Revised Selected Papers

Accademia Musicale Studio Musica
Michele Della Ventura, *editor*

2024

Proceedings of the International Conference on New Music Concepts

Vol. 11

Accademia Musicale Studio Musica

International Conference on New Music Concepts

Proceedings Book
Vol. 11

Accademia Musicale Studio Musica
Michele Della Ventura
Editor

Published in Italy
First edition: April 2024

www.studiomusicatreviso.it
Accademia Musicale Studio Musica – Treviso (Italy)
ISBN: 978-88-944350-5-4

Preface

This volume of proceedings from the conference provides an opportunity for readers to engage with a selection of refereed papers that were presented during the International Conference on New Music Concepts. The reader will sample here reports of research on topics ranging from mathematical models in music to pattern recognition in music; symbolic music processing; music synthesis and transformation; learning and conceptual change; teaching strategies; e-learning and innovative learning.
This book is meant to be a *textbook* that is suitable for courses at the advanced undergraduate and beginning master level. By mixing theory and practice, the book provides both profound technological knowledge as well as a comprehensive treatment of music processing applications.
The goals of the Conference are to foster international research collaborations in the fields of Music Studies and Education as well as to provide a forum to present current research results in the forms of technical sessions, round table discussions during the conference period in a relaxable and enjoyable atmosphere.
61 papers from 15 countries were received. All the submissions were reviewed on the basis of their significance, novelty, technical quality, and practical impact. After careful reviews by at least three experts in the relevant areas for each paper, 18 papers were accepted for presentation or poster display at the conference.

I want to take this opportunity to thank all participants who have worked hard to make this conference a success. Thanks are also due to the staff of "Studio Musica" for their help with producing the proceedings. I am also grateful to all members of Organizing Committee, Local Arrangement Committee and Program Committee as well as all participants who have worked hard to make this conference a success.
Finally, I want to appreciate all authors for their excellent papers to this conference.

March 2024 Michele Della Ventura

International Conference on New Music Concepts 2024

Conference Chair

Michele Della Ventura, Accademia Musicale Studio Musica, Treviso, Italy

Keynote Speakers

Gus Xia, Music X Lab, Abu Dhabi
Zhengshan (Kitty) Shi, Stanford University, UK
Matteo Farnè, University of Bologna, Italy
Luca Turchet, University of Trento, Italy

International Scientific Committee

Islah Ali-Maclachlan, Birmingham City University, UK
Jioanne Armitage, University of Leeds, UK
Suzanne Aspden, Faculty of Music, University of Oxford, UK
Jean-Julien Aucouturier, IRCAM, Paris, France
Freya Bailes, University of Leeds, UK
Oded Ben-Tal, Kingston University, UK
Gilberto Bernardes, INESC TEC, University of Porto, Portugal
Per Bloland, Miami University, Ohio, USA
Jeffrey Boehm, Bath Spa University, UK
Eliot Britton, UTEMS Electronic Music Studio, University of Toronto, Canada
Jorge Calvo-Zaragozza, University of Alicante, Spain
David Carabias Galindo, University of Segovia, Spain
Shreyan Chowdhury, Johannes Kepler University Linz, Austria
Marko Ciciliani, University for Music and Performing Arts Vienna, Austria
Ana Clemente, University of Barcelona, Spain
Sally Jo Cunningham,University of Waikato, New Zealand
Ching-Hua Chuan, University of North Florida, U.S.A.
Roger Dannenberg, School of Computer Science, Carnegie Mellon University, USA
Darryl N. Davis, University of Hull, UK
Marlo De Lara, University of Leeds, UK
Elga Dorner, Central European University, Budapest, Hungary
Simon Emmerson, De Montfort University, Leicester, UK

Morwaread M. Farbood, New York University, USA
Travis Garrison, University of Central Missouri, USA
Inés María Monreal Guerrero, University of Valladolid, Spain
Andrew Hankinson, Bodleian Libraries, University of Oxford, UK
Joseph Hyde, Bath SPA University, UK
Wladyslaw Homenda, Warsaw University of Technology, Poland
Florent Jacquemard, Inria, CNAM-Cédric lab, Paris, France
Kelly Jakubowski, Durham University, UK
Orestis Karamanlis, Bournemouth University, UK
Kaustuv Kanti Ganguli, New York University Abu Dhabi, UAE
Alexandros Kontogeorgakopoulos, Cardiff Metropolitan University, UK
Amanda Krause, James Cook Unkiversity, Australia
Steven Jan, University of Huddersfield, UK
Audrey Laplante, Université de Montréal, Canada
Shengchen LI, Xi'an Jiaotong - Liverpool University, UK
Cory McKay, Marianopolis College, Canada
Emilia Parada-Cabaleiro, University of Linz, Austria
Tae Hong Park, New York University Steinhardt, USA
Geoffroy Peeters, LTCI - Télécom Paris, France
Lorenzo Porcaro, Universitat Pompeu Fabra, Spain
Rudolf Rabenstein, University Erlangen-Nuremberg, Erlangen, Germany
Silvia Rosani, Goldsmiths, University of London, UK
Robert Rowe, New York University, USA
Federico Simonetta, Gran Sasso Science Institute, Italy
Nikos Stavropoulos, Leeds Beckett University, UK
Jacob David Sudol, Florida International University, U.S.A.
Duncan Williams, University of Plymouth, UK

Contents

Sphere of Influence: Art, Music, or Science?

David Allan

Marketing
dallan@sju.edu

Abstract. For once, what happened in Vegas did not stay in Vegas. It exploded on the Sin City strip like an asteroid. It's called the Sphere and by most accounts it's a spectacle to be both seen and experienced. It is the latest attempt to align the planets of art, music and science. The following is a perspective piece to provide academic inspiration to add to the popular press. Some considerations include art v. commerce, atmospherics, performance, plus spectacle and intimacy. Let this article begin the show.

Keywords. Art, Las Vegas, Music, Science, Sphere

1 Introduction

It opened on September 29, 2023 at the Venetian Resort in Las Vegas. It is 366 feet high and 516 feet wide. It is the largest spherical building in the world at 875,000 square feet. It seats over 18,000 people and all seats have high speed internet access. It was built by Madison Square Garden Entertainment over five years at a cost of $2.3 billion. It was the Sphere, "a hyperstimulating new performance venue in which the whole exterior is a screen, and the whole interior as well" [1]. U2 was the first performer in a residency titled U2: UV Achtung Baby Live at Sphere (Figure 1).

Fig. 1. The Sphere.

The city unofficially known as “Sin City” where “What happens in Vegas, stays in Vegas” was founded in 1905. Las Vegas (“the meadows” in Spanish) is city of 2,899,000 people. “Las Vegas’ transformation from local watering hole to world renowned fame can be attributed to a series of turning points in its history, with influences ranging from geology, government, organized crime, entertainers, the economy, business visionaries and large corporations” [2]. Las Vegas is home to 175 casinos, 100 sports betting parlours, 95 casino hotels, 349 venues and 664 restaurants not to theme and amusement parks. It has been called a city where “the simulacrum of glamour available to everyone ensures that no one gets the real thing [3].

Formed in 1976, U2[1] are an Irish rock band from Dublin, Ireland consisting of Bono (lead vocals and rhythm guitar), the Edge (lead guitar, keyboards, and backing vocals), Adam Clayton (bass guitar), and Larry Mullen Jr. (drums and percussion). They have recorded 15 studio albums, one live album, three compilation albums, 84 singles, and nine extended plays (EPs). They have won more Grammys for a rock band than any other band (22). They have currently have performed live 2132 times including the most recent performance in Las Vegas, Nevada. The New York Times review of their performance on opening night at the Sphere said: “a band unafraid of pomp and spectacle was sometimes out-pomped and out-spectacled” [4].

2 Themes for consideration

The goal of this section is to stimulate some theoretical thought about the Sphere by suggesting literature that has historically been applied to music including art vs. commerce, atmospherics, performance, spectacle, and intimacy each with my own contributions for future consideration.

2.1 Art v. Commerce

It has been and will continue be the first consideration whenever art and commerce (and in this case “science & art”) collide. The Sphere not only does not deny it, it gives it a spotlight.[2] It is a bit curious that “Science” (Table 1) is much larger than “Art-ists” (Table 2) on its website. Of course, the “artists” part is music. Music is after all, a social and political indicator that mirrors and influences the society we live in [5]. Whenever music and commerce are mentioned together there will popular culture implications. To its proponents, it is a cultural product that entertains and inspires large segments of society by providing meaningful and chronological reference points. To its opponents, it is part of a vast economic system that hypnotizes and massifies segments of

[1] U2.com

[2] https://www.thespherevegas.com/

consumers through manipulation and commodification. It is, after all, popular music [6]. But, in this instance, is the science the show and not the music? It seems like U2 may think so, as U2's creative director Willie Williams said "if the audience gets bored they can always look at the band" [7]. Future Research:

1) Art, Music, Science positive vs. negative considerations.
2) Is Science v. Commerce better or worse than Art v. Commerce?

TABLE 1: SCIENCE[3]

Area of Sphere
Volume of Sphere
Finite Element Method
Geodesic Math
Laws of Sines
Pi
Stereographic Projection
Visual Acuity
Snell's Law
Lens Projection Formulas
Fanger's Equation
Linear Stress Constitutive Equation
Venturi Effect
Navier-Stokes Equations
Shannon-Hartley Theorem
Huygens-Fresnel Principle
Helmholtz Equation
Kirchhoff-Helmholtz Integral
Kirchhoff Integral Theorem
Wave Equation
Hartley's Law
Laplacian
Cauchy Momentum Equation
Sellmeier Equation

[3] https://www.thespherevegas.com/science

TABLE 2: ART: ARTISTS[4]
Refik Anadol
Darren Aronofsky
U2

2.2 Atmospherics

What if music in this case is just an atmospheric and not the main attraction? Is that a bad thing? Certainly, popular culturalists would think so since it would suggest that music is again a tool of commerce. Kotler [8] introduced the concept of *atmospherics* (including sound) as a controlled marketing tool in relation to indoor space. Bitner [9] developed this theme further with the *servicescape* concept, highlighting music as one of many physical environment dimensions (including temperature, air quality, scent, etc.) that impact upon cognitive, affective, and behavioural responses of consumers and employees within service environments. Music does play an important role. Kozinets [10] suggested that background music contributes to environmental aesthetics by transforming shopping environments and making them more attractive to consumers.

However, while the servicescape and its environmental aesthetics are relatively controllable indoors, many servicescape variables (e.g. temperature) are much less controllable when aspects of the service offer are performed outdoors. Oakes [11] has coined the term 'musicscape' to describe a framework summarizing the benefits of high levels of congruity between musical genre and other servicescape variables. to enhance evaluation of the service environment Oakes [12]. However, ensuring congruity between music and other servicescape variables is much more problematic in an outdoor, urban context. The concept of heterogeneity (drawn from the services marketing literature) provides useful theoretical grounding to link and discriminate between live, pre-recorded, indoor, and outdoor music. Live concerts are high in heterogeneity due to the inevitable unpredictability of live performance, whereas pre-recorded servicescape music is low in heterogeneity. The role of music in the marketing of outdoor, urban spaces in terms of three distinct continua: managed/spontaneous, spectacular/mundane, and exclusive/inclusive. The impact of music on their perceptions of the nature of urban space [13] However, the value and potential impact of music (and other performing arts) in terms of changing the character of outdoor urban space, as highlighted by Whyte [14], has been relatively neglected by the marketing literature.

Future Research:

1) Is music just an atmospheric at the Sphere?
2) Can science be an atmospheric?

[4] https://www.thespherevegas.com/artists

3) The Sphere has science both indoors and outdoors, what are the implications?

2.3 Performance

Performance has always been the driving force for the success of live music. It defines a social or communicative process that requires an audience and is dependent on interpretation; depends on an audience which can interpret work through its own experience of performance; listening" is a performance [15]. But to think of music as performance is therefore to focus on how meaning is created in real time-in the act of performing it, and equally in the act of hearing it, whether live or on a recording [16]. There are lots of ways of making sense of music as performance, and lots of senses there for the making. It really is as simple and a complicated as that [17]. For it to be considered "excellent" it involves a genuine understanding of what the music is about, its structure and meaning. Therefore, evaluation of performance at the Sphere must at least consider the music along with the science.

Future Research:

1) Is the performance at Sphere music or science or both?
2) Are art, music and science one at the Sphere?

2.4 Intimacy and Spectacle

There is probably no better example and/or collision of intimacy and spectacle than at the Sphere. They are traditionally considered as seemingly opposite methods by which artists can affect their audiences. They are "grandeur-a principle of force based on the sheer size and power of the impression; and Intimacy-a principle of attraction that demands a closeness of association between subject and audience and necessitates identification and involvement [19]. Of course, Las Vegas has always been an environment of non-verbal spectacle, providing visual stimulation through lights, animation, colors, symbols, and images [20]. To evaluate the Sphere in this regard perhaps we need to use an interdisciplinary model to the concert spectacle analysis that includes content, structure, and design as they intersect to communicate cultural messages and emotional narratives thus creating a sense of intimacy [21]. But, it is conceivable that music can never explain the visual spectacle alone. It only contributes to it by form and content, subjectivity and objectivity, and self and other [22].

Future Research:

1) Does Sphere science facilitate the perfect balance of music intimacy/spectacle?
2) Are art, music and science one at the Sphere?

3 Conclusion

The Sphere has already garnered a good deal of global popular press attention. It needs more academic attention. It has great potential for influence. Is it good for the art and music or just the science? Hopefully, this piece will stimulate thought and scholarship.

References

[1] J. Caramanica, “U2 Returns, in Las Vegas Limbo,” *The New York Times*, https://www.nytimes.com/2023/09/30/arts/music/u@-sphere-las-vegas.html, September 2023.

[2] B. Richard, Las Vegas: Past, Present, and Future. *Journal of Tourism Futures*, vol. 4, 3 pp. 182-192, 2018.

[3] J. Caramanica, “U2 Returns, in Las Vegas Limbo,” *The New York Times*, https://www.nytimes.com/2023/09/30/arts/music/u@-sphere-las-vegas.html, September 2023.

[4] https://www.nytimes.com/2023/09/30/arts/music/u2-sphere-las-vegas.html, 2023.

[5] J. Caramanica, “U2 Returns, in Las Vegas Limbo,” *The New York Times*, https://www.nytimes.com/2023/09/30/arts/music/u@-sphere-las-vegas.html, September 2023.

[6] R. Garofalo, Pop goes to war, 2001–2004: US popular music after 9/11. In: Ritter J, Martin Daughtry J (eds) Music in the Post-9/11 World. New York; London: Routledge, pp. 3–26, 2013.

[7] D. Allan, “An Essay on Popular Music in Advertising: Bankruptcy of Culture or Marriage of Art and Commerce,” *Advertising and Society* vol. 6, 1 2005.

[8] CBS Sunday Morning https://www.cbs.com/shows/video/jvGAJO263P7UXqTSqP9fCVjGlT1eL7tQ/ , September 30, 2023.

[9] P. Kotler, Atmospherics as a Marketing Tool. *Journal of Retailing*, 49, 48-64, 1974.

[10] M. J. Bitner, Servicescapes: The impact of physical surroundings on customers and employees, *Journal of Marketing*, vol. 56, 2, pp. 57–71, 1992.

[11] R. V. Kozinets, “The field behind the screen: Using netnography for marketing research in online communities,” *Journal of Marketing Research*, vol. 39, 1, pp. 61-72, 2002.

[12] S. Oakes, “The Influence of the Musicscape within Service Environments” *Journal of Services Marketing* vol. 14: pp. 539–56, 2000.

[13] S. Oakes, “Musical Tempo and Waiting Perceptions,” *Psychology & Marketing* vol. 20, 8, pp. 685–705, 2003.

[14] S. Oakes and G. Warnaby, “Conceptualizing the management and consumption of live music in urban space,” *Marketing Theory*, vol. 11, 4), pp. 405-418, 2011.

[15] W. H. Whyte, The Social Life of Small Urban Spaces. Washington, DC: Conservation Foundation, 1980.

[16] S. Frith, Performing Rites: On the Value of Popular Music. Cambridge, Mass.: Harvard University Press, 1996.

[17] N. Cook, “Between art and science: Music as performance,” *Journal of the British Academy*, vol. 2, pp. 1-25, 2014.

[18] N. Cook, Beyond the Score. New York: Oxford University Press, 2013.

[19] A. Gabrielsson, The Performance of Music, In Cognition and Perception, The Psychology of Music (Second Edition), Chapter 14, pp. 501-602, ed. Diana Deutsch, Cambridge Mass: Academic Press, 1999.

[20] J. S. Siegel, Grandeur-Intimacy: The Dramatist’s Dilemma. Diderot Studies, 4, 2 pp. 47-260, 1963.

[21] J. M. Luo and C.F. Lam, Entertainment Tourism. Abingdon: Routledge, 2018.

[22] L. Burns, and J. Watson, Spectacle and intimacy in live concert film: lyrics, music, staging, and film mediation in P!nk's Funhouse Tour (2009). *Music, Sound, and the Moving Image*, vol. 7, 2, pp. 103-140, 2013.

[23] S. Hawkins, Virtual insanity or the real thing? In “The British Pop Dandy”, Chapter 3. London: Routledge, 201

Sonification of 3D Protein Structures Using Supervised Machine Learning

Isabel Ronan, Yanlin Mi, Venkata Vamsi Bharadwaj Yallapragada, Cárthach Ó Nuanáin and Sabin Tabirca

University College Cork
firstinitial.lastname@cs.ucc.ie
Munster Technological University
firstname.lastname@mtu.ie

Abstract. Proteins are intricate structures that can be analysed by biologists and presented to the public using visualisations. However, with an increase in the amount of readily available protein-related information, new forms of data representation are needed. Sonification offers multiple advantages when conveying large amounts of complex data to interested audiences. Previous attempts have been made to sonify protein data; these techniques mainly focus on using amino acid sequences and secondary structures. This paper proposes a novel protein sonification algorithm involving atomic coordinates, B-factors, and occupancies to investigate new ways of displaying 3D protein structure data. This study culminates in creating a cultural showcase involving some of nature's most significant molecular structures. Results of both musical analysis and the showcase indicate that protein sonification has the potential to act as a helpful outreach and engagement tool for biologists while also helping experts in the field glean new insights from complex data.

Keywords: machine learning, knn, protein, sonification, music

1 Introduction

Sonification is the transformation of data into non-speech audio to facilitate communication and interpretation [1]. Sonification is naturally interdisciplinary; it can merge the arts with a plethora of different sciences. Recently, sonification has grown in popularity. With the advent of big data, sonification is used to make large volumes of information comprehensible [2]. Additionally, sonification can make complex, non-textual data accessible for those who are visually impaired [3].

Human beings are inherently able to respond to sound [3]. Due to the subjective nature of the communication medium, we can find patterns in the audio environments presented to us. This inherent ability to glean new insights from otherwise unremarkable data makes sound an ideal communication medium for multidimensional, complex data. Sonification can also be used as an outreach tool for education and public engagement (EPE) to bring complex protein structures to a larger audience.

