & Pearce, 2018), and its *identification* (Bruner et al., 1959). Patel (2003b) claims that when the order of elements (parts) is altered, a piece of music loses its identity, which corresponds to how sequential information and structure are processed in non-musical domains (Garner, 1974).

When a piece of music has a strong structure, with recurring dominant musical elements (patterns), and strong relationships between these patterns, it is deemed to be *regular* (from the perspective of musical syntax and/or listener perception) (Manjunath et al., 2000). On the other hand, an *irregular* piece has a non-structured or weakly structured texture, with a small number of recognizable patterns and weak interrelationships (Kramer, 1988). Klein and Jacobsen (2014) point out that for each compositional rule in a given tonal style, a complementary regularity is generated: a structure that appears to be syntactically regular in one piece may appear to be irregular in another (e.g., frequent endings on dominant triads in Romanian folk songs, which are perceived as syntactically "regular", however as "irregular" in children's folk songs from the same country, as these songs end on the tonic triad).

Several studies have demonstrated that a stronger sense of regularity is perceived when similar parts are repeated on a periodic basis, either in an "absolute repetition condition", where parts are repeated identically, or in a "relative repetition condition," where parts are considered conditionally identical (Bader et al., 2017).[3] Relative repetition is important in the formation of motifs that serve as the structural framework for larger works (Cambouropoulos et al., 2001; Deliège, 1987). Thus, the predominance of repetition in music, found in all known human cultures, is unsurprising, as is the fact that listeners frequently return to familiar musical compositions (Margulis, 2014). Repetition is a fundamental characteristic of music, "design feature" of music (Fitch, 2006), and it contributes considerably to our understanding of music (Schoenberg et al., 1967).

Bruner et al. (1959) point out that perception of recurrent regularities (absolute or relative repeated parts/patterns) is impacted either by elements that mask the identification of recurrent regularities (e.g., an input-stimulus, which does not conform the recurrent

[3]An example is the *transposition* of a fragment where only the pitch intervals in a part are repeated, while the entire fragment is shifted in pitch.

series in a sequence) or by the regularity itself, in the case where it "exceeds the memory span" of an observer (Bruner et al., 1959, p. 84). To identify recurrent regularity in a musical structure, a listener must either develop a model to represent the regularity or deploy a previously constructed model. In either scenario, identification success is dependent upon separating recurrent regularities from interfering structures. The more elements that interfere with recurring regularities, the more noise is detected in the stimulus, and thus the more difficult it is to identify recurrent regularities (Bruner et al., 1959).

Perception and identification of recurrent regularities and structural peculiarities in a musical structure are dependent upon the listener's internalization of the musical structure and its syntax. Internalization of music, which can occur implicitly through mere exposure to music without listening or explicitly through interaction with music (M. T. Pearce & Wiggins, 2012), is dependent on a number of factors, including the listener's musical experience and/or training (Bangert & Altenmüller, 2003; Lappe et al., 2013). Lu and Vicario (2014) demonstrates that human infants and adults recognize recurring sound patterns even when musical sequences are presented passively. Regardless of how internalization happens, internal models of (permissible) musical structures are generated and applied whenever new music is heard (Agres, 2019; Deliège, 1987). If the internal models come into conflict with the (prescribed) rules for the arrangement of the constituent parts of a musical structure, the structure is perceived as not only irregular, but also as more complex and less enjoyable (Mihelač & Povh, 2020b; Sauvé & Pearce, 2019), because more effort has to be put into the listening process in order to fully understand the novel structures.

Musical genres differ in terms of recurrent regularities in musical structure and musical elements that are emphasized in an absolute or relative repetition condition, thereby contributing more or less to the regularity. Some musical genres (for example, modernist and expressly avant-garde approaches) avoid repeating (Margulis, 2014),and any sense of surface regularity is purposefully missing. Regular structure is achieved in several other genres, for example, minimalist music, by accentuating the rhythm and employing recurring rhythmic patterns (Johnson, 1994). Children's (folk) songs are another example of music with recurring regularities in its structure, utilising both absolute and relative repetition requirements (Jožef-Beg & Mihelač, 2019).

Several approaches have been used to measure (ir)regularity in musical structure, including subjective evaluation by listeners (e.g., Deutsch, 1980; Mihelač et al., 2018; Mihelač & Povh, 2020b; Tillmann & Bigand, 1996), measurement of neural response (mismatch negativity, MMN) and functional magnetic resonance imaging (fMRI) (e.g., Grahn & Rowe, 2012; Ulanovsky et al., 2003; Yu et al., 2015), or simulation of human perception of (ir)regularity using a computational system (e.g., Hansen & Pearce, 2014; Mihelač & Povh, 2020a).

6.2 Exploring Vertical and Horizontal (Ir)regularity in Musical Structure

Music is multidimensional, and its dimensions are never found in isolation; rather, they are continually interacting (more or less) with one another (Prince, Thompson, et al., 2009). These dimensions are perceived differently depending on how they are presented in the musical structure, which can be either *vertical*, when the relationships between notes are presented simultaneously (for example, harmony) or *horizontal* (for example, melody), in which notes are presented sequentially. Pitch is unique in that it can be displayed vertically (as chord) or horizontally as a sequence of notes (Loui, 2012).

However, when both vertical and horizontal dimensions of music are used within a musical piece (encompassing all issues of what exactly should be considered vertical or horizontal), it forms a unit in which the musical content is stored (Busch & Graubart, 1986). Both of these dimensions are an extension of one another, with vertical being an extension of horizontal and vice versa (Williams, 2005). In spite of simply being given one dimension (such as melody), listeners frequently infer structures from another dimension (such as harmony) (Butler & Brown, 1994; Lerdahl & Jackendoff, 1983a; Platt & Racine, 1994), which (partially) explains why it is not always sufficient to capture only one musical dimension when seeking to understand what causes listeners' perception of (ir)regularity to be higher or lower.

In a study that examined the impact of the complexity of harmony on the acceptability of music (Mihelač & Povh, 2020b), thus simply the vertical dimension, 53 out of 160 musical excerpts were found to have a more complicated and irregular musical structure, even when the harmony was not complex. In addition to the vertical dimension, the horizontal dimension (melody) was also examined to identify what precisely contributes to the feeling of irregularity in the same 53 musical excerpts (Mihelač & Povh, 2020a). Melodies were extracted from all 160 musical excerpts in order to obtain pure monophonic musical excerpts, i.e., only the very first upper lines in each melody were used.[4] Using IDyOM to simulate the listener's perception of music and eight viewpoints [5] to examine the entropy and information content of the musical structure, significantly different distributions of pitch and *implied harmony*, examined with the viewpoints `IC_cpintfref` and `E_cpintfref` were discovered on the sets of regular and irregular musical excerpts.

According to Sloboda and Parker (1985), each tone in a single melodic line can *imply a harmony* as a mental model of the underlying structure, a conclusion that has been replicated in additional research by Holleran et al. (1995), Platt and Racine (1994), and Thompson and Cuddy (1989). When melody and harmony are combined in a musical example (vertical and horizontal dimensions), a harmonic frame, which can have two dimensions, is established: a *global* dimension (key and mode) and a *local* dimension defined as a region within the key that is assigned to a harmony and defined as a function (e.g., tonic, subdominant, dominant, etc.) (Povel & Jansen, 2002). As per Povel and Jansen (2002), a listener establishes global and then local aspects, although the mechanisms underlying the development of these two aspects, which are generally conceptualized as hierarchical are not well understood (J. J. Bharucha, 1987; Tillmann et al., 2000).

Applying the concepts of "global" and "local" establishment of the harmonic frame from Povel and Jansen (2002) to the data used in the study in 2020 (Mihelač & Povh, 2020a), it was found that while listening to a particular monophonic musical excerpt, listeners first generate a key and mode, after which implied harmonies are "created" for each note. In some cases, these implied harmonies do not "fit" within the existing harmonic framework, which is (re)created when the same melody is combined with its official underlying harmony.

Therefore, when "horizontal" (melody) and "vertical" (harmony) musical content are presented together, a "fusion" of different tones occurs (Huron, 2001; Parncutt, 1989) to generate different harmonies. These harmonies can be emphasized depending on the information value either in melody or harmony, which could explain the perception of higher complexity in musical excerpts with a simple harmonic progression (e.g., tonic - dominant - tonic), as the focus is placed on the melody and its harmonies. This agrees with previous findings reported by Melara and Algom (2003) and Prince, Schmuckler, et al. (2009).

Significant differences in pitch perception between regular and irregular musical ex-

[4] This was done because some melodies within this dataset were composed in a polyphonic or chordal manner.

[5] `IC_cpitch`, `IC_cpint`, `IC_cpintfref`, `IC_cpitch⊗dur`, `E_cpitch`, `E_cpint`, `E_cpintfref`, and `E_cpitch⊗dur`.

cerpts in Mihelač and Povh (2020a) were explained by the higher degree of pitch diversity and higher proportion of *non-chordal tones* in 53 irregular musical excerpts. Specifically, non-chordal tones appear to affect listener's enjoyment of music, as well as the expectation for forthcoming events, according to findings in Mihelač et al. (2018) and Mihelač and Povh (2020b). The successful harmonic analysis of a musical piece by a listener is clearly dependent on how successfully the non-chordal tones are resolved and assigned to a harmony and how the tones distributed in a melody are perceived as an implied harmony, which agree with findings from a previous study (Povel & Jansen, 2002).

To summarize, the perception of a musical dimension alone can differ from the perception of the same musical dimension in conjunction with another dimension (Prince, Thompson, et al., 2009). In the latter case, seemingly minor changes to the structure of one dimension (such as melody) may affect how the structure of another dimension is perceived (e.g., harmony). Depending on stimulus and task characteristics, the relevance of a certain musical dimension can be magnified, and a dimension with a higher informative value is more likely to dominate other dimensions, corresponding with the findings of Mihelač and Povh (2020a).

Exploring Vertical and Horizontal (Ir)regularity in Children's Folk Songs

Algorithm Ir_Reg

To examine the horizontal and vertical (ir)regularity in children's folk songs, the algorithm Ir_Reg was used. Algorithm Ir_Reg takes as an input dataset D (Corpus II), a subset $\bar{D}$ of D, and a set V of viewpoints that we want to analyze. Firstly, Ir_Reg computes by IDyOM information content and entropy for all viewpoints from V, for all songs from $\bar{D}$, using the entire dataset D, to train the long-term model. Next, Ir_Reg calculates for each viewpoint the maximum values across the dataset ($\bar{D}$). Finally, for each song from the subset $\bar{D}$ the algorithm counts the number of viewpoints above the 75 % of the maximum value of the viewpoint (upper threshold) or below the 25 % of the maximum value of the viewpoint (lower threshold). The Ir_Reg algorithm is actually defined for general set of information content and entropy viewpoints V, but in this chapter, it was applied with V consisting of the chosen ten viewpoints. If a given song has at least 5 values (entropy or information content) above the upper threshold it is classified as *irregular*. Otherwise, if it has at least 5 values below the lower threshold, it is classified as *regular*. If neither of these cases happens, it is labelled as *unclassified*. The number 5 in the algorithm was determined empirically. Ir_Reg is specified in Algorithm 6.1.

Exploring Vertical and Horizontal (Ir)regularity in Three Children's Folk Songs with Entropy and Information Content

Primary dimensions of children's folk songs are pitch and time, which are represented in IDyOM by the basic viewpoints `cpitch` (chromatic pitch), `dur` (duration) and `bioi` (basic inter-onset interval). To investigate horizontal and vertical (ir)regularity in children's folk songs, entropy (E) and information content (IC) were measured with the same viewpoints as in Mihelač and Povh (2020a)'s study, as they have been demonstrated to be relevant to the capture of (ir)regularity. Specifically, the basic viewpoint `cpitch`, and the derived viewpoints `cpint`, and `cpintfref` to explore the vertical dimension, were employed. Pitch and duration were used as linked viewpoint `cpitch`⊗`dur` because there is no consensus on whether pitch and duration are processed separately or simultaneously (more about this topic in Boltz, 1999; Jones & Boltz, 1989; C. L. Krumhansl, 2000; Volk, 2016). In addition to these four viewpoints, the derived `contour` viewpoint was utilized.

```
Algorithm 6.1: IR_REG
  Input: D, D̄ ← full and partial dataset of songs, i.e., D̄ ⊂ D
  Input: V ← viewpoint set
  Result: Classification of members of set of songs as (IR)REGULAR
  /* Import function to compute IDyOM mean values for IC and E for each
     song a dataset                                                    */
  Import IDyOM(Measure,Song,Viewpoint,Dataset);
  /* Step 1:  Find maxima for mean IC and mean E across D̄             */
  for v ∈ V do                                  /* ...for each viewpoint */
  |  max_{v,D̄,IC} ← max_{s∈D̄} IDyOM(IC, s, v, D)) ;   /* Find max mean IC */
  |  max_{v,D̄,E} ← max_{s∈D̄} IDyOM(E, s, v, D)) ;     /* Find max mean E */
  end
  /* Step 2:  Classify Songs                                              */
  for s ∈ D̄ do                                       /* Consider each song */
  |  nMax = nMin = 0;                                /* Reset the counters */
  |  for v ∈ V do                               /* Consider each viewpoint */
  |  |  if IDyOM(IC, s, v, D) > max_{v,D̄,IC} ×0.75 then      /* Test max IC */
  |  |  |  nMax++;                                     /* Count if high IC */
  |  |  end
  |  |  if IDyOM(IC, s, v, D) < max_{v,D̄,IC} ×0.25 then      /* Test min IC */
  |  |  |  nMin++;                                      /* Count if low IC */
  |  |  end
  |  |  if IDyOM(E, s, v, D) > max_{v,D̄,E} ×0.75 then         /* Test max E */
  |  |  |  nMax++;                                      /* Count if high E */
  |  |  end
  |  |  if IDyOM(E, s, v, D) < max_{v,D̄,E} ×0.25 then         /* Test min E */
  |  |  |  nMin++;                                       /* Count if low E */
  |  |  end
  |  end
  |  if nMax ≥ 5 then
  |  |  Classify(s,IRREGULAR);
  |  else
  |  |  if nMin ≥ 5 then
  |  |  |  Classify(s,REGULAR);
  |  |  else
  |  |  |  Classify(s,UNCLASSIFIED);
  |  |  end
  |  end
  end
```

As an example, three children's folk songs from Corpus II are presented in the continuation. They have been classified as "regular", "irregular", and "neutral" by the algorithm IR_REG. The information content (unpredictability) and entropy (uncertainty) are shown by examining the viewpoint `cpitch`. The viewpoint `cpitch` was chosen, as it has shown that pitch significantly contributes to the perception of regular and irregular musical excerpts (Mihelač & Povh, 2020a). For each event, e_i (in our case a tone in a sequence/melody), the information content and entropy is calculated from a probability distribution. In Figure 6.1, the Turkish children's folk song, *Cumhuriyet Çocuklarıyız*, classified by algorithm IR_REG as "regular" is presented. It consists of 78 events, in which motifs, hereafter defined as *patterns* (p), are denoted by p_1, p_2, and p_3. From the perspective of pitch span (within a perfect fifth), rhythm, and meter, the song is simple. Furthermore, the song has a clear and understandable musical syntax, and with a structure that is strong, consisting of three dominant and often repeated patterns with substantial relationships, as p_2 always follows p_1, and p_3 follows p_2.

Figure 6.1: An example of a regular, structured melody. The Turkish children's folk song *Cumhuriyet Çocuklarıyız*, consists of three repeated motifs (the motifs are marked with p), which are repeated throughout the entire song. From the perspective of their frequency of appearance, these motifs can be considered as "recurrent", "dominant", and the melody appears to be "regular".

In Table 6.1, the information content (IC), and entropy (E), obtained for the viewpoint `cpitch`, and assigned to each event in pattern p_1 is shown. The pattern p_1 appears at the beginning of the song (from e_1 to e_4), and is repeated four times, in $e_{25–28}$, $e_{35–38}$, $e_{59–62}$, and $e_{69–72}$. The order of the events (from e_1 to e_4) is always the same, which means, that each repeated version of pattern p_1 starts with the same note presented as event e_1, and afterwards as e_{25}, e_{35}, e_{59}, and e_{69} in the repeated patterns.

Tables 6.2 and 6.3 present the IC and E values for the viewpoint `cpitch`, for pattern p_2 and p_3 respectively. As in pattern p_1, the events in these two patterns always appear in the same order in the repeated versions. A tendency to decrease is visible in the IC of the events in all three patterns, from their first to their last (repeated) appearance.

The second song, *Schneeglöggli*, which is from Switzerland (see Figure 6.2) consists of 50 events and 11 patterns, again labeled as p_i. IR_REG classifies this song as "irregular". It has a weak structure, without any dominant or repeated patterns, and it has large intervals (e.g., event $e_{35–36}$) and no special relations between the patterns. The information content (IC) and entropy (E) obtained for the viewpoint `cpitch` are presented for each event in this song (in the musical score, in the boxed text IC and E). We can see clearly very high IC values in each of the 11 patterns, giving the impression of highly unexpected and complex musical content.

The third song, *Doidas andam as galinhas*, from Portugal, has been classified by

Table 6.1: Information content (IC) and entropy (E) obtained for the viewpoint `cpitch`, defined for each event (from the distribution at that point in the sequence), for the pattern p_1 in the Turkish song *Cumhuriyet Çocuklarıyız*. The IC values in the first and second appearance of this pattern are more or less similar, but then decrease in the third, fourth, and fifth appearance.

Event	1st (IC/E)	Event	2nd (IC/E)	Event	3rd (IC/E)	Event	4th (IC/E)	Event	5th (IC/E)
e_1	3.62/3.87	e_{25}	3.96/2.56	e_{35}	2.90/2.50	e_{59}	1.24/2.38	e_{69}	1.05/2.20
e_2	1.91/3.60	e_{26}	2.44/2.76	e_{36}	2.88/1.80	e_{60}	1.05/2.29	e_{70}	1.04/2.20
e_3	2.99/3.37	e_{27}	1.66/2.62	e_{37}	2.05/2.42	e_{61}	1.41/2.29	e_{71}	1.43/2.30
e_4	0.56/1.62	e_{28}	0.51/1.66	e_{38}	0.44/1.52	e_{62}	0.37/1.38	e_{72}	0.33/1.27

Table 6.2: Information content (IC) and entropy (E) obtained for the viewpoint `cpitch`, assigned to each event in the pattern p_2 in the Turkish song *Cumhuriyet Çocuklarıyız*. The IC values and entropy values in this pattern decrease after its first appearance in each further repetition.

Event	1st (IC/E)	Event	2nd (IC/E)	Event	3rd (IC/E)	Event	4th (IC/E)	Event	5th (IC/E)
e_5	1.42/2.13	e_{29}	1.20/2.09	e_{39}	0.63/1.62	e_{63}	0.62/1.65	e_{73}	0.45/1.37
e_6	6.89/1.51	e_{30}	2.03/2.44	e_{40}	1.85/2.05	e_{64}	0.69/1.82	e_{74}	0.57/1.68
e_7	2.35/3.02	e_{31}	1.74/2.47	e_{41}	0.97/2.12	e_{65}	0.73/1.87	e_{75}	0.59/1.68
e_8	2.16/2.79	e_{32}	1.09/2.35	e_{42}	0.74/1.96	e_{66}	0.59/1.74	e_{76}	0.49/1.57
e_9	1.81/2.14	e_{33}	0.98/2.07	e_{43}	0.68/1.78	e_{67}	0.54/1.58	e_{77}	0.45/1.42
e_{10}	6.16/2.45	e_{34}	1.67/2.54	e_{44}	0.89/2.10	e_{68}	0.66/1.83	e_{78}	0.53/1.63

Table 6.3: Information content (IC) and entropy (E) assigned to each event obtained for the viewpoint `cpitch`, in the pattern p_3 in the Turkish song *Cumhuriyet Çocuklarıyız*. The IC values in this pattern (exception is the second repetition compared to the first appearance) decreases, as in pattern p_2, with each new repetition.

.

Event	1st (IC/E)	Event	2nd (IC/E)	Event	3rd (IC/E)	Event	4th (IC/E)
e_{11}	4.13/2.94	e_{18}	4.78/2.54	e_{45}	2.03/2.41	e_{52}	1.56/2.46
e_{12}	1.65/2.67	e_{19}	1.81/2.72	e_{46}	1.03/2.19	e_{53}	0.92/2.17
e_{13}	1.51/2.33	e_{20}	1.28/2.45	e_{47}	0.58/1.69	e_{54}	0.43/1.45
e_{14}	2.16/1.97	e_{21}	1.46/2.34	e_{48}	0.84/1.98	e_{55}	0.60/1.75
e_{15}	1.63/2.04	e_{22}	1.37/2.04	e_{49}	0.80/1.82	e_{56}	0.63/1.70
e_{16}	3.74/2.06	e_{23}	2.28/2.39	e_{50}	0.89/2.06	e_{57}	0.64/1.79
e_{17}	2.53/2.54	e_{24}	1.53/2.52	e_{51}	0.88/2.08	e_{58}	0.67/1.83

IR_REG as "unclassified". It consists of 48 events with three distinct patterns (p_1, p_2, p_3), and three relative repeated patterns, of which the pattern t_1 is a relative repetition of pattern p_1, t_2 a relative repetition of pattern p_2, and pattern t_3, a relative repetition of pattern p_3. As can be seen in Figures 6.3 and 6.6, the information content of the patterns t_1,t_2, and t_3 suggests that the content of these patterns is perceived as unexpected, even though they are a (relative) repetition of the three main patterns.

Table 6.4 presents the mean values for information content and entropy (the mean value of all the events in a particular song), for each viewpoint (ten variables) and song separately. The mean values of IC and E are taken as a measure of (ir)regularity: songs with a high IC and E, are classified as more irregular than those with low values.

Figure 6.2: An example of a structure with weak relationships between the patterns (motifs). This Swiss children's folk song *Schneeglöggli* has even 11 motifs, and none of them is dominant or repeated, therefore giving the impression of an irregular, non-structured melody.

Table 6.4: Average values for information content and entropy, for each viewpoint (ten variables in total) used in the observation of the three presented songs.