This paper proposes a novel sonification method for protein data involving supervised machine learning. The algorithm proposed in this paper is one of the first to determine musical pitch based on 3D protein structures in relation to their centres of mass. Additionally, B-factors and occupancies of each atom within a protein are used. B-factors measure atom oscillation amplitudes around their equilibrium [4]. Occupancies monitor atom presence at its mean position [5]. As B-factors and occupancies are related, our proposed algorithm aims to link both factors to determine note duration [4]. Furthermore, this algorithm is used to generate pieces for an artistic showcase to demonstrate the enjoyable and educational nature of protein data exploration.

The main contributions of this paper are the definition of a novel sonification algorithm, the musical analysis of resultant sounds to discern structural features of proteins and the use of protein sonification in a real-life outreach application to assess its effectiveness. The rest of this paper is divided into sections regarding related work, methodology, algorithmic results, and conclusions.

2 Related Work

Sonification has many applications and can be used in various fields. Sonification has been explored in depth within the context of sports, select forms of scientific communication, and healthcare [6], [7], [8]. The Sonification Handbook [9] gives a comprehensive overview of such auditory processes; this book defines sonification as the transformation of data into acoustic signals to facilitate communication or interpretation. Additionally, the aforementioned book outlines various types of sonification and musification, such as parameter mapping [9]. With regards to parameter mapping, there often exists a compromise between accuracy and aesthetics [9]; this compromise is of particular concern when dealing with complex data. Dubus and Bresin's systematic literature review on physical quantity sonification [10] revealed that the most common motivations for data sonification include data exploration, aesthetics, accessibility, motion perception, monitoring, complements to visualisation, and psychoacoustic study. Low-level synthesis, sample-based synthesis, musical sounds, and MIDI are the most common musical products of sonification processes.

Many preexisting studies regarding protein sonification focus heavily on using amino acids to generate music. Dunn and Clark [11] proposed a sonification method involving analysis of the primary amino acid sequence and the folding patterns of proteins. They used changes in instrumentation and pitch to mark region turns within a protein. They arranged protein amino acids in order of water solubility and subsequently created solubility scales to map each amino acid to a pitch. Their project resulted in an artistic CD to provide empirical proof of the aesthetic patterning of proteins. Their protein music contains rich counter melodies and tune offsets, which create deeply intricate sonic motifs.

Yu, Qin, Martin-Martinez and Buehler [12] demonstrated the use of the natural vibrational frequencies of amino acids in creating playable musical instruments. These amino acids were subsequently mapped to piano keyboard notes to create musical sequences. They used Python scripts to translate any sequence into a musical score. Additionally, the secondary structure of the protein geometry was computed using DSSP; this secondary structural information is used to determine note duration and volume. The hierarchical nature of a protein is subsequently captured by mapping these features to music; this process is furthered by training a recurrent neural network (RNN) on this protein music. They highlight the need for further work to generate new proteins. They also report on the outreach potential found in such music; a broader range of individuals could understand proteins and their importance by using sound-based representations.

Martin, Meagher, and Barker [13] developed five different sonification algorithms for protein sequences. Their algorithms used amino acid, entropy, and hydrophobicity features when mapping protein data to musical notes. They also used a reduced alphabet of four letters to decrease the complexity of the amino acid sequences. They subsequently tested their algorithms for effectiveness with a range of bioinformatics participants. The results of their usability tests indicated that such sonification can convey particular protein features; these representations lead to heightened interest in proteins.

Gena and Strom [14] created musical pieces from DNA files by converting the list of 64 codons into musical events according to their physical properties. The dissociation constant of each amino acid was used to define pitch. The hydrogen bonding in each codon determined the velocity of each note. Dissociation constants, along with the atomic weights of amino acids, are used to determine note duration. Their work describes one of the most intricate mappings of protein properties to music.

Meytin [15] proposed three sonification methods involving amino acid molecular vibrations and manual pitch assignments. They found that it is possible to sonify proteins consistently for future scientific applications. They also highlighted artistic implications for protein structure sonification; the highly patterned nature of the results can be applied to musical composition to create novel forms of scientific entertainment.

Bouchara and Montès [16] have proposed an immersive sonification method involving protein surfaces. Their method involves decimating protein point clouds and using immersive spatialisation to place users at the centre of the protein for analysis. Their work consists of the use of interactive features to filter data. Subsequent results find consistent point position and density-related audio patterns.

Although amino acid sequences are often used to sonify proteins, their geometrical structures have yet to be explored in depth. While surface-based mapping techniques have been developed, further work is needed to analyse these structures thoroughly. Using machine learning to enable efficient parameter mapping has yet to be thoroughly investigated. Additionally, while many studies discuss the outreach capabilities of such sonification, there is a lack of work surrounding the use of public engagement to assess general enthusiasm for such projects.

3 Methodology

The parameter-mapping algorithm developed for this project involves transforming atom coordinate points within a protein structure into meaningful measurements using the protein's centre of mass. A protein's centre of mass is an artificial point for detecting structural features, shapes, and associations [17]. This paper outlines one of the first sonification methods, which uses centre of mass to discern meaning from atomic coordinates within a 3D protein structure. Ten different biologically significant protein structures are chosen for testing and presentation purposes.

Music generation is carried out using Python scripts. Post-processing is subsequently carried out using Logic Pro. Protein PDB files downloaded from the RCSB Protein Data Bank were used in the music generation process [18]. Biopython, pretty midi, mir eval, scikit-learn and Matplotlib libraries were used to organise protein material, visualise algorithmic strategies, analyse resultant sounds and generate appropriate musical outputs for further aesthetic processing.

The protein-to-music conversion process is outlined as follows:

- Step 1: Extracting Protein Data,
- Step 2: Labelling Training Data,
- Step 3: Training KNN Model,
- Step 4: Predicting Notes and Creating MIDI,
- Step 5: Analysing Music and Selecting MIDI,
- Step 6: Post-Processing.

3.1 Extracting Protein Data

The algorithm developed involves extracting the centre of mass, atomic coordinates, B-factors, and occupancy measures from a protein PDB file. Note durations are calculated based on the B-factors and occupancies of the protein points used. Specifically, B-factors are normalised with respect to their occupancies. Table 1 outlines the quantisation thresholds for B-factor occupancies.

TABLE I. B-FACTOR VALUES.

B-Factor Threshold	Mapped Amount
b < 25	0.25
25 <= b =< 50	0.5
50 < b < 75	0.75
b >= 75	1.0

Once the centre of mass and all of the coordinates of the protein are calculated, a standard 3-dimensional distance formula is used to discern the distances between all points. The distance measurements calculated are used for both pitch predictions and velocity note parameter mappings. This distance formula for d is outlined below where (x2, y2, z2) is an atomic coordinate and (x1, y1, z1) is the center of mass of the protein.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

3.2 Labelling Training Data

The protein point distances are then divided into eight different distance bands representing the eight notes of a C major scale. The eight musical bands are evenly arranged around the centre of mass; the furthest point from the centre of mass is used to calculate the distribution. The points at each of the eight distances from the centre of mass are mapped to their corresponding notes. This labelled data is then used to train a KNN machine-learning model to assign notes to the rest of the points in the protein. The KNN-supervised machine-learning technique is used due to its simplicity. Each protein analysed and converted into music will have its own KNN model trained specifically for note-mapping purposes. Figure 1 shows the labelled protein data and the data for KNN categorisation.

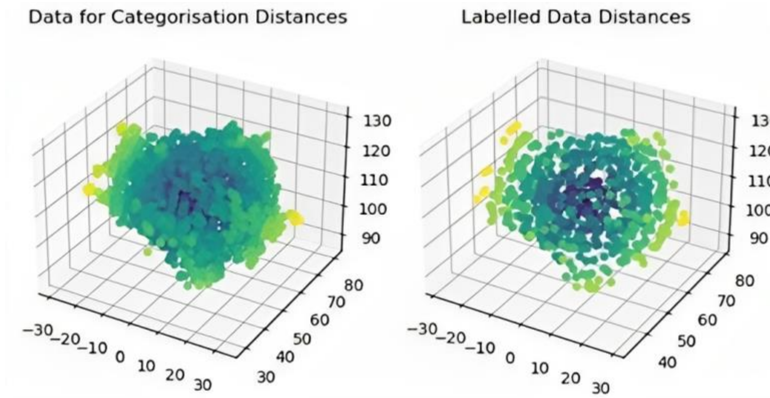


Fig. 1. Unclassified Versus Labelled Data.

3.3 Training KNN Model

The labelled data is divided into training and testing sets. Subsequently, the KNN model is trained; two neighbours were found to be the optimal number for this algorithm after multiple training and testing cycles. The average accuracy of the KNN models trained for each protein was 99.3%. Results for an exemplary selection of proteins mapped are outlined in Table 2.

TABLE II. KNN ACCURACIES.

Protein	KNN Accuracy
1N8Z	99.19%
3BIK	99.13%
3J6R	99.37%
6PV7	99.51%

3.4 Predicting Notes and Creating MIDI

Once all protein points have been assigned a note, the protein is traversed in 20 arbitrary directions in a scan-like fashion. These 20 random scans are used to organise the notes as they are written to MIDI files. Protein distance values related to each assigned pitch are used as velocity values for the MIDI notes. Additionally, the B-factor-occupancy values calculated during the earlier stages of the algorithmic process are used to assign durations to notes. Figure 2 provides a conceptual representation of how this process was undertaken.

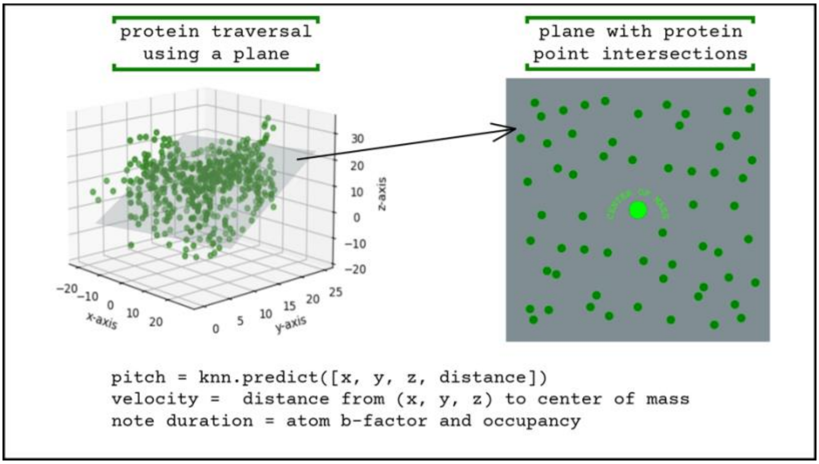


Fig. 2. Protein Point Plane Traversal.

3.5 Analysing Music and Selecting MIDI

Each of the 20 MIDI files is musically analysed based on duration, tempo, rhythm, and pitch stability. Duration is assessed using the pretty-midi library's get end time() function. Piece length is categorised as 'appropriate' when it is of one to four minutes in duration. The piece's tempo is calculated using prettymidi's estimate tempo() function. This function returns a beats-per-minute amount; the higher the score, the quicker the piece and the more exciting the music. Additionally, pretty-midi's get pitch class histogram() function calculates the strongest pitch within a piece. The strongest pitch number is then added to the total score with respect to its strength. The higher the pitch score, the more stable and aesthetically pleasing the pitch centre is in the piece. Rhythmic consistency is calculated using the mir eval library's information gain() function.

The resultant scores for each musical feature are combined to create an overall score. This overall score determines which MIDI file is the most 'musical' and should subsequently be subject to post-processing. The MIDI file with the highest overall score is used to represent the protein.

3.6 Post-Processing

The MIDI files chosen during the musical analysis phase are inputted into Logic Pro for further processing. Post-processing ensures that aesthetically appealing soundscapes are generated from the protein MIDI data. The pitch and note duration values outputted during the MIDI generation phase remain untouched during post-processing; however, each MIDI file is truncated to facilitate optimal performance length. All post-processing decisions are made to facilitate the optimal mix of aesthetic appeal and accuracy. Instruments and sound effects are all chosen with the relevant protein in mind. For example, as shown in Figure 3, breath sound effects are added to the music of the Human Oxyhaemoglobin protein 1HH0 [19] to facilitate an oxygen-focused soundscape. Additionally, reverb, compression, and EQ plugins were added to create ambient atmospheres.

4 Results and Discussion

Both musical analysis and public performance assessments support the feasibility of the proposed sonification algorithm. Musical analysis revealed insightful patterns within the proteins. A public performance undertaken as part of this research showed the use of biological music within an outreach context. Results provide justification for future work, both within scientific and artistic contexts.

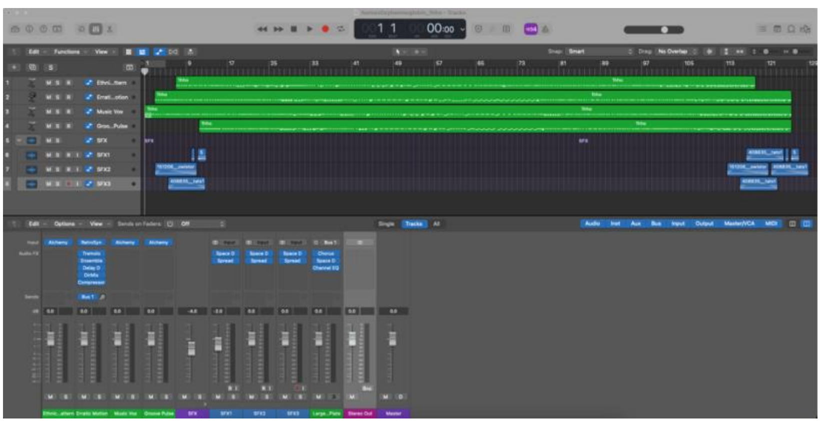


Fig. 3. Human Oxyhaemoglobin in Logic Pro

4.1 Musical Analysis

The results of this algorithmic process were empirically analysed by the authors of this paper to discern meaning from the music. Observations highlighted the underlying natures of the proteins themselves. For example, the tobacco mosaic virus [19] generated a strong pulsing rhythm; one can coincidentally observe the mottled mosaic-like appearance of infected plants in the music.

All protein music was generated using MIDI notes related to the C major scale. The musical notes used, along with their corresponding MIDI numbers, are outlined in Table 3.

TABLE III. MIDI MUSICAL NOTES.

Musical Note	C_4	D_4	E_4	F_4	G_4	A_4	B_4	C_5
MIDI Number	60	62	64	65	67	69	71	72

Each protein used within this study generated music with a different key centre. The key centres discerned from the post-processed protein tracks are outlined in Table 4. From these results, one can see that C minor is the most popular key centre. Figure 4 shows the C minor pitch histogram of the Immunoglobulin protein molecule.

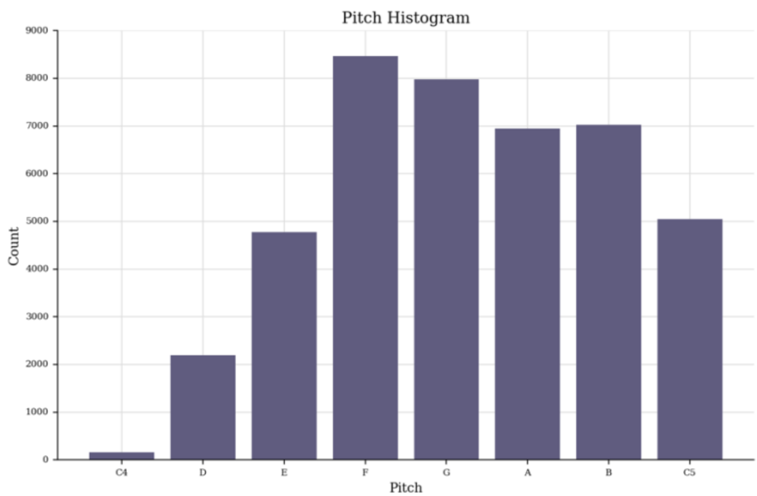


Fig. 4. Pitch Histogram of Immunoglobulin.

Additionally, many human-related proteins tend to focus on the same key; for example, the music generated from human alpha thrombin and genome proteins results in the same pitch centre. Key signatures involving accidentals, such as Pig Insulin's F# minor, are attributed to post-processing effects and not to failures on the part of the algorithm itself.

The structure of all MIDI files generated revealed the symmetrical nature of many proteins. As each piece progresses, the different musical lines get increasingly more complex. After intense sonic complexity, the sounds become more sparse and aesthetically spacious. Additionally, there are pitch descent and ascent-related patterns present in each file; these patterns vary according to the file chosen during the music analysis section of the proposed algorithm.

Most pieces assessed as part of this project follow a similar density and pitch trajectory. One can infer from the musical representations of structures analysed that most proteins are densest at their centre of mass. One can also infer that proteins have fewer atoms around the edges of their structure. Figure 5 exhibits the symmetrical MIDI file shape of 1FQY as seen in Logic Pro. This figure clearly shows that the protein structure grows in density as the music moves closer to the centre of mass of the molecule.

In summary, the protein sounds generated exhibit clear musical patterns. The significance of this work lies within the use of a novel structure-focused sonification algorithm; this algorithm has clearly exhibited the ability to extract patterns from protein data. From these patterns, one can recognise 3D structural features of proteins. Future work could focus on establishing links between structures based on their musical features. Additionally, the parameter mapping techniques explored in this study could be further investigated within the context of "de novo" protein design; if proteins can be made into music, music can be made into proteins. Further investigation could lead to significant advances within the field of molecular biology. Both within educational and innovative contexts, sonification could help scientists glean new insights into pre-existing protein data and generate new proteins using music.

TABLE IV. PROTEIN KEY CENTRES.

Protein Name	Protein Code	Key Centre
Immunoglobulin	1igt	C minor
Human Oxyhaemoglobin	1hho	C minor
DnaB Helicase	4esv	C minor
Tobacco Mosaic Virus	2tmv	E minor
Serotonin	4iar	E minor
Human Alpha Thrombin	1ppb	A minor
Genome	4un3	A minor
Zika Virus	5ire	G minor
Pig Insulin	4ins	F# minor
Water Channel	1fqy	C major

Fig. 5. 1FQY MIDI in Logic Pro.

4.2 Public Performance

An artistic showcase was prepared to present the results of the developed algorithm to interested members of the public. This outreach event assessed the significance of algorithmically-generated soundscapes within a public engagement context. The event took place during Culture Night 2023, an annual, nationwide event developed by The Arts Council that celebrates cultural institutions and their activities[1]. The "Molecular Melodies" exhibition comprised three components:

- A quadraphonic audiovisual installation of a selection of the musical pieces accompanied by 3D visuals
- Physical 3D printed models of the protein structures along with descriptive materials
- Interactive virtual reality experiences with 3D models

The protein sounds presented at the showcase were chosen based on their importance in nature. Each protein was a well-known molecule from biological history or from key functional processes within human beings and animals. Each piece was roughly three to four minutes in duration. All generated music was made available on SoundCloud, so interested members of the general public could access the pieces [20]. The event was open to all members of the public; interested observers were free to view 3D models of the proteins in virtual reality to accompany the sonic soundscapes. Figure 6(a) shows the layout of the installation space for the event, and Figure 6(b) shows some interested audience members experiencing protein in VR to accompany the music.

(a) (b)

Fig. 6. 3D Soundscape Event.

The audience response to the showcase was very positive. People of all ages attended the event. Members of the public engaged actively with those involved in the project and were eager to learn more about the same. Those involved in the project presentation

[1] https://culturenight.ie/about/

were present to answer questions and assist members of the public in understanding the showcase. Informal feedback was gleaned from audience members through the use of a small comment book placed at the entrance of the room. 46 responses were collected from the comment book. Some of the most detailed responses highlighted future applications of the showcase.

"*I'm amazed by how science and music can be brought together so skillfully. One has to be probably very passionate about both. I'd love to have this introduced to schools as well to make biochemistry more relatable and appealing*".

Additionally, audience members focused on the comprehensibility of the showcase; they outlined the positive impact the generated music had on understanding the protein visualisations.

"*I don't have a science background and this made it much easier for me to visualise proteins*".

In summary, the showcase was very successful, emphasising the future educational and entertainment-related advantages of such outreach activities. The use of such an event is significant within the field of protein sonification as, to the best of our knowledge, no prior related work has implemented an artistic showcase utilising experimental sonic results. This paper suggests that through artistic events, proteins can be made more appealing, entertaining, and understandable to the general public. Future work could focus on conducting more rigorous usability tests within a similar installation-based context; these usability tests could further establish the viability of such events.

5 Conclusion

Protein-generated music and biologically-influenced soundscapes can be used in a variety of ways. Outreach activities can be envisaged involving these structures, and educational programs can be developed to help increase the comprehensibility of complex biological forms. This paper proposes a novel music-generation algorithm using supervised machine learning. Atomic coordinates within protein structures are mapped to musical notes using distance calculations; these distance calculations are also used to determine velocities. B-factors and occupancies are also used to determine note durations.

This algorithm has proven useful both in the discernment of new patterns within protein data and in the development of new outreach activities. Outreach activities can make protein structures more accessible while simultaneously allowing artistic insights to be

gleaned from complex biological data. In-depth analysis of such music can also be used as a tool for molecular biologists. Visualisations of proteins can be augmented by the use of sound, helping those who study such structures to discover new features of well-known forms.
Protein data has been turned into music; the next steps in this work involve refining the algorithm with more detailed parameter mapping; this refinement will lead to decorative and elaborate musical arrangements along with deeper data insights. Additionally, future work will include inverting the proposed algorithm to turn music into protein data; this backward biological mapping could be used to discover new proteins through the power of sound.
Visualisation is one of our most common forms of data representation. However, visual depictions of proteins can be insufficient when dealing with intricate structures. Sonification offers scientists the ability to examine complex data using a sense other than sight. These auditory representations can not only augment the visuals we are already so familiar with, but can lead to new and intuitive insights into proteins. The more we sense, the more we understand. If there is meaning in music and power in proteins, protein music holds the ability to be as meaningful as it is powerful; such sounds can help us sense new potentials for the building blocks of life and new futures for both science and art.

Notes

The website dedicated to housing this project's materials can be found via the following link: https://cs1.ucc.ie/~imr1/ampc/.

Acknowledgements

This project is funded by the MTU Arts Office Create Le Chéile Arts Project Award.
This publication has emanated from research supported in part by a grant from Science Foundation Ireland under Grant number 18/CRT/6222. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] Gregory Kramer, Bruce Walker, Terri Bonebright, Perry Cook, John Flowers, Nadine Miner, and John Neuhoff. Sonification Report: Status of the Field and

Research Agenda. Faculty Publications, Department of Psychology, March 2010. URL https://digitalcommons.unl.edu/psychfacpub/ 444.

[2] Jeffrey Rimland, Mark Ballora, and Wade Shumaker. Beyond visualization of big data: A multi-stage data exploration approach using visualization, sonification, and storification, volume 8758. May 2013. https://doi.org/10.1117/12.2016019. Journal Abbreviation: Proceedings of SPIE - The International Society for Optical Engineering Publication Title: Proceedings of SPIE - The International Society for Optical Engineering.

[3] Stephen Barrass and Gregory Kramer. Using sonification. Multimedia Systems, 7(1):23–31, January 1999. ISSN 1432-1882. https://doi.org/10.1007/s005300050108

[4] Oliviero Carugo. How large B-factors can be in protein crystal structures. BMC Bioinformatics, 19(1):61, February 2018. ISSN 1471-2105. https://doi.org/10.1186/s12859-018-2083-8

[5] O. Carugo. Correlation between occupancy and B factor of water molecules in protein crystal structures. Protein Engineering, Design and Selection, 12(12):1021–1024, December 1999. ISSN 1741-0126. https://doi.org/10.1093/protein/12.12.1021

[6] Vincent Van Rheden, Thomas Grah, and Alexander Meschtscherjakov. Sonification approaches in sports in the past decade: a literature review. In Proceedings of the 15th International Audio Mostly Conference, pages 199–205, Graz Austria, September 2020. ACM. ISBN 978-1-4503-7563- 4. https://doi.org/10.1145/3411109.3411126

[7] Nik Sawe, Chris Chafe, and Jeffrey Trevi˜no. Using Data Sonification to Overcome Science Literacy, Numeracy, and Visualization Barriers in Science Communication. Frontiers in Communication, 5, 2020. ISSN 2297- 900X. URL https://www.frontiersin.org/articles/10.3389/fcomm. 2020.00046.

[8] J. Guerra, L. Smith, D. Vicinanza, B. Stubbs, N. Veronese, and G. Williams. The use of sonification for physiotherapy in human movement tasks: A scoping review. Science & Sports, 35(3):119–129, June 2020. ISSN 0765- 1597. https://doi.org/10.1016/j.scispo.2019.12.004
URL https://www. sciencedirect.com/science/article/pii/S0765159720300046.

[9] The Sonification Handbook | edited by Hermann, Hunt, Neuhoff. URL https://sonification.de/handbook/. Sonification of 3D Protein Structures Using Supervised Machine Learning 15

[10] Ga¨el Dubus and Roberto Bresin. A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities. PLOS ONE, 8(12):e82491, December 2013.
ISSN 1932-6203. https://doi.org/10.1371/journal.pone.0082491.

[11] John Dunn and Mary Anne Clark. The Sonification of Proteins.

[12] Chi-Hua Yu, Zhao Qin, Francisco J. Martin-Martinez, and Markus J. Buehler. A Self-Consistent Sonification Method to Translate Amino Acid Sequences into Musical Compositions and Application in Protein Design Using Artificial Intelligence. ACS Nano, 13(7):7471–7482, July 2019. ISSN 1936-0851, 1936-086X.
https://doi.org/10.1021/acsnano.9b02180
https://pubs.acs.org/doi/10.1021/acsnano.9b02180

[13] Edward J. Martin, Thomas R. Meagher, and Daniel Barker. Using sound to understand protein sequence data: new sonification algorithms for protein sequences and multiple sequence alignments. BMC Bioinformatics, 22(1): 456, September 2021. ISSN 1471-2105. https://doi.org/10.1186/s12859- 021-04362-7. URL https://doi.org/10.1186/s12859-021-04362-7.

[14] Peter Gena. Musical Synthesis of DNA Sequences.

[15] Sophia Meytin. A Novel Method for Protein-Protein Interface Analysis Using Sonification. In 2021 IEEE MIT Undergraduate Research Technology Conference (URTC), pages 1–5, October 2021.
https://doi.org/10.1109/URTC54388.2021.9701622.
https: //ieeexplore.ieee.org/document/9701622.

[16] Tifanie Bouchara and Matthieu Montes. Immersive sonification of protein surface. In 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pages 380– 383, Atlanta, GA, USA, March 2020. IEEE. ISBN 978-1-72816- 532-5.
https://doi.org/10.1109/VRW50115.2020.00082.
https:// ieeexplore.ieee.org/document/9090531/.

[17] Susan Costantini, Antonella Paladino, and Angelo M Facchiano. CALCOM: A software for calculating the center of mass of proteins. Bioinformation, 2(7):271–272, February 2008. ISSN 0973-2063.
https://www.ncbi. nlm.nih.gov/pmc/articles/PMC2374369/.

[18] RCSB PDB: Homepage, . URL https://www.rcsb.org/.

[19] RCSB Protein Data Bank. RCSB PDB - 2TMV: VISUALIZATION OF PROTEIN-NUCLEIC ACID INTERACTIONS IN A VIRUS. REFINED STRUCTURE OF INTACT TOBACCO MOSAIC VIRUS AT 2.9 ANGSTROMS RESOLUTION BY X-RAY FIBER DIFFRACTION. URL https://www.rcsb.org/structure/2tmv.

[20] Soundscapes From 3D Protein Structures, September 2023. URL https://soundcloud.com/user-455170818/sets/ soundscapes-from-3d-protein-structures.

Unsupervised Learning of Harmonic Analysis Based on Neural HSMM with Code Quality Templates

Yui Uehara

Kanagawa University
yuiuehara@kanagawa-u.ac.jp

Abstract. This paper presents a method of unsupervised learning of harmonic analysis based on a hidden semi-Markov model (HSMM). We introduce the chord quality templates, which specify the probability of pitch class emissions given a root note and a chord quality. Other probability distributions that comprise the HSMM are automatically learned via unsupervised learning, which has been a challenge in existing research. The results of the harmonic analysis of the proposed model were evaluated using existing labeled data. While our proposed method has yet to perform as well as existing models that used supervised learning and complex rule design, it has the advantage of not requiring expensive labeled data or rule elaboration. Furthermore, we also show how to recognize the tonic without prior knowledge, based on the transition probabilities of the Markov model.

Keywords. Automatic chord recognition, Harmonic analysis, Hidden Semi-Markov Models, Neural network.

1 Introduction

Harmonic analysis is the process of representing a musical piece as a sequence of chord labels, which facilitates understanding the structure of the piece. In tonal music, the chord label is called chord degree. The chord degree represents chords with a position (denoted by a Roman numeral) on the scale of a local key. This approach to labeling chords in the context of the key is based on the idea that chord progressions play a crucial role in establishing a tonality, which is fundamental in tonal music. Therefore, harmonic analysis can be applied to various tasks such as composition [6, 28] and higher-level music analysis [5, 20].

Several studies have explored automated harmonic analysis [1, 2, 12, 14, 15, 19, 22-26, 31]. Among them, unsupervised learning is advantageous as it does not require expensive labeled data. However, there have been few works on unsupervised harmonic analysis [19, 31]. The challenge with harmonic analysis is that it involves simultaneously identifying both keys and chords. Since many possible combinations exist, obtaining an optimal result through unsupervised learning is challenging. Therefore, in order to make the unsupervised learning tractable, previous studies relied on manually designed model parameters [19, 31]. In this sense, the models were not fully unsupervised.

This study presents a new method for unsupervised harmonic analysis[1]. The proposed model is based on a hidden semi-Markov model. Most model parameters can be learned automatically using non-labeled data, which is a departure from the previous approaches. As an exception, code quality templates are set manually. The chord quality templates correspond to chord labels in supervised learning, which allow comparison with existing harmonic analysis. We construct the semi-Markov model with the technique of deep latent variable models [9]. The technique is to employ neural networks to approximate probability distributions that comprise a targeted latent variable model, which helps unsupervised learning.

Although the experimental results show that our model still has room to be improved, we exemplify the potential of our model with automatic evaluations and discussions on the obtained harmonic analysis. We also discuss how transition probabilities of the model obtained by unsupervised learning can find the tonic without prior knowledge.

This paper is organized as follows. In Section 2, we review related studies. We introduce the proposed model in Section 3. We describe experimental setups and results in Section 4. Then, we summarize our contributions in Section 5.

2 Related Work

2.1 Automated harmonic analyses

2.1.1 Rule-based harmonic analysis

The Melisma Music Analyzer is one of the leading harmonic analysis models, comprising the Meter, Grouper, Streamer, Harmony, and Key programs [22-26]. Since the chord labels (Roman numerals) in harmonic analysis identify chords within the context of keys, they require information about both keys and chords. The pipeline analyses using Meter, Harmony, and Key programs can provide Roman numerals. The system is rule-based, and utilizes musical knowledge to determine chord tones. Additionally, it provides various criteria for dealing with ambiguous events, for example, preferring chord changes on strong beats and root progression on the line of fifth (circle of fifth) [24]. However, writing down all the preference rules and their

[1] https://github.com/yui-u/harmonic-analysis-chorales

priorities to deal with this ambiguity takes much work. This could pose a challenge when expanding the system to cover different types of music.

2.1.2 Harmonic analyses with supervised learning

On the other hand, supervised learning relies on high-quality labeled data to learn chord discriminators automatically. Masada and Bunescu employed a semi-Markov Conditional Random Field (semi-CRF) for supervised harmonic analysis [12]. In Conditional Random Fields (CRFs), feature functions act like preference rules. In addition, priorities (weights) for each feature function are automatically learned with the labeled data.

Furthermore, Chen and Su developed a harmonic analyzer called the Harmony Transformer that used a Transformer encoder-decoder as an input feature extractor [1, 2]. More recently, Micchi et al. proposed a model with a Convolutional Recurrent Neural Network (CRNN) encoder and Neural Autoregressive Distribution Estimator (NADE) that outperformed the Harmony Transformer [14, 15]. However, even with recent neural network-based supervised learning, the performance of harmonic analysis still has room to be improved. In addition, the dataset is biased toward piano pieces in the classical era and vocal pieces with relatively clear harmony.

2.1.3 Unsupervised harmonic analysis

Increasing the amount of training data is a way to improve performance in supervised learning. However, preparing the data can be expensive. Alternatively, unsupervised learning does not require labeled training data. However, there have been few works on unsupervised harmonic analysis [19, 31]. The challenge with harmonic analysis is that it involves simultaneously identifying both keys (determined by the combination of modes and tonics) and chords. Since many possible combinations exist, obtaining an optimal result through unsupervised learning is challenging. Consequently, past studies had to rely on manually set model parameters [19, 31]. For instance, Wang and Wechsler proposed using an Infinite Gaussian Mixture Model (IGMM) for harmonic analysis [31]. In this model, keys and chords were considered as hidden variables. Since IGMM is a type of Bayesian model, optimal keys and chords for targeted musical notes are obtained through the sampling process. One limitation of their model was that the model parameters were given manually based on musical knowledge. For example, they provided the Gaussian mean and covariance for the major and minor scales, which worked like key profiles. In addition, the IGMM could not consider chord transitions since it was a note clustering.

Raphael and Stoddard also proposed an unsupervised harmonic analysis approach based on Hidden Markov Models (HMMs) [19]. The HMM had the advantage of being able to account for chord transitions. However, according to Raphael and Stoddard, while the chord degree transition probability was learnable, the key transition probability and the probability of the chord degree after a modulation were difficult to learn [19]. This difficulty is probable because the chord degrees appear relatively evenly, whereas keys and modulations strongly depend on individual pieces.

2.2 Deep latent variable models

Deep latent variable models are methods in which deep neural networks work as the approximators of the probability distributions that make up the latent variable models [9, 13, 27]. In recent years, deep latent variable models have attracted attention as a new method for improving unsupervised learning. For example, Tran et al. proposed unsupervised neural Hidden Markov Models for the part-of-speech tag induction [27], and Miao et al. introduced neural topic models [13].

In the music field, Uehara and Tojo extended the unsupervised neural HMMs to semi-Markov models (HSMMs) and performed recognition of chord segments and their transitions [30]. However, their model was not able to recognize keys and chords simultaneously, and thus could not perform harmonic analysis. Our model is based on Uehara's Neural HSMMs but with significant changes for harmonic analysis, including output emission probability with chord quality templates, simultaneous recognition of keys and roots, and dynamic modulation detection.

3 Methodology

The proposed model for harmonic analysis requires no labeled data for parameter training. In this sense, we describe the model as "unsupervised," which is a significant step forward, given that conventional models require all or part of the model parameters to be designed manually [19, 31]. However, our model incorporates simplifications proposed in previous methods, such as the assumption of transpositional equivalency, a predefined set of chord qualities, and simplified conditional probabilities of the stochastic model.

The hidden semi-Markov model [32] forms the core of the proposed model. Unlike Markov models, semi-Markov models equip a notion of the duration of a state, thus more suitable for segment-level recognition. This property of the HSMM is advantageous since chords are recognized as a result of score segmentation [12]. Among several variants of the hidden semi-Markov model, we utilize the "Residential-time HMM" [33] that assumes independence between the duration of the current and the previous hidden states. The EM algorithm is a widely known method for learning the parameters of an HSMM [32, 33]. However, it has been reported that Neural HSMM, a type of deep latent variable model, could achieve better marginal likelihood than the EM algorithm [29]. As described in Section 2, the deep latent variable model predicts model parameters through neural networks [9]. In the following sections, we will first overview the proposed HSMM for harmonic analysis. Then, the neural networks that approximate probability distributions that constitute HSMM are described. Table I shows the notations that will be used in the rest of this paper.

TABLE I. NOTATIONS.

L	Output sequence length	$\in \mathbb{N}$
t	Time step	$\in \{0, \ldots, L-1\}$
$\boldsymbol{x}_t$	An observation at t	$\in \{0, 1\}^{12}$
pc	Pitch class[2]	$\in \{0, \ldots, 11\}$
m	Mode index	$\in \{0, 1\}$
r	Root pitch class index[3,4]	$\in \{0, \ldots, 12\}$
i, j	Root pitch class indices that distinguish before/after transition.	$\in \{0, \ldots, 12\}$
s	Shift value. (m, s) specifies a key k.	$\in \{0, \ldots, 11\}$
k	Key index	$\in \{0, \ldots, 23\}$
q	Chord quality index	$\in \{0, \ldots, 6\}$
d	Hidden state duration index	$\in \{0, \ldots, 15\}$
$N_{(\cdot)}$	The number of indices $(\cdot)$	$\in \mathbb{N}$
$v_{(\cdot)}$	*Logit*[5] for the probability of $(\cdot)$	$\in \mathbb{R}$
$\boldsymbol{v}_{(\cdot)}$	Vector of logits for the probability of $(\cdot)$	$\in \mathbb{R}^{N_{(\cdot)}-1}$
$\boldsymbol{z}_{(\cdot)}$	Index vector with the *1 of N representation*[6]	$\in \{0, 1\}^{N_{(\cdot)}}$
$z_{(\cdot)}$	The $(\cdot)$-th component of the *1 of N* index vector.	$\in \{0, 1\}$
$\boldsymbol{e}_m$	Mode embedding	$\in \mathbb{R}^{12}$

3.1 The hidden semi-Markov model for harmonic analysis

The proposed model is based on a hidden semi-Markov model comprising initial hidden-state, hidden-state transition, hidden-state duration, and output emission distributions [30, 32, 33]. The output for the model is a sequence of pitch classes represented in binary twelve-dimension vectors. Since the hidden states are not

[2] The pitch classes are the 12 numbered notes in an octave: {C, C#/Db, D, ..., B} are numbered {0, 1, 2, ..., 11} respectively.

[3] In our design, the root pitch class represented by the index *r* indicates the root note of the chord, independent of the key.