Viewpoint	Cumhuriyet Çocuklarıyız (regular)	Doidas andam as galinhas (unclassified)	Schneeglöggli (irregular)
IC_cpitch	1.36	2.46	4.40
E_cpitch	1.99	2.49	3.37
IC_cpint	1.70	2.12	3.73
E_cpint	2.12	2.41	3.03
IC_cpintfref	2.05	3.15	3.83
E_cpintfref	2.61	3.18	3.68
IC_cpitch⊗dur	1.46	2.81	4.80
E_cpitch⊗dur	1.78	2.41	3.24
IC_contour	4.15	4.02	4.96
E_contour	3.68	3.72	3.90

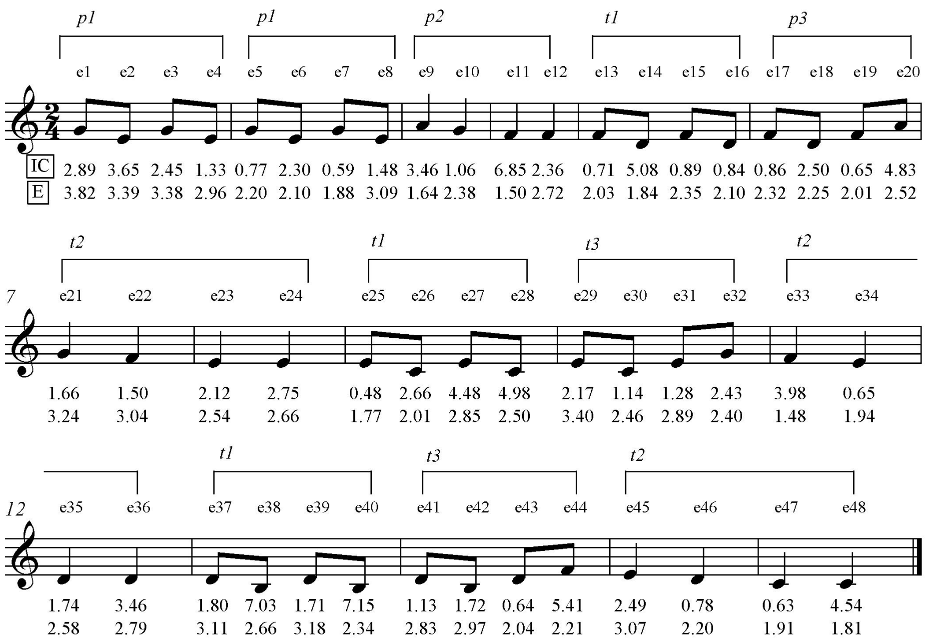


Figure 6.3: An example with relative repeated patterns in the Portuguese children's folk song *Doidas andam as galinhas.*

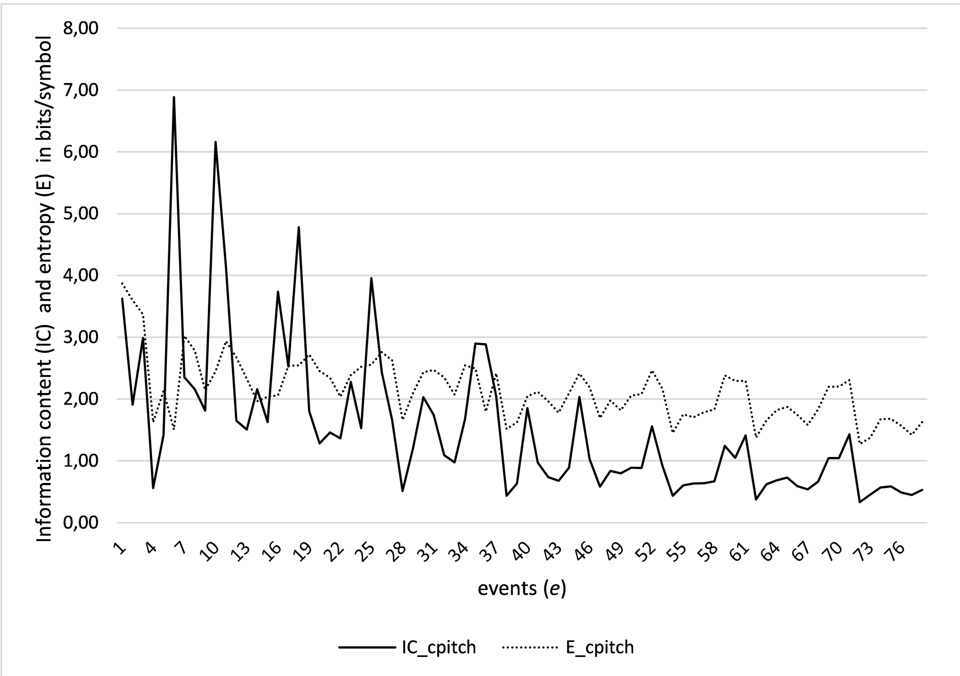


Figure 6.4: Information content (IC) and entropy (E) for the viewpoint cpitch in the Turkish children's folk song *Cumhuriyet Çocuklarıyız.* Both values decrease through this song.The peaks in the graph (event $e6$, $e10$, $e18$, and $e25$), are indicating low-probable notes.

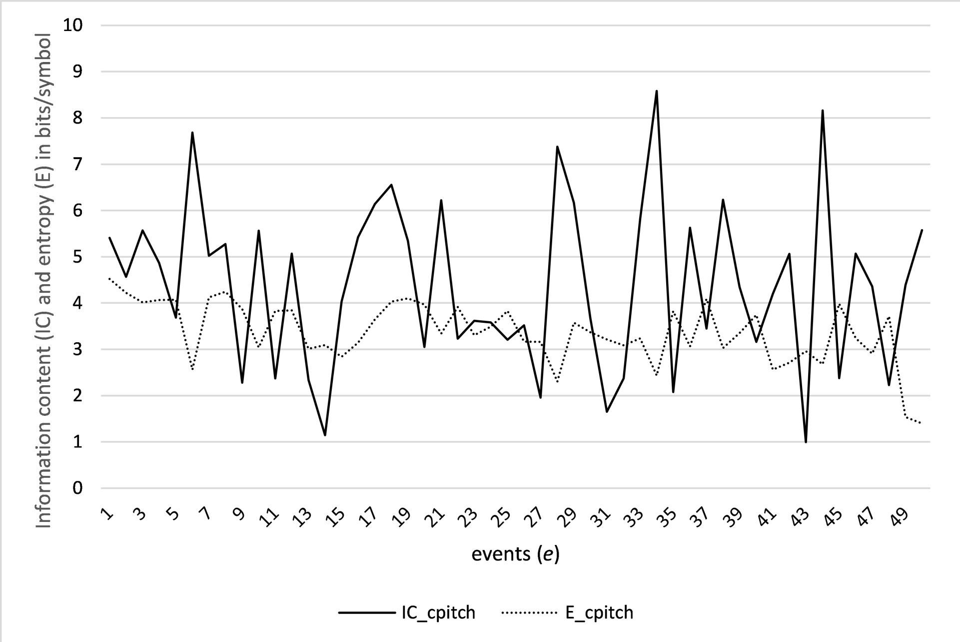


Figure 6.5: Information content (IC) and entropy (E) obtained for the viewpoint `cpitch` in the Swiss children's song *Schneeglöggli*. The entropy values (E) are slightly decreasing. The values for the information content remain very high, showing that the events in this song are highly unexpected.

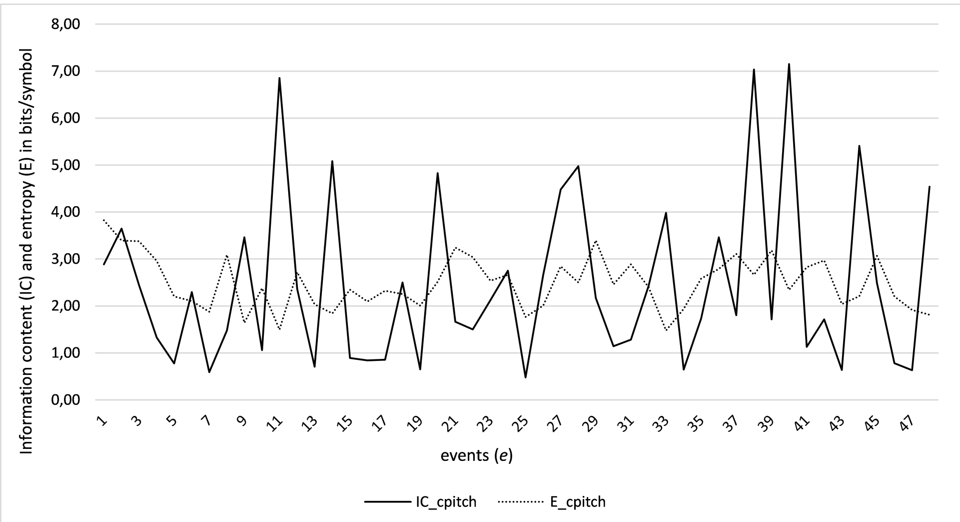


Figure 6.6: Information content (IC) and entropy (E) obtained for the viewpoint cpitch in the Portugal children's song *Doidas andam as galinhas*.

6.3 The (Ir)regularity in Children's Folk Songs

Hypotheses

This chapter tests two hypotheses:

(i) that the frequency of repeating patterns correlates with the irregularity of musical structure in children's folk songs, and

(ii) that repeated patterns contribute to regularity when presented at the same pitch.

Methodology

Stimuli

The Corpus II (see Subsection 2.2) consisting of monophonic 736 children's folk songs from 22 European countries was used to simulate and identify (ir)regularities in the musical structure.

Procedure

IDyOM and ten viewpoints (`IC_cpitch, E_cpitch, IC_cpint, E_cpint, IC_cpintfref, E_cpintfref, IC_cpitch⊗dur, E_cpitch⊗dur, IC_contour, E_contour`) were utilized to simulate the (ir)regularity in Corpus II. In addition, the algorithm Ir_Reg was used to examine the horizontal and vertical (ir)regularity in children's folk songs.

Results

To examine the (ir)regularity of the musical structure of children's folk songs, the full data set (736 children's folk songs from Corpus II) was analyzed by assessing the information content and entropy with the algorithm Ir_Reg on the set of viewpoints V. IDyOM calculated the mean values of `IC_cpitch, E_cpitch, IC_cpint, E_cpint, IC_cpintfref, E_cpintfref, IC_cpitch⊗dur, E_cpitch⊗dur, IC_contour`, and `E_contour` for each song from D. Next, the songs were classified as regular, irregular, or unclassified based on the global threshold values (i.e., the values computed across D), and the numbers of songs classified as regular, irregular, or unclassified for each country were counted. These results are reported in Tables 6.5 and 6.6, and visualized in Figure 6.7.

Regarding the relative frequencies, the most irregular in the structure are the songs from Norway, since 17 out of 23 (74 %) are classified as irregular. The most regular songs are the songs from Serbia, where 11 songs out of 13 (85 %) were labelled as regular. Other countries with no regular songs in our dataset are Switzerland and United Kingdom, while the countries with no irregular songs are Croatia, Hungary, Portugal, Romania, and Turkey.

Principal component analysis (PCA) was also performed using the same variables. Figure 6.8 shows irregular, regular and unclassified children's folk songs, projected to the 2-dimensional subspace spanned by the first two principal components, of which the first explains 54,7 % variance, and the second explains 18.4 % of total variance. We can actually observe that these three groups could be well separated also using only the first principal component. Table 6.7 contains data about importance of the principal components. We can see that we need five principal components to explain 94 % of variability (conventional goal in PCA is 95 %), which reveals that the data has five important dimensions.

Using the same algorithm on data sets of songs from each country, the information content and entropy were also used to examine the (ir)regularity of the musical structure in data from each country separately. Thus, the algorithm Ir_Reg was performed separately

Table 6.5: Number of irregular, regular, and unclassified children's folk songs found in 736 songs.

Country	Irregular	Regular	Unclassified	Country	Irregular	Regular	Unclassified
Bulgaria	2	3	13	Croatia	0	12	4
Denmark	4	2	9	France	16	9	46
Germany	16	28	80	Great Britain	23	0	15
Greece	3	10	13	Hungary	0	12	15
Italy	3	6	13	Latvia	2	4	17
Netherlands	6	14	38	Norway	17	0	6
Poland	1	5	16	Portugal	0	8	19
Romania	0	5	13	Russia	4	5	12
Serbia	0	11	2	Slovenia	8	7	29
Spain	16	13	25	Sweden	17	2	10
Switzerland	18	0	5	Turkey	0	16	8

Table 6.6: Upper (75%) and lower (25%) threshold for each viewpoint, and mean values for all 10 viewpoints, separately for regular, irregular, and unclassified children's folk songs, computed over the entire dataset of 736 songs.

Viewpoint	Irregular (mean)	Regular (mean)	Unclassified (mean)	Treshold (25%)	Treshold (75%)
`IC_cpitch`	3.06	1.84	2.37	2.00	2.72
`E_cpitch`	2.79	2.41	2.61	2.45	2.75
`IC_cpint`	3.07	1.97	2.49	2.13	2.82
`E_cpint`	2.66	2.41	2.57	2.40	2.71
`IC_cpintfref`	3.45	2.61	3.04	2.71	3.73
`E_cpintfref`	3.39	3.15	3.30	3.05	3.57
`IC_cpitch`⊗`dur`	3.17	1.77	2.41	1.93	2.87
`E_cpitch`⊗`dur`	2.71	2.18	2.43	2.21	2.63
`IC_contour`	4.23	3.63	3.86	3.51	4.27
`E_contour`	3.60	3.48	3.50	3.32	3.73

for each country with the inputs D - for the entire dataset and $\bar{D}$ - for the dataset of songs from a specific country. More exactly, taken were the values of viewpoints computed over the entire data set D, but the threshold values were computed using only the viewpoints from the songs of a specific country $\bar{D}$. These threshold values served as the basis for categorizing songs from $\bar{D}$. This resulted in a new classification, which is displayed in Table 6.8. Using this method, 158 irregular and 173 regular children's folk songs were identified out of a total of 736.

Using the musical score, two additional and independent musical experts examined patterns (motifs) in 331 children's folk songs. Each pattern (p), absolute or relative (t), was annotated in the musical score first. Second, based on the frequency distribution of pattern occurrences in each song, the entropy of unigrams for each song was calculated (More details about computing the unigrams in: Mihelač & Povh, 2020b).

Table 6.7: Results of Principal component analysis (PCA) of viewpoints computed over the complete dataset of 736 songs.

component	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Standard deviation	2.457	1.324	0.938	0.757	0.567	0.364	0.334	0.295	0.267	0.189
Proportion of Variance	0.604	0.175	0.088	0.05724	0.032	0.013	0.011	0.009	0.007	0.004
Cumulative Proportion	0.604	0.779	0.867	0.924	0.956	0.969	0.981	0.989	0.996	1.000

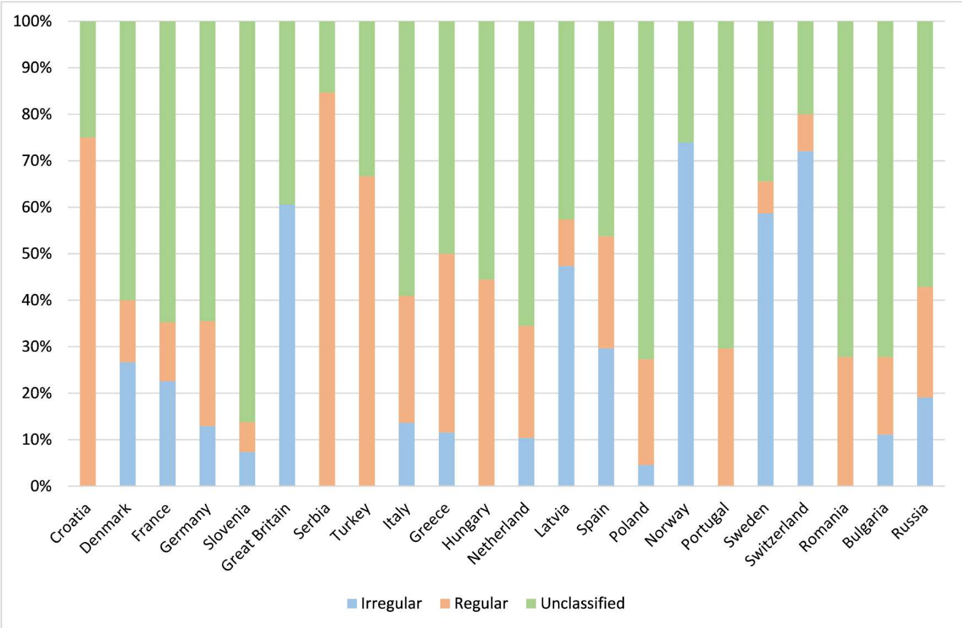


Figure 6.7: The percentage of irregular, regular, and unclassified children's folk songs in each country. Total number of children's folk songs is 736.

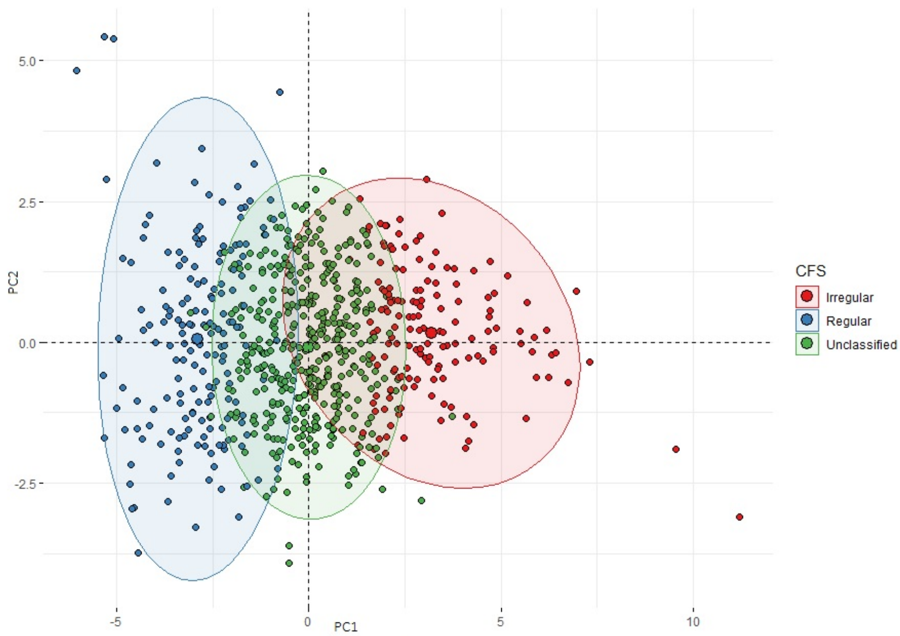


Figure 6.8: Visualization of irregular, regular, and unclassified children's folk songs with the first two principal components.

The Welch two-samples t-test was used to compare the difference in mean values of unigrams between the regular and irregular children's folk songs. The difference is statistically significant, ($p < .001$), i.e., the mean value of unigram was significantly smaller on the set of on regular children's folk songs, mainly because the repeated patterns at pitch p

Table 6.8: Number of irregular, regular, and unclassified children's folk songs found in 736 songs, examined in each country separately.

Country	Irregular	Regular	Unclassified	Country	Irregular	Regular	Unclassified
Bulgaria	5	3	10	Croatia	4	4	8
Denmark	3	4	8	France	15	16	40
Germany	29	27	68	Great Britain	7	10	21
Greece	5	8	13	Hungary	6	5	16
Italy	6	7	9	Latvia	3	4	16
Netherlands	12	16	30	Norway	5	4	14
Poland	3	5	14	Portugal	3	7	17
Romania	4	5	9	Russia	3	6	12
Serbia	3	3	7	Slovenia	11	10	23
Spain	11	11	32	Sweden	7	8	14
Switzerland	7	5	11	Turkey	6	5	13

were more recurrent in regular children's folk songs compared to irregular.

A chi-square test of independence was performed to examine the relation between relative repeated patterns t and the (ir)regularity in musical structure. Moreover, only the classified songs (regular and irregular) were considered, and a new binary variable with the value 1 was introduced if the song featured transposed repeated patterns and 0 otherwise. There was a significant relation between the two variables, $X^2(1,N{=}331) = 27.09$, $p < .001$, indicating that transposed repeated patterns t were more prevalent in irregular children's folk songs (80 out of 158) than in regular ones (only in 39 songs out from 173).

Discussion

According to the findings (see Table 6.5), children's folk songs with high or low irregularity were not found in all countries. A plausible explanation for a higher or lower presence of (ir)regular songs in some countries could be that children's folk songs exist in each of the 22 countries, whose origin can be traced to folk songs. Folk songs are a true amalgam of different historical, cultural, and musical processes (Golež Kaučič, 2003), thus carrying all the specificities of a (musical) culture from a particular country. This means that all the diversity found in folk songs, passed on to children due to their (suitable) content, is also present in children's folk songs, presumably without any simplification of the musical structure from the perspective of musical dimensions or musical elements. This could clarify either a very high or low information content and entropy found in children's folk songs in some countries.

The hypotheses, (i) that the frequency of repeating patterns correlates with the irregularity of musical structure in children's folk songs was confirmed, and (ii) that repeated patterns contribute to a stronger feeling of regularity if presented at same pitch were partially confirmed. According to the results, repeated patterns contribute to a more regular musical structure and to a stronger feeling of regularity, if the patterns are repeated at the same pitch, which is in accordance with the findings from Bader et al. (2017). Relative repeated patterns can be perceived as different, depending on the context which precedes or follow it (Margulis, 2014, p. 27–54). Musical structure is perceived as more regular, if the recurrence of a pattern is not "masked" (More about it in the study from: Bruner et al., 1959), as it was the case with the relative repeated patterns found in irregular and regular children's folk songs.

As children's folk songs are approximately 8–12 bars long, the conjecture is, that the perception and identification of relative repeated patterns is somehow affected by the duration of songs. If the children's folk songs were longer, the relative repeated patterns could be easier identified, and contribute more significantly to the regularity of the musical structure (Cambouropoulos, 2001; Deliège, 1987). Due to the duration of a children's folk song, the "mixture" of absolute and relative repeated patterns in a very short time period is actually interfering with the recurrent regularities, i.e., it is not contributing to a more regular structure, but rather to more noise, which is in accordance with the findings of Bruner et al. (1959).

The principal component analysis (PCA) performed on all the variables (`IC_cpitch, E_cpitch, IC_cpint, E_cpint, IC_cpintfref, E_cpintfref, IC_cpitch⊗dur, E_cpitch⊗dur, IC_contour, E_contour`), has shown that regular and irregular musical songs were well separated in the two dimensional space, spanned by the first two principal components. The salience of pitch in the perception of (ir)regular structure, can be understood, that each pitch, even when being acoustically identical to another one, can be perceived as different, depending on the context, as each pitch can differ in its tonal function, i.e., each pitch can have assigned a unique function in two different tonalities (C. L. Krumhansl, 1979; Mihelač & Povh, 2020a).

Small pitch intervals are prevailing in children's folk songs, as well as in a considerable number of folk songs across the world (Huron, 2001; von Hippel, 2000), of which some, included in the data, were found to exist as children's folk songs in different countries. Some of these folk/children's folk songs in the data were assigned as extremely irregular and complex (especially the songs from Great Britain, Norway, and Switzerland, according to the results), not only because of highly unexpected distributions of pitch in the melody, but also because of the use of large leaps (intervals) between successive notes. Thus, not only pitch, but also pitch intervals are affecting the regularity of a musical structure, which was found also in the study of Beauvois (2007).