[4] The last 13th dimension is used as the *Rest* state.

[5] In this paper, we use the term *logit* as a value that is derived from a neural network and parameterize a probability. The logit is fed into the sigmoid or softmax function to produce the parameter of a targeted probability distribution.

[6] The *1 of N representation* is an *N*-dimensional vector representation of a category index. For example, for chord quality index $q = 3$, the corresponding vector of *1 of N representation* is {0, 0, 0, 1, 0, 0, 0}.

observed and no labeled data is used, what the hidden states represent is not specified. However, simplifications of the model, described later in this paper, make the hidden states correspond to combinations of keys and root pitch classes.

Fig. 1 shows an example of possible paths of hidden states. Note that the numbers of keys, root pitch classes, and the maximum duration are less than the actual settings for ease of reading the figure. At each time step, the hidden states are represented by a combination of a key and a root pitch class. In addition, possible remaining times of the hidden states are combined with each hidden state. Then, triplets of (key, root, duration) represent all possible states. Key or root transitions are only permitted when the remaining duration time is zero in a semi-Markov model. Otherwise, the remaining duration times decrease by one at each time step. Note that the Residential-time HMM, a variant of HSMMs, assumes a hidden state transition is independent of the duration (residential-time) of the previous hidden state [32, 33]. Thus, the hidden state transition is represented as follows.

$$p(k_t, r_t, d_t | k_{t-1}, r_{t-1}, d_{t-1}) = \begin{cases} p(k_t, r_t, d_t | k_{t-1}, r_{t-1}) & (\text{if } d_{t-1} = 0) \\ \mathbb{1}(d_t = d_{t-1} - 1) & (\text{if } d_{t-1} > 0) \end{cases} \quad (1)$$

Note that transition probabilities do not include self-transitions, as illustrated in Fig. 1. In addition, we simplify the hidden-state transition probability $p(k_t, r_t, d_t | k_{t-1}, r_{t-1})$ as follows.

$$p(k_t, r_t, d_t | k_{t-1}, r_{t-1}) = \begin{cases} p(d_t) p(r_t | k_t) p(k_t | k_{t-1}) & (\text{if } k_t \neq k_{t-1}) \\ p(d_t) p(r_t | r_{t-1}, k_t) p(k_t | k_{t-1}) & (\text{if } k_t = k_{t-1}) \end{cases} \quad (2)$$

In the rest of the paper, we call $p(k_t | k_{t-1})$ as the key transition probability, $p(r_t | k_t)$ as the initial root probability, $p(r_t | r_{t-1}, k_t)$ as the root transition probability, and $p(d_t)$ as the duration probability.

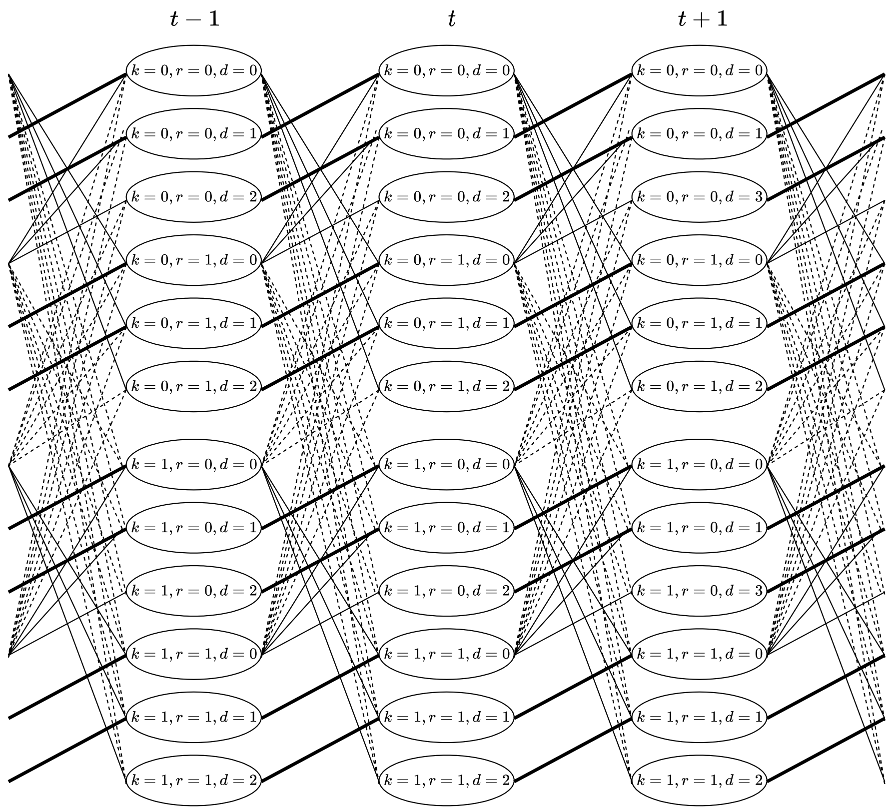


Fig. 1. An example of possible paths of hidden states where the number of keys is 2, the number of roots is 2, and the maximum duration is 3. For simplicity, the number of keys, roots, and the maximum duration are less than the actual settings. Thin solid lines represent root transitions, dotted lines represent key transitions (modulations), and bold solid lines represent the consumption of the remaining duration times.

Since there is no previous hidden state when $t = 0$, the probability of the first hidden state is computed as follows.

$$p(k_0, r_0, d_0) = p(d_0)p(r_0|k_0)p(k_0) \tag{3}$$

The distribution for $p(d_0)$ is the same as $p(d_t)$, and $p(r_0|k_0)$ is the same as $p(r_t|k_t)$. $p(k_0)$ is the initial key probability.

At each time step, the output $\boldsymbol{x}_t$ depends only on the hidden state at the same time step. We decompose the emission probability $p(\boldsymbol{x}_t|k_t, r_t)$ into a chord quality probability

$p(q_t|k_t, r_t)$ and the pitch class emission probability $p(\boldsymbol{x}_t|q_t, r_t)$, where q_t denotes a chord quality.

$$p(\boldsymbol{x}_t|k_t, r_t) = \sum_{q_t} p(\boldsymbol{x}_t|q_t, r_t)p(q_t|k_t, r_t) \tag{4}$$

Among all the probability distributions in the proposed HSMM, we only set manually $p(\boldsymbol{x}_t|q_t, r_t)$, which we call *chord quality templates*[7]. The details of the chord quality templates are described in 3.1.1.

With the initial (3) and transition (1)(2) probability of the hidden states and the emission probability of the observations (4) described above, the generative process of the HSMM is represented as follows, where *L* is the sequence length.

$$\begin{aligned} &p(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{L-1}, k_0, \ldots, k_{L-1}, r_0, \ldots, r_{L-1}, d_0, \ldots d_{L-1}) \\ &= P(k_0, r_0, d_0) \prod_{t=0}^{L-1} P(\boldsymbol{x}_t|k_t, r_t) \prod_{t=1}^{L-1} P(k_t, r_t, d_t|k_{t-1}, r_{t-1}, d_{t-1}) \end{aligned} \tag{5}$$

3.1.1 Emission distribution with chord quality templates

In this section, we first describe the details of the calculation of output emission distribution. As described above, we decompose the output emission probability into the chord quality and pitch class emission probability. We use seven chord qualities: {major triad, minor triad, diminish triad, dominant seventh, major seventh, minor seventh, and diminish seventh}. Then, we formulate the probability of pitch class emission given the chord quality and the root as follows.

$$\begin{aligned} p(\boldsymbol{x}_t|q_t, r_t) &= \prod_{pc=0}^{11} \mathrm{Bernoulli}(x_{t,pc}|q_t, r_t) \\ &= \prod_{pc=0}^{11} (\mu_{pc|q,r})^{x_{t,pc}} (1 - \mu_{pc|q,r})^{1-x_{t,pc}} \end{aligned} \tag{6}$$

$$\mu_{pc|q,r} = \mathrm{sigmoid}(v_{pc|q,r}) \tag{7}$$

$x_{t,pc}$ represents the value at the pitch class (*pc*) of an output binary vector $\boldsymbol{x}_t$. For example, if $\boldsymbol{x}_t = \{0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0\}$, the values of $x_{t,2}$, $x_{t,6}$, and $x_{t,9}$ are 1, and the remaining are 0.

[7] More precisely, in addition to this, the maximum inter-key transition probability limit is set manually.

TABLE II. SETTINGS OF CHORD QUALITY TEMPLATES.

quality (q)	pc=0	1	2	3	4	5	6	7	8	9	10	11
major triad (M)	1				1			1				
minor triad (m)	1			1				1				
diminish triad (d)	1			1			1					
dominant (7)	1				1			1			1	
major seventh (M_7)	1				1			1				1
minor seventh (m_7)	1			1				1			1	
diminish seventh (d_7)	1			1			1			1		

We set the logit $v_{pc|q,r}$ manually, which we call *chord quality templates*, equivalent to setting chord labels in supervised learning. In particular, we set the quality template for the case root pitch class of 0, as in Table II. $v_{pc|q,0} = w$ when marked by one(1) in Table II, and $v_{pc|q,0} = -w$ otherwise. The weight w is set to 5.0. As described in Table I, we use a bold symbol to represent a vector of logits, for example,

$$\boldsymbol{v}_{pc|q,r} = (v_{pc=0|q,r}, v_{pc=1|q,r}, \ldots, v_{pc=11|q,r})^\top \in \mathbb{R}^{12}. \tag{8}$$

If the root pitch class is not 0, the chord templates for a specific root pitch class are obtained simply by shifting the code templates in Table II. For example, when the root pitch class is 5, the chord template of minor seventh
$\boldsymbol{v}_{pc|q=\mathrm{m}_7,r=5} = \{w, -w, -w, w, -w, w, -w, -w, w, -w, -w, -w\}$, where
$v_{pc=0|q=\mathrm{m}_7,r=5}$, $v_{pc=3|q=\mathrm{m}_7,r=5}$, $v_{pc=5,|q=\mathrm{m}_7,r=5}$ and $v_{pc=8|q=\mathrm{m}_7,r=5}$ are w and the remaining are $-w$.

On the other hand, the probability of chord quality given a key and a root pitch class is calculated via neural networks as follows[8].

$$p(\boldsymbol{z}_{q|m,r}|m,r) = \mathrm{Categorical}(\boldsymbol{z}_{q|m,r}|m,r) = \prod_{q=0}^{N_q-1} \pi_{q|m,r}^{z_{q|m,r}}$$

$$\pi_{q|m,r} = p(q|m,r) = \mathrm{softmax}_q(V_{pc,q|r}^\top \boldsymbol{e}_{m,r}) \tag{9}$$

$$\boldsymbol{e}_{m,r} = (\boldsymbol{e}_{m\times r})_r \in \mathbb{R}^{12} \tag{10}$$

$$\boldsymbol{e}_{m\times r} = W_{m\to m\times r}^\top \boldsymbol{e}_m \in \mathbb{R}^{12\times 12} \tag{11}$$

Here, $V_{pc,q|r} \in \mathbb{R}^{12\times N_q}$ is a set of chord templates $\boldsymbol{v}_{pc|q,r} \in \mathbb{R}^{12}$, represented with a tensor. Equation (11) represents the operation of expanding the dimension of mode embedding by a trainable linear transformation to the number of modes times the number of roots, then, eq. (10) represents extracting the components corresponding to the targeted root. The computation of the mode embeddings $\boldsymbol{e}_m$ will be explained in the following paragraph.

8 $\mathrm{softmax}_j(\boldsymbol{x}) = \frac{\exp(x_j)}{\sum_{j'}\exp(x'_j)}$

Since the probability $p(q|m,r)$ is independent on the time step *t*, the subscript *t* is omitted in (9). Hereafter, we sometimes omit the subscript *t* when the variable is time-independent. Note that the chord quality probability is conditioned on the mode *m* and root *r* instead of key *k* and root *r*. In this paper, we set the number of modes to two. However, the two modes are not restricted to major and minor modes but are automatically learned by unsupervised learning.

$$\begin{aligned} \boldsymbol{h}_m &= \mathrm{LSTM}(\boldsymbol{e}_{m-1}, \boldsymbol{h}_{m-1}) \\ \boldsymbol{e}_m &= \boldsymbol{h}_m \end{aligned} \tag{12}$$

The $\boldsymbol{e}_m \in \mathbb{R}^{12}$ in equation (11) is the mode embedding, which works as a latent feature vector of a mode. This $\boldsymbol{e}_m$ can be a learnable vector but is generated through a Recurrent Neural Network (RNN) in this study, as shown in (12). In particular, we utilize the Long-Short Term Memory [7] as an RNN. In this way, any number of mode embeddings can be generated with the same RNN. In this study, the number of modes is fixed to two, so there is no direct advantage to using an RNN, but this setting is advantageous if the number of modes is larger in the future (e.g., considering church modes) or if it is difficult to fix the number of modes in advance.

Once $p(q|m,r)$ is calculated, $p(q|k,r)$ is obtained by shifting $p(q|m,r)$, where we assume transpositional equivalence in keys. For example, a key with (mode index = 1, shift = 3) is assumed to represent a key, each pitch class of which is shifted by three from (mode index = 1, shift = 0). We let the symbol *m* denote the shift(*s*) = 0 case, and let *k* denote $k = 0 : (m = 0, s = 0)$, $k = 1 : (m = 0, s = 1)$, ..., $k = 11 : (m = 0, s = 11)$, $k = 12 : (m = 1, s = 0)$, $k = 13 : (m = 1, s = 1)$, ..., $k = 23 : (m = 1, s = 11)$. Note that we do not restrict the tonic pitch class of a mode as *r* = 0. We will describe how we get the tonic pitch class later. The shift of $p(q|m,r)$ is done without the tonic information. For example, when $k = 14,\ p(q|k = 14, r = 2) = p(q|m = 1, r = 0)$.

As we have seen, the parameters in equations (11) and (12) are learnable and are utilized to generate the chord quality distribution. Thus, the neural networks generate the distributions that make up the model, which is a characteristic feature of the deep latent variable model [9]. The mode embedding $\boldsymbol{e}_m$ is also used to calculate the initial hidden state and the hidden state transition probabilities described below. The results of previous studies suggest that using such features and network elaborations can lead to better convergence than conventional methods [13, 27, 30].

3.1.2 Hidden state transition distribution

A hidden state is represented in a combination of key, root pitch class, and remaining duration. As shown in equations (1) and (2), the hidden state transition probability is decomposed in the duration probability $p(d_t)$, root transition probability $p(r_t|r_{t-1}, k_t)$, and key transition probability $p(k_t|k_{t-1})$.

In this work, we assume $p(d)$ is independent of keys and roots since coherent chord length would be preferred. Therefore, the logit for $p(d)$ is a simple learnable vector, the dimension size of which is the maximum duration length. Thus, the duration distribution is obtained as follows, where $\boldsymbol{v}_d \in \mathbb{R}^{N_d}$ is the learnable vector.

$$p(\boldsymbol{z}_d) = \text{Cat.}(\boldsymbol{z}_d) = \prod_{d=0}^{N_d-1} \pi_d^{z_d} \tag{13}$$

$$\pi_d^{z_d} = p(d) = \text{softmax}_d(\boldsymbol{v}_d) \tag{14}$$

For the root transition probability $p(r_t = j|r_{t-1} = i, k)$, we first compute $p(r_t = j|r_{t-1} = i, m)$ as follows, then shift them under the assumption of transpositional equivalency.

$$p(\boldsymbol{z}_{j|i,m}|r_{t-1} = i, m) = \text{Cat.}(\boldsymbol{z}_{j|i,m}|r_{t-1} = i, m) = \prod_{j=0}^{12} \pi_{j|i,m}^{z_{j|i,m}} \tag{15}$$

$$\pi_{j|i,m}^{z_{j|i,m}} = p(r_t = j|r_{t-1} = i, m) = \text{softmax}_j(\boldsymbol{v}_{j|i,m}) \tag{16}$$

$$v_{r_t=j|r_{t-1}=i,m} = \text{MLP}_2(\boldsymbol{e}_m \oplus \boldsymbol{v}_{pc|i,m} \oplus \boldsymbol{v}_{pc|j,m}) \tag{17}$$

$$\boldsymbol{v}_{pc|r,m} = \sum_q p(q|m, r)\boldsymbol{v}_{pc|q,r} \tag{18}$$

The equation (18) computes a marginal emission pitch class logit of the root pitch class *r*. The marginal emission pitch class logit $\boldsymbol{v}_{pc|r,m} \in \mathbb{R}^{12}$ is used as a feature of a chord whose root is *r* given the mode *m*. We consider the last root index (*r* = 12) as *Rest* and set $\boldsymbol{v}_{pc|r,m}$ for it to a 12-dimensional vector with each dimension's value as $-w$. Then, equation (17) computes a root-transition logit $v_{r_t=j|r_{t-1}=i,m}$ where $\oplus$ denotes a vector concatenation. The root transition probability $p(j|i, m)$ is then computed by the softmax on the set of logits $\boldsymbol{v}_{j|i,m} = (v_{j=0|i,m}, \ldots, v_{j=12|i,m})$.

Once $p(r_t|r_{t-1}, m)$ is computed, $p(r_t|r_{t-1}, k)$ is obtained by shifting $p(r_t|r_{t-1}, m)$, similar to the above discussion of the chord quality probability case. For example, when *k* = 14, which means (*m* = 1, *s* = 2), a component of the probability $p(r_t = 9|r_{t-1} = 2, k = 14)$ equals $p(r_t = 7|r_{t-1} = 0, m = 1)$.

We use the same emission distribution, duration distribution, and root transition distribution for all observed data. On the other hand, the key distributions described hereafter can take different values for each observation sequence. Changing the probability distribution of the model depending on the observation is not possible with a conventional HSMM, but is possible with deep latent variable models [9, 13, 27, 30]. We first obtain the embedding of an observed sequence to compute the key transition distribution.

$$\boldsymbol{h}_t^{\text{obs.}} = \text{LSTM}^{\text{obs.}}\left(\tanh(W^\top \boldsymbol{x}_t), \boldsymbol{h}_{t-1}^{\text{obs.}}\right) \tag{19}$$

$$\hat{a}_t = \text{MLP}_2(\boldsymbol{h}_t^{\text{obs.}} \oplus t^{ratio}) \tag{20}$$

$$\alpha_t = \frac{\exp\left(\hat{a}_t\right)}{\sum_{t'}^{L-1} \exp\left(\hat{a}_{t'}\right)} \tag{21}$$

$$\hat{\boldsymbol{h}}^{\text{obs.}} = \sum_t \alpha_t \boldsymbol{h}_t^{\text{obs.}} \tag{22}$$

The observed data at each time step $\boldsymbol{x}_t$ s first embedded by an LSTM, as eq. (19). Next, the weight for each latent feature $\boldsymbol{h}_t^{\text{obs.}} \in \mathbb{R}^{12}$ is calculated by (20) and (21), where the scalar value $t^{ratio} = \frac{t}{L}$ is an additional information of the time step. Then, the embedding of the observed sequence is obtained by the weighted summation of the latent features, as shown in (22). Using the resulting $\hat{\boldsymbol{h}}^{\text{obs.}} \in \mathbb{R}^{12}$, we compute the mode probability as follows.

$$p(\boldsymbol{z}_m) = \text{Cat.}(\boldsymbol{z}_m) = \prod_{m=0}^{N_m-1} \pi_m^{z_m} \tag{23}$$

$$\pi_m^{z_m} = p(m) = \text{softmax}_m(E^\top \hat{\boldsymbol{h}}^{\text{obs.}}) \tag{24}$$

Here, $E \in \mathbb{R}^{12 \times N_m}$ is a set of mode embeddings $\boldsymbol{e}_m$, represented with a tensor.
As described in Section 3.1.1, a key is defined by specifying a combination of mode index *m* and shift value *s*. The distribution of modes is obtained by (24), and the shift value is obtained as follows.

$$p(\boldsymbol{z}_{s|m}|m) = \text{Cat.}(\boldsymbol{z}_{s|m}|m) = \prod_{s=0}^{11} \pi_{s|m}^{z_{s|m}} \tag{25}$$

$$\pi_{s|m}^{z_{s|m}} = p(s|m) = \text{softmax}_s(\text{MLP}_2(\boldsymbol{e}_m \oplus \hat{\boldsymbol{h}}^{\text{obs.}})) \tag{26}$$

Then, the following equation gives the key probability.

$$p(k) = p(m, s) = p(s|m)p(m) \tag{27}$$

Note that the key probability $p(k)$ is not conditioned on the previous key; therefore, it is used as the initial key probability. There are two situations where the initial key probability is used, one at the beginning of the sequence and the other immediately after the modulation. Although the transition probabilities between keys should be considered, the number of parameters to be estimated was reduced by computing $p(k)$ instead of $p(k_t|k_{t-1})$, as the proportions are more important than the transitions for the keys. The key transition probability is simplified as follows.

$$p(k_t|k_{t-1}) = \begin{cases} 1-\beta & (\text{if } k_t = k_{t-1}) \\ \beta\ \hat{p}(k) & (\text{if } k_t \neq k_{t-1}) \end{cases} \tag{28}$$

β is a learnable value. However, we set the upper limit of β to 0.01, which means the model penalizes modulations. $\hat{p}(k)$ is the key probability modified so that $k_t \neq k_{t-1}$. At the beginning of the sequence, the key probability $p(k)$ is combined with the initial root probability given a key in the following equation. Here, again, we first compute the initial root probability given a mode $p(r|m)$, and the initial root probability given a key is obtained automatically by shifting $p(r|m)$.

$$p(\boldsymbol{z}_{r|m}|m) = \text{Cat.}(\boldsymbol{z}_{r|m}|m) = \prod_{r=0}^{12} \pi_{r|m}^{z_{r|m}} \tag{29}$$

$$\pi_{r|m}^{z_{r|m}} = p(r|m) = \text{softmax}_r(\boldsymbol{v}_{r|m}) \tag{30}$$

$$v_{r|m} = \text{MLP}_2(\boldsymbol{e}_m \oplus \boldsymbol{v}_{pc|r,m}) \tag{31}$$

In the above equation, $\boldsymbol{e}_m$ is the mode embedding, and $\boldsymbol{v}_{pc|r,m}$ is the marginal emission pitch class logit, the same as those used in (18) and (17). The Multi-Layer Perceptron (MLP) with one hidden layer maps the concatenated vector of $\boldsymbol{e}_m \oplus \boldsymbol{v}_{pc|r,m}$ to a scalar logit $v_{r|m}$. The initial probability given a mode $p(r|m)$ is then computed by applying the softmax function to the set of logits $\boldsymbol{v}_{r|m} = (v_{r=0|m}, \cdots, v_{r=12|m})$. The initial root probability given a key $p(r|k)$ can be automatically computed by shifting the $p(r|m)$. Finally, we can obtain the initial probability by $p(k, r, d) = p(d)p(r|k)p(k)$.

3.2 Training

As a deep latent variable model, the proposed model is trained by optimizing the network parameters that produce the probability distributions of the HSMM. The loss function here is the negative log-likelihood (NLL) of the observed sequence $-\log p(\boldsymbol{x}_0, \cdots, \boldsymbol{x}_{L-1})$ [9], obtained by marginalizing all possible paths of hidden states. The technique to marginalize all possible paths of hidden states for an H(S)MM is known as the *forward algorithm* [17, 32]. The details of the forward algorithm for the HSMM, especially for the Residential-time HMM, can be found in [33] and [30]. In the proposed model, a hidden state is represented in a combination of key and root, as illustrated in Fig. 1. Then, the model can directly apply the forward algorithm for the Residential-time HMM, described in [30].

We perform unsupervised training in two phases. The proposed model is not given a feature of keys like the *Key Profile* [11]. Therefore, in the first phase, we train the score

[9] More precisely, the loss is the average of the NLL of all the sequences in the mini-batch.

without key signatures to obtain features of two modes. After that, additional training is performed using the original key signature.

Phase 1:

Training with the *normalized* data. *Normalized* means transposing a score to a key without a key signature. If the key signature information has been lost, the score is transposed to maximize the percentage of pitch classes {0, 2, 4, 5, 7, 9, 11}. In Phase 1 training, keys are limited to the two modes only. These two modes are not necessarily limited to C major and A minor keys, but are learned under the condition that the shift value is always 0. In addition, modulation is disabled by setting the maximum inter-key transition probability limit to zero.

Phase 2:

Training with the original (not transposed) data. In the Phase 2 training, the shift values and the inter-key transition are enabled. Phase 2 training is performed as additional training using the result of Phase 1 as the initial value.

3.3 Inference (harmonic analysis)

After the training, the proposed model performs harmonic analysis by finding the most likely sequence of hidden states (combinations of keys and roots) and their residential times from an observed sequence. Thus, the harmonic analysis is the inference problem of an HSMM, and the method for it is well known as the Viterbi algorithm [4, 17, 32]. After the hidden state inference, the most likely chord quality (q_t^*) for each timestep is obtained by taking the argmax of the probability of output emission given a key and a root.

$$q_t^* = \arg\max_q p(\boldsymbol{x}_t|q, r_t)p(q|k_t, r_t) \tag{32}$$

Thus, we obtain the sequence of keys, root pitch classes, and qualities. Both chord name analysis and Roman numeral analysis are possible based on them. In particular, a chord name is derived from the information of a root pitch class[10] and a quality. The Roman numeral is derived from a key, root, and quality.

To obtain the chord degree (Roman numeral), we first convert the root pitch class by (r - s) mod 12. Here, s is the shift value that is combined with mode m to specify a key (k = (m, s)). Although our model does not explicitly know the diatonic scale, the transition probabilities of the learned model indicate that the chords with 0 and 7 as the roots are the most dominant. This result will be discussed later, but based on this observation, the correspondence between pitch class and degree is given in Table III. Here, the pitch classes corresponding to I and V are 0 and 7, respectively, and the rest

[10] → footnote 3

are assigned one degree per two successive roots. After the conversion, the degree is denoted in uppercase if the predicted quality is major; otherwise, it is denoted in lowercase.

TABLE III. CONVERSION TABLE OF SHIFTED ROOT PITCH CLASSES TO DEGREES.

Shifted root pc.	0	1	2	3	4	5	6	7	8	9	10	11
Degree	I	II	II	III	III	IV	IV	V	VI	VI	VII	VII

Our model does not predict the chord inversion. However, we determine chord inversions in Table IV, where the blank components are assumed to be the basic chords.

TABLE IV. CONVERSION TABLE FOR DETERMINING CHORD INVERSIONS.

(bass pc. - r) mod 12	0	1	2	3	4	5	6	7	8	9	10	11
major triad (M)						6/4			6			
minor triad (m)						6/4				6		
diminish triad (d)							6/4			6		
dominant (7)			2			4/3			6/5			
major seventh (M_7)		2				4/3			6/5			
minor seventh (m_7)			2			4/3				6/5		
diminish seventh (d_7)				2			4/3			6/5		

4 Experiments

4.1 Datasets

We used two different sets of J.S. Bach's four-part chorales: a dataset of 60 chorales formatted by Radicioni and Esposito, and a set of 371 chorales in MusicXML format from the Music21 Library. In this paper, we denote the former dataset as "JSBChorales60" and the latter as "JSBChorales371".

Radicioni and Esposito provided preprocessed scores and human-annotated chord labels in the JSBChorales60 dataset[11] [18]. The preprocessed scores include pitch classes, bass pitch classes, and metrical accents computed by the Meter program of the Melisma Analyzer [25]. However, the preprocessing lost some information in the original scores, for example, beat positions, key signatures, time signatures, and parts (except bass pitch classes). The dataset contains 60 chorales in total.

On the other hand, The JSBChorales371 dataset from the Music21 Library [3] is a set of scores in MusicXML format, and it retains all information as a score. In addition, the dataset contains 371 chorales, some of which are not included in the JSBChorales60 dataset. One of the reasons the JSBChorales371 dataset has not been used much is that it is just a collection of scores and does not provide labeled data. Since our model is unsupervised, we can utilize the JSBChorales371 dataset. For evaluation, we consult

[11] https://archive.ics.uci.edu/dataset/298/bach+choral+harmony

the human analysis of 20 pieces publicly available in the Music21 Library[12] [3]. We used fixed train, validation, and test splits for the JSBChorales371 dataset. The testing data is the set of 20 pieces where the human annotations are available. We preprocess the MusicXML scores into sequences of pitch classes. The length of each event of pitch classes is 16th-note width, which is the minimum duration of the dataset (except for very few exceptions).

4.2 Experimental setups

For the JSBChorales60 dataset, we followed the original 10-fold cross-validation splits provided by Masada and Bunescu. However, unlike the previous work, we used one fold for testing, another for development, and the remaining 8-folds for training. The random seed was fixed to 123. To train the proposed model, we only used the preprocessed pitch classes. The minibatch size was 2. Since only the annotation of chord names is provided, we performed chord name analysis for the JSBChorales60 dataset. In addition, the chord root names in the annotation were converted to pitch classes before the evaluation since our model did not distinguish enharmonic notes. However, we do not consider this a serious limitation since enharmonic distinctions are possible if the original score information has not been lost in preprocessing.

For the JSBChorales371 dataset, we used fixed train, validation, and test splits for the JSBChorales dataset. The testing data was the set of 20 pieces where the human annotations were available. However, three pieces were excluded because of the collapsed format or inconsistency of key signatures between the annotation and the original score. Thus, the resulting number of testing data was 17. In addition, we used 62 pieces randomly selected as the development data and the remaining 243 pieces for training. The total number of pieces used was 322, where 49 pieces were excluded because of the collapsed format or duplication. In both training and testing, we separated a piece into segments by the *fermata*[13] positions and treated the segments as independent sequences. We trained with three random seeds: 123, 456, and 789. The minibatch size was 8.

Other settings were the same between JSBChorales60 and JSBChorales371. The number of training epochs was 480 for phase 1 and 240 for phase 2. The best model was chosen by the negative log-likelihood on the development data. The training was stopped prematurely if the best model was not updated in 80 consecutive epochs. The optimizer was Adam [10], and the learning rate was 1e-3.

[12] https://github.com/cuthbertLab/music21/tree/master/music21/corpus/bach/choraleAnalyses

[13] A *fermata* represents a full-stop marker in a lyric in chorales.

4.3 Automatic evaluation results of harmonic analysis

TABLE V. EVALUATION RESULTS OF ACCURACY ON JSBCHORALES60.

model	method	Full Chord	Root Chord
HMPerceptron [18]	supervised	80.1	-
HMPerceptron (re-experimented [12])	supervised	77.2	84.8
Semi-CRF [12]	supervised	83.2	88.9
Melisma [25] (reported by [12])	rule-based	-	84.3
Ours	unsupervised	66.8	79.2

The *Accuracy*[14] of our model on the JSBChorales60 dataset is shown in Table V. Note that the proposed model did not consider enharmonic notes. Therefore, we converted the root names to pitch classes when evaluating our model. In this sense, the evaluation is not under exactly the same conditions as the other models in the table. However, as mentioned earlier, we do not consider the enharmonic issue a limitation since it can be resolved if the information in the original score remains.

Table V shows that our model underperformed compared to supervised learning and sophisticated rule-based models. However, we have contributed to the advancement of harmonic analysis with unsupervised learning by enabling unsupervised learning of model parameters and presenting evaluation scores to demonstrate the current performance of unsupervised learning.

TABLE VI. EVALUATION RESULTS OF ACCURACY ON JSBCHORALES371.

model	Key	Full RN	Root RN
Ours	74.2	61.6	66.9

As mentioned in Section 4.2, we evaluated chord names on JSBChorales60, whereas we can evaluate harmonic analysis using Roman numerals (RNs) on JSBChorale371. Analysis with Roman numerals requires simultaneous recognition of the keys and roots; therefore, it is more complex than recognizing chord symbols. The *Accuracy* scores for JSBChorales371 with our model are shown in Table VI[15]. To the best of our knowledge, this is the first report comparing the manual and unsupervised harmonic analyses attached to the Riemenschneider numbers 1-20 of J. S. Bach's chorales[16].

[14] *Accuracy* is the percentage of correctly predicted labels from the total number of events in the dataset.

[15] The human annotation sometimes gives multiple interpretations of a single chord at the start or end of modulation, but in this evaluation, the last label was used.

[16] → footnote 12

However, our model still has challenges, as seen in the scores obtained. In the next section, we will discuss these issues by presenting the results of the analysis.

4.4 Discussion on examples of the obtained analyses

Fig. 2 shows the harmonic analyses of BWV269, bars 13-20. The labels are displayed below the bass notes, if present, or at the places where the note exists, in preference of lower parts. A typical error in our model, exemplified by dotted circles in Fig. 2, is the extra annotation of passing tones. As shown in Fig. 2, the gold annotation distinguished V6-V6/5 (bar 18) and V-V7 (bar 19), while it did not distinguish vi-vi2 (bar 20), unlike our model's prediction. The difference between the V and vi cases may be suggested by whether they were on the Bass passing note or not; however, these differences were difficult to detect in our model, which is based on statistical unsupervised learning.
Another issue with our model is that it does not support borrowed chords; therefore, as shown by the dotted square (bars 15 and 16) in Fig. 2, borrowed chords are sometimes detected as modulations.

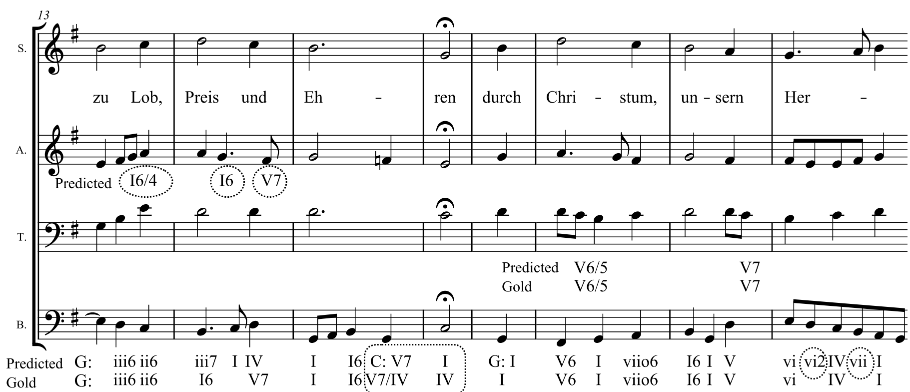


Fig. 2. Harmonic analysis of BWV269 (Riemenschneider No.1), bars 13 - 20.

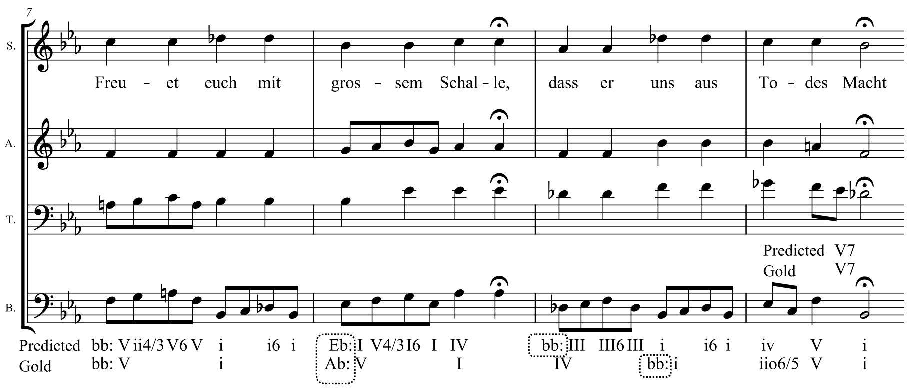


Fig. 3. Harmonic analysis of BWV40.8 (Riemenschneider No.8), bars 7 - 10.

The process of identifying keys is more complex than commonly thought. Fig. 3 shows an example. As shown by dotted circles, both gold and predicted analyses detect a modulation at bar 8. However, a consistent four-degree ascending root progression (F → Bb → Eb → Ab → Db) can be observed in bars 7-9. Therefore, it is possible to interpret the progression as (V → i → IV → VII → III) without any modulation.

4.5 Tonality derived from the learned root-transition probabilities

As noted earlier, we did not assume that the two modes we set were major and minor, nor did any assumption to tonic notes. Many works have considered that the pitch class with the highest percentage of notes is the tonic [8, 11, 23, 31]. We argue, however, that the tonic should be recognized in the context of harmonic progressions. In particular, we consider that tonic is the largest component of the stationary distribution of the root-transition probability matrix. The stationary distribution of a Markov chain is also used in tasks such as measuring the importance of a Web site [16].

Here, we briefly describe the stationary distribution of a Markov chain [21]. If the initial state vector is $\boldsymbol{\pi}_0$ and the transition probability matrix is M, the state probability vector after the *t* steps is as follows.

$$\boldsymbol{\pi}_t^\top = \boldsymbol{\pi}_0^\top M^t \tag{33}$$

The following equation holds if $\boldsymbol{\pi}_t$ approaches a constant value $\boldsymbol{\pi}$ when $t \to \infty$; this $\boldsymbol{\pi}$ is called the stationary distribution in a Markov chain.

$$\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top M, \quad \sum_l \pi_l = 1 \tag{34}$$

The equation (34) means that the stationary distribution is an eigenvector for eigenvalue 1 of the matrix M.

In our study, the matrix of root transition probabilities corresponds to M in (34). However, since our model is a semi-Markov model and the root transition probabilities do not include self-transitions, we modify the root-transition matrix by combining it with the inverse of the average duration probability $a^{\mathrm{dur.}}$ to construct M.

$$(a^{\mathrm{dur.}})^{-1} = \frac{1}{\sum_{d=0}^{N_d-1}(d+1)\pi_d^{z_d}} \tag{35}$$

$$M_m = (1-(a^{\mathrm{dur.}})^{-1})\boldsymbol{I} + (a^{\mathrm{dur.}})^{-1}\begin{pmatrix} \pi_{j=0|i=0,m} \cdots \pi_{j=12|i=0,m} \\ \ddots \\ \pi_{j=0|i=12,m} \cdots \pi_{j=12|i=12,m} \end{pmatrix} \tag{36}$$

$\pi_d^{z_d}$ is the same duration probability in eq. (13), and $\pi_{j|i,m}$ are root-transition probabilities (16).