Differences in rhythm (duration) were found between irregular and regular children's folk songs (see Table 6.6). The regularity in metrical structure, is according to different studies affected by the absence/presence of a regular beat, which groups different rhythms in time intervals (Bouwer et al., 2018). However, with the exception of some children's folk songs found in e.g., Bulgaria, Romania and Turkey, rhythm, based on recurring pulse/meter was found more or less in the majority of children's folk songs included in the data. Recurrent pulse/meter has obviously contributed to a more regular musical structure, in which the events are more predictable, which is also in accordance with the findings in the study of Lappe et al. (2013).

No significant impact on regularity was found using `IC_cpintfref` and `E_cpintfref` in the observation of (implied) harmony (Table 6.6). This result was expected, as basic harmonic progressions are prevalent in the melodies, either in irregular or regular children's folk songs (e.g., progressions I-V-I, I-IV-V-I, I-IV-I, ...). The use of dominant/subdominant functions and triads at the beginning and end of songs is to be understood more as a rare exception rather than a rule in children's folk songs. Furthermore, three basic harmonic functions (tonic, dominant, subdominant) are predominantly found in children's folk songs (Berget, 2017), suggesting that harmony, as a secondary parameter, is in this genre established and heavily dependable on the syntactical constraints and rules formed by the primary parameter pitch (Bauer, 2001; Meyer, 1989).

Contour has not been shown to affect the regularity of the musical structure in children's folk songs (Table 6.6). A plausible explanation could be that the information about contour is present in other viewpoints (for example in `cpitch` and `cpint`). Furthermore, in the majority of children's folk songs, ascending and descending directions of the melodies (arch-like structure) have been found. This is probably related to the origins of some children's folk songs traced to folk songs with an arch-like melody.

The same tendency of ascending-descending (convex) melodies has been found in a computerized analysis from Huron (1996) in 40 % of approximately 10,000 phrases (5–11 notes in length), and also in combined phrases (two phrases together, with a low midpoint, producing an overall convex shape) in over 6000 European folk songs. Furthermore, contour has proven to be a powerful identifying factor in melodic recognition, and to contribute to a better memorization of melodies (Bartlett & Dowling, 1980; Dowling, 1978; Trehub et al., 1984), which could also explain the prevailing arch-shaped contour in children's folk songs, which is more or less predictable and probably contributing to a higher feeling of regularity.

6.4 Summary

The purpose of this chapter was to demonstrate that irregularity is present in children's folk songs (Corpus II), that this genre can be complex, and considered as an alone standing genre. IDyOM and ten viewpoints were used to simulate the listener's perception of (ir)regularity in melodies and to identify patterns that contribute to a higher or lower (ir)regularity in this genre. Additionally, the algorithm Ir_Reg was utilized to assess the horizontal and vertical (ir)regularity of children's folk songs.

In this chapter, two hypotheses were tested: (i) that the frequency of repeating patterns correlates with the irregularity of musical structure in children's folk songs, and (ii) that relative repeated patterns contribute to a stronger sense of regularity if they are given at the same pitch. The hypotheses were partially confirmed. Repeated patterns were found more frequent in songs with a regular musical structure, while relative repeated patterns that did not use the same pitch were found to contribute to a stronger sense of irregularity.

In some countries, the absence of children's folk songs with (very) high or low irregularity suggests a plausible relationship between folk songs and children's folk songs, in which the content of folk songs has been transferred to children's folk songs without any simplification (or variation) of the musical structure. Future research with additional examples of children's folk songs in each country, and with more musical feature variables in defining (ir)regular songs could extend the explanations of (ir)regularity. An in-depth analysis of the origin of children's folk songs would certainly contribute to the understanding of (ir)regularities found in the musical structure of this genre, and why irregular structure is more frequently found in some European countries than in other European countries.

Rhythm, implied harmony, and contour have not shown to affect the regularity of the musical structure in this genre.

Chapter 7

Computational Cross-Cultural Study of (Dis)similarities in Musical Features and Dimensions Between and Within 22 European countries

Though music be a universal language,
it is spoken with all sorts of accents.

George Bernard Shaw

Poet Henry Wadsworth Longfellow once said, that "Music is the universal language of all mankind." Whether or not music is a universal language depends on how the terms "universal" and "language" are defined. As a tool for communication, language has meaningful symbols (words) that, when joined in a more complicated structure (following particular rules for combining — syntax), facilitate communication. Similarly, music as a specific "language" has its own symbols (notes) that may be combined into lower-order structures, (e.g. motifs, phrases, sentences/periods, etc.), and lower-order into higher-order structures (e.g. two-part, three-part songs…), using musical syntax which can differ in Western and Eastern music tradition. As opposed to language, none of the lower/higher-order musical elements has a meaning on its own, but acquires one in the context of complex musical structures, e.g., melody, (Ludden, 2015), which may (or may not) be understood in the same manner across cultures and even within cultures. Given this, it is difficult to claim that music is a "language" and "universal."

What is *universal* is the fact that music is widespread and may be found even in cultures where music is not referred to as "music" but rather as a musical activity (Mehr et al., 2019). However, even if music is viewed as a universality from this perspective, it differs not only between but also within musical cultures. Cross-cultural studies reveal that, despite more than a century of intensive research into the reasons why music differs between/within cultures, there is (still) no consensus (e.g., Brown & Jordania, 2011; Harwood, 1976; Henry, 1976; Higgins, 2012; List, 1971; Lomax, 1976; Savage et al., 2014; Savage et al., 2012; Serrà et al., 2012). Is it the comprehension and processing of music in an individual and/or society, the manner in which music has been passed over time, the social and cultural milieu that influences music production, or a combination of these factors? In addition, there are continuous discussions regarding what to seek for in music to examine the diversity between/within cultures, which genre (style of music) and which approach to use (manual, semi-automatic, automatic).

This chapter seeks to determine which musical features and dimensions contribute to the (dis)similarity between and within 22 European countries[1] by analyzing a sample of 2,184 homophonic folk songs, children's folk songs, and children's songs. Some of the selected countries (see Subsection 2.3) share a common cultural, historical, political, musicological, and sociological background; therefore, we could agree to define European folk music as a "single corpus of musical style" (Nettl & Béhague, 1980, p. 37). However, in spite of the fact that each culture shares some elements of its music with another, particularly with a neighboring country, each country possesses important and specific qualities such as pitch, interval, contour patterns, etc. (Savage et al., 2012).

Emphasizing Western folk music tradition could be interpreted as limiting the potential to answer broader or more specific issues concerning the diversity and scope of human music. To date, there are no cross-cultural research studies that compare folk songs, children's folk songs, and children's songs from the Western musical tradition. A thorough comparison of countries using these three genres could illuminate the impact of variability on the (dis)similarity of countries and reveal how children's folk songs and children's songs as musical genres contribute to the cultural identity of a country.

According to Nettl and Béhague (1980), music from different cultures should be analyzed from two perspectives: its style and structure, and its cultural context. The structure of a piece of music reveals how musical features and dimensions are utilized within a culture (Mihelač & Povh, 2021), and thus reflects (as a "material artefact") a specific culture. Nonetheless, the society, whether on a micro (as a person) or macro (as a population) level within a territory or region, processes music. By structuring musical features and dimensions in musical artifacts *dynamically* across time, a society influences the dynamic and ongoing transformation of a culture based on shared or individual beliefs (Wesch, 2018).

Thus, the existence of music should not be understood as existence "per se," but rather as an essential component of society and as the totality of human behavior within a culture. Consequently, a multiple approach is utilized in this chapter, which combines the analysis of the structure of music, the cultural/political/historical implications on music of a specific country, and the impact of the society at micro and macro level on the processing of music, by simulating the listeners' perception of music and their enculturation using the computational model IDyOM.

In this chapter, two hypotheses are tested, that there are substantial variances in the use of musical features and dimensions between European countries that are regarded to share a single musical style, and that the musical features and dimensions used in the representative music of a certain country are more similar in countries that share a similar cultural, political, historical, economic background, and are geographically close.

The structure of this chapter is as follows. Exposition of the related literature and theoretical foundations for the current research is presented in Section 7.1. A short portrayal of folk songs, children's folk songs and children's songs is provided in Section 7.2. The investigation of (dis)similarities between countries is presented in Section 7.3, followed by the investigation of (dis)similarities within countries in Section 7.4. Concluding remarks and future work are presented in Section 7.5.

7.1 Scientific Background and Related Work

Despite the absence of archaeological evidence that could shed light on the origins of music, the truth remains that it has not developed in a single culture, but simultaneously in numerous cultures around the world (Mehr et al., 2019; Peretz, 2006). Throughout

[1] Although culture is not synonymous with country or continent, and most people do not belong to a particular culture, the terms "culture" and "country" will be used interchangeably in this work.

human history and throughout cultures, music has been discovered in every society (Mehr et al., 2019), making it a global *art form.* Consequently, if music is universal, why is there musical diversity between cultures throughout the world and even within a culture? Several plausible explanations have been proposed, with the *cultural* and *biological* explanations appearing to be the most prevalent in recent decades. Music can be regarded a biological production (Peretz, 2006), since people are biological entities and everything generated by the human mind can be considered as biological. Various studies demonstrate the existence of similar universals in the creation and processing of music by humans (e.g., Drake & Bertrand, 2001; Trehub, 2000; Zatorre, 2001), while also highlighting the existence of biases and limitations in music processing and production (e.g., Gingras et al., 2015; Honing et al., 2015). From this perspective, considering the differences in human perception, production, and appreciation, it is expected to find cultural diversity, as it is extremely unlikely to find exactly the same processing, appreciation, or creation of music within or between cultures.

As a social construct, music *varies* not only between cultures (Blacking, 1990; Higgins, 2012; Lomax, 1977), but even within a specific culture (Henry, 1976; Rzeszutek et al., 2012). Each culture interacts (more or less) with the outside world and experiences various external and internal pressures (e.g., with subgroups with its own tastes and preferences). The function and presentation of music (within or between cultures) can be identical to various degrees, yet the *ideas* of musical sounds and *forms* are different (Friedmann, 1980). Due to the fact that the fundamental components of music, such as timbre, pitch, and rhythm, vary in frequency, form, and how they are perceived across and within cultures, it is important to also consider the behavior that is connected to music when examining the diversity in music that can be found in different cultures around the world.

The findings of the study by Lumaca et al. (2018) demonstrate that the cultural and biological explanations for the diversity of music are insufficient when viewed independently, as music is neither a purely cultural nor biological product. Music (as language) can be defined as a system with structured symbols and syntax, as well as a collection of behaviors that are transmitted from one generation to the next (Le Bomin et al., 2016; Morley, 2013), horizontally or vertically (Savage, 2019). Which cultural properties are transmitted to the next generation depends on the "memory bottleneck" (Deacon, 1997), i.e., the "human neurobiological filter" (Lumaca et al., 2018, p. 3), a collection of all the limitations (e.g., motoric, motoric-expressive, physiological, cross-modal, and semantic) that may affect musical structures. Information is more likely to be transmitted to the next generation when (i) it is easier to encode, understand, process, and memorize, and (ii) it is useful and attachable (Gladwell, 2002). As each individual is unique, a "neuronal niche" (Dehaene & Cohen, 2007, p. 384), it follows that there is a high possibility of inter-individual variances in the processing of music within a group. Eventually, the neurobiological heterogeneity of individuals should show at the population level, as variances in musical behavior can have substantial effects on musical systems, especially if magnified by cultural transmission (Lumaca et al., 2018).

Multiple studies have demonstrated that music, culture, and society are intricately intertwined, and that cultural context influences (to a greater or lesser extent) music created within cultural boundaries (Brenna, 1992; Lomax, 1976; Lundquiest & Szego, 1998; Merriam, 1960; Nettl, 1992). From this perspective, the best way to analyze (dis)similarities between/within cultures is through the use of music, which best *reflects a culture.* However, a song produced lately in a "traditional manner" (folk revival) might be considered to "represent" a culture, as can a song whose roots date back more than 300 years. Art music, children's folk songs, and children's songs, of which the latter two are generally neglected as a genre (Jožef-Beg & Mihelač, 2019; Mihelač, 2021; Mihelač & Panić Grazio, 2021), also reflect a particular culture, since many composers have included traditional folk components into their works.

The review of cross-cultural studies reveals that (folk) songs are predominantly used. Mehr et al. (2019) uses songs (with lyrics) to examine (non)universality in societies around the globe. Their research demonstrates that songs (as musical forms) are a "human universal," as songs can be found all over the world, that songs are associated with similar behavior in different societies, and that songs from different societies share certain musical characteristics, such as tone, pitch, and rhythm. Songs are also utilized by Lomax (1976) to demonstrate that songs identify and depict the fundamental social systems. Due to the wide variety of instruments, acoustic peculiarities, tuning systems, and production techniques, songs are also used by Savage et al. (2012) since they are simpler to compare cross-culturally than instrumental music (Ellis, 1885).

Unlike instruments, there are no examples of songs that have been fossilized. The hypothesis, however, is that even before the advent of musical instruments, a song-like communication mechanism, a "protolanguage" (Fitch, 2004), preceded human language (e.g., Darwin, 1871; List, 1973; Marler, 1976; Mithen, 2005), and that modern music is a sort of "behavioral fossil" of this communication system. The use of a song to analyze (dis)similarity between/within civilizations is logical and acceptable when seen from this perspective and in light of the fact that songs are a universal human trait prevalent in all cultures.

In cross-cultural studies, corpora consisting of folk (traditional) music are mostly used, while "non-representative" and "weird" corpora are often avoided. However, Heinrich et al. (2010) show strong evidence that "weird societies" (compared to other societies) are not just outliers, but "may represent the worst population on which to base our understanding of Homo sapiens" (Heinrich et al., 2010, p. 80), i.e., that even (supposedly) non-representative societies contribute to the study of human music (Savage, 2018), which may also be true for (apparently) "non-representative" corpora consisting of music from subcultures (such as the children's society).

According to a vast number of studies, the search for musical universals to explain cultural diversity has been and continues to be a topic of the utmost interest for decades (e.g., Brown & Jordania, 2011; Harwood, 1976; Lomax, 1976; Savage et al., 2012, ...). Lomax's search for a "typical" song, a "modal profile" for each culture by employing musical universals has induced countless discussions about the (non)existence of musical universals (Harwood, 1976; Justus & Bharucha, 1998; Meyer, 1998), as well as similar studies from Brown and Jordania (2011), proposing a detailed list of 70 putative musical universals, and later Savage et al. (2015), Savage et al. (2012), revealing that there are no absolute music universals between cultures, but only *statistical* ones.

These results are consistent with Steven's study, in which he explains that the same musical components, assumed to be musical universals and found in musical structures around the world, imply "static features of static environments" and neglect "informative differences" (C. J. Stevens, 2012, p. 654). Rather than searching for generalities, a more logical approach could be to search for particular instances of *diversity* at different levels, i.e. to search for "anti-patterns," patterns that are rare or absent in a dataset compared to patterns to be found as frequent (Conklin, 2013), or to recognize the musical *processing* over musical features and content (Harwood, 1976), a complementary approach.

Depending on the task, e.g., if the task is a computational modelling of melodic similarity, processing of musical features, or simply a classification problem, where only a few melodic features may be sufficient, the (dis)similarity in music between/within cultures is examined using audio recordings or/and notation, a manual approach (e.g., Le Bomin et al., 2016; Lomax, 1976; Rzeszutek et al., 2012; Savage et al., 2015; Volk et al., 2008), a semi-computational approach (e.g., Bronson, 1949; Rhodes, 1965), and, beginning around

1960, a computational approach (e.g., Gómez et al., 2009; Juhász, 2006, 2009; M. T. Pearce, 2018; Serrà et al., 2012; Tzanetakis et al., 2007).

7.2 Short Portrayal of Folk Songs, Children's Folk Songs and Children's Songs

Folk Songs

The definition and acceptance of the folk song have been the subject of numerous doubts and speculative discussions throughout history, ranging from its denial of existence, which is related to the denial of the folk culture by some ethnologists (Kumer, 1988), to the endless discussions about the most appropriate term, opposing the use of the term "folk song" and introducing other, more appropriate terms, such as "national song", "group song", "traditional song" or even "creation of the people" (Mihelač, 2008, 2012).

Folk song is defined in this thesis as a song made by an unknown, gifted member of a non-literate, rural part of a society (Bohlman, 1988), where the song is kept, transferred, and altered through an oral tradition. It is simultaneously a product of the individual *and* the society, as many people are continually making changes. This process can be termed as "communal re-creation" (Nettl & Béhague, 1980) owing to the song's constant refinement to sound like previously heard music or due to the fact that certain parts of the song have been forgotten.

As folk song (sung or played) is traditionally transmitted orally, thus not in a written form as the newly created folk songs (Kumer, 1988), and as notation is not the traditional "medium" of folk songs, we can expect variations of the same folk song, depending on all the changes the song may be exposed to, as well as the communal re-creation (e.g., transformations in a society, wars, compulsory schooling, cultural migrations, ...). Depending on the isolation of a particular culture, some songs may have been unaffected by these changes and passed in their original form (from the perspective of structure and performance) from one generation to the next, by individuals with "good memory" (Ling, 1997), to the present day, and then finally archived as recordings and/or transcriptions.

Despite the fact that various studies demonstrate that each culture's folk music has its own distinctive expression (e.g., Jing, 2017; Lomax, 1976), it is predicted that each culture's folk music shares features (musical universals) with other cultures (Nettl & Béhague, 1980). This is evident in cultures with a shared political/historical/economical/cultural past (e.g., the Habsburg Monarchy: the territories and provinces controlled by the junior Austrian branch of House of Habsburg between 1526 and 1780) (Kann, 1974).

A folk song differs from the music of trained musicians in terms of its structure and syntax. It follows the imagination of an uneducated composer, and the musical-aesthetic sensibility of the composer's society, as described by Stockmann (1985) and Vodušek (1980). In general, the structure of folk songs is simple. The structural forms AABB, ABBA, and ABCA are quite common (particularly in European folk music), and melodies are repeated multiple times with varied words. Repetition is also evident in the repetition of rhythmic patterns and identical motifs at varying pitch levels. The simplistic tonal and form structure, as well as the frequent repeating of lower as well as higher structural levels in songs, can be explained by a lack of formal musical training as well as by the usage of the musical material as a type of "mnemonic device" (Nettl & Béhague, 1980).

Children's Folk Songs

John Blacking states that music-making is "an inherited biological predisposition which is unique to the human species" (Blacking, 1973, p. 7), and that it emerges early in children's and adults' lives (Trehub, 2000). Despite the fact that music "happens" to children all the time (P. S. Campbell, 1998), manifesting itself (among other ways) in singing tunes during various activities, it appears that adults in the past frequently failed to acknowledge it. Due to this, many traditional songs and contributions by and for children that could contribute to a deeper knowledge of children's songs across countries may not have been gathered.

The collections of children's songs from around the world demonstrate that these collections contain songs written by adults specifically for children, folk songs, and songs that children have produced and passed down to one another (Mihelač, 2021; Mihelač & Panić Grazio, 2021). Summarizing all of these songs, it can be very challenging to identify their original creators because children's songs frequently undergo changes and adaptations over the course of generations, or are even published by adults due to children's lack of musical training and inability to correctly notate songs. [2] Folk songs (which can be found among children's songs) have a similar issue because it is unclear whether they were written by children or by adults for children. The only exception are lullabies, which are songs written by adults for children and are found in almost every culture (Jožef-Beg & Mihelač, 2019).

Regardless of who composed these songs, a number of them share many characteristics with folk songs and are passed down orally from generation to generation, much as folk songs are. These children's songs, which could be defined as "children folk songs" for which there is no clear definition (Sutton-Smith, 1999), are tonal and structurally simple, even more so than folk songs, syllabic, free of ornamentation, and full of repetition, with a content close to the children's world, likely to make them as adaptable as possible for children (Nograšek & Virant Iršič, 2005; Pond, 1981; Romet, 1980; Voglar & Nograšek, 2009).

Despite the fact that adults and children of a culture share the same cultural milieu, the question is how (dis)similar children's folk songs are to folk songs, even if traditional folk songs are considered "children folk songs." Children's societies are frequently conservative, ritualistic, and governed by routine, exhibiting moments of high fantasy and silly innovation, with their own "group traditions" (McDowell, 1999), altering songs created by themselves or by adults, repeatedly, based on their abilities, needs, games, and so on.

Nonetheless, it is anticipated that children's folk songs and folk songs would share many similarities, as children's societies reflect in their songs (regardless of who wrote them) the folklore of their family, neighborhood, and other groups and folk activities they participate in. This process can be more or less dynamic according to the age of the children, bearing in mind that the older the children are, the more external influences from outside their cultural milieu will be reflected in their songs (Ling, 1997).

Children's Songs

How is a children's song defined? Is it an adult's composition, a child's simple play song, or a commercial children's song? The most difficult aspect of categorizing children's songs is that this genre grew out of a changing conception of childhood in societal norms (Lawson, 2011). Childhood was not understood and recognized as a stage of life until the eighteenth century (Lowe, 2004), and it became an intriguing object of study at the end of the nineteenth century.

[2] Sometimes, the origin of children's songs can be documented. The documentary fieldwork of Charles L. Todd and Robert Sonkin, who identified child composers whose families had fled the Dust Bowl in the 1930s and subsequently resided in the Shafter Farm Security Administration (FSA) Camp in Shafter, California, is an example. The "Government camp song," composed by Betty and Mary Campbell, along with Margaret Treat, depicts what occurred in the lives of the younger children of migrant workers. More in (Todd & Sonkin, 1940–41)

Since 1940, ethnomusicologists, sociologists, educators, and folklorists have taken an increased interest in children's songs, focusing their studies on musical content, on the social and cultural significance of children's songs and their relationships to the music of adults, on how children's songs contribute to the preservation of a culture, on how children's songs are transmitted from one generation to the next, etc. (e.g., Blacking, 1973; Brailoiu, 1954; Herzog, 1944; Nettl & Béhague, 1980; Nettl, 1983; Newell, 1963; Pond, 1981). If children's culture was unimaginable for centuries, then the last several decades have witnessed a gradual increase in interest in this field (P. Campbell, 2010, p. 6). As music is significant to children and a "childhood constant," studies from ethnomusicologists, folklorists, educators, anthropologists, sociologists, and more recently also musicologists, are focusing more than ever on the relationship between music and children, covering a variety of topics such as children's musical behavior, musical activities, music-making, music performance, music perception, etc.

According to Kartomi (1999), children's songs written by children are worlds apart from those developed by adults for children, since the songs created by children follow different rules and have a unique approach to rhythm, structure, text, and substance, possessing a childlike aspect. Nonetheless, authors who have devoted themselves to composing entirely or primarily children's songs use the aforementioned traits, writing the children's songs in a way that children would comprehend and that is content- and ability-appropriate. These songs are based on words that come from the child's world, which is full of toys, plays, games, imagination, and activities that primarily occupy adults (e.g., cleaning, sawing, building...). From the standpoint of the musical construction, they have a strong rhyme, a multiverse text, a simple structure, syncopated rhythms, and unusual rhythm and pitch patterns (Jožef-Beg & Mihelač, 2019).