The obtained stationary distributions are shown in Fig. 4. In mode 0, the pitch class with the highest percentage was 2, which made D the tonic. On the other hand, in mode 1, C was the tonic. Thus, $m = 0$ can be interpreted as the d-minor key and $m = 1$ as the C-major key. The reason why the d-minor was learned instead of the a-minor can be interpreted as follows. Several pieces in the JSBChorales371 dataset were written in the Dorian mode, which can be identified by having one less key signature than modern notation. Hence, in the Phase 1 training (Section 3.2), the learned mode seemed to be a mixture of d-minor and a-minor. After that, mode 0 converged into d-minor in the Phase 2 training. The learned mode 0 had nearly equal proportions of 0 and 1 pitch classes corresponding to degree VII. Thus, the proposed model learned the ambiguity of degree VII in the minor key. It is also interesting to note that the importance of I and V was similar for both modes, but in d-minor, IV was more important than II.

Furthermore, the pitch class probability can be obtained using the stationary distribution, and the logits of the pitch class given a root computed in eq. (18).

$$p(pc|m,r) = \mathrm{sigmoid}(v_{pc|r,m}) \tag{37}$$

$$p(pc|m) = \sum_r p(pc|m,r)p(r) \tag{38}$$

The $p(r)$ is the root probability calculated as the stationary distribution of the root-transition matrix described above.

Obtained pitch class probabilities are shown in Fig. 5. Unlike the *Key Profiles* [8, 11, 23], the pitch class probability of V could be larger than I. This may be partly due to the output representation as a binary vector of pitch classes. However, the stationary distribution in Fig. 4 appears to express the importance of each degree more clearly than the pitch class probabilities in Fig. 5.

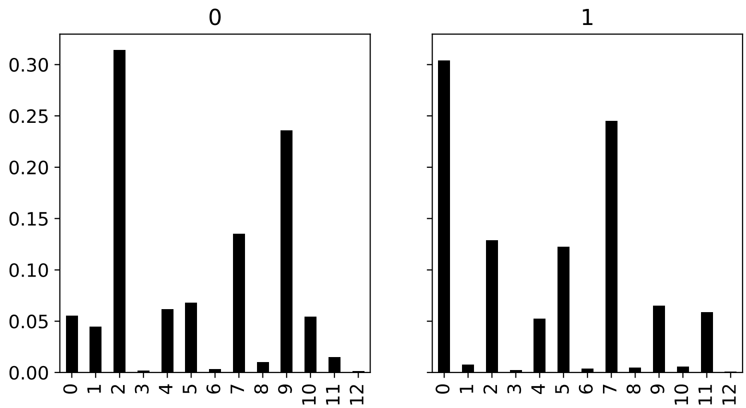


Fig. 4. Result of the stationary distribution of root pitch classes (JSBChorales371, seed=789). The left side is mode *m* = 0, and the right is mode *m* = 1.

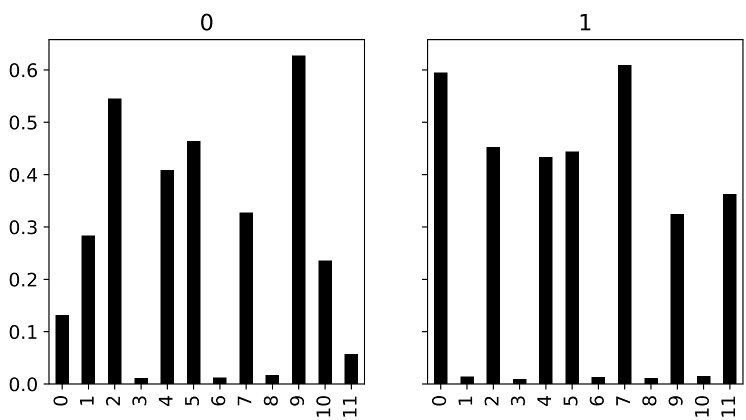


Fig. 5. Result of the pitch class probabilities (JSBChorales371, seed=789). The left side is mode *m* = 0, and the right is mode *m* = 1.

5 Conclusion

This paper proposed an unsupervised harmonic analysis based on the neural hidden semi-Markov model (HSMM). The model was constructed with neural networks that approximate probability distributions in the HSMM. This technique allowed feasible unsupervised learning of model parameters, which has been difficult in previous studies. In addition, we introduced chord quality templates, which enabled the harmonic analysis with known chord labels such as chord names and Roman numerals. Although we presupposed that the number of modes was two and keys were equivalent if transposed, we did not make any other assumption for the two modes. Nevertheless, our model could find the minor and major modes and their tonic notes properly, as was discussed in Section 4.5.

However, the *Accuracy* results on labeled data and examples of obtained analyses suggest that our model still has room to improve. An important future work is to distinguish passing tones. Since our model is a generative model, it assigns probability to all notes in a score. This may cause excess labeling of passing notes, which would be more severe for instrumental music. We may consider a method of probabilistic estimation of passing notes and changing the generation mechanism depending on whether they are passing notes or not.

If the issue of passing tones is resolved, the proposed method should apply to a wider range of music. The true value of unsupervised learning would be demonstrated for music after the late Romantic period, for which there is little labeled data. Furthermore, the model may be effective for pre-Renaissance music since our model could learn modes and tonics unsupervised.

Acknowledgments

This research has been supported by JSPS KAKENHI No. 23K20011. Computational resource of AI Bridging Cloud Infrastructure (ABCI) provided by National Institute of Advanced Industrial Science and Technology (AIST) was used.

References

[1] T.-P. Chen and L. Su. Harmony Transformer: Incorporating Chord Segmentation into Harmony Recognition. In *Proceedings of the 20th International Society for Music Information Retrieval Conference*, pages 259–267, 2019.

[2] T.-P. Chen and L. Su. Attend to chords: Improving harmonic analysis of symbolic music using transformer-based models. *Transactions of the International Society for Music Information Retrieval*, 4(1):1–13, 2021.

[3] M. S. Cuthbert and C. Ariza. music21: A toolkit for computer-aided musicology and symbolic music data. In *Proceedings of the 11th International Society for Music Information Retrieval Conference*, 2010.

[4] G. D. Forney. The viterbi algorithm. In *Proceedings of the IEEE*, 61(3):268–278, 1973.

[5] M. Granroth-Wilding and M. Steedman. Statistical parsing for harmonic analysis of jazz chord sequences. In *International Computer Music Conference*, pages 478–485, 2012.

[6] R. Groves. Automatic harmonization using a hidden semi-markov model. In *AIIDE Workshop*, pages 48–54, 2013.

[7] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[8] D. Hu and L. K. Saul. A Probabilistic Topic Model for Unsupervised Learning of Musical Key-Profiles. In *Proceedings of the 10th International Society for Music Information Retrieval Conference*, pages 441–446, 2009.

[9] Y. Kim, S. Wiseman, and A. M. Rush. A tutorial on deep latent variable models of natural language. *arXiv preprint arXiv:1812.06834*, 2018.

[10] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations*, 2015.

[11] C. L. Krumhansl. *Cognitive foundations of musical pitch*. Oxford University Press, 2001.

[12] K. Masada and R. Bunescu. Chord recognition in symbolic music: A segmental crf model, segment-level features, and comparative evaluations on classical and popular music. *Transactions of the International Society for Music Information Retrieval*, 2(1):1–13, 2019.

[13] Y. Miao, E. Grefenstette, and P. Blunsom. Discovering discrete latent topics with neural variational inference. In Proceedings of the 34th International Conference on Machine Learning, volume 70 of *Proceedings of Machine Learning Research*, pages 2410–2419, 06–11 Aug 2017.

[14] G. Micchi, M. Gotham, and M. Giraud. Not all roads lead to rome: Pitch representation and model architecture for automatic harmonic analysis. *Transactions of the International Society for Music Information Retrieval*, 3(1):42–54, 2020.

[15] G. Micchi, K. Kosta, G. Medeot, and P. Chanquion. A deep learning method for enforcing coherence in Automatic Chord Recognition. In *Proceedings of the 22nd International Society for Music Information Retrieval Conference*, pages 443–451, 2021.

[16] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bring order to the web. Technical report, Stanford University, 1998.

[17] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, 77(2):257–286, 1989.

[18] D. P. Radicioni and R. Esposito. BREVE: An HMPerceptron-Based Chord Recognition System. In *Advances in Music Information Retrieval*, pages 143–164. Springer Berlin Heidelberg, 2010.

[19] C. Raphael and J. Stoddard. Harmonic analysis with probabilistic graphical models. In *Proceedings of the 4th International Conference on Music Information Retrieval*, 2003.

[20] M. Rohrmeier. Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1):35–53, 2011.

[21] R. Serfozo. *Basics of applied stochastic processes*. Springer, 2009.

[22] D. Temperley. An Algorithm for Harmonic Analysis. *Music Perception*, 15(1):31–68, 10 1997.

[23] D. Temperley. What's Key for Key? The Krumhansl-Schmuckler Key-Finding Algorithm Reconsidered. *Music Perception*, 17(1):65–100, 10 1999.

[24] D. Temperley. *The Cognition of Basic Musical Structures*. The MIT Press, 2004.

[25] D. Temperley and D. Sleator. The melisma music analyzer. https://www.link.cs.cmu.edu/music-analysis/.

[26] D. Temperley and D. Sleator. Modeling meter and harmony: A preference-rule approach. *Computer Music Journal*, 23(1):10–27, 1999.

[27] K. Tran, Y. Bisk, A. Vaswani, D. Marcu, and K. Knight. Unsupervised neural hidden markov models. In *Proceedings of the Workshop on Structured Prediction for NLP*, pages 63–71, 2016.

[28] H. Tsushima, E. Nakamura, K. Itoyama, and K. Yoshii. Function- and rhythm-aware melody harmonization based on tree-structured parsing and split-merge sampling of chord sequences. In *Proceedings of 18th International Society for Music Information Retrieval Conference*, pages 502–508, 2017.

[29] Y. Uehara. *Unsupervised Recognition of Chords, Functions*, and Tonality. PhD thesis, Japan Advanced Institute of Science and Technology, 2022.

[30] Y. Uehara and S. Tojo. Chord function recognition as latent state transition. *SN Computer Science*, 3:508, 2022.

[31] Y.-S. Wang and H. Wechsler. Musical keys and chords recognition using unsupervised learning with infinite gaussian mixture. In *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval*. Association for Computing Machinery, 2012.

[32] S.-Z. Yu. Hidden semi-Markov models. *Artificial intelligence*, 174(2):215–243, 2010.

[33] S.-Z. Yu and H. Kobayashi. An efficient forward-backward algorithm for an explicit-duration hidden Markov model. *IEEE signal processing letters*, 10(1):11–14, 2003

The Interactive Digital Transcription and Analysis Platform (IDTAP): Enabling the Computational and Heuristic Analysis of Sound, Music, and the Social

Dard Neuman, Jonathan B. Myers

University of California, Santa Cruz
DNeuman@ucsc.edu, JBMyers@ucsc.edu

Abstract. This paper outlines the historical background, motivation, development process, and potential impact of a new interactive digital transcription and analysis platform (IDTAP): The IDTAP is a web-based application that enables users to digitally transcribe, archive, share, and analyze audio recordings of oral melodic traditions. In 2023, authors and principal investigators Dard Neuman and Jon Myers launched version 1.0 of the platform, disseminated it globally, and cultivated an initial base of users. Whereas the platform has been built around Hindustani music (i.e., North Indian classical music), the principles, melodic contour archetypes, technologies, and methodologies behind the platform are designed to expand to other traditions where continuous melodic contour movements are accentuated. This expansion of the IDTAP, in turn, opens multiple recorded sound collections and archives to digital preservation, musical creation, as well as statistical, quantitative, and interpretive analysis, equipping scholars from a range of disciplinary backgrounds to apply the power of twenty-first-century computational methodologies and large datasets to humanistic endeavors.

Keywords. Digital Archive, Hindustani Music, Non-Western Music, Notation, Software, Transcription.

1 Introduction

The IDTAP is an interactive digital archive for oral melodic expressive traditions. Specifically, it is an open-source and open-access, multi-layered and interactive platform that allows users to upload audio recordings, transcribe melodic sound intuitively, efficiently, and accurately from those recordings; test the accuracy of their transcription through synthesized audio playback; and analyze the transcriptions qualitatively, quantitatively, and computationally. The larger goal of the IDTAP is to de-center music scholarship and composition from a western theory and notational framework while centering the place of music (or melodic expressive traditions) in humanistic and social scientific studies. Staff notation is the primary, predominant, if not hegemonic, tool for graphically representing music in academia world-wide.

Whereas staff notation has a number of advantages—i.e. a spatially efficient method of representing rhythmic sequences; relative ubiquity among musical practitioners throughout the world; easy transferability to extant notation software and data analysis tools—it does not accurately represent various aspects of musical information, particularly those that are found in many non-western music traditions. Some musicologists and ethnomusicologists have, since the 1950s, argued against the utilization of staff notation for the transcription, composition, and analysis of non-European art music traditions [1]-[4]. Even so, many researchers and artists continue to use it when engaging with non-western repertoires. This presents problems that are ethical, empirical, and epistemological. Ethically, staff notation imposes a Eurocentric representational understanding of musical form. Empirically, staff notation flattens out important aspects of many oral-melodic idioms. Epistemologically, it discourages scholars without knowledge of staff notation, let alone other methodological tools to include melody and rhythm in their research. This effectively removes from humanistic and social scientific inquiry a central aspect of everyday life—the relationship of organized sound to a range of communities, be they small scale proximate communities or large-scale religious, dynastic, or national imagined communities. These relationships are even more elusive in oral cultures and their secular and spiritual expressive traditions. Lastly, extant alternative forms of musical representation, including those that are digital, exist within a two-dimensional framework.

By contrast, the IDTAP provides a multi-layered and interactive platform and corresponding store of knowledge: it is a transdisciplinary research archive that bridges computational media, linguistics, comparative literature, statistics, history, folklore, religious studies, cultural and comparative musicology, ethnomusicology, and music composition. The goal of the IDTAP research team is to grow and develop the platform so that recorded music from a diversity of oral cultures can be both preserved and made available for music making and research in empirically analyzable and comparable data formats.

The IDTAP software is able to represent a diversity of melodic contours due to a unique rethinking of a widely held music-theoretical tenet. Instead of taking a fixed-pitch note as the only basic unit of structure, the IDTAP is organized around a succession of "trajectories": formally/mathematically specified archetypal paths from one pitch to another, among a series of pitches, or on a fixed-pitch. These trajectories are able to represent a range of finely calibrated glissandi that appear in many musical traditions outside of the keyboard and staff-oriented Eurocentric musical paradigm.

2 Disciplinary Interventions

The fundamental principle of the IDTAP is to diversify music research, representation, creation, criticism and theory through more expansive transcription forms and corresponding analytics. This principle represents a commitment to empirical and comparative research that is untethered to 19th-century colonial investments in taxonomic schemas that plot European racial and cultural measures as the apex around which other human and cultural creations are assessed and judged. We believe the IDTAP and the resultant notations and datasets will be pertinent to many areas of research/scholarship and teaching/learning.

First, the IDTAP is organized to diversify humanistic inquiry and to help professionalize the next generation of researchers. The dual roles of music transcription and the empirical analysis of music, once distinctive methodological features in the fields of musicology and ethnomusicology, are currently in a state of flux. The musicological turn in U.S. music departments in the 1980s and '90s decentered the role of music theory and other empirical approaches in favor of critical and cultural theoretical approaches [5]-[8]. This shift occurred, in part, because music transcription and analysis relied on staff notation, regardless of whether the music in question resembled the European classical music tradition within which staff notation was originally developed.

There have been two broad correctives to this situation: the abandonment of transcription and empirical methods altogether, and the creation of custom notation systems. These solutions, however, raise additional challenges. The turn away from transcription-based research moves the domain of melodic/rhythmic sound analysis away from quantitatively measurable evidence-based research, as well as cutting edge computational technologies based on digital audio research. This effectively removes a central part of everyday life from humanistic and social scientific inquiry: the interconnected relationship between melodic expressive traditions and affective and emotional communities [9]. The second corrective, custom transcription methods, lack commensurate fields of data for quantitative and/or computational analysis, let alone a common language for scholars to engage a broader community of readers and researchers with music analysis.

Scholars of Hindustani music based in India and the UK have worked to address some of these issues. Music in Motion is a web archive that contains videos that display melodic contour plots with textual annotations of lyrical content scrolling in real time with audio playback [10]. A research group at Durham University has been conducting empirical research into rhythmic-event micro-timing differences in Hindustani music [11] and has published a dataset for rhythmic events that is openly available for researchers to use in timing- and rhythm-focused empirical studies [12]. The growing interest in this empirical vein is also evinced by the recent appearance of a new journal, *Analytical Approaches to Music of South Asia*. These projects, however, are close-ended: They do not provide an interactive platform that allows others to upload, transcribe, query, and analyze the data beyond the melodic contour graphs or the already produced datasets. The IDTAP web app, by contrast, is an interactive platform—in a sense, a kind of scholarly crowdsourcing.

The IDTAP is intended, then, as a research tool for scholars of oral melodic traditions and expressive cultures. Statistically oriented researchers can use the data that undergirds this visual representation to ask new, quantitative kinds of research questions. Ethnomusicologists can use "big data" computational approaches to query this archive at a large scale across a range of performances (diachronically or synchronically), or at a small scale through specific measurements within a single performance, or even a single phrase. Music theorists can use the data as a platform for the discovery/construction of new kinds of theoretical knowledge regarding melodic contour, macro- and micro-rhythm, and performer style. Social historians, folklorists, religious studies scholars, and musicologists can supplement their research of (affective) communities, classes, castes, or individuals through the treatment of recorded melodic sound as historical evidence.

Second, the IDTAP is intended as a pedagogical tool for undergraduate and graduate students, as well as independent musicians and music students who want to teach themselves. Whereas generations of practitioners have learned in the "traditional" way, i.e. under the direct tutelage of a hereditary or discipular forebear, the IDTAP has the potential to engender a more diverse community of practitioners, including those who weren't born into a musical family and who may not have privileged access to a master-teacher, due to socio-cultural, economic, and/or geographical factors.

Lasly, the IDTAP is intended as a diversifying corrective to a current bias in Music Information Retrieval (MIR), Music Artificial Intelligence (MAI) research, and corpus studies, all of which are focused on western musical corpora [13]. These epistemological and data-borne biases may have far-reaching consequences over the coming decades, as AI systems play an exponentially increasing role in the processes by which music is generated, disseminated, and consumed. To incorporate the important bodies of non-western / non-notated musical knowledge, their unique forms and structures must first be made computationally legible by projects such as the IDTAP.

3 IDTAP Framework and Theoretical Foundation

The IDTAP has been designed, developed, and implemented with the following components: 1) the archive of recordings, including tools for users to upload, document, and listen to the recordings; 2) the editor interface, which includes a suite of tools for users to compose or transcribe melodic contours and instrumental, syllabic, and/or textual articulations, as well as to segment and annotate the transcriptions; 3) the archive of transcriptions/notations; 4) data-visualization and data-query functions; and

5) the collections of different users, organized by function: research, pedagogy, appreciation, and creative production. The IDTAP user interface and analysis suite are meant to enable and make available digital transcriptions for scholars and students who have archival, language, linguistic, musical, or literary research skills found in humanities-based disciplines but are not necessarily trained in programming or computational methodologies.

The IDTAP framework is organized around two core concepts, articulations and trajectories: Articulations are instrumental and vocal onsets while trajectories are formally/mathematically specified archetypal straight or curved paths on one pitch, from one pitch to another, or among a series of pitches. The articulation structure distinguishes vocal from instrumental music, with further hierarchical groupings that draw on two transferrable and generalizable systems: the International Phonetic Alphabet (IPA) for vocal expressions and a modification of Bharata's system of instrument classification based on articulation types (i.e., whether the instrument attack is plucked, bowed, blown, pressed, or hammered) for instrumental expressions. For vocal transcriptions, the IDTAP editor allows for the inclusion of syllabic articulations: for every trajectory, the user can mark a starting consonant, vowel, and ending consonant. The synthesizer for the vocal transcriptions is rendered via an implementation of the Klatt Synthesizer, a powerful dual cascade and parallel formant model of human speech production [14]. The synthesizer for plucked string instrumental transcriptions is rendered via an implementation of the Karplus-Strong string synthesis algorithm [15]. The articulation framework for vocal and instrumental music is organized for the expansion of the IDTAP to other traditions. At present, the user can select between the following alphabets: IPA, Latin, and Devanagari. As the IDTAP develops to other traditions and languages it will add corresponding alphabets. Similarly, while the instrumental trajectories are currently modeled on a plucked string instrument, they will be expanded to include other instrument types.

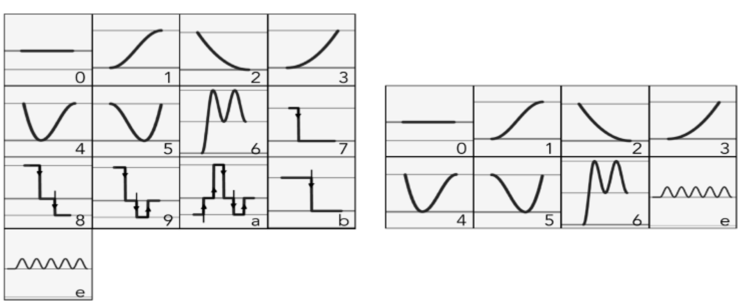


Fig. 1. The thirteen trajectories used in the plucked instrument tradition of the sitar and the eight trajectories used in vocal music.

To transcribe, users trace trajectories on top of the spectrogram and/or melograph. For certain trajectories, the slope of the curve can be adjusted to match the melodic contour being transcribed. Trajectories can also be "offset" to represent microtones (*shrutis*)—or spaces between scale tones—and movements between scale tones and microtones (*andolans*). Fig. 2 shows two transcriptions of a section of a Hindustani composition, "Babul Mora," performed by the artist Begum Akhtar. Note how the IDTAP is as empirically attuned to the nuances of movements between scale tones as on them.

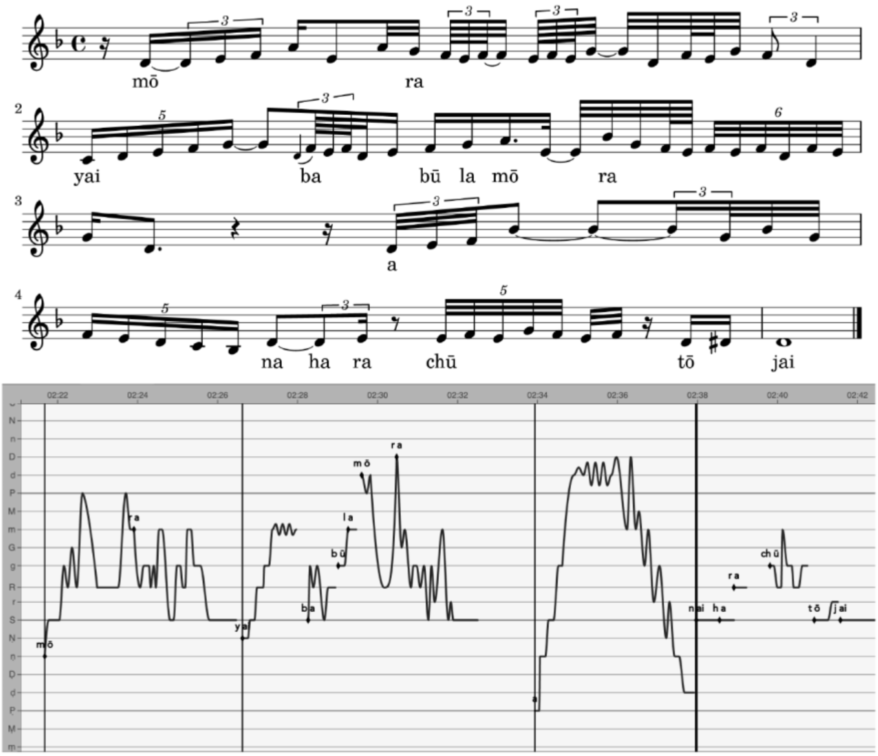



Fig. 2. Transcription of the same selection of the song, *Babul Mora*, using staff and IDTAP notations.

4 IDTAP Upload and Set Up Process

The set-up process for the transcription files includes five actions. The first involves specifying general and idiomatically specific metadata. The next two actions involve 1) adjusting the tonic to that which is used in the performance; and 2) establishing which scale tones are used in the performance. The final two actions involve music information retrieval to aid the transcription process: the generation of a constant Q spectrogram [16] and a "melograph" [17], two-dimensional graphical representations of audio information that map frequency on the Y axis against time on the X axis. Fig. 3 shows three representations of the same section of a performance: the first pane shows an algorithmically generated melograph; the second pane a logarithmically scaled spectrogram, and the third pane a manually engraved IDTAP transcription that can be plotted/traced on top of the spectrogram and/or melograph.

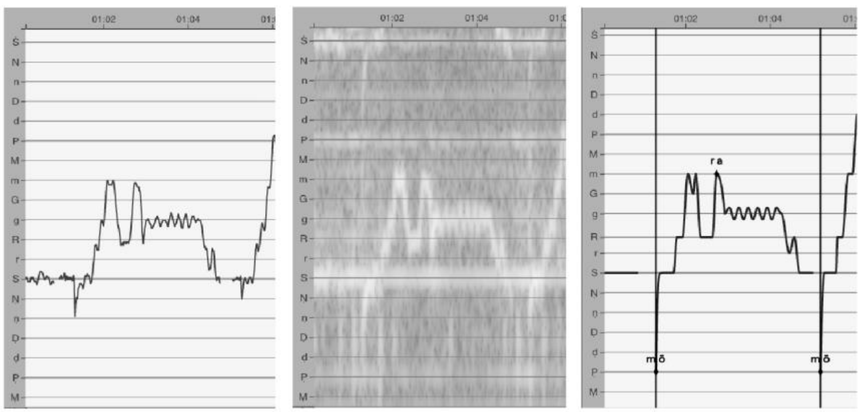


Fig. 3. (Left) Algorithmically generated "Melograph" / pitch trace of excerpted performance by Begum Akhtar. (Middle) Logarithmically scaled Spectrogram (specifically, a Constant Q Transform) of the same excerpt. (Right) IDTAP transcription of the same excerpt, by D. Neuman.

5 IDTAP Transcription Process

The "Editor" interface allows users to transcribe a particular recording (or compose/make notations unrelated to any particular recording) by inserting trajectories onto a two-dimensional time-by-pitch graph. The transcription process is facilitated by: the user-controlled overlay of a spectrogram and/or melograph; custom synthesis software that sonifies data fed from user-generated transcriptions; affordances that allow for the adjustment of tuning at both the macro-scale (adjusting the tuning system associated with the entire transcription) and the micro-scale (adjusting the frequency of individual pivot points associated with particular trajectories); the ability to select a certain region of audio, loop it, and/or adjust its playback speed; a flexible interface for assigning, displaying, and adjusting hierarchically nested systems of temporal division (that is, *meter*); and the ability to shift the pitch of the entire recording and synthesis to match the user's physical instrument or vocal range.

This process is simplified by tools for 1) adjusting the proportions of the editor's tuning-system to the pitches in the performance; 2) pitch shifting the recording and synthesis in tandem to match a user's voice or physical instrument tuning and register; 3) transcribing melodic contours; 4) transcribing articulations (vocal vowel and consonant syllable/syllable clusters, and instrumental articulations—plucked, hammered, bowed, blown, etc; 5) segmenting a transcription into hierarchically-nested collections of sections, phrases, and groups; 6) specifying metrical templates, adding them at the appropriate time in a transcription, and adjusting their micro-rhythmic placement to match the natural temporal fluctuation of performance practice (i.e. rubato); 7) playing back the synthesis of the transcription in relation to the recording; 8) looping specified regions and adjusting playback speed; 9) annotating the transcriptions at different levels of form and content, choosing from a set of IDTAP-provided options.

6 IDTAP Analysis, Data Visualization, and Query Process

The "Analysis" interface includes interactive tools for users to engage with transcription data heuristically. These include 1) a query system for users to isolate and view particular sequences from a transcription according to an unlimited number of filtering criteria; 2) a pitch occurrence data-visualization tool that displays the relative proportion of time spent on particular pitches, and segmented by section, phrase, or durational window; 3) a "pitch pattern" visualization tool that shows the relative frequencies of particular patterns of pitches that appear throughout a transcription. This interface includes parameters for emic and etic inquiries, the first enabling the analyst to annotate and analyze according to "insider" sensibilities, categories and concepts, and the latter enabling the analyst to analyze according to "neutral" parameters.

Pitch Prevalence

This tool allows users to generate visualizations that represent the prevalence of particular pitches over the course of a transcription. This is achieved by splitting the transcription into durational windows according to one of three user-selected segmentation schemes. One segmentation scheme—by duration—is considered etic, as it splits the transcription into windows of particular user-defined duration. This approach aims to overcome various biases: both insider knowledge of traditional conceptions of phrase structure as well as outsider intuitive perceptions of phrase structure. Two of the segmentation schemes—by section and by phrase—are considered emic, relying on an analyst's insider knowledge of the musical tradition and their analytical intentions in explicitly entering phrase and section divisions throughout the process of transcription. For the emic segmentation approaches, the user is also able to filter the segments to include or exclude particular section or phrase types, as specified by the user during transcription. With the section segmentation, a user may choose to include or exclude section types: in the case of Hindustani music, different forms of composition and improvisation. The *phrase* segmentations are subdivided into: phrase-types (compositional forms); elaboration-types (improvisational forms); articulation-types (e.g, if a vocalist is singing lyrics (*bols*), solfege (*sargam*), non-lexical syllables (*aakar*, *nom-tom*); and incidentals (e.g., if the musician speaks, tunes an instrument, etc.).

For example, fig. 4 displays the pitch prevalence and duration of the first six sections and eleven phrases of the *thumri* song, “Babul Mora,” performed by Begum Akhtar. It also shows the coded description of each section and phrase type so that a user can see at a glance the formal sections and pitch structure of an entire performance.

Pitch Range and Percentage of Duration on each Fixed Pitch, Segmented by Phrase

Begum Akhtar - Babul Mora Naihar Chhooto Jaaye

Section #		1		2	3			4	5			6
Section		Pre-Chiz Alap		Composition	Improvisation			Composition	Improvisation			Composition
Phrase #		1	2	3	4	5	6	7	8	9	10	11
Start		0:00	0:06	0:26	0:37	0:44	0:47	0:50	0:57	1:01	1:05	1:09
Duration		0:06	0:20	0:11	0:07	0:03	0:03	0:07	0:05	0:04	0:04	0:06
Phrase Type				Mukra	Mukra	Asthai	Asthai	Mukra				Antara
Elaboration			Bol Banao	Bol Banao	Bol Banao	Bol Banao	Bol Banao		Bol Banao	Bol Banao	Bol Banao	
Articulation		Aakar	Bol	Bol	Bol	Bol	Bol	Bol	Bol	Bol	Bol	Bol
Incidental												
0	d											
	P											
	M											
	m						9%				24%	
	G											
	g		3%	6%			14%		39%	58%		12%
	R		14%	5%			21%	4%	11%	18%	10%	
	r		7%	10%	6%	6%		12%			18%	9%
	S	83%	60%	80%	73%	5%	56%	71%	50%	24%	48%	80%
-1	N		2%		6%							
	n		6%		7%	15%						
	D		1%									
	d	17%	1%		6%	5%						
	P		5%		3%	70%		13%				

Fig. 4. Pitch Prevalence Data Visualization.

In addition to the type of segmentation, users may also toggle a) for their visualization to tabulate the prevalence of each pitch chroma (also known as pitch class: the quality of a pitch abstracted from its octave registration) instead of the pitch; b) for the display to be vertically condensed, leaving out the empty vertical space representing pitches not included in the transcription; or c) for the visualization to display the relative prevalence of elements as a “heatmap” grayscale gradient, as opposed to just the most common (mode) element in dark gray and the rest in light gray.

Query Display

This tool enables users to construct a multi-part query that filters a transcription, displaying only those segments that match the user's specifications. Fig. 5 shows the query function and query output for every time the vocalist, Begum Akhtar, sings the word "*babul*" (father). The query function in this case is threefold: looking for a phrase with starting consonant "ba," the vowel "u" and the ending consonant "l." This particular query would allow for either a pedagogical or analytical investigation on how an artist expresses a composition—where word and melody are fixed—in multiple ways, allowing one to examine the dialectic of composition and improvisation visually.

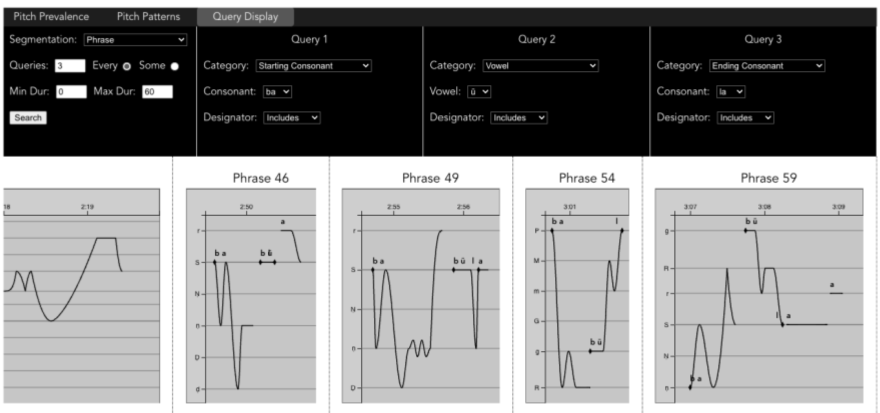


Fig. 5. Query functions and display example.

The following categories are available for users to construct queries: Pitch, Strict Pitch Sequence (pitches must appear in this exact order without any gaps between them), Loose Pitch Sequence (pitches must appear in this order, but may have gaps between them), Trajectory, Strict Trajectory Sequence, Loose Trajectory Sequence, Section Type, Alap Type, Composition Type, Tempo, Tala (descriptor of metric structure and hierarchy, something like a Time Signature in western music), Phrase Type, Elaboration Type, Incidental Type, or Articulation Type.

For each of these categories, users can select a designator: "includes," "excludes," and, if applicable, "starts with," or "ends with." For example, a user might construct a query that includes all phrases that contain the pitch "Sa" in octave 0, or another query that excludes all phrases that contain a particular trajectory, i.e. a "krintin-slide." This querying system is especially powerful when the user combines multiple of them in order to hone in on a particular analytical idea or hypothesis. Users can also choose to display those segments that satisfy all of the queries, or just at least one of the queries. With the desired phrases now displaying, the analyst can make discoveries based on observing the filtered results.

Pitch Patterns

This tool enables users to select a specific type of segmentation and pattern size. It then counts and records the number of occurrences of every unique pitch sequence within each segment, based on the chosen pattern size, and displays this information along with a graphical depiction of the contours of each sequence. This kind of display allows the analyst to intuitively search for common patterns of various sizes in different segments throughout a performance, thus revealing some essence of the style of the melodic elaboration: See fig. 6.

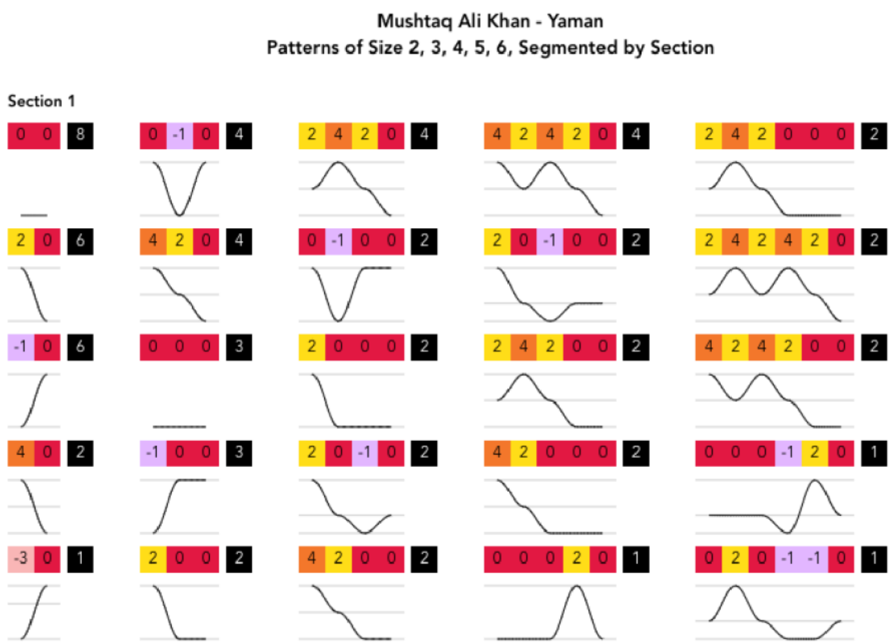


Fig. 6. Pitch Patterns analysis tool

Finally, the IDTAP allows users to generate datasets to conduct quantitative statistical analysis: see fig. 7.

The types of analysis enabled by these queries, data visualizations, and datasets are multiple. To name just a few, users can examine a particular recording, section from a recording, or body of recordings associated with an individual or group of musicians from a range of traditions. Research topics could include: relationships between lyric-text and melody, or syllable and melodic movement; empirically measurable changes of a musician's performance style over time: throughout an individual performance or across many performances; analysis across instruments, venues, styles, performance sections, time-periods, geographic areas, recording media, or other clearly specified method of slicing a corpus. These three analysis tools will be enhanced and added to as our project team continues to engage with and respond to the needs of our research colleagues.