Children's songs are frequently influenced by folk elements and folk music of the composer's native country (e.g., use of intervals, motifs, melodic figure, and ancient tonal patterns), while globalization, a dynamic force in contemporary societies (Trask, 2010), influences (among other things) the creation of new children's songs (e.g., using contemporary cultural and social values, integrating folk elements with popular music).

Ethnomusicologists are divided as to whether or not children's songs reflect or contribute to the development of a culture. The opinion of Nettl is that music reflects culture (Nettl, 1983). Rogoff (2003) and Corsaro (2005) perceive the creation of a culture in a more active manner (e.g., the contribution of children to the creation of their cultures). Slobin (2000) describes the existence of multiple levels of culture, including super-cultures, sub-cultures, inter-cultures, and micro-cultures. For instance, children's musical cultures are sub-cultures inside the super-culture of a contemporary country, as well as micro-cultures and inter-cultures, due to the effect of musical cultures in a community that merge with global and childhood cultures.

7.3 (Dis)similarities Between Countries

Hypotheses

Two hypotheses are tested:

(i) There are substantial variances in the use of musical features and dimensions between European countries that are regarded to share a single musical style.

(ii) The musical features and dimensions used in the representative music of a certain country[3] are more similar in countries that share a similar cultural, political, historical, economic background, and are geographically close.

Methodology

Stimuli

Corpus III (see Subsection 2.3) was used to explore the variety between countries by analyzing in depth the use of musical features and dimensions. The decision to include Russia is based on geography (approximately 23 percent of western Russia can be considered a European part) and shared history between former and present Russia and Europe (Graney, 2019). Turkey's songs have been added to this data set because a small portion of the country is regarded to be in Europe (about 3 percent) and because the Ottoman Empire shared a historical, political, and cultural background with many European countries (Hurewitz, 1961; Kostopoulou, 2016). Some European countries, such as Austria and the Czech Republic, were omitted from Corpus III due to insufficient data for a particular genre, and the decision to include a proportional number of countries from West, East, North, and South Europe.

Procedure

In order to explain the (dis)similarities between the countries and genres using musical features and dimensions, IDyOM was utilized to simulate the listener's processing and perception of songs. Because empirical findings from numerous studies have shown that the information content is effective in capturing stylistic statistical patterns and psychological processes in listeners' music perception (Hansen & Pearce, 2014; Omigie et al., 2013; M. T. Pearce, 2005, 2018), the average information content (IC), as estimated by the model for each target viewpoint, was used.

The musical features and dimensions were observed using 16 viewpoints (see Table 7.2) at two different levels of granularity: low (song-based) and high (genre- and country-based).[4] Precisely, the following were computed for each of the 22 countries and each of the three genres:

- the mean values of all viewpoints for each song in order to simulate the listener's perception of each song and collect as much information as possible on the events in a particular song (the song based granularity)
- the mean values for all viewpoints for a group of chosen songs from a given genre (the genre based granularity)
- the mean values for all three genres combined from a particular country (the country based granularity)

Multivariate Analysis of Variance (MANOVA) was complemented with a series of independent univariate Analysis of Variance (ANOVA) models to study how viewpoints vary across genres and countries. Agglomerative hierarchical algorithms, Euclidean distance as a similarity measure, and the Ward agglomeration method (Ward, 1963) were utilized to find clusters in the multidimensional data.

[3]This thesis defines "representative music" as music that symbolizes the culture of a certain country, such as folk music, children's folk music, or any other music including musical features that are exclusive to that country.

[4]For a comprehensive list and description of each viewpoint, see the subsection 3.2.2.

Results

First, the hypothesis that there are substantial differences in the use of musical features and dimensions among European countries considered to have a single musical style was tested with Wilks' Λ statistic.[5] When 16 viewpoints were utilized to examine the musical features and dimensions, the following was found:

- There are significant differences between vectors of mean values of viewpoints computed across the countries (Wilk's $\Lambda = 0.223$, $F(336, 27473) = 10.18$, $p < 2.2\text{e-}16$).
- There are significant differences between vectors of mean values of viewpoints computed across the genres FS, CFS, CS (Wilk's $\Lambda = 0.502$, $F(32, 4332) = 55.63$, $p < 2.2\text{e-}16$).

Next, a one-way Multivariate Analysis of Variance (MANOVA) was performed to determine if there are statistically significant differences in the vectors of mean values of viewpoints across genres for each country. As a result, 22 MANOVAs were conducted, and the findings, which are presented in Table 7.1, indicate that there are significant differences between the mean values of viewpoints used to observe the musical features and dimensions in songs belonging to the FS, CFS, and CS song groups, for each country.

Table 7.1: Results of the series of MANOVA, conducted within each of the countries separately.

country	Wilks Λ	F-value	df_1	df_2	p-value
Bulgaria	0.074	6.843	32	82	0.000000
Croatia	0.043	17.170	32	144	0.000000
Denmark	0.076	4.925	32	60	0.000000
France	0.053	43.211	32	416	0.000000
Germany	0.045	61.988	32	532	0.000000
Great Britain	0.047	17.545	32	156	0.000000
Greece	0.331	2.307	32	100	0.000873
Hungary	0.028	18.283	32	118	0.000000
Italy	0.050	8.428	32	78	0.000000
Latvia	0.367	2.889	32	142	0.000009
Netherlands	0.071	19.991	32	232	0.000000
Norway	0.020	20.965	32	110	0.000000
Poland	0.075	8.136	32	98	0.000000
Portugal	0.073	6.581	32	78	0.000000
Romania	0.015	18.617	32	82	0.000000
Russia	0.042	12.550	32	104	0.000000
Serbia	0.071	7.208	32	84	0.000000
Slovenia	0.069	36.560	32	416	0.000000
Spain	0.056	14.166	32	140	0.000000
Sweden	0.124	6.688	32	116	0.000000
Switzerland	0.052	20.296	32	192	0.000000
Turkey	0.066	8.645	32	96	0.000000

ANOVAs were performed for each viewpoint to see which viewpoints have significantly different mean values across countries and genres. Table 7.2 contains the results. There are significant differences between the mean values of all viewpoints when the groups are defined by the countries and the genres, respectively, with the exception of the `tessitura` viewpoint, for which significant differences between genres cannot be confirmed.

Figure 7.1 illustrates how the mean values of viewpoints in different countries differ from the overall mean values. The boxplots display the mean values for each viewpoint across 22 countries as well as the lowest and highest mean values for each country.The fact that each viewpoint's intervals of values varies widely across the countries further proves

[5] Only the results for the Wilks' Λ statistic are reported. The other standard statistics (Hotelling-Lawley, Roy, Pillai) imply the same conclusions.

Table 7.2: Results of ANOVAs for each viewpoint. The columns 2–5 correspond to ANOVA tests across the countries and the four rightmost columns to the ANOVA tests across the genres. The first columns in both groups report the values of F statistic, the second and the third columns report the corresponding degrees of freedom and the fourth column in both groups reports the p-values.

viewpoint	F	df_1	df_2	p-value	F	df_1	df_2	p-value
cpitch	10.197	21	2162	0.000000	38.761	2	2181	0.000000
cpitch_class	34.802	21	2162	0.000000	181.110	2	2181	0.000000
tessitura	17.799	21	2162	0.000000	1.083	2	2181	**0.338810**
cpint	9.349	21	2162	0.000000	40.472	2	2181	0.000000
cpint_size	15.324	21	2162	0.000000	144.053	2	2181	0.000000
cpcint	26.319	21	2162	0.000000	122.552	2	2181	0.000000
cpcint_size	42.263	21	2162	0.000000	257.634	2	2181	0.000000
contour	18.666	21	2162	0.000000	353.957	2	2181	0.000000
newcontour	33.296	21	2162	0.000000	357.952	2	2181	0.000000
cpintfip	7.141	21	2162	0.000000	21.314	2	2181	0.000000
cpintfref	38.978	21	2162	0.000000	197.064	2	2181	0.000000
inscale	77.834	21	2162	0.000000	372.920	2	2181	0.000000
cpitch_dur	8.917	21	2162	0.000000	58.367	2	2181	0.000000
dur_ratio	13.226	21	2162	0.000000	40.492	2	2181	0.000000
ioi_ratio	7.736	21	2162	0.000000	55.663	2	2181	0.000000
cpitch_ioi_ratio	7.117	21	2162	0.000000	54.191	2	2181	0.000000

that the differences are indeed significant. The countries that differ most from the mean value, for each viewpoint, are also identified by the whiskers of each boxplot.

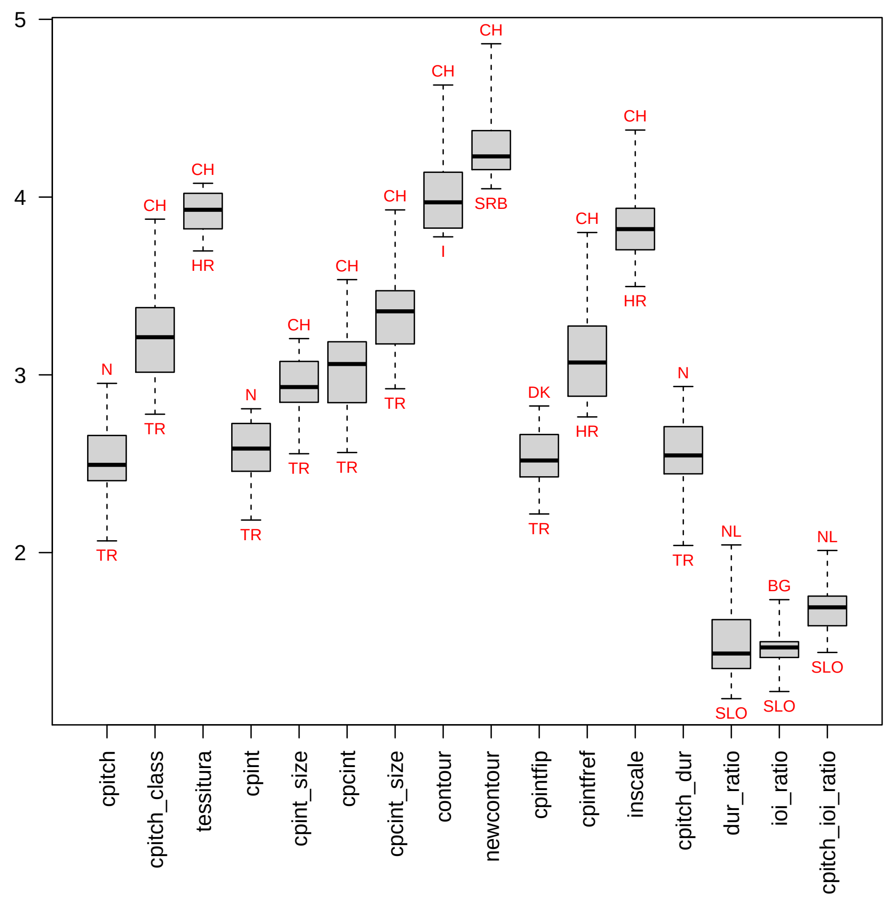


Figure 7.1: Boxplots for the countries' means values of viewpoints

The hypothesis that musical features and dimensions used in a country's national representative music are more similar in countries with similar cultural, political, historical, and economic backgrounds and that are geographically close to one another is tested in the continuation using clustering analysis. For each country is used:

(1) a vector of length 16, containing mean values of each viewpoint computed across the songs belonging to the set of FS, CFS, and CS songs

(2) a vector of length 48, obtained by stacking the three vectors of length 16, which are described in the previous item

The results of clusters are summarized in Table 7.3, and visualized in Figures 7.6–7.9.[6] Only a basic hierarchical clustering was performed, in which point-to-point dissimilarity is estimated using Euclidean distance and agglomeration is performed using the Ward method (Ward, 1963). Dendrograms were utilized to determine the optimal number of groups based on the results of this procedure.

Figure 7.2 is a dendrogram illustrating the bottom-up clustering procedure, where merging levels correspond to the measure of dissimilarity between two groups of countries. This dendrogram was obtained by representing the 22 countries with stacked FS, CFS, and CS vectors of viewpoints. The vertical dashed lines represent the most significant clusters in this dendrogram, which are comprised of the most related countries. Figures 7.3–7.5 depict the dendrograms for each genre (FS, CFS, CS) independently.

When employing the FS, CFS, and CS genres separately and together, significant differences between groupings of countries can be observed. Additionally, some geographically distant countries without a common political, cultural, economic, or historical background (such as the UK, Switzerland, and the Netherlands or Latvia and Slovenia) use musical features and dimensions similarly, while other geographically close countries with a common past use them differently.

[6] The countries in grey are those that were excluded from Corpus III.

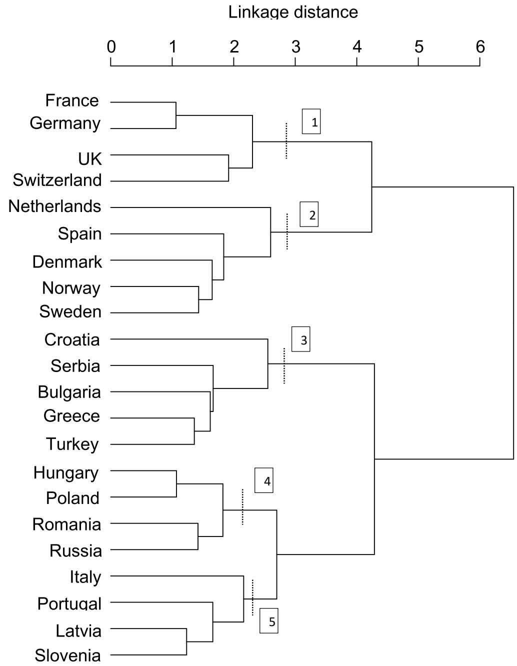


Figure 7.2: Dendrogram, based on stacked vectors of viewpoints.

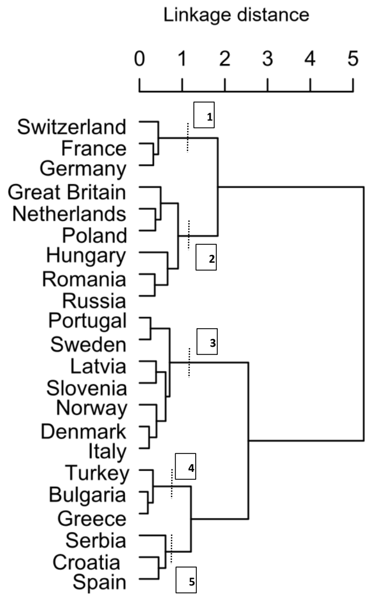


Figure 7.3: Dendrogram, based on the mean viewpoints computed over FS genre.

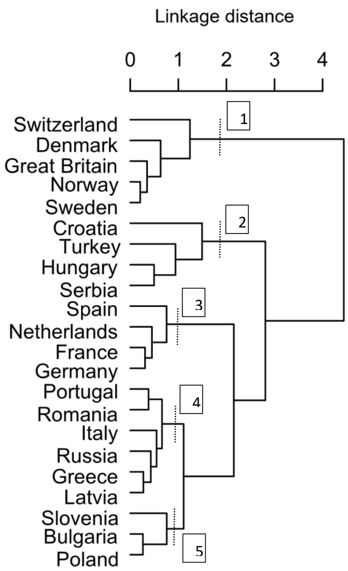


Figure 7.4: Dendrogram, based on the mean viewpoints computed over CFS genre.

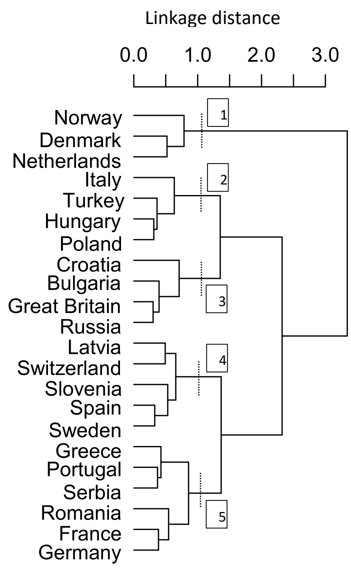


Figure 7.5: Dendrogram, based on the mean viewpoints computed over CS genre.

Table 7.3: 22 European countries grouped by genres.

All genres	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Latvia	Greece	Romania	Norway	Switzerland
	Portugal	Bulgaria	Russia	Sweden	Germany
	Italy	Serbia	Poland	Netherlands	France
	Slovenia	Turkey	Hungary	Denmark	UK
		Croatia		Spain	
Only FS	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Germany	Great Britain	Portugal	Turkey	Serbia
	France	Netherlands	Sweden	Bulgaria	Croatia
	Switzerland	Poland	Latvia	Greece	Spain
		Hungary	Slovenia		
		Romania	Norway		
		Russia	Denmark		
			Italy		
Only CFS	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Switzerland	Croatia	Spain	Portugal	Slovenia
	Denmark	Turkey	Netherlands	Romania	Bulgaria
	Great Britain	Hungary	France	Italy	Poland
	Norway	Serbia	Germany	Russia	
	Sweden			Greece	
				Latvia	
Only CS	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Norway	Italy	Croatia	Latvia	Greece
	Denmark	Turkey	Bulgaria	Switzerland	Portugal
	Netherlands	Hungary	Great Britain	Slovenia	Serbia
		Poland	Russia	Spain	Romania
				Sweden	France
					Germany

Figure 7.6: Clusters obtained by using all genres.

Figure 7.7: Clusters obtained by using FS only.

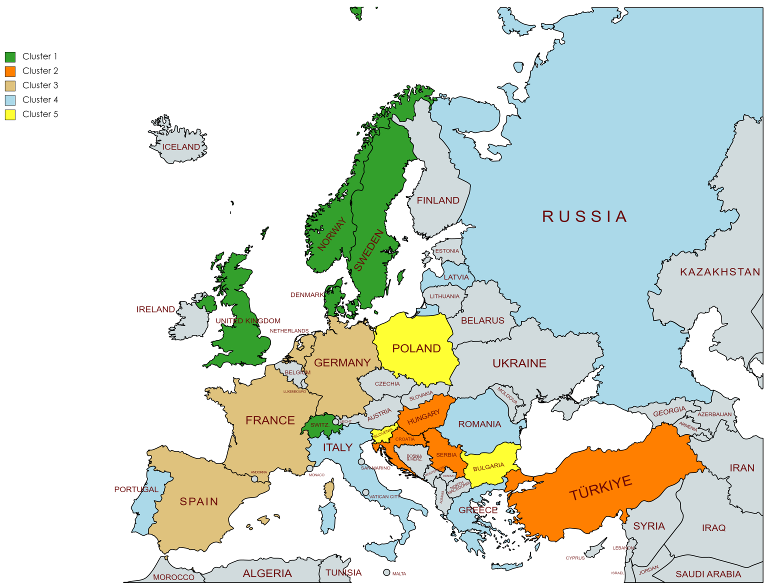


Figure 7.8: Clusters obtained by using CFS only.

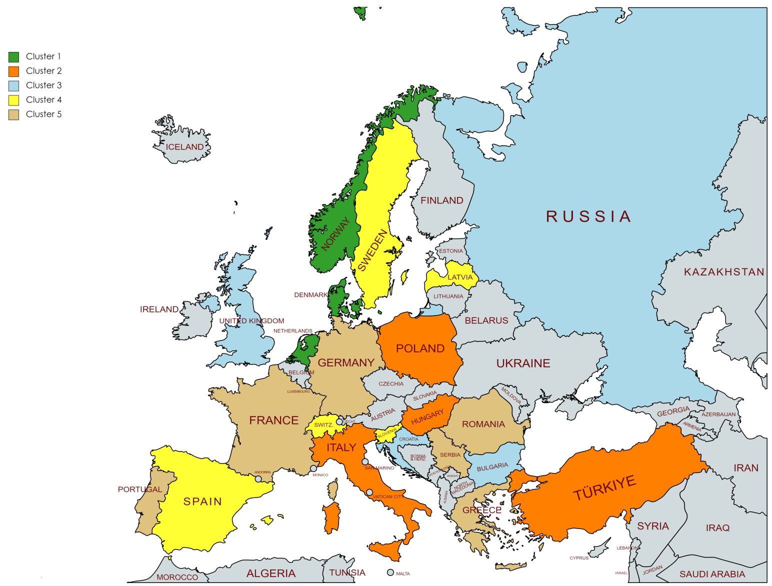


Figure 7.9: Clusters obtained by using CS only.

7.4 (Dis)similarities Within Countries

Hypotheses

One hypothesis is tested, that there are differences in the use of musical features and dimensions in genres considered to belong to the representative music of a particular country, however depending on cultural forces which homogenize and diversify these genres.

Methodology

Stimuli

In order to explore the variety of musical features and dimensions in genres within countries, the same Corpus III was employed.

Procedure

IDyOM and information content (IC) for the same 16 target viewpoints that were employed in the observation of musical features and dimensions between countries were also used within each of the 22 countries to simulate the listener's processing and perception of songs in each genre. To identify for which country and viewpoint there are significant differences in the mean values between genres, a series of ANOVA tests were conducted for each country and viewpoint across genres. MANOVA could not be applied because the samples for the countries were not sufficiently large.

Results

To test the hypothesis that there are substantial differences in the use of musical features and dimensions across genres within a country, a set of ANOVA tests were run for each country and for each viewpoint across the genres. The results (p-values) of ANOVAs are shown in Table 7.4 and in Figure 7.10. We can see that Greece and Denmark have the smallest differences in the mean values across genres (significant differences only for 2 and 4 viewpoints, respectively), but France, Germany, Hungary, The Netherlands, Portugal, and Romania have significant differences for at least 8 viewpoints, suggesting, (i) that there are significant dissimilarities in the way musical features and dimensions are understood and used in three genres within countries.

Similar results were discovered in the study conducted by Rzeszutek et al. (2012). Using only one genre (folk songs), significant differences were found in the use of musical features/dimensions within populations (countries), but smaller differences between populations (countries), validating and quantifying the criticisms of ethnomusicologists that the musical diversity within a particular culture is underestimated when searching for musical diversity in cultures (Feld, 1984; Henry, 1976).

In a recent research by Daikoku et al. (202) cross-cultural music similarity ratings were analyzed on a global song sample from 110 participants (from Japan, north and south India). Greater musical diversity was found within cultures than between them.

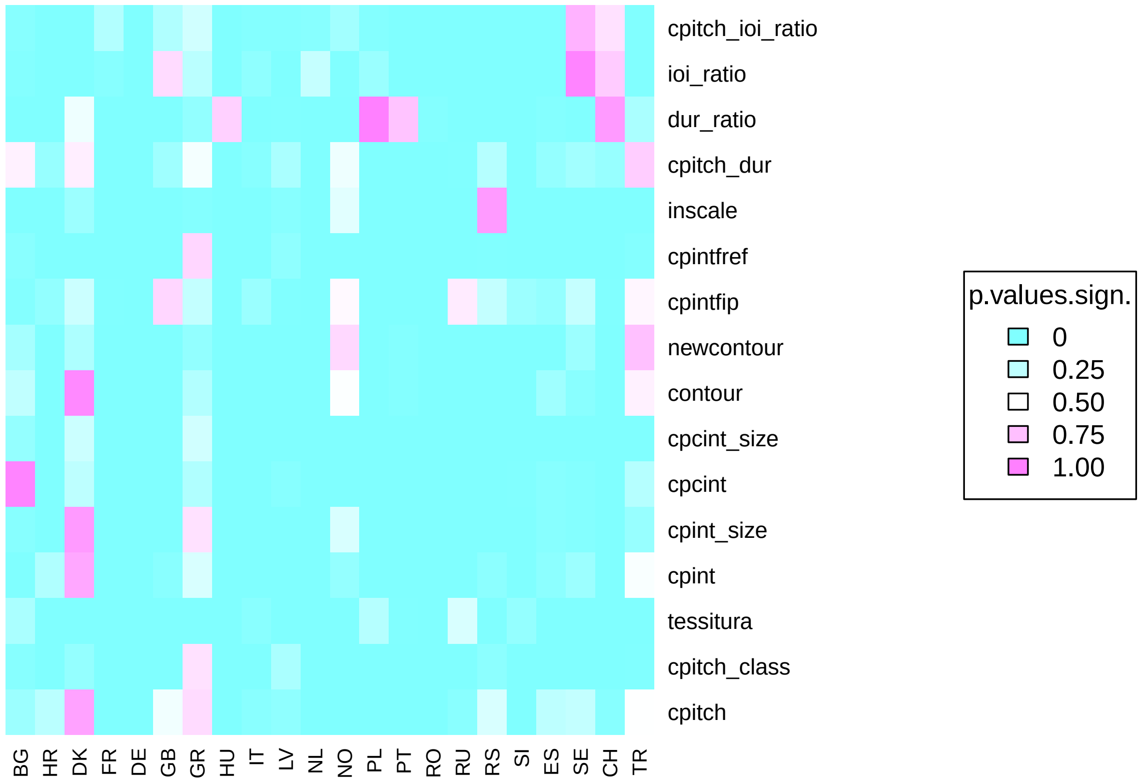


Figure 7.10: Heatmap showing differences in the use of musical features across genres within countries.

Table 7.4: Results (p-values) of ANOVAs, computed for each viewpoint, across the genres within each of the countries. No significance is presented with bold values.

viewpoint	BG	HR	DK	FR	DE	GB	GR	HU	IT	LV	NL	NO	PL	PT	RO	RU	RS	SI	ES	SE	CH	TR
cpitch	**0.11**	**0.22**	**0.81**	0.00	0.00	**0.42**	**0.60**	0.00	0.04	**0.06**	0.00	0.00	0.00	0.00	0.00	0.04	**0.33**	0.00	**0.23**	**0.25**	0.03	**0.47**
cpitch_class	0.03	0.00	**0.08**	0.00	0.00	0.00	**0.58**	0.00	0.00	**0.16**	0.00	0.00	0.00	0.00	0.00	0.00	**0.05**	0.00	0.00	0.00	0.00	0.01
tessitura	**0.16**	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.04	0.00	0.00	0.00	**0.20**	0.01	0.00	**0.33**	0.00	**0.08**	0.00	0.00	0.00	0.00
cpint	0.00	**0.18**	**0.79**	0.00	0.00	0.04	**0.33**	0.00	0.00	0.00	0.00	**0.08**	0.00	0.00	0.00	0.00	**0.05**	0.00	**0.05**	**0.11**	0.00	**0.45**
cpint_size	0.03	0.00	**0.84**	0.00	0.00	0.00	**0.58**	0.00	0.00	0.00	0.00	**0.33**	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.02	0.00	**0.09**
cpcint	**0.92**	0.00	**0.23**	0.00	0.00	0.00	**0.18**	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.02	0.00	**0.20**
cpcint_size	**0.08**	0.00	**0.28**	0.00	0.00	0.00	**0.30**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
contour	**0.24**	0.00	**0.90**	0.00	0.00	0.00	**0.19**	0.00	0.00	0.00	0.00	**0.46**	0.00	0.02	0.00	0.00	0.00	0.00	**0.12**	0.04	0.00	**0.52**
newcontour	**0.14**	0.00	**0.17**	0.00	0.00	0.00	**0.07**	0.00	0.00	0.00	0.00	**0.61**	0.00	0.02	0.00	0.00	0.00	0.00	0.00	**0.11**	0.00	**0.70**
cpintfip	0.02	**0.07**	**0.28**	0.01	0.00	**0.62**	**0.25**	0.00	**0.10**	0.01	0.00	**0.49**	0.00	0.00	0.00	**0.54**	**0.25**	**0.11**	**0.08**	**0.26**	0.01	**0.50**
cpintfref	0.04	0.00	0.00	0.00	0.00	0.00	**0.62**	0.00	0.00	**0.06**	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.02
inscale	0.00	0.00	**0.11**	0.00	0.00	0.00	0.02	0.00	0.00	0.03	0.00	**0.36**	0.00	0.00	0.00	0.00	**0.84**	0.00	0.00	0.00	0.00	0.00
cpitch_dur	**0.52**	**0.09**	**0.53**	0.00	0.00	**0.12**	**0.43**	0.00	0.03	**0.16**	0.00	**0.41**	0.00	0.00	0.00	0.01	**0.20**	0.00	**0.08**	**0.13**	**0.09**	**0.65**
dur_ratio	0.00	0.00	**0.41**	0.00	0.00	0.00	**0.07**	**0.64**	0.00	0.01	0.00	0.00	**0.94**	**0.69**	0.02	0.00	0.00	0.00	0.02	0.01	**0.84**	**0.16**
ioi_ratio	0.02	0.00	0.00	0.03	0.00	**0.60**	**0.22**	0.00	**0.06**	0.00	**0.26**	0.01	**0.10**	0.00	0.00	0.00	0.00	0.00	0.00	**0.92**	**0.66**	0.01
cpitch_ioi_ratio	0.03	0.00	0.00	**0.19**	0.00	**0.18**	**0.30**	0.00	0.02	0.02	0.03	**0.13**	0.02	0.00	0.00	0.00	0.00	0.00	0.00	**0.75**	**0.58**	0.01

Discussion

In this chapter, folk songs, children's folk songs, and children's songs were used to study the diversity of musical features and dimensions across and within 22 European countries. The results of the study of these three genres between and within 22 countries have revealed substantial differences, highlighting the distinctive manifestations of these genres and the manner in which musical features and dimensions are incorporated into the musical structure, confirming the first hypothesis, that European countries cannot be considered as a single musical style.

The fundamental question that arises is why similar relationships have been discovered between countries compared to other cross-cultural research studies that deal with only one genre, namely folk songs (e.g., Juhász, 2006; Panteli, 2018)? There is always the chance of genres "assimilating" or "merging," much as (sub)cultures adapt to a culture that exerts strong dominance over other cultures, and after a given period of time, behaviours/cultural variances within a subculture are driven towards fixity (Foley & Mirazón Larh, 2011; Newson et al., 2007). For instance, what is considered a folk song representative of a particular country may actually be a children's folk song, or even a children's song that was labelled a folk song after a certain period of time, perhaps due to insufficient attention to the origins and circumstances under which a song was created. This could explain the same outcomes in several cross-cultural research, regardless of whether only one genre (folk music) or multiple genres were employed.

From a historical, geographical, cultural, social, and musical perspective better results of hierarchical clustering were achieved by combining the three genres (FS, CFS, CS) than utilizing them independently. This should be interpreted similarly to the contribution of various (sub)cultures to the national identity of a multi-national country, without losing their primary identity (Huntington & Dunn, 2004). There is the possibility that specific songs belonging to a particular genre that are transmitted from generation to generation in a certain group of people with shared cultural norms and beliefs contribute to the final formation of a distinctive musical identity at "top-level" - the population level of a country, while retaining some of their original characteristics and expressions.

In the hierarchical cluster analysis-derived dendrogram (Figure 7.2), countries are grouped according to their similarities. The partition of the dendrogram into two main clusters, the first relating to a group of countries merged into clusters no. 1 and no. 2, and the second relating to all the other countries merged into clusters no. 3 - 5, reveals the historical, political, cultural, economical, and social roots of the 22 countries, as well as their very specific and complex relationships, which can be traced over a long period of time (see Figure A.1 and A.2). The two major clusters show the conceivable coexistence of two opposites from roughly the 15th to the 20th centuries: the Habsburg Empire (later Austro-Hungarian) and the Ottoman Empire; Western and Eastern; Catholic and Muslim; which is consistent with the findings of Juhász (2006, 2009).

The second hypothesis, that musical features and dimensions used in a country's national representative music are more similar in countries with similar cultural, political, historical, and economic backgrounds and that are geographically close to one another was only partially confirmed. Similarities have been discovered between geographically distant countries (e.g., between Slovenia and Latvia, between Sweden and Spain). Mehr et al. (2019, p. 1) suggest that a plausible explanation could be that "behavioural patterns once considered arbitrary cultural products may exhibit deeper, abstract similarities across societies". The existence of dissimilarity between geographically close countries (e.g., the Netherlands and Germany or Spain and Portugal) indicates either (i) the absence (or abundance) of a common musical core (Juhász, 2006) or (ii) the fact that the variability of individuals has manifested itself more strongly at the population level, consistent with the findings of Lumaca et al. (2018).

The comparison of all the three genres (CFS, CS, and FS) within countries using 16 viewpoints, has shown statistically significant differences, confirming the importance of observing the diversity in music not only between, but also within cultures(e.g., Feld, 1984; Henry, 1976; Rzeszutek et al., 2012). The differences between genres are important, as neither children's folk songs nor children's songs are (to date) considered to be a genre on its own. The analysis of these three genres within countries, and the differences found in the way how musical features and dimensions are used, reveals how music is understood and processed in the children's society or in authors writing children's music.

Small differences in the use of musical features and dimensions across genres within some countries (e.g., Bulgaria and Denmark) and vast differences in others (e.g., France, Germany, Portugal, Romania, etc.) can be understood in the sense that songs belonging to a particular genre may undergo more or less random changes, depending on the cultural forces that diversify or homogenize musical genres within a country. (Rzeszutek et al., 2012). These "cultural forces" might be regarded as globalization, digitization, the cultural policy enforced by a country's government, a strong sense for the preservation of cultural heritage, etc. For instance, despite the fact that socialist guidelines and socialist governments support ethnic variety, implementing a policy of monoethnicism or ignoring ethnic differences in order to establish a "unified national socialist identity" (Rees, 2011) can contribute to a decrease in genre diversity in a country. Furthermore, the results of a study conducted by Bourreau et al. (2022) analyzing the impact of digitization on the homogenization of music content in ten countries reveal that digitization has increased acoustic disparity among music charts. It is possible that digitization has had an impact on newly composed children's songs, where songs may be more (or less) influenced by globalized music than by traditional music, resulting in fewer or greater differences between genres of representative music in a given country.

The diversity found between the three genres within countries confirms partially the hypothesis, that there are differences in the use of musical features and dimensions in genres considered to belong to the representative music of a country, as it is not clear how cultural forces, which homogenize and diversify genres, are impacting these genres in a particular country.

7.5 Summary

Using Corpus III, which includes 2,184 songs from three genres (folk songs, children's folk songs, and children's songs), this chapter examined the diversity of musical features and dimensions between and within 22 European countries. Using the computational model IDyOM and 48 different perceptual models for the simulation of the human perception of the musical structure, as well as a multiple approach to examine the social, historical, cultural, geographical, and political background in 22 European countries, differences in the way musical features and dimensions are used in all three genres were discovered between and within countries. A detailed analysis of the (dis)similarities between and within 22 European countries shown that within-country variability outweighs between-country variability, which is in accordance with findings from other studies (e.g., Feld, 1984; Henry, 1976; Rzeszutek et al., 2012).

There were expected similarities in close distanced countries (e.g., France and Germany) and dissimilarities in distant countries (e.g., Bulgaria and Denmark), as well as unexpected similarities in far countries (e.g., Slovenia and Latvia) and dissimilarities in close distanced countries (e.g., Portugal and Spain). Plausible explanations include (i) the degree of dominance of two mighty Empires (the Ottoman and Habsburg Empires) in countries included in these state formations, (ii) the absence (or abundance) of a common musical core, or (iii) individuals exerting a strong influence on the culture of a particular country, probably more so in countries where "individualism" tends to predominate (Hofstede, 2011). Similarities seen in distant countries point to either (i) a probable "abstract" acceptance of behavioral (cultural) products/patterns across societies, or (ii) similar cultural behavior and practices that already exist in these cultures.

The best results were obtained when employing five clusters and combining all three genres, using k-means and hierarchical clustering, taking into account also the distinctive development of the 22 European countries over a period of about five centuries. The most comparable clusters when using each genre independently were those when using only folk songs, followed by fairly similar clusters when using children's folk songs, and the least similar clusters when using only children's songs. This is not really surprising given that folk songs have all the distinctive elements of a particular culture and that children's folk songs tend to contain more of these elements than children's songs do (Mihelač, 2021; Mihelač & Panić Grazio, 2021). As evidence of the growing impact of globalization, the overview of children's music literature, and particularly children's (commercialized) music, reveals shared expression in both close and distant countries (Hassi & Storti, 2012). The extent to which globalization is influencing this genre has not been confirmed in this study and will be covered in a future research.

The findings indicate that each of the three genres used in this chapter can be regarded as a separate genre with its own characteristics and forms of expression. The best results (clusters) were produced when all three genres were combined, confirming their unique contribution to the identification of a given culture without necessarily sacrificing their individual identities, which can be seen when all three genres are combined or used separately. It is still unclear, however, how precisely each genre affects a country's identity at the "top-level," i.e., to what extent and under what conditions, and how intra-variability in countries affects the (dis)similarities between countries. These issues are still up for debate at the present and will be covered in an upcoming study.

Chapter 8

Conclusions and Future Work

If you want the answer — ask the question.

Lorii Myers

8.1 Conclusions

In this thesis, three main objectives have been addressed: (i) the segmentation of music in children of various ages, (ii) (ir)regularity in children's folk songs, and (iii) (dis)similarities in the use of musical features and dimensions across and within 22 European Countries.

Segmentation of Music in Children of Various Ages: Experimental Approach

The perception and processing of melody have been examined experimentally in a specific population, aged 5–6 years, and two older groups: middle, aged 8–12 years, and adolescents, aged 14–16 years, both with and without musical knowledge. Two experiments, The Game Experiment and The Breathing Experiment, conducted in 2018 and 2020 were designed to address a need in the literature, as there are currently no studies that explore how these age groups perceive musical phrases.

The Game Experiment, conducted in 2018 and 2020, examined whether a subject's memory of a song is solely dependent on recalling the entire song's notes or on recalling specific phrases. Using both normative (the most probable interpretation according to music theory and musical experience), and non-normative segments of melody to recreate songs, participants were more successful in reconstructing melodies with normative phrasing, independent of age or musical knowledge. The results indicate that neither the teaching approach (listening to sung or instrumental versions) nor the random order of songs in both different versions (normative and non-normative) affect the success of recreating the songs. The significance of prior phrase memory in the transition from one phrase to the next has been proven, providing support for the hypothesis that children and adolescents recall melodies through the usage of phrases.

The Breathing Experiment conducted in 2018 and 2020 examined two hypotheses, that agreement within age groups grows with age, and that similarity between the normative and participant segmentations also increases with age. Hypotheses were validated by the results in both experiments.The Breathing Experiment in 2018, revealed that the absence of meaningful lyrics pushes linguistic components into the background, and that the music and its phrases and phrase cues become the dominant focus, since phrase cues in the lyrics cannot be used. The 2020 Breathing Experiment investigated two additional issues, how the method of learning affects the perception of phrases, and whether the random order of songs influences the results of song segmentation. As with The Game Experiment, there were no statistically significant differences between the 2018 and 2020 Breathing Experiments. Involving children of varying ages in two distinct experiments, The Game Experiment and The Breathing Experiment yielded fruitful results, revealing possible causes for the observed over-segmentation, particularly in the youngest population.

Segmentation of Music in Children of Various Ages: Computational Approach

To determine which musical features and dimensions are used at different ages and how they vary depending on the musical structure, normative phrasings of the same three songs used in The Game and The Breathing Experiments, as well as Corpus 2.1, consisting of 155 children's folk songs and children's songs from Slovenia, were utilized to train the IDyOM computation model. The decision to use a computational model was based on the absence of studies comparing the segmentation of music using computational models to children of varying ages and musical experts.

In addition, preliminary data from a study comparing the segmentation of children, adolescents, and musical experts to IDyOM's automatic segmentation indicated that IDyOM may be capable of simulating human perception of music and its segmentation across ages. Thus, broad use of a computer model in the segmentation process would not only lead to a significant knowledge of how children of different ages interpret phrases and phrase boundaries, but also to the identification of the most effective computational models for the segmentation task.

IDyOM's capability to capture the human perception of musical segments at different ages was demonstrated by an increase in measurement similarity between human and computer segmentation from the youngest group (5–6 years old children) to the oldest (musical experts). The findings revealed that each musical feature can play a different role in phrase identification when employed alone or in combination with other musical features, depending on factors such as experience, age, and feature combination in a particular section of the song. The results also demonstrated that oversegmentation, which was most prevalent in the youngest group of children, might be caused not only by individual interpretation, but also by extremely contradicting perceptual cues present in the musical structure.

(Ir)regularity in Children's Folk Songs from 22 European Countries

Numerous studies highlight the significance of musical syntax and its contribution to the concept of musical structure that adheres to the norms of musical syntax. In general, these studies employ a conventional empirical approach (including listeners in experiments and music rating). In recent decades, computational approaches to music analysis that compare human and computer perceptions of the regularity in musical structure have enabled a more objective examination of musical structure.

The fact that expert-based detection of (ir)regularity in musical structure can be effectively replaced by a suitable computational model was the motivation for employing the computational model IDyOM and ten viewpoints to simulate and detect the (ir)regularity in children's folk songs (Corpus 2.2), which are commonly assumed to have a simple and regular structure. In addition, the algorithm IR_REG, which categorizes melodies based on the regularity of their musical structure, was implemented.

Children's folk songs with high or low irregularity have not been discovered in each of the 22 European countries. This can be explained by the fact that the origins of some children's folk songs can be traced back to folk songs and that the content has been passed down without any simplification or variation of the musical structure in terms of musical dimensions or musical features.

Two hypotheses were confirmed: (i) that the frequency of repeating patterns correlates with the irregularity of musical structure in children's folk songs, and (ii) that repeated patterns contribute to a stronger sense of regularity if they are given at the same pitch. Children's folk songs with a regular structure have a higher frequency of repeated patterns, which contributes to a greater sense of regularity, as indicated by the results. However, if patterns are repeated at a different pitch (as transposed patterns), they can be perceived differently depending on the preceding and following context. As children's folk songs are approximately 8–12 bars long, it is also possible that the duration of the songs influences the perception of these transposed recurring patterns.

Furthermore, observing the musical structure with ten different viewpoints has revealed the importance of pitch in the perception of regularity in children's traditional folk songs. Some of the children's folk songs in the data were classified as extremely irregular and complex (especially the songs from Great Britain, Norway, and Switzerland, according to the results), not only because of highly unexpected pitch distributions in the melody, but also because of the use of large leaps (intervals) between successive notes. The regularity of a musical structure is thus affected not just by pitch, but also by pitch intervals.

There were differences in rhythm (duration) between irregular and regular children's folk songs; however, with the exception of a few from Bulgaria, Romania, and Turkey, the majority of children's folk songs had recurrent pulse/meter, contributing to a more regular musical structure in which the events are more predictable. Analyzing the implied harmony to investigate the vertical dimension in children's folk songs showed no meaningful results. This was expected, given basic harmonic progressions are widespread in the melodies of both irregular and regular children's folk songs, indicating that harmony, as a secondary parameter, is established and largely dependent on the syntactical constraints and rules formed by the primary parameter pitch.

(Dis)similarities in the Use of Musical Features and Dimensions Across and Within 22 European Countries

Despite more than a century of extensive research into the reasons why music differs between/within cultures, there is no consensus, according to cross-cultural studies. In addition, there are ongoing discussions over what to look for in music to examine cultural diversity, which genre (type of music) to use, and which method to employ (manual, semi-automatic, automatic). In this thesis, a computational approach was used to test two hypotheses: (i) that there are substantial differences in the use of musical features and dimensions between European countries considered to share a single musical style, and (ii) that the musical features and dimensions used in the representative music of a country (music that best symbolizes the culture of a country) are more similar in countries that share a similar cultural, political, and historical background.

There were three justifications for employing not only traditional folk songs but also children's traditional folk songs and children's songs. The first was that 'non-representative' songs, such as children's folk songs and children's songs, are typically ignored in cross-cultural music studies because (it is considered) they add little to the study of human music. The second aim was to determine how and to what extent these two genres contribute to (dis)similarity between and within countries. The third reason was to demonstrate that children's folk songs and children's songs are a distinct genre based on the musical features and dimensions they employ.

(Dis)similarities Between 22 European Countries

Using IDyOM and 16 viewpoints to examine three different genres, folk songs (FS), children's folk songs (CFS), and children's songs (CS), similar relationships have been observed between countries compared to other cross-cultural research studies that only examine one genre. A feasible explanation could be the 'assimilating' or 'merging' of genres in a manner analogous to how (sub)cultures adapt to a dominant culture.

Using hierarchical clustering, better results were obtained by combining the three genres (FS, CFS, and CS) as opposed to employing them separately. The possibility exists that particular songs belonging to a particular genre and transmitted from generation to generation in a certain group of people with shared cultural norms and beliefs contribute to the final formation of a distinctive musical identity at the 'top-level' - the population level of a country, while retaining some of their original characteristics and expressions.

In the hierarchical cluster analysis-derived dendrogram (Figure 7.2), it was shown that countries are grouped according to their similarities. The partition of the dendrogram into two main clusters, revealed the historical, political, cultural, economic, and social roots of the 22 countries, as well as their very specific and complex relationships, which can be traced over a long period of time. The two major clusters show the conceivable coexistence of two opposites from roughly the 15th to the 20th centuries: the Habsburg Empire (later Austro-Hungarian) and the Ottoman Empire; Western and Eastern; Catholic and Muslim.

Surprisingly, similarities have been discovered between geographically distant countries (e.g., between Slovenia and Latvia, between Sweden and Spain). It is possible that behaviors that were once thought to be the result of random cultural products reveal deeper, more abstract connections between cultures. The existence of dissimilarity between geographically close countries (e.g., the Netherlands and Germany or Spain and Portugal) indicates either (i) the absence (or abundance) of a common musical core or (ii) the fact that the variability of individuals has manifested itself more strongly at the population level.

According to the results, the 22 European countries, can not be considered as a 'single musical culture' which is in accordance with the first hypothesis. The second hypothesis was only partially confirmed, as similarities have been found between geographically distant countries, and dissimilarities between geographically close countries.

(Dis)similarities Within 22 European Countries

The comparison of all the three genres (CFS, CS, and FS) within countries using 16 viewpoints, has shown statistically significant differences, confirming the importance to observe the diversity in music not only between, but also within cultures. A detailed analysis of the (dis)similarities between and within 22 European countries has shown that within-country variability outweighs between-country variability.

The diversity found between these three genres is important, as neither children's folk songs nor children's songs are (to date) considered to be a genre on its own. The detailed analysis of these three genres within countries, and the differences found in the way how musical features and dimensions are used, reveals how music is understood and processed in the children's society or in authors writing children's music.

Small differences in the use of musical features and dimensions across genres within some countries (e.g., Bulgaria and Denmark) and vast differences in others (e.g., France, Germany, Portugal, Romania, etc.) can be understood in the sense that songs belonging to a particular genre may undergo more or less random changes, depending on the cultural forces that diversify or homogenize musical genres within a country. These 'cultural forces' might be regarded as globalization, digitization, the cultural policy enforced by a country's government, a strong sense for the preservation of cultural heritage, etc. As it is (still) unclear how cultural forces homogenize and diversify these genres, the hypothesis that there are differences in the use of musical features and dimensions in genres regarded to belong to the representative music of a given country was partially proven.

8.2 Future Directions

Segmentation of Music in Children of Various Ages: Experimental Approach

It is unclear how motifs with a high frequency of repetition vs distinguishing motifs that are not repeated have an effect on the memorization of tones/phrases. It could be interesting to compare motifs that are repeated frequently to those that are repeated infrequently in more songs.

Future research could investigate why the breathing experiment's issues (implicit harmony and sensitivity to contour on phrasing) appear to impair the perception of phrases and phrase boundaries in the youngest group of children. In addition, it should be examined whether over-segmentation is associated with the fact that different age groups interpret the same sound pattern (or song) differently.

Regardless of the fact that the songs used in this experiment were selected on purpose and with utmost care due to their unique musical structure, having only three songs was a certain limitation, as more songs with similar features could have shed more light on previously mentioned topics. As an experimental approach is used with participants, another limitation in future research would be the limited number of songs due to the varying abilities (e.g., song memorization, concentration, etc.) of children of different ages participating in the experiment.

Segmentation of Music in Children of Various Ages: Computational Approach

More research is needed to identify what factors influence how certain events are used in the partitioning of music in different age groups, and why some musical features (higher or lower order) are perceived to be more essential than others. The impact of song familiarity on the perception of phrases and phrase boundaries, independent of age or experience, is of special interest, as is the analysis of the (potential) impact of song familiarity on over-segmentation. Future research is also intended to focus on all children's (folk) songs containing anacrusis, as it is unclear how anacrusis affects the perception of musical structure and its segmentation.

(Ir)regularity in Children's Folk Songs from 22 European Countries

Future research with more musical feature variables could contribute to the explanations of (ir)regularity in musical structure. As each genre is "not born in an empty space but in a musical system that is already structured" (Fabbri, 1981, p. 57) means that a considerable part of the rules that define e.g., folk songs, could be common to children's folk songs and children's songs, and that individuals and/or community accepting (or denying) these rules in a period of time contributes more or less to the (i)rregularity of the musical structure giving more or less importance to various musical features. An in-depth analysis of the origin of children's folk songs would certainly contribute to the understanding of (ir)regularities found in the musical structure of this genre, and why irregular structure is more frequently found in some European countries than in other European countries.

The (ir)regularity was evaluated using Corpus II. Due to the limited availability of children's folk songs from the remaining 21 countries, this Corpus contains 124 (out of 736) German children's folk songs. To achieve better statistical outcomes and representation of (ir)regularity and to minimize statistical biases, future work will focus on adding more children's folk songs to this Corpus in order to train the model with a proportional number of songs.

(Dis)similarities in the Use of Musical Features and Dimensions Across and Within 22 European Countries

The best results (clusters) were produced when all three genres (folk songs, children's folk songs, and children's songs) were combined, confirming their unique contribution to the identification of a given culture without sacrificing their individual identities, as is evident when all three genres are combined or used separately. In addition, the results demonstrated that each of the three genres can be considered a distinct genre with its own features and forms of expression. It remains unknown, however, how precisely each genre impacts a country's identity at the 'top-level,' i.e. to what extent and under what conditions, and how intra-variability in countries affects the (dis)similarities between countries. These concerns are still under discussion and will be the subject of a subsequent study.

In all three musical genres throughout the 22 European nations, there were differences in the manner that musical features and dimensions were used, though to varying degrees in some of these countries. An additional investigation of cultural forces (such as globalization, digitization, a country's government-enforced cultural policy, the preservation of cultural heritage, etc.) may provide more insight.

Appendix A

Supplementary Material

A.1 Additional Figures

Figure A.1: Ottoman Empire. Copyright 2021 by Swanston Map Archive Limited. Reprinted with permission.

Figure A.2: Habsburg Empire. Copyright 2021 by Swanston Map Archive Limited. Reprinted with permission.

References

Agres, K. R., Abdallah, S., & Pearce, M. T. (2018). Information-theoretic properties of auditory sequences dynamically influence expectation and memory. *Cognitive Science*, *42*(1), 43–76.

Agres, K. R. (2019). Change detection and schematic processing in music. *Psychology of Music*, *47*(2), 173–193. https://doi.org/10.1177/0305735617751249

Ahlbäck, S. (2004). *Melody beyond notes: A study of melody cognition.* Goteborgs Universitet.

Ahlbäck, S. (2007). Melodic similarity as a determinant of melody structure. *Musicae Scientiae*, *11*(1_suppl), 235–280. https://doi.org/10.1177/102986490701100110

Altenmüller, E. (1993). Psychophysiology and eeg. In F. Niedermeyer & L. da Silva (Eds.), *Electroencephalography* (pp. 597–614). Williams Wilkins.

Bader, M., Schröger, E., & Grimm, S. (2017). How regularity representations of short sound patterns that are based on relative or absolute pitch information establish over time: An eeg study. [[Online; accessed 2019-10-22]].

Balkwill, L. L., & Thompson, W. F. (1999). The cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, *17*, 43–64.

Bamberger, J. (1982). Revisiting children's drawings of simple rhythms: A function of reflection-in-action. In S. Strauss (Ed.), *U-shaped behavioral growth.* New York: Academic Press.

Bangert, M., & Altenmüller, E. (2003). Mapping perception to action in piano practice: A longitudinal dc-eeg study. *BMC Neuroscience*, *4*, 26.

Bartlett, J. C., & Dowling, W. J. (1980). Recognition of transposed melodies: A key–distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 501–515.

Bauer, A. (2001). Composing the sound itself: Secondary parameters and structure in the music of ligeti. *Indiana Theory Review*, *22*(1), 37–64.

Beauvois, M. W. (2007). Quantifying aesthetic preference and perceived complexity for fractal melodies. *Music Perception: An Interdisciplinary Journal*, *24*(3), 247–264. https://doi.org/10.1525/mp.2007.24.3.247

Beekes, R. (2010). *Etymological dictionary of greek.* Leiden, Brill.

Begleiter, R., El-Yaniv, R., & Yona, G. (2004). *Journal of Artificial Intelligence Research*, *22*, 385–421.

Bell, T. C., Cleary, J. G., & Witten, I. H. (1990). *Text compression.* Prentice Hall.

Bennet, P. D. (1990). Children's perceptions of anacrusis patterns within songs. In T. W. Tunks (Ed.), *Reports of research in music education* (pp. 1–8). Texas Music Educators Association, Austin, Texas.

Berezovsky, J. (2019). The structure of musical harmony as an ordered phase of sound: A statistical mechanics approach to music theory. *Science Advances*, *5*(5). https://doi.org/10.1126/sciadv.aav8490

Berget, G. E. (2017). *Using hidden markov models for musical chord prediction.* (Master's thesis). Norwegian University of Science and Technology, Norway.

Berland, A., Gaillard, P., Guidetti, M., & Barone, P. (2015). Perception of everyday sounds: A developmental study of a free sorting task. *PLOS ONE*, *10*(2), 1–21.

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart*, *92*(4), 445–452.

Berrar, D. (2019). Cross-validation. In S. Ranganathan, M. Gribskov, K. Nakai, & C. Schönbach (Eds.), *Encyclopedia of bioinformatics and computational biology* (pp. 542–545). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Bertrand, D. (1999). *Groupement rhythmique et representation mentale de melodies chez l'enfant* (Doctoral dissertation). Liège University, Belgium. Unpublished doctoral dissertation.

Besson, M., Faïta, F., Peretz, I., Bonnel, A.-M., & Requin, J. (1998). Singing in the brain: Independence of lyrics and tunes. *Psychological Science*, *9*, 494–498.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, *5*, 1–30.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and performance*, *5*, 403–410.

Bigand, E. (2003). More about the musical expertise of musically untrained listeners. *Annals of the New York Academy of Sciences*, *999*(1), 304–312.

Bigand, E., Delbé, C., Poulin-Charronnat, B., Leman, M., & Tillmann, B. (2014). Empirical evidence for musical syntax processing? computer simulations reveal the contribution of auditory short-term memory. *Frontiers in Systems Neuroscience*, *8*, 94. https://doi.org/10.3389/fnsys.2014.00094

Bigand, E., & Poulin-Charronat, B. (2006). Are we "experienced listeners"? a review of the musical capacities that do not depend on formal musical training [[Online; accessed 2018-7-27]].

Bigand, E. (1993). The influence of implicit harmony, rhythm and musical training on the abstraction of "tension–relaxation schemas" in a tonal musical phrase. *Contemporary Music Review*, *9*, 128–139.

Blacking, J. (1973). *How musical is man?* Seattle, London:University of Washington Press.

Blacking, J. (1990). Transcultural communication and the biological foundations of music. In R. Pozzi (Ed.), *La musica come linguaggio universale* (pp. 179–188). Olschki.

Bohlman, P. V. (1988). *The study of folk music in the modern world.* Indiana University Press.

Boltz, M. G. (1999). The processing of melodic and temporal information: Independent or unified dimensions? *Journal of New Music Research*, *28*, 67–79.

Bourreau, M., Moreau, F., & Wikstrom, P. (2022). Does digitization lead to the homogenization of cultural content? *Economic Inquiry*, *60*, 427–453.

Bouwer, F. L., Burgoyne, J. A., Odijk, D., Honing, H., & Grahn, J. a. (2018). What makes a rhythm complex? the influence of musical training and accent type on beat perception. *PLoS ONE*, *13*(1). https://doi.org/doi.org/10.1371/journal.pone.0190322

Bowling, D. L., Sundarajan, J., Han, S., & Purves, D. (2012). Expression of emotion in eastern and western music mirrors vocalization. *PloS one*, *7*(3), e31942. https://doi.org/https://doi.org/10.1371/journal.pone.0031942

Brailoiu, C. (1954). Le rythme enfantin. *Colloques de Wegimonts, Cercle International d'Etude Ethnomusicologique, Brussels*, 64–96.

Brenna, P. S. (1992). Design and implementation of curricula experiences in world music: A perspective. *Music Education: Sharing Musics of the World. Seoul, Korea: Conference proceedings from World Conference International Society of Music Education*, 221–225.

Bresson, M. (1997). Electrophysiological studies of music processing. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 217–250). Hove: Psychology Press.

Brochard, R., Dufour, A., Drake, C., & Scheiber, C. Functional brain imaging of rhythm perception (C. Woods, R. Luck, R. Brochard, J. Seddon, & A. Sloboda, Eds.). In: In *Proceedings of the sixth international conference of music perception and cognition* (C. Woods, R. Luck, R. Brochard, J. Seddon, & A. Sloboda, Eds.). Ed. by Woods, C., Luck, R., Brochard, R., Seddon, J., & Sloboda, A. Keele: University of Keele, 2000, August.

Bronson, B. H. (1949). Mechanical help in the study of folk song. *Journal of American Folklore*, *62*(244), 81–86.

Brown, S., & Jordania, J. (2011). Universals in the world's musics. *Psychology of Music*, *41*(2), 229–248.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain: A pet study of the generation of melodies and sentences. *Journal of Neuroscience*, *23*(10), 2791–2803.

Bruner, J. S., Wallach, M. A., & Galanter, E. H. (1959). The identification of recurrent regularity. *The American Journal of Psychology*, *72*(2), 200–209. https://doi.org/10.2307/1419364

Busch, R., & Graubart, M. (1986). On the horizontal and vertical presentation of musical ideas and on musical space (ii). *Tempo*, (156), 7–15. https://doi.org/10.1017/S0040298200022063

Butler, D., & Brown, H. (1994). Describing the mental representation of tonality in music. In R. Aiello & J. A. Sloboda (Eds.), *Musical perceptions* (pp. 191–212). New York: Oxford University Press.

Cabredo, R. A., Legaspi, R. S., & Numao, M. (2011). Identifying emotion segments in music by discovering motifs in physiological data. [[Online; accessed 2022-18-8]].

Calvert, S. L., & Billingsley, R. L. (1998). Young children's recitation and comprehension of information presented by songs. *Journal of Applied Developmental Psychology*, *19*(1), 97–108.

Cambouropoulos, E. (2001). The local boundary detection model (lbdm) and its application in the study of expressive timing. *Proceedings of the International Computer Music Conference, Havana, Cuba.*

Cambouropoulos, E. (2006). Musical parallelism and melodic segmentation: A computational approach. *Music Perception: An Interdisciplinary Journal*, *23*(3), 249–268.

Cambouropoulos, E., Crawford, T., & Iliopoulos, C. S. (2001). Pattern processing in melodic sequences: Challenges, caveats and prospects (P.-Y. Rolland, E. Cambouropoulos, & G. A. Wiggins, Eds.) [Special Issue on Pattern Processing in Music Analysis and Creation]. *Computers and the Humanities*, *35*, 9–21.

Campbell, D. (2002). *The mozart effect for children. awakening your child's mind, health, and creativity with music.* HarperCollins Publisher Inc.

Campbell, P. (2010). *Songs in their heads. music and its meaning in children's lives.* Oxford, University Press.

Campbell, P. S. (1998). *Songs in their heads : Music and its meaning in children's lives.* Cary, N, USA:Oxford University Press.

Cancino-Chacón, C., Grachten, M., & Agres, K. (2017). From bach to the beatles: The simulation of human tonal expectation using ecologically-trained predictive models. *ISMIR*.

Chew, E. (2000). *Towards a mathematical model of tonality* (Doctoral dissertation). Cambridge, MA.

Chiappe, P., & Schmuckler, M. A. (1997). Phrasing influences the recognition of melodies. *Psychonomic Bulletin & Review*, *4*(2), 254–259. https://doi.org/10.3758/BF03209402

Cholin, J., Schiller, N. O., & Levelt, W. J. M. (2004). The preparation of syllables in speech production. *Journal of Memory and Language*, *50*, 46–61.

Christ, G. (1983). The cognitive development sequence of music skills in elementary school aged children.

Cleary, J. G., & Witten, I. (1984). Data compression using adaptive coding and partial string matching. *IEEE Transactions on Communications*, *32*, 396–402.

Cohen, A. J. (2000). Development of tonality induction: Plasticity, exposure and training. *Music Perception*, *17*(4), 437–459.

Cohen, D. (2003). Perception and responses to schemata in different cultures: Western and arab music. *Proceedings of the 5th Triennial ESCOM Conference*, 519–522.

Cohen, D., & Katz, R. L. (2013). Toward a musical analysis of world music. *Israel Studies in Musicology Online*, *13*.

Cohrdes, C., Grolig, L., & Schroeder, S. (2016). Relating language and music skills in young children: A first approach to systemize and compare distinct competencies on different levels. *Frontiers in Psychology*, *7*, 1616. https://doi.org/10.3389/fpsyg.2016.01616

Cohrdes, C., Platz, F., & Kopiez, R. (2014). Der körper als mediator: Möglichkeiten einer unvermittelten beschreibung von musik(-präferenzen) im grundschulalter [the body as a mediator: Opportunities to describe music (preferences) in primary-school age directly]. In W. Auhagen, C. Bullerjahn, & R. von Georgi (Eds.), *Jahrbuch musikpsychologie* (pp. 169–197). Göttingen, Germany: Hogrefe.

Conklin, D. (1990). *Prediction and entropy of music* (Master's thesis). Department of Computer Science, University of Calgary, Canada.

Conklin, D. (2013). Antipattern discovery in folk tunes. *Journal of New Music Research*, *42*(2), 161–169. https://doi.org/10.1080/09298215.2013.809125

Cooksey, J. M. (1992). *Working with adolescent voices.* Missouri, MO: Concordia Publishing House.

Corrigall, K. A., & Schellenberg, G. E. (2015). Music cognition in childhood. [[Online; accessed 2018-4-3]].

Corsaro, W. (2005). *The sociology of childhood (second edition).* Thousand Oaks, CA: Pine Forge Press.

Costa-Giomi, E. (2003). Young children's harmonic perception. *Annals of the New York Academy of Sciences*, *999*(1), 477–484.

Cox, S. M. O. (1970). *A descriptive analysis of the response to beat, meter, and rhythm pattern by children, grades on to six* (Doctoral dissertation). University of of Wisconsin-Madisson. ProQuest Dissertations Publishing.

Cremin, L. A. (1959). John dewey and the progressive-education movement, 1915-1952. *School Review*, *67*(2), 160–173.

Crowder, R. G., Serafine, M. L., & Repp, B. (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory & cognition*, *18*(5), 469–476.

Cuddy, L. L., Cohen, A., & J, M. D. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology Human Perception & Performance*, *7*(4), 869–883.

Daikoku, S., Shimozono, T., Fujii, S., Hegde, S., & Savage, P. E. (202). Cross-cultural perception of musical similarity within and between india and japan. [[Online; accessed 2022-12-7]].

Dal, E., & Pihl, L. (1956). Scandinavian folk music: A survey. *Journal of the International Folk Music Council.*, *8*, 6–11.

Darwin, C. (1871). *The descent of man and selection in relation to sex.* John Murray.

Deacon, T. (1997). *The symbolic species.* W.W.Norton.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*, 384–398.

Deliège, I. (1987). Grouping conditions in listening to music: An approach to lerdahl and jackendoff's grouping preference rules. *Music Perception*, *4*, 325–360.

Demorest, S. M. (2000). The influence of phrase cues on children's melodic reconstruction [[Online; accessed 2018-4-6]].

de Nooijer, J., Wiering, F., Volk, A., & Tabachneck-Schijf, H. J. M. (2008). An experimental comparison of human and automatic music segmentation. *Proceedings of the 10th International Conference on Music Perception and Cognition.*

Despić, D. (2007). *Harmonija sa harmonskom analizom.* (N. Jovanović, Ed.). Zavod za udžbenike.

Deutsch, D. (1979). Binaural integration of melodic patterns. *Perception and Psychophysics*, *25*, 399–405.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. gestalt perception in music. *Perception & Psychophysics*, *28*, 381–389.

Deutsch, D. (Ed.). (1982). Structural representations of musical pitch. In *Psychology of music* (pp. 343–390). New York: Academic Press.

Deutsch, D. (1983). Auditory illusions and audio. *Journal of the Audio Engineering Society*, *31*, 606.

Deutsch, D., Hamaoui, K., & Henthorn, T. (2007). The glissando illusion and handedness. *Journal Neuropsychologia*, *45*(13), 2981–2988.

Deutsch, D. (2013). Grouping mechanisms in music [[Online; accessed 2018-4-2]].

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341–354.

Dowling, W. J. (1999). The development of music perception and cognition. In D. Deutsch (Ed.), *The psychology of music* (pp. 603–625). New York, NY: Oxford University Press.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception & psychophysics*, *53*(1), 25–33.

Drake, C., & Bertrand, D. (2001). The quest for universals in temporal processing in music. *Annals of the New York Academy of Sciences*, *930*(1), 17–27. https://doi.org/10.1111/j.1749-6632.2001.tb05722.x

Drinic, M., Kirovski, D., & Potkonjak, M. (2003). Ppm model cleaning. *Data Compression Conference, 2003. Proceedings. DCC 2003*, 163–172.

Dunhill, T. F. (1907). The evolution of melody. *Proceedings of the Musical Association*, 101–122.

Eitan, Z., & Granot, R. Y. (2008). Perception of rhythmic grouping depends on auditory experience. *Music Perception*, *25*, 397–418.

Ellis, A. J. (1885). On the musical scales of various nations. *Journal of the Society of Arts*, *33*(1), 485–527.

Fabbri, F. (1981). A theory of musical genres: Two applications. In D. Horn & P. Tagg (Eds.), *Popular music perspectives* (pp. 52–81). Götebort; Exeter: International Association for the Study of Popular Music.

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Musical taste and ingroup favouritism. *Mem Cognition*, *37*(1), 1–9.

Feld, S. (1984). Sound structure as social structure. *Ethnomusicology*, *28*, 383–409.

Feldman, J. (1997). Regularity-based perceptual grouping. *Computational Intelligence*, *13*(4), 582–623.

Fitch, W. T. (2004). The evolution of language. In M. Gazzaniga (Ed.), *The cognitive neurosciences iii.* (pp. 873–883). Cambridge, MA: MIT Press.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, *100*, 173–215.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*, 378–382.

Foley, R., & Mirazón Larh, M. (2011). The evolution of the diversity of cultures. *Phil. Trans. R. Soc. B*, *366*, 1080–1087.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 680–698.

François, C., Chobert, J., Besson, M., & Schö, D. (2012). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043. https://doi.org/10.1093/cercor/bhs180

Friberg, A., & Battel, G. U. (2002). Structural communication (chapter 13). In R. Parncutt & G. E. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning.* (pp. 453–492). Oxford, OUP.

Friedmann, J. L. (1980). *Music in biblical life: The roles of song in ancient israel.* McFarland&Company, Inc. Publishers.

Fujioka, T., Trainor, L. J., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*(6), 1010–1021.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Gelbart, M., & Rehding, A. (2011). Riemann and melodic analysis: Studies in folk-musical tonality. In E. Gollin & A. Rehding (Eds.), *The oxford handbook of neo-riemannian music theories.* (pp. 140–165). New York, NY: Oxford University Press.

Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2015). Review defining the biological bases of individual differences in musicality. *Phil. Trans. R. Soc. B.*, *370*(1664).

Gingras, B., Pearce, M. T., Goodchild, M., Dean, R. T., Wiggins, A. G., & McAdams, S. (2016). Linking melodic expectation to expressive performance timing and perceived musical tension. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(4), 594–609.

Gladwell, M. (2002). *The tipping point: How little things can make a difference.* New York: Little, Brown Company.

Gobet, F., Lane, P., Croker, S., Cheng, P. C.-H., Jones, G., Oliver, I., & Pine, J. (2001). Chunking mechanisms in human learning. *TRENDS in Cognitive Sciences*, *5*(6), 236–243.

Godt, I. (2005). Music: A practical definition. *The Musical Times*, *146*(1890), 83–88. http://www.jstor.org/stable/30044071

Gold, B. P., Pearce, M. T., Mas-Herrero, E., Dagher, A., & Zatorre, R. J. (2019). Cultural evolution of music. *Palgrave Commun*, *5*, 16.

Golež Kaučič, M. (2003). *Ljudsko in umetno. dva obraza ustvarjalnosti.* Ljubljana: Založba ZRC.

Gómez, E., Haro, M., & Herrera, P. (2009). Music and geography: Content description of musical audio from different parts of the world. *Proceedings of the International Society for Music Information Retrieval Conference. Kobe.*

Gómez, E., Klapuri, A., & Beno$^{i}t, M.$. (2003). Melody description and extraction in the context of music content processing. *Journal of New Music Research*, *32*(1), 422–71.

Gooding, L., & Standley, J. M. (2011). Musical development and learning characteristics of students: A compilation of key points from the research literature organized by age. update applications of research in music education. *Update Applications of Research in Music Education*, *30*(1), 32–45.

Gordon, E. (1979). *Primary measures of music audiation.* Chicago: GIA.

Gordon, R. L., Schön, D., Magne, C., Astésano, C., & Besson, M. (2010). Words and melody are intertwined in perception of sung words: Eeg and behavioral evidence [[Online; accessed 2018-7-30]].

Grahn, J. A., & Rowe, J. B. (2012). Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral Cortex*, *23*(4), 913–921. https://doi.org/10.1093/cercor/bhs083

Graney, K. (2019). *Russia, the former soviet republics, and europe since 1989.* Oxford, University Press.

Graves, J. E., & Oxenham, A. J. (2017). Familiar tonal context improves accuracy of pitch interval perception. *Frontiers in psychology*, *8*, 1753.

Han, J., & Pei, J. (2012). Getting to know your data. *Data mining: Concepts and techniques* (pp. 39–82). Elsevier.

Handel, S. (1989). *Listening.* MIT Press, Cambridge, MA.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*(11), 466–472.

Hansen, N. C., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology*, *5*, 1052. https://doi.org/10.3389/fpsyg.2014.01052

Hartmann, M., Lartillot, O., & Toiviainen, P. (2016). Multi-scale modelling of segmentation : Effect of music training and experimental task. *Music Perception*, *34*(2), 192–217.

Harwood, D. L. (1976). Universals in music: A perspective from cognitive psychology. *Ethnomusicology*, *20*(3), 521–533.

Hassi, A., & Storti, G. (2012). Globalization and culture: The three h scenarios. In H. Cuadra-Montiel (Ed.), *Globalization - approaches to diversity.* London: IntechOpen.

Heinrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*, 61–135.

Henry, E. (1976). The variety of music in a north indian village: Reassessing cantometrics. *Ethnomusicology*, *20*(1–2), 49–66.

Herzog, G. (1944). Some primitive layers in european folk music. *Bulletin of the American Musicological Society*, *9–10*, 11–14.

Herzog, G. (1947). Some primitive layers in european folk music. *Bulletin of the American Musicological Society*, (9/10), 11–14. http://www.jstor.org/stable/829223

Higgins, K. M. (2012). *The music between us: Is music a universal language?* University of Chicago Press.

Hoeschele, M., Weisman, R. G., & Sturdy, C. B. (2012). Pitch chroma discrimination, generalization, and transfer tests of octave equivalence in humans. *Attention, Perception and Psychophysics*, *74*(8), 1742–1760.

Hofstede, G. (2011). Dimensionalizing cultures: The hofstede model in context. online readings in psychology and culture, 2(1). [[Online; accessed 2021-8-15]].

Holleran, S., Jones, M. R., & Butler, D. (1995). Perceiving implied harmony: The influence of melodic and harmonic context. *Journal of Experimental Psychology: LMC*, *21*, 737–753.

Honing, H., ten Cate, C., Peretz, I., & E, T. S. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *370*, 1664.

Honing, H., Bouwer, F. L., Prado, L., & Merchant, H. (2018). Rhesus monkeys (macaca mulatta) sense isochrony in rhythm, but not the beat: Additional support for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *12*, 475.

Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *41*(1), 47–56.

Huntington, S. P., & Dunn, S. (2004). *Who are we? america's national identity and the challenges it faces.* New York: Simon & Schuster, Inc.

Hurewitz, J. C. (1961). Ottoman diplomacy and the european state system. *Middle East Journal*, *15*(2), 141–152.

Huron, D. (1996). The melodic arch in western folksongs. *Computing in Musicology*, *10*, 3–23.

Huron, D. (1997). Humdrum and kern: Selective feature encoding. In E. Selfridge-Field (Ed.), *Beyond midi: The handbook of musical codes* (pp. 375–401). MIT Press.

Huron, D. (2001). What is a musical feature? forte's analysis of brahms's opus 51, no. 1., revisited. *MTO a journal of the Society for Music Theory*, *7*(4), 1–12.

Hutchins, S., & Palmer, C. (2008). Repetition priming in music. *Journal of Experimental Psychology*, *34*(3), 693–707.

Ireland, K., Parker, A., Foster, N., & Penhune, V. (2018). Rhythm and melody tasks for school-aged children with and without musical training: Age-equivalent scores and reliability. *Frontiers in Psychology*, *9*, 426. https://doi.org/10.3389/fpsyg.2018.00426

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, *124*(4), 2263–2271.

Iwanaga, M., & Moroki, Y. (1999). Subjective and physiological responses to music stimuli controlled over activity and preference. *Journal of Music Therapy*, *36*, 26–38.

Jackendoff, R. (1987). *Consciousness and the computational mind.* Cambridge, MA: MIT Press.

Janata, P., Birk, J. L., Horn, J. D. V., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying western music. *Science*, *298*, 2167–2170.

Jing, L. (2017). A comprehensive study on the development of folk music tourism culture in jiangxi. *Proceedings of the International Conference on Social Science, Management and Economics*, 182–186.

Johnson, T. A. (1994). Minimalism: Aesthetic, style, or technique? *The Musical Quarterly*, *78*(4), 742–773. https://doi.org/10.2307/742508

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychology Revue*, *96*, 459–491.

Jouste, M. (2009). Traditional melodic types in the music of the sámi in finland. In J. Niemi (Ed.), *Perspectives on the song of the indigenous peoples of northern eurasia: Performance, genres, musical syntax, sound.* (pp. 240–266). Oxford University Press.

Jožef-Beg, J., & Mihelač, L. (2019). *Na začetku je bila pesem : Medpredmetni samostojni delovni zvezek za književno in glasbeno vzgojo (at the beginning was a poem: A cross-curricular independent workbook for literary and musical education).* Šolski center Novo mesto, Novo mesto.

Juhász, Z. (2006). A systematic comparison of different european folk music traditions using self- organizing maps. *Journal of New Music Research*, *35*, 95–112.

Juhász, Z. (2009). Automatic segmentation and comparative study of motives in eleven folk song collections using selforganizing maps and multidimensional mapping. *Journal of New Music Research*, *38*(1), 71–85.

Justus, T., & Bharucha, J. (1998). Universals in music and musical experiences. In C. Miereanu & X. Hascher (Eds.), *Les universaux en musique, actes du 4e congrès international sur la signification musicale.* (pp. 657–666). Paris: Publications de la Sorbonne.

Justus, T., & Bharucha, J. (2003). Music perception and cognition. In A. Yantis & H. Pasler (Eds.), *Stevens handbook of experimental psychology, volume i: Sensation and perception.* (pp. 453–492). New York: Wiley.

Kandel, E., Schwartz, J., & Jessel, T. (1991). *Principles of neural science.* New Jersey: Prentice-Hall.

Kann, R. A. (1974). *A history of the habsburg empire.* Berkeley, California: University of California Press.

Kartomi, M. (1999). Play songs by children and their educational implications. *Aboriginal History*, *23*, 61–71.

Kim, J. (2000). Children's pitch matching, vocal range, and developmentally appropriate practice. *Journal of Research in Childhood Education*, *14*(2), 152–160.

Kim, Y. E., Chai, W., Garcia, R., & Vercoe, B. Analysis of a contour-based representation for melody. In: *Proceedings of the international symposium on music information retrieval.* 2000.

Klein, J., & Jacobsen, T. (2014). Music is not a language [[Online; accessed 2020-08-19]].

Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., & Friederici, A. D. (2005). Perception of phrase structure in music. human brain mapping. *Human brain mapping*, *24*(4), 259–273.

Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: An erp study. *Brain Research*, *1212*, 55–62.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9*, 578–584.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschkea, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), 15443–15448.

Koniari, D., Predazzer, S., & Méélen, M. (2001). Categorization and schematization processes used in music perception by 10- to 11-year-old children. *Music Perception*, *18*(3), 297–324.

Kostopoulou, E. (2016). Autonomy and federation within the ottoman empire: Introduction to the special issue. *Journal of Balkan and Near Eastern Studies*, *18*(6), 525–532. https://doi.org/10.1080/19448953.2016.1196039

Kragness, H., & Trainor, L. (2017). Young children pause on phrase boundaries in self-paced music listening: The role of harmonic cues [[Online; accessed 2018-4-6]].

Kramer, J. D. (1988). *The time of music: New meanings, new temporalities, new listening strategies.* Schirmer Books.

Krumhansl, C. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, *51*(4), 336–353.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*, 159–179.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive psychology*, *11*(3), 346–374.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* Oxford University Press.

Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science*, *1*(1), 70–73.

Kumer, Z. (1988). *Etnomuzikologija. razgled po znanosti o ljudski glasbi.* Ljubljana, Univerza Edvarda Kardelja v Ljubljani, Filozofska fakulteta.

Lamont, A. (1998). Music, education, and the development of pitch perception: The role of context, age and musical experience. *Psychology of Music*, *26*(1), 7–25. https://doi.org/10.1177/0305735698261003

Lappe, C., Steinsträter, O., & Pantev, C. (2013). Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection. *Frontiers in human neuroscience*, *7*, 260. https://doi.org/doi.org/10.3389/fnhum.2013.00260

Lawson, J. (2011). Poetry. In P. Nel & L. Paul (Eds.), *Keywords for children's literature: Mapping the critical moment.* (pp. 306–314). New York: New York University Press.

Le Bomin, S., Lecointre, G., & Heyer, E. (2016). The evolution of musical diversity: The key role of vertical transmission. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(3), e0151570. https://doi.org/10.1371/journal.pone.0151570

Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune: Infants' perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant Behavior Development*, *33*(4), 419–430.

Lerdahl, F., & Jackendoff, R. (1977). Toward a formal theory of music. *Journal of Music Theory*, *21*, 111–172.

Lerdahl, F., & Jackendoff, R. (1983a). An overview of hierarchical structure in music. *Music Perception*, *1*(2), 229–252.

Lerdahl, F., & Jackendoff, R. (1983b). *A generative theory of tonal music.* MIT Press, Cambridge, MA.

Levitin, D. J. (1999). Memory for musical attributes. In P. R. Cook (Ed.), *Music, cognition and computerized sound : An introduction to psychoacoustics* (pp. 1617–1625). Massachusetts: MIT Press.

Levitin, D. J. (2000). The neural locus of temporal structure and expectancies in music: Evidence from functional neuroimaging at 3 tesla. *Music Perception*, *22*, 563–575. https://doi.org/10.1525/mp.2005.22.3.563

Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music. *Annals of the New York Academy of Sciences*, *1156*(1), 211–231.

Liegeoise-Chauvel, C., Peretz, I., Babai, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Psychological Science*, *121*, 1853–1867.

Ling, J. (1997). *A history of european folk music.* University of Rochester Press.

List, G. (1971). On the non-universality of musical perspectives. *Ethnomusicology*, *15*, 399–402.

List, G. (1973). Did the australopithecines sing? *Current Anthropology*, *14*(1–2), 25–29.

Lomax, A. (1976). *Cantometrics: An approach to the anthropology of music.* Berkeley: University of California Extension Media Center.

Lomax, A. (1977). Les universaux dans le chant. *The World of Music*, *19*(19–20), 131–141.

Longuet-Higgins, H. C. (1962). Letter to a musical friend. *Music Review*, *23*, 244–248.

López, M. R., & Volk, A. (2012). *Automatic segmentation of symbolic music encodings: A survey* (tech. rep.). Department of Information and Computing Sciences,Utrecht University, Utrecht.

Loui, P. (2012). Learning and liking of melody and harmony: Further studies in artificial grammar learning. *Topics in Cognitive Science*, *4*(4), 554–567.

Lowe, R. (2004). Childhood through the ages. In P. Nel & L. Paul (Eds.), *An introduction to early childhood studies* (pp. 64–74). London: SAGE Publications Ltd.

Lu, K., & Vicario, D. S. (2014). Statistical learning of recurring sound patterns encodes auditory objects in songbird forebrain. *Proceedings of the National Academy of Sciences*, *111*(40), 14553–14558. https://doi.org/10.1073/pnas.1412109111

Ludden, D. (2015). Is music an universal language? [[Online; accessed 2018-19-11]].

Lumaca, M., Ravignani, A., & Baggio, G. (2018). Music evolution in the laboratory: Cultural transmission meets neurophysiology. *Frontiers in Neuroscience*, *12*, 246. https://doi.org/10.3389/fnins.2018.00246

Lundquiest, B., & Szego, C. K. (1998). *Musics of the world's cultures: A source book for music educators.* Nedlands, Western Australia, CIRCME.

MacKay, D. (2003). *Information theory, inference, and learning algorithms.* Cambridge, UK: Cambridge University Press.

Manjunath, B. S., Wu, P., Newsam, S., & Shin, H. D. (2000). A texture descriptor for browsing and similarity retrieval. *Signal Processing: Image Communication*, *16*(1-2), 33–43.

Manuel, P. (1989). Modal harmony in andalusian, eastern european, and turkish syncretic musics. *Yearbook for Traditional Music*, *21*, 70–94.

Marcus, S. (2003). The eastern arab system of melodic modes: A case study of maqam bayyati. *The garland encyclopedia of world music.* (pp. 33–44). New York: Routledge.

Margulis, E. H. (2014). *On repeat: How music plays the mind.* New York: Oxford University Press.

Margulis, E. H., Levine, W. H., Simchy-Gross, R., & Kroger, C. (2017). Expressive intent, ambiguity, and aesthetic experiences of music and poetry [[Online; accessed 2019-7-1]].

Marler, P. (1976). An ethological theory of the origin of vocal learning. *Annals of the New York Academy of Sciences*, *280*, 386–395.

Marmel, F., & Tillmann, B. (2008). Tonal expectations influence pitch perception. *Perception & Psychophysics*, *70*(5), 841–852.

Marolt, M. (2008). A mid-level representation for melody-based retrieval in audio collections. *IEEE Transactions on Multimedia*, *10*(8), 1617–1625. https://doi.org/10.1109/TMM.2008.2007293

Marvin, E. W. (1991). The perception of rhythm in non-tonal music: Rhythmic contours in the music of edgard varèse. *Music Theory Spectrum*, *13*, 61–78.

Massaro, D. W., Kallman, H. J., & Kelly, J. L. (1980). The role of tone height, melodic contour, and tone chroma in melody recognition. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(1), 91–105.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science*, *19*(12), 1263–1271.

McDowell, J. H. (1999). The transmission of children's folklore. In B. Sutton-Smith, J. Mechling, T. W. Johnson, & F. R. McMahon (Eds.), *Children's folklore*. Utah: USU Press.

McNemar, Q. (1944). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153–157.

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(6468). https://doi.org/10.1126/science.aax0868

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of stroop effects. *Psychological review*, *110*(3), 422–71.

Melen, M., & Wachsmann, J. (2013). Categorization of musical motifs in infancy. *Music Perception*, *18*, 325–346.

Melucci, M., & Orio, N. (2002). A comparison of manual and automatic melody segmentation. *ISMIR*.

Merriam, A. P. (1960). Ethnomusicology: Discussion and definition of the field. *Ethnomusicology*, *4*(3), 107–114.

Meyer, L. B. (1956). *Emotion and meaning in music.* University of Chicago Press.

Meyer, L. B. (1989). *Style and music. theory, history and ideology.* Chicago: The University of Chicago Press.

Meyer, L. B. (1998). A universe of universals. *Psychological Science*, *16*(1), 3–25.

Mihelač, L. (2008). Glasbena podoba slovenske ljudske pesmi. *Ampak : mesečnik za kulturo, politiko in gospodarstvo*, *9*(12), 50–53.

Mihelač, L. (2012). *Nacionalna identiteta in glasba prišoloobveznih mladostnikih/national identity and music in primary school adolescents* (Doctoral dissertation). FPŠ - Fakulteta za podiplomski študij, Nova Gorica.

Mihelač, L. (2021). The role of songbooks in the preservation of childrens folk songs in kindergartens. *Journal of Elementary Education.*

Mihelač, L., & Panić Grazio, J. (2021). The classification of children's songs with the classification model cmcs. *The Journal of Music Education of the Academy of Music in Ljubljana*, *17*(35), 41–58.

Mihelač, L., Panićć Grazio, J., & Stefanija, L. (2018). An analysis of the slovenian terminology of musical forms. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, *44*(2), 551–566. https://doi.org/doi.org/10.31724/rihjj.44.2.15

Mihelač, L., & Povh, J. (2020a). Ai based algorithms for the detection of (ir)regularity in musical structure. *AMCS*, *30*(4), 761–772.

Mihelač, L., & Povh, J. Computational analysis of the musical diversity in 22 european countries. (L. Zadnik Stirn, M. Kljajić Borštar, J. Žerovnik, S. Drobne, & J. Povh, Eds.). In: In *Sor'21 proceedings* (L. Zadnik Stirn, M. Kljajić Borštar, J. Žerovnik, S. Drobne, & J. Povh, Eds.). Ed. by Zadnik Stirn, L., Kljajić Borštar, M., Žerovnik, J., Drobne, S., & Povh, J. Bled, 2021, September, 691–696.

Mihelač, L., Povh, J., & Wiggins, G. A. (2021). Segmentation of melody: A comparison of segmentation between children, musical experts and idyom. *Proceedings of ESCOM*, 963–964.

Mihelač, L., Povh, J., & Wiggins, G. A. (2022). Melody segmentation and breathing in children and adolescents. *Music & Science.*

Mihelač, L., Povh, J., & Wiggins, G. A. (2023). A computational approach to the detection and prediction of(ir)regularity in children's folk songs. *Empirical Musicology Review*, *16*(2), 205–230.

Mihelač, L., & Povh, J. (2020b). The impact of the complexity of harmony on the acceptability of music. *ACM Trans. Appl. Percept.*, *17*(1). https://doi.org/10.1145/3375014

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev*, *63*, 81–97.

Mithen, S. (2005). *The singing neanderthals: The origins of music, language, mind, and body.* London: Weidenfeld Nicolson.

Montagu, J. (2017). How music and instruments began: A brief overview of the origin and entire development of music, from its earliest stages. *Frontiers in Sociology*, *2*. https://doi.org/10.3389/fsoc.2017.00008

Moore, R. S. (1991). Comparison of children's and adults' vocal ranges and preferred tessituras in singing familiar songs. *Bulletin of the Council for Research in Music Education*, *107*, 13–22.

Morley, I. (2013). *The prehistory of music: Human evolution, archaeology, and the origins of musicality.* Oxford University Press.

Morris, R. D. (1993). New directions in the theory and analysis of musical contour. *Music Theory Spectrum*, *15*, 205–228.

Morrongiello, B. A., & Roes, C. L. (1990). Children's memory for new songs: Integration or independent storage of words and tunes? *Journal of experimental child psychology*, *33*(1), 25–38.

Morrongiello, B. A., Trehub, S. E., Thorpe, L. A., & Capodilupo, S. (1985). Integration of melody and text in memory for songs. *Journal of Experimental Child Psychology*, *40*(2), 279–292.

Müllensiefen, D. Exact measures of musical structure for predicting mem- exact measures of musical structure for predicting memory for melodies. (M. Baroni, A. R. Addessi, R. Caterina, & M. Costa, Eds.). In: In *Proceedings of the 9th international conference of music perception and cognition (icmpc9)* (M. Baroni, A. R. Addessi, R. Caterina, & M. Costa, Eds.). Ed. by Baroni, M., Addessi, A. R., Caterina, R., & Costa, M. Manchester, UK, 2006.

Nan, Y., Knösche, T. R., & Friederici, A. D. (2006). The perception of musical phrase structure: A cross-cultural erp study. *Brain research*, *1094*(1), 179–191.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication–realisation model.* University of Chicago Press.

Nettl, B. (1992). Ethnomusicology and the teaching of world music. *International Journal of Music Education*, *20*, 3–7.

Nettl, B., & Béhague, G. (1980). *Folk and traditional music of the western continents.* Englewood Cliffs, N.J., Prentice-Hall.

Nettl, B. (1983). *The study of ethnomusicology: Twenty-nine issues and concepts.* University of Illinois Press.

Nettle, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In B. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 463–472). MIT-Press, Cambridge, MA.

Neumann, F., & Stevens, J. R. (1993). *Performance practices of the seventeenth and eighteenth centuries.* Schirmer Books, New York.

Nevers, B., & Versace, R. (2003). Word frequency effects on repetition priming as a function of prime duration and delay between the prime and target. *British Journal of Psychology*, *94*, 389–408.

Newell, W. W. (1963). *Games and songs of american children.* New York: Dover Publications, Inc.

Newson, L., Richerson, P. J., & Boyd, R. (2007). Cultural evolution and the shaping of cultural diversity. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 454–476). New York, NY: Guildford Press.

Nograšek, M., & Virant Iršič, K. (2005). *Piške sem pasla*. Ljubljana: DZS.

Omigie, D., Pearce, M. T., & Williamson, V. J. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia*, *51*, 1749–1762.

Pachet, F. (2009). Description-based design of melodies. *Computer Music Journal*, *33*(4), 56–68.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 116–126.

Panteli, M. (2018). *Computational analysis of world music corpora* (Doctoral dissertation). School of Electronic Engineering and Computer Science, Queen Mary University of London.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach.* Springer Verlag, Berlin.

Patel, A. D. (2003a). Language, music, syntax and the brain. *Nature Neuroscience*, *6*, 674–681.

Patel, A. D. (2003b). Language, music, syntax and the brain. *Nature Neuroscience*, *6*, 674–681.

Patel, A. D. (2014). The evolutionary biology of musical rhythm: Was darwin wrong? *PLOS Biology*, *12*(3), 1–6. https://doi.org/10.1371/journal.pbio.1001821

Pauly, M., Mitra, N. J., Wallner, J., Pottmann, H., & Guibas, L. J. (2008). Discovering structural regularity in 3d geometry. [[Online; accessed 2018-10-20]].

Pearce, M., & Wiggins, G. A. (2004). Improved methods for statistical modelling of monophonic music. *Journal of New Music Research*, *33*(4), 367–385. https://doi.org/10.1080/0929821052000343840

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition* (Doctoral dissertation). Department of Computing City University, London.

Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences*, *1423*(1), 378–395. https://doi.org/10.1111/nyas.13654

Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2010a). The role of expectation and probabilistic learning in auditory boundary perception: A model comparison. *Perception*, *39*(10), 1365–1389.

Pearce, M. T., & Wiggins, G. A. (2006a). The information dynamics of melodic boundary detection. *Proceedings of the Ninth International Conference on Music Perception and Cognition, Bologna*.

Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, *4*, 625–652.

Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2010b). Melodic grouping in music information retrieval: New methods and applications. *Advances in music information retrieval* (pp. 364–388). Springer.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, *50*(1), 302–313.

Pearce, M. T., & Wiggins, G. A. (2006b). Expectation in melody: The influence of context and learning. *Music Perception*, *23*(5), 377–405.

Pearce, M., Ruiz, M., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, *50*(1), 301–313.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*, 1185–1205.

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition*, *100*(1), 1–32.

Peter, B., Stoel-Gammon, C., & Kim, D. (2008). Octave equivalence as an aspect of stimulus–response similarity during nonword and sentence imitations in young children. [[Online; accessed 2019-11-02]].

Petric, J. (2018). Music memory game [[Online; accessed 2018-4-9]].

Petzold, R. G. (1963). The development of auditory perception of musical sounds by children in the first six grades. *Journal of Research in Music Education.*, *11*, 21–43.

Petzold, R. G. (1966). *Auditory perception of musical sounds by children in the first six grades* (tech. rep.) [Cooperative Research Project No.1051]. The University of Wisconsin.

Phillips, K. H., & Vispoel, W. P. (1990). The effects of class voice and breath management instruction on vocal knowledge, attitudes, and vocal performance among elementary education majors. *The Quarterly*, *1*(152), 96–105.

Phillips, M., Stewart, A. J., Wilcoxson, J. M., Jones, L. A., Howard, E., Willcox, P., du Sautoy, M., & De Roure, D. (2020). What determines the perception of segmentation in contemporary music? *Frontiers in Psychology*, *11*, 1001. https://doi.org/10.3389/fpsyg.2020.01001

Platt, J. R., & Racine, R. J. (1994). Detection of implied harmony changes in triadic melodies. *Music Perception*, *11*, 243–264.

Pole, W. (2014). *The philosophy of music*. Routledge, Taylor & Francis.

Pond, D. (1981). A composer's study of young children's innate musicality. *Bulletin of the Council for Research in Music Education.*, *68*, 1–12.

Potter, K., Wiggins, G., & Pearce, M. (2007). Towards greater objectivity in music theory: Information-dynamic analysis of minimalist music. *Musicae Scientiae*, *11*(2), 295–322.

Povel, D.-.-J., & Jansen, E. (2002). *Music Perception*, *20*(1), 51–85.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009). The effect of task and pitch structure on pitch-time interactions in music. *Memory and Cognition*, *37*, 368–381.

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology Human Perception & Performance*, *35*(5), 1598–1617.

Raphael, C. (2001). Automated rhythm description. *Proceedings of the 2nd ISMIR, Bloomington*, 99–107.

Rees, H. (2011). *Echoes of history: Naxi music in modern china*. Oxford University Press.

Rhodes, W. (1965). He use of computer in the classification of folk tunes. *Studia Musicologica*, *7*, 339–343.

Rice, J. A. (2005). *Mathematical statistics and data analysis*. Belmont, CA: Cengage Learning.

Riemann, H. (1877). *Musikalische syntaxis*. Breitkopf und Härtel.

Rodriguez, J. D., Perez, A., & Lozano, J. A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 569–575. https://doi.org/10.1109/TPAMI.2009.187

Rogoff, B. (2003). *The cultural nature of human development.* New York: Oxford University Press.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, *5*(1), 35–53. https://doi.org/10.1080/17459737.2011.573676

Rohrmeier, M., & Pearce, M. (2018). Musical syntax i: Theoretical perspectives [[Online; accessed 2018-9-16]].

Rojek, C. (2011). *Pop music, pop culture.* Cambridge: Polity Press.

Romet, C. (1980). The play rhymes of children: A cross cultural source of natural learning materials for music education. *Australian Journal of Music Education*, *27*, 27–31.

Rzeszutek, T., Savage, P. E., & Brown, S. (2012). The structure of cross-cultural musical diversity. *Proc. R. Soc. B*, *279*, 1606–1612.

Saari, P., Burunat, I., Brattico, E., & Toiviainen, P. (2018). Decoding musical training from dynamic processing of musical features in the brain. *Scientific reports*, *8*(1), 708. https://doi.org/doi.org/10.1038/s41598-018-19177-5

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.

Sauvé, S. A., & Pearce, M. T. (2019). Information-theoretic modeling of perceived musical complexity. *Music Perception: An Interdisciplinary Journal*, *37*(2), 165–178. https://doi.org/10.1525/mp.2019.37.2.165

Savage, P. E. (2018). An overview of cross-cultural music corpus studies. In D. Shanahan, A. Burgoyne, & I. Quinn (Eds.), *Oxford handbook of music and corpus studies.* New York: Oxford University Press.

Savage, P. E. (2019). Cultural evolution of music. *Palgrave Commun*, *5*(16). https://doi.org/10.1057/s41599-019-0221-1

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2014). Statistical universals reveal the structures and functions of human music. *PNAS*, *112*(29), 8987–8992.

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. Statistical universals reveal the structures and functions of human music. (J. Ginsborg, A. Lamont, M. Phillips, & S. Bramley, Eds.). In: In *Proceedings of the national academy of sciences of the united states of america* (J. Ginsborg, A. Lamont, M. Phillips, & S. Bramley, Eds.). Ed. by Ginsborg, J., Lamont, A., Phillips, M., & Bramley, S. *112.* (29). Manchester, UK, 2015, 8987–8992.

Savage, P. E., Merritt, E., Rzeszutek, T., & Brown, S. (2012). Cantocore: A new cross-cultural song classification scheme. *Analytical Approaches to World Music*, *2*(1), 87–137.

Schaefer, R. S., Murre, J. M. J., & Bod, R. (2004). Limits to universality in segmentation of simple melodies. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *International conference on music perception and cognition (icmpc8).* Evanston, USA, Causal Productions.

Schaffrath, H. (1995). The essen folksong collection in the humdrum kern format.

Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General*, *131*, 511–537.

Schenker, H. (1935). *Free composition (der freie satz).* Universal Edition A. G.

Schoenberg, A., Strang, G., & Stein, L. (1967). *Fundamentals of musical composition.* London: Faber & Faber London.

Schoen-Nazzaro, M. B. (1978). Plato and aristotle on the ends of music. *Laval théologique et philosophique*, *34*(3), 261–273.

Schön, D., & François, C. (2011). Musical expertise and statistical learning of musical and linguistic structures. *Frontiers in psychology*, *2*, 167.

Selfridge-Field, E. (1998). Conceptual and representational issues in melodic comparison. In W. B. Hewlett & E. Selfridge-Field (Eds.), *Melodic similarity – concepts, procedures, and applications* (pp. 64–74). Cambridge, Massachusetts: MIT Press.

Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for songs. *Cognition*, *16*, 285–303.

Serafine, M. L., Davidson, J., & Crowder, R. G. (1986). On the nature of melody-text integration in memory for songs. *Memory and Language*, *25*, 123–135.

Serra, X. (2014). The computational study of a musical culture through its digital traces. *Acta Musicologica*, *89*(22-44).

Serrà, J., Corral, Á., Boguñá, M., Haro, M., & Arcos, J. L. (2012). Measuring the evolution of contemporary western popular music. *Scientific Reports*, *7*, 109–150.

Shaffer, D. R., & Kipp, K. (2009). *Developmental psychology: Childhood and adolescence.* Belmont:Wadsworth Publishing Company.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana, IL: University of Illionis Press.

Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, *5*(1), 3–55.

Silva, S., Branco, P., Barbosa, F., Marques-Teixeira, J., Magnus-Petersson, K., & Castro, S. L. (2014). Musical phrase boundaries, wrap-up and the closure positive shift. *Brain Research*, *1585*, 99–107.

Slobin, M. (1996). *Tenement songs:the popular music of the jewish immigrants.* Champaign, Illinois: University of Illinois Press.

Slobin, M. (2000). *Subcultural sounds: Micromusics of the west.* Hanover NH: University Press of New England.

Sloboda, J. A., & Gregory, A. H. (1980). The psychological reality of musical segments. *Canadian Journal of Psychology*, *34*(3), 274–280.

Sloboda, J. A., & Parker, D. H. H. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, mind and brain: The neuropsychology of music* (pp. 321–351). Plenum, New York.

Sloboda, J. A., & Parker, D. H. H. (1985). Immediate recall of melodies. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 143–167). Academic Press, London.

Smaill, A., Wiggins, G. A., & Harris, M. (1993). Hierarchical music representation for analysis and composition. *Computers and the Humanities*, *27*, 7–17.

Spiro, N. (2007). *What contributes to the perception of musical phrases in western classical music?* (Doctoral dissertation). Institute for Logic, Language and Computation Universiteit van Amsterdam. Universiteit van Amsterdam.

Stadler Elmer, S. Song singing: How children apply musico-linguistic rules or a grammar (J. Ginsborg, A. Lamont, M. Phillips, & S. Bramley, Eds.). In: In *Proceedings of the ninth triennial conference of the european society for the cognitive sciences of music* (J. Ginsborg, A. Lamont, M. Phillips, & S. Bramley, Eds.). Ed. by Ginsborg, J., Lamont, A., Phillips, M., & Bramley, S. Manchester, UK, 2015, August.

Stefani, G. (1987). Melody: A popular perspective. *Popular Music*, *6*(1), 21–35.

Steinruecken, C., Ghahramani, Z., & D. MacKay, D. (2015). Improving ppm with dynamic parameter updates. *2015 Data Compression Conference*, 193–202.

Stevens, C., & Gallagher, M. (2004). The development of mental models for auditory events: Relational complexity and discrimination of pitch and duration. *British Journal of Developmental Psychology*, *22*, 569–583.

Stevens, C. J. (2012). Music perception and cognition: A review of recent cross-cultural research. *Topics in Cognitive Science*, *4*, 653–667.

Stockmann, D. (1985). Musica vulgaris im französichen hochmittelalter. *Musikethnologische sammelbände 7* (pp. 163–180). Graz.

Sutton-Smith, B. (1999). What is children's folklore? In B. Sutton-Smith, J. Mechling, T. W. Johnson, & F. R. McMahon (Eds.), *Children's folklore*. Utah: USU Press.

Swain, J. (1997). *Musical languages*. New York: W.W. Norton; Company.

Temperley, D. (2001). *The cognition of basic musical structures*. MIT Press.

Tenney, J., & Polansky, L. (1980). Temporal gestalt perception in music. *Contemporary Music Review*, *24*(2), 205–241.

The Meertens Tune Collections. (2019). Dutch song database [[Online; accessed 2018-4-9]].

Thom, B., Spevak, C., & Höthker, K. (2002). Melodic segmentation: Evaluating the performance of algorithms and musical experts. *Proceedings of the International Computer Music Conference (pp. 65–72).*

Thompson, W. F., & Cuddy, L. L. (1989). Sensitivity to key change in chorale sequences: A comparison of single voices and four-voice harmony. *Music Perception*, *7*, 151–168.

Thorpe, L. A., Trehub, S. E., Morrongiello, B. A., & Bull, D. (1988). Perceptual grouping by infants and preschool children. *Developmental Psychology*, *24*(4), 484–491.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*(4), 885–913.

Tillmann, B., & Bigand, E. (1996). Does formal musical structure affect perception of musical expressiveness? *Psychology of Music*, *24*(1), 3–17. https://doi.org/10.1177/0305735696241002

Todd, C. L., & Sonkin, R. (1940–41). Voices from the dust bowl: The charles l. todd and robert sonkin migrant worker collection, 1940–41. [[Online; accessed 2023-5-5]].

Trask, B. S. (2010). *Globalization and families*. New York: Springer.

Trehub, S. E. (2000). Human processing predispositions and musical universals. In B. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 427–448). MIT-Press, Cambridge, MA.

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *J Exp Psychol Hum Percept Perform.*, *25*(4), 965–975.

Trehub, S. E. (1987). Infants' perception of musical patterns. *Perception and Psychophysics*, *41*(6), 635–641.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, *55*(3), 821–830.

Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, *100*, 73–99.

Trehub, S. E., Thorpe, L. A., & Trainor, L. J. (1990). Infant's perception of good and bad melodies. *Psychomusicology*, *9*, 5–19.

Tzanetakis, G., Kapur, A., Schloss, W. A., & Wright, M. (2007). Computational ethnomusicology. *Journal of Interdisciplinary Music Studies*, *1*(2), 1–24.

Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat Neuroscience*, *6*, 391–398.

Umemoto, T. (1990). The psychological structure of music. *Music Perception*, *8*(2), 115–127.

Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music*, *2*(1), 15–28.

van Kranenburg, P., Volk, A., & Wiering, F. (2012). On identifying folk song melodies employing recurring motifs. *Proceedings of the 12th International Conference on Music Perception and Cognition*, 1057–1062.

van Kranenburg, P. (2010). A computational approach to content-based retrieval of folk song melodies.

Vodušek, V. (1980). O evropski etnomuzikologiji 1. In J. Bogataj & S. Kremenšek (Eds.), *Poglavja iz metodike etnološkega raziskovanja* (pp. 375–401). Ljubljana: Filozofska fakulteta.

Voglar, M., & Nograšek, M. (2009). *Majhna sem bila*. Ljubljana:DZS.

Volchegorskaya, E., & Nogina, O. (2014). Musical development in early childhood. *Procedia - Social and Behavioral Sciences*, *146*, 364–368.

Volk, A. (2008). Persistence and change: Local and global components of meter induction using inner metric analysis. *Journal of Mathematics and Music*, *2*, 99–115.

Volk, A. (2016). Computational music structure analysis: A computational enterprise into time in music. In M. Müller, A. Chew, & J. P. Bello (Eds.), *Computational music structure analysis* (p. 159). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. https://doi.org/10.4230/DagRep.6.2.147

Volk, A., van Kranenburg, P., Garbers, J., Wiering, F., Veltkamp, R. C., & Grijp, L. P. (2008). A manual annotation method for melodic similarity and the study of melody feature sets. *Proceedings of the Ninth International Conference on Music Information Retrieval*, 101–106.

von Hippel, P. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception*, *17*, 1–13.

Vos, P. G. (2000). Tonality induction: Theoretical problems and dilemmas. *Music Perception: An Interdisciplinary Journal*, *17*(4), 403–416.

Vuvan, D. T., & Hughes, B. (2019). Musical style affects the strength of harmonic expectancy. *Music & Science*, *2*, 2059204318816066. https://doi.org/10.1177/2059204318816066

Wang, Y., Green, J. R., Nip, S. B., Kent, R. D., & Kent, J. F. (2010). The language of song: Some recent approaches in description and analysis. *Australian Journal of Linguistics*, *30*(1), 1–17.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*, 236–244.

Warner, T. (2003). *Pop music - technology and creativity (trevor horn and the digital revolution).* Oxfordshire, UK: Routledge.

Weiss, M., Trehub, S., & Schellenberg, E. (2012). Something in the way she sings: Enhanced memory for vocal melodies. *Psychological Science*, *23*(10), 1074–1078.

Welch, G. F. (1979). Vocal range and poor pitch singing. *Psychology of Music*, *7*(2), 13–31.

Wesch, M. (2018). The art of being human [[Online; accessed 2021-6-4]].

Wheeldon, L. R., & Smith, M. C. (2003). Phrase structure priming: A short-lived effect. *Language and Cognitive Processes*, *18*, 431–442.

Widmer, G. (2016). Getting closer to the essence of music: The con espressione manifesto. *ACM Transactions on Intelligent Systems and Technology*, *8*(2), Article 19. https://doi.org/DOI:10.1145/2899004

Wiggins, A. G., Pearce, M. T., & Müllensiefen, D. (2009). Computational modelling of music cognition and musical creativity. In T. D. Roger (Ed.), *The oxford handbook of computer music.* (pp. 383–420). Oxford University Press.

Wiggins, G. A. (2010). Cue abstraction, paradigmatic analysis and information dynamics: Towards music analysis by cognitive model. *Musicae Scientiae*, *Special Issue: Understanding musical structure and form: papers in honour of Irène Deliège*, 307–322.

Wiggins, G. A., & Sanjekdar, A. (2019). Learning and consolidation as re-representation: Revising the meaning of memory. *Frontiers in Psychology: Cognitive Science*, *10*(802). https://doi.org/10.3389/fpsyg.2019.00802

Williams, L. R. (2005). Effect of music training and musical complexity on focus of attention to melody or harmony. *Journal of Research in Music Education*, *53*(3), 210–221.

Yu, X., Liu, T., & Gao, D. (2015). The mismatch negativity: An indicator of perception of regularities in music. *Behavioural Neurology*, *2015*, 12.

Zatorre, R. J. (2001). Neural specializations for tonal processing. *Annals of the New York Academy of Sciences*, *930*(1), 193–210. https://doi.org/10.1111/j.1749-6632.2001.tb05734.x

Zimmerman, M. P. (1971). Musical characteristics of children. *Music Educators National Conference*.

Zonis, E. (1983). Classical iranian music. In E. May & M. Hood (Eds.), *Musics of many cultures: An introduction.* Berkeley: University of California Press.

Bibliography

Publications Related to the Thesis

Journal Papers

- Mihelač, L., Grazio, J., & Stefanija, L. (2018). An analysis of the Slovenian terminology of musical forms. *Rasprave Instituta za hrvatski jezik i jezikoslovlje.* 2018, vol. 44, no. 2, str. 551-566.
- Mihelač, L. & Povh, J. (2020). The impact of the complexity of harmony on the acceptability of music. *ACM TAP, 17*(1), 1-27.
- Mihelač, L. & Povh, J. (2020). AI based algorithms for the detection of (ir)regularity in musical structure. *International journal of applied mathematics and computer science, 30* 4), 761–772.
- Mihelač, L. & Panić Grazio, J. (2021). The Classification of Children's Songs with the Classification Model CMCS. *The Journal of Music Education of the Academy of Music in Ljubljana, 17*(35), 41–58.
- Mihelač, L. (2021). The role of songbooks in the preservation of childrens folk songs in kindergartens. *Journal of Elementary Education*, 15(3), 301–315.
- Mihelač, L., Povh, J. & Wiggins, G. A. (2021). A computational approach to the detection and prediction of(ir)regularity in children's folk songs. *Empirical Musicology Review*, 16(2), 205–230.
- Mihelač, L., Wiggins, G. A. & Povh, J. (2022). Melody segmentation and breathing in children and adolescents. Submitted to journal *Music & Science* and accepted.

Textbook

- Jožef-Beg, J. & Mihelač, L. (2019). *Na začetku je bila pesem/In the beginning there was a song.* Novo mesto: Šolski center Novo mesto.

Proceedings Papers

- Mihelač, L. & Povh, J. (2017). Predicting the acceptability of music with entropy of harmony. In *Proceedings SOR'17* (pp. 371-375).
- Mihelač, L., Wiggins, G. A., Lavrač, N., & Povh, J. (2018). Entropy and acceptability: information dynamics and music acceptance. In *Proceedings of ICMPC15/ ESCOM10* (pp. 313-317).
- Mihelač, L. & Povh, J. (2019). The impact of harmony on the perception of music. In *Proceedings SOR'19* (pp. 360-365).

- Mihelač, L. & Povh, J. (2021). Computational Analysis of the Musical Diversity in 22 European Countries. In *Proceedings SOR'21* (pp. 691-696).
- Mihelač, L., Povh, J. & Wiggins, G. A.(2021, July, 28-31). Segmentation of Melody: A Comparison of Segmentation Between Children, Musical Experts and IDyOM. *ICMPC16-ESCOM11.* https://icmpc2021.sites.sheffield.ac.uk/home

Presentations on International Conferences

- Mihelač, L. (2021, April, 16-18). *(Re)making tradition: the untold story of children's folk songs in Slovenian kindergartens.* Vlado S.Milošević: Tradition as inspiration, Banja Luka. https://au.unibl.org/index.php/sr-latn/novosti/naucni-skup-vlado-s-milosevic-etnomuzikolog-kompozitor-pedagog-tradicija-kao-inspiracija
- Mihelač, L. & Stefanija, L. (2021, March, 26–27) *Musical universals : the history of an idea.* Workshop : Life-world and musical form - concepts, models, and analogies, Salzburg. http://historiography-of-musical-form-through-mir.sbg.ac.at/cfp/
- Mihelač, L., Povh, J, & Wiggins, G. A. (2022). The detection of (ir)regularity in music and why it matters. ITIS 2022, Novo mesto.

Other

- Mihelač, L. (2018, June, 20).*Presentation: Napovedovanje slušne sprejemljivosti glasbe na osnovi entropije harmonije/Predicting the auditory acceptability of music based on harmony entropy.* IBMI (Institute for Biostatistics and Medical Informatics, University of Ljubljana, Faculty of Medicine), Ljubljana. https://ibmi.mf.uni-lj.si/sl/node/339
- Mihelač, L. (2023, January, 20).*Presentation: Percepcija (i)regularnosti v glasbi/Perception of (ir)regularity in music.* IBMI (Institute for Biostatistics and Medical Informatics, University of Ljubljana, Faculty of Medicine), Ljubljana. https://ibmi.mf.uni-lj.si/sl/node/432

Biography

Lorena Mihelač was born in 1963. She completed primary school in Rotterdam, Netherlands and then went on to start secondary school at Sint Montfort College (Program in Natural Sciences) in Rotterdam. She graduated from secondary school (General High school for Languages) and Music secondary school in Zagreb in 1981. In 1980, while still in secondary school, she enrolled at the Academy of Music in Zagreb, from which she graduated in 1984. In 1986, she began her postgraduate studies for a master's degree at the Academy of Arts in Novi Sad, Serbia. She completed the program in 1988, and went on to earn her Ph.D. in Intercultural Studies (Slovene Studies Program) in 2012 from ZRC SAZU in Ljubljana, Slovenia. In 2012, she began a new course of study at the Faculty of Information Studies in Novo mesto, Slovenia, and in 2015 she completed her diploma thesis. She continued her studies at the same faculty in 2015, earning her master's degree in 2017. In 2017, she decided to begin a new Ph.D. at the International Postgraduate School at the Jožef Stefan Institute, Slovenia (Program "Information and Communication Technologies") under the supervision of Prof. Dr. Dr. Geraint A. Wiggins and Prof. Dr. Janez Povh.

In 1985, she was employed as a junior researcher under the supervision of Prof. Dr. Smiljka Horga at the Department of Kinesiology Intonation – Applied music in Sport and Information Technology in Sport in Zagreb, Croatia. She remained in this position until 1991, when she returned to Slovenia due to war conditions.

Her research covers diverse fields, including sport and music, education science, and music psychology. From 2015 on, the focus is on music perception, particularly on the modelling of human music perception in children and adolescents.

She has coordinated various international projects for nearly two decades, focusing primarily on mechatronics, robotics, and electronics.

She is a partner and supervisor in SciDrom – IoT, ŠC Novo mesto, a scientific lab within the School center ŠC Novo mesto, addressing the topic "measuring of loudness" (in preparation for the Dolenjska region) and investigating the effect of music on human health (music therapy).