Phrases																	
	Start Time	Duration	Trajectories														Chikaris
									Articulations			Pitches					
			Num	Id	Id Name	Duration	Start Time*	Slope	time**	type	stroke	swara	raised	octave	Proportional Duration Within Phrase	Proportional Duration within Trajectory	Timing within Phrase
0	0	10.1	0	6	Bend: Simple Multiple	0.41	0		0	pluck	d	0	TRUE	0	0.04	0.53	2.39
												6	TRUE	-1		0.47	5.2
												0	TRUE	0			8.38
			1	0	Fixed	2.36	0.41					0	TRUE	0	0.23	1	9.6
			2	0	Fixed	2.88	2.77		0	pluck	d	0	TRUE	0	0.29	1	
			3	0	Fixed	4.44	5.66		0	pluck	d	0	TRUE	0	0.44	1	
	Start Time	Duration	Trajectories														Chikaris
									Articulations			Pitches					
			Num	Id	Id Name	Duration	Start Time*	Slope	time**	type	stroke	swara	raised	octave	Proportional Duration Within Phrase	Proportional Duration within Trajectory	Timing within Phrase
1	10.1	10.32	0	8	Krintin Slide	0.31	0		0	pluck	d	1	TRUE	0	0.03	0.32	2.22
									0.32	hammer-off		0	TRUE	0		0.32	7.54
									0.64	slide		6	TRUE	-1		0.36	9.84
												6	TRUE	-1			
			1	0	Fixed	2.4	0.31					6	TRUE	-1	0.23	1	
			2	0	Fixed	1.21	2.71		0	pluck	d	6	TRUE	-1	0.12	1	
			3	12	Silent	0.11	3.92								0.01	1	
			4	2	Bend: Sloped Start	0.3	4.04	2	0	pluck	d	0	TRUE	0	0.03	1	
												6	TRUE	-1			
			5	0	Fixed	0.46	4.34					6	TRUE	-1	0.04	1	
			6	4	Bend: Ladle	0.97	4.8	2	0	pluck	d	2	TRUE	0	0.09	0.51	
												1	TRUE	0		0.49	
												2	TRUE	0			
			7	0	Fixed	2.09	5.77		1	dampen		2	TRUE	0	0.2	1	
			8	12	Silent	0.17	7.86								0.02	1	
			9	0	Fixed	2.28	8.03		0	pluck	d	2	TRUE	0	0.22	1	

Fig. 7. Excel sheet representation of transcription data.

Next Steps

We have identified the following five components of research as next steps in the upcoming two years.

1) Enhancing the current version of the notation/ transcription web application. These include: developing synthesis engines for bowed-string, reed, wind, hammered, and percussion instruments; attending to timbre and non-melodic parameters such as amplitude and dynamics for appropriate instruments; and expanding the query functions to produce datasets and data visualizations across transcriptions (intra in addition to inter performance analysis). The percussion components require the development of a novel visual representation scheme that will exist alongside the two-dimensional pitch-by-time graphic strategy of our melodic instrumental representations.

2) Expanding the representational capacities to Carnatic music (South Indian classical music) and traditional Korean court music. We have already beta-tested the IDTAP with a variety of vocal and instrumental traditions and will work with scholars, musicians, and doctoral students on the advisory board who specialize in Korean and Carnatic traditions.

3) Conducting video, audio, and ultrasound recording sessions with visiting musicians from South Asia and Korea at the electronic music studio lab at UCSC. These data-collection events will provide important materials for the development of audio-synthesis algorithms and machine-learning approaches to auto-transcription, and for the understanding of the organological/ phonological particularities of sound traditions and their related spoken languages.

4) Transcribing recordings, which serve two purposes: 1) building the digital archive for users for their research, pedagogical, and music composition, and 2) serving as training data for machine-learning models.

5) Developing a methodology for automating some of the transcription process. Under the guidance of our advisory board, we draw on the work of Google Magenta researchers who have demonstrated that a well-curated dataset of approximately two hundred hours of audio recordings can be used to train a sequence-to-sequence Transformer neural network model to make accurate transcriptions [18].

7 Conclusion

There are many possible encoding schemes and graphical models for representing musical information. Different representations are more or less suitable for different kinds of tasks, be they descriptive or prescriptive, exacting (along a particular subset of parameters of discernibility) or ambiguous, archival or generative, textual or graphic, abstract or concrete. Extant visual representations are socio-historical artifacts, outgrowths of complex social feedback systems with particular historically-informed quirks and affordances. Each such schema stratifies around and is constitutive of particular social practices (artistic, scientific, commercial, spiritual). Like languages, they persist as long as a *fluent* community continues to engage with and develop corresponding activities. Practitioners conversant in any of these systems will have molded themselves into a bodily/psychological configuration with such symbol systems, and with the acts of reading, interpreting, engraving, and sonifying. Just as this training extends their ability to affect and perceive the sonic, it also necessarily establishes a limit, confining them within circumscribed habits of thought and activity. Indeed, experience with any such engraving technology cultivates a conceptual world and a corresponding technological apparatus for exploring it, but there are many possible worlds. The IDTAP aims to engender one such world for a 21st century community of practitioners and researchers to engage with oral traditions and their corresponding archives, and to allow these traditions to flourish among the hybrid AI artistic praxes to come.

References

[1] C. Seeger, "Toward a Universal Music Sound-Writing for Musicology," *Journal of the International Folk Music Council*, vol. 9, pp. 63-66, Jan. 1957.

[2] N. M. England et al., "Symposium on Transcription and Analysis: A Hukwe Song with Musical Bow," *Ethnomusicology*, vol. 8, no. 3, pp. 223-233, Sep. 1964.

[3] M. Hood, *The Ethnomusicologist*. New York, 1971.

[4] T. Ellingson, "Transcription," in *Ethnomusicology: An Introduction*, H. Myers, Ed. Basingstoke, 1992, pp. 153-164.

[5] S. Mclary, *Feminine Endings: Music, Gender, and Sexuality*. Minneapolis: Univ. Minn. Press, 1991.

[6] J. Kerman, *Contemplating Music: Challenges to Musicology*. Cambridge, MA: Harvard University Press, 1985.

[7] L. Goehr, *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford: Clarendon Press, 1992.

[8] M. Citron, *Gender and the Musical Canon*. Cambridge: Cambridge University Press, 1993.

[9] L. Gandhi, *Affective Communities: Anticolonial Thought, Fin-de-siècle Radicalism, and the Politics of Friendship*. Durham: Duke University Press, 2006.

[10] S. Rao and W. van der Meer, "Music in Motion: The automated transcription for Indian music (AUTRIM) project by NCPA and UvA," 2017. [Online]. Available: https://autrimncpa.wordpress.com/. [Accessed Jan. 16, 2024].

[11] M. Clayton et al., "Interpersonal Entrainment in Indian Instrumental Music Performance: Synchronization and Movement Coordination Relate to Tempo, Dynamics, Metrical and Cadential Structure," *Musicae Scientiae*, vol. 23, no. 3, pp. 304-331, 2019.

[12] M. Clayton, L. Leante, and S. Tarsitani, "IEMP North Indian Raga," Database of annotated audiovisual recordings, 2018. [Online]. Available: https://doi.org/10.17605/OSF.IO/KS325

[13] P. E. Savage, "An overview of cross-cultural music corpus studies," in *Oxford Handbook of Music and Corpus Studies*, D. Shanahan, A. Burgoyne, and I. Quinn, Eds. New York: Oxford University Press, [In press]. Available: http://doi.org/10.31235/osf.io/nxtbg

[14] D. H. Klatt, "Software for a cascade/parallel formant synthesizer," in *J. Acoust. Soc. Am.*, vol. 67, no. 3, pp. 971-995, 1980.

[15] K. Karplus and A. Strong, "Digital synthesis of plucked-string and drum timbres," *Comput. Music J.*, vol. 7, no. 2, pp. 43-55, 1983.

[16] C. Schörkhuber, A. Klapuri, N. Holighaus, and M. Dörfler, "A MATLAB toolbox for efficient perfect reconstruction time-frequency transforms with log-frequency resolution," in *Proc. 53rd International Conference: Semantic Audio*, Audio Engineering Societ*y*, Jan. 2014.

[17] J. Salamon and E. Gómez, "Melody extraction from polyphonic music signals using pitch contour characteristics," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 6, pp. 1759–1770, 2012.

[18] C. Hawthorne et al., "Enabling Factorized Piano Music Modeling and Generation with the MAESTRO dataset," in Proc. International Conference on Learning Representations, 2019. [Online]. Available: https://openreview.net/forum?id=r1lYRjC9F7

Crafting the "elevator pitch": joining rule-based and stochastic methods for musical procedural rhetoric in computer-aided composition

Juan S. Vassallo

University of Bergen
Juan.vassallo@uib.no

Abstract. This paper examines the compositional process of "Elevator Pitch", a piece for cello and electronics. The work comes as a response to an artistic inquiry delving into how computational methods from the field of symbolic artificial intelligence can influence the development of compositional frameworks, in particular the combination of constraint algorithms with machine learning and stochastic techniques. From an artistic perspective, these methods are conceived as analogous to classic rhetorical formulas and serve as tools to organize both the local scope and overall form of the piece. I have termed this framework as *musical procedural rhetoric*. The idea of rhetoric formulations embodied in musical procedural rules strongly relates to the piece's conceptual backdrop, as a sardonic analogy of a political speech. The discussion on the constructive aspects of the piece further addresses aspects regarding the formalization of qualitative musical dimensions using an experimental expansion of the library Cluster-Engine in its implementation for Max.

Keywords. Computer-aided composition, constraint algorithms, Markov chains, rhetoric, procedural rhetoric.

1 Introduction

The Basics of Rhetoric

Since antiquity, the study of rhetoric has been considered as a set of skills aimed at achieving a more persuasive manner of speaking and writing. The effectiveness of the rhetorical situation is usually facilitated through several types of linguistic artifacts, known as *rhetorical figures* or *figures of speech*. Examples of these figures of speech as *simile, metaphor, allegory, hyperbole, irony, antithesis, metonym, synecdoche, prosopopoeia, apostrophe, climax, incrementum, erclamation, and interrogation, among many others*. The essence of these figures of speech relies mainly on ornamental discursive techniques and entails an alternative meaning or function apart from its literal content.

The classic *Quadripartita Ratio* has historically been used as a framework for manipulating and varying discourse across various levels of linguistic structure [2]. In it, the figures of speech are classified within the four fundamental rhetorical operations: *adiectio*, *detractio*, *immutatio*, and *trasmutatio*. *Adiectio* involves adding or incorporating additional information to an argument or statement to reinforce a particular point. Most of the time refers to a form of arranging words or clauses in a sequence of increasing force. *Detractio* involves leaving out information or certain details to emphasize other aspects of an argument. *Immutatio* involves rearranging the order of words or phrases. Finally, *transmutatio* involves changing the form or structure of a word or phrase while preserving its meaning. In contemporary times, the *Quadripartite Ratio* has been used as a point of departure for the study and development of generative strategies for constructing narratives, and as a method of imitation by which one could transform a speech model into something new and original [3].

Rhetoric and Music

Historically, music has been influenced by rhetoric theory and terminology [4-6]. During the Renaissance and Baroque, composers often incorporated elements of rhetoric in their works, as they viewed an intrinsic connection between the rhetoric of language and the capability of music to stir and evoke emotions in listeners through the use of *affects*[1]. Composers such as Claudio Monteverdi displayed a profound comprehension of the relationship between music and lyrics in vocal music. In 1605, Monteverdi introduced the distinction between the *prima prattica* and the *seconda prattica*, or first and second practices, something that marked a shift in composition, moving from an established style to a more modern approach. In addition, composers of instrumental music also drew from the field of rhetoric to articulate their composition techniques [7]. Proficiency in using these figures emerged as a standard for measuring the aesthetic appeal and value of a musical piece during this time. Musical features such as melodic repetition, fugal imitation, dissonance or consonance, intervallic movement patterns, and even silence were formulated into processes widely applicable in diverse musical contexts (see Table 1).

In the 18th and 19th centuries, newer generations of composers began to view aesthetic ideals based on rhetorical formulas as outdated. Despite this shift, conventional musical forms, such as the sonata, continued to be analyzed using terms of rhetoric by writers and theorists of the time, in particular Johann Matheson, who highlighted the

[1] Affects have been defined as *emotional abstractions* or *idealized emotional states* [6, p. 12] that a composer was expected to arouse in the listener through a musical work.

importance of the *sentence* structure in shaping a rhetorical musical style, stating its intrinsic link to the musical idea itself. Mattheson viewed the need to employ rhetorical concepts to compose music, by claiming that *a good composition should have the same form as a good speech* [8, p. 48], and these views extended into the early 20th century [9]. However, as traditional elements structuring musical discourse started to dissolve, new forms of musical organization took shape, in open criticism of the rhetoric language-based ideas. Ultimately, the scientificist discourse of avant-garde movements in Europe further deviated from the idea of music as a speech-based art.

TABLE I. MUSICAL EQUIVALENT FOR RHETORICAL FIGURES.

Rhetoric figure	Musical equivalent	As described by:
Anaphora	The repetition of a melodic statement in different notes in different parts.	A. Kircher (1601-1680)
Auxesis	Ascending consecutive repetitions of a musical passage.	J. G. Walther (1684-1748)
Catabasis	Descending musical passages expressing negative images or affections.	A. Kircher (1601-1680)
Climax	A gradual increase or rise in sound and pitch, creating a growth in intensity.	J. A. Scheibe (1708-1776)
Epizeuxis	Emphatic repetition of a motif or melodic phrase.	J. G. Walther (1684-1748)

Musical procedural rhetoric

The discipline of rhetorical code studies (RCS) has developed recently [11,13]. This area of study combines the classic understanding of rhetoric—how language can persuade and influence—with the study of software and computer code, essentially treating computational algorithms as rhetorical agents that function persuasively to induce or influence an audience to act in response to a particular message. In RCS, an algorithm is viewed as a system of signs, just like language, with its own grammar and rhetorical tools. An important concept is the idea of procedural rhetoric, developed by Ian Bogost [12]. This term refers to the idea that processes, especially the ones used in computing, are a means of persuasion. In this light, computational algorithms represented in computer code are viewed as rhetoric agents from a perspective based on procedure as a form of facilitating action.

The conceptual link between constrained computational methods in computer-aided composition and rhetorical principles, particularly *procedural rhetoric*, stems from the use of rhetoric-inspired formulations expressed as musical constraint rules in computer code, to generate a musical narrative. While I take a somewhat loose and metaphorical approach to this translation process, in my recent compositional output, the concept of musical procedural rhetoric has provided me with interesting considerations for delving into diverse facets of musical organization.

The compositional methodology for this piece explores a branch of computational methods for artificial intelligence known as Constraint Satisfaction Programming (C.S.P). In the field of computer-assisted composition, C.S.P has a long history [14-16], as it has served as a powerful tool for creating musical narratives by implementing musical rules expressed as logical statements, a process that is at the core of the concept of computational algorithms and algorithmic composition. In my recent work, the method of C.S.P is used in classical ways related to the organization of conventional parameters such as pitch and rhythm, as well as in what I see as not-so-well-explored implementations of it for contemporary music composition, something that I will explain further in this text.

2 Conceptual Backdrop of the Piece

Social Acceleration and newer communication dynamics: The *elevator pitch*

The philosopher Hartmut Rosa has proposed that modern society can be described as an *acceleration society*, by observing the evidence of technological acceleration and an increase in the pace of life due to a chronic shortage of time resources [17]. As contemporary society embraces a fast-moving dynamic where instant updates are consumed through smartphones and social media engagement is a cornerstone of a technology-interconnected world, everyday communication becomes *accelerated* and *temporarily fragmented* [18, p. 1]. In light of this new paradigm, newer forms of communication that emphasize the briefness and succinctness of information delivery thrive. One of these new types of discourse is the *elevator pitch*. An elevator pitch is a type of summary speech that defines an idea, product, or company and a value proposition for it. The name reflects the idea that it should be possible to deliver a summary of the topic in the duration of an elevator ride, usually between approximately 30 seconds and two minutes. For an elevator pitch, it is important to communicate things in a manner that is understandable for large segments of the population in a short time, and it is also important to increase the communicational effectiveness, to induce a listener to make a decision sought by the speaker.

It is not until recently that the high effectiveness of these discursive methods has been studied from an empirical perspective [19], and this has probably contributed to these pitches landing recently in the world of political communication. The phenomenon of the professionalization of political communication, mainly represented by the rise of *political marketing*, has been studied as a means of social manipulation in contemporary Western democracies [20-22]. Some of the techniques of these new political communication strategies and discursive modes rely on the generation of strong and brief key messages where, by employing a directed selectivity on the supporting information presented, the full disclosure of the facts is avoided. Thus, a

political discourse ruled by the dynamics of these communication strategies is usually deprived of a meaningful message and converted into a performative act aimed solely at exacerbating passions, where audiences are targeted for plain opinion driving on complex social issues where subjectivities and hate discourses are prone to arise, such as those regarding immigration, minorities rights, among other topics. In many cases, the use of these communication strategies has been linked to authoritarian and intolerant forces that use them as the entry door to step into the political scene.
The piece is a sardonic analogy to a political speech, which is portrayed here as empty of substance, and as an artificial construct of a "laboratory" rhetoric relying purely on historical conventional rhetorical strategies. The piece thus poses an artistic focus on these contemporary methods of communication -such as an *elevator pitch*- and the potential for manipulation of sound-bite content by political figures. The title suggests an ironic connection between a content-lacking political speech and an algorithmic musical organization. However, despite the potential analogy between the creation of musical narratives through constraint computation and a real-world phenomenon, the composition itself does not intend to represent this idea. Rather, the chosen compositional methodology serves the purpose of creating an artistic allegory of *emptiness* in political discourse. In addition to the political or politicized message that the piece aims to convey, the study and development of the methods employed here are aimed at exploring and creating valuable artistic works.

3 Compositional Methods

Two overall main inquiries were explored when creating this piece. Firstly, the constraint rules that govern the local organization of the musical narrative focus mainly on idiomatic sound production types for the violoncello. An important motivation for this has to do with exploring ways of overcoming biases in computer-assisted composition that often view qualitative or *weak*[2] musical parameters as secondary to pitch and rhythm[3]. For that, some of the compositional work was concerned with

[2] The concept of weak musical parameters is not universally accepted. However, the parameters often considered most potent are pitch, rhythm, and dynamics, with articulation sometimes viewed as a more subtle or nuanced aspect of music expression.

[3] Initially, this sentence refers to the dominance of pitch and rhythm as the primary musical compositional domains through the history of Western music. This was addressed critically by James Tenney in the early 60s, as he proposed that every musical parameter could hold structural importance within a musical phrase [10]. However, in computer-assisted composition, I believe that the issue of hierarchical formalizations persists. At first glance, an observation of the established theory and practice in the field seems to consider mainly those musical domains directly related to pitch (note, pitch class, frequency, MIDI cents, etc.) and duration (milliseconds, rational values, onsets, etc.) as the only valid for formalizations and compositional operations. In this sense, I believe that further compositional research towards formalizations of diverse musical parameters is desired.

investigating diverse ways of how these can be formalized and operated as a musical *search space*[4]. Secondly, the implementation of a *steered* Markovian generation process[5], facilitated by the combination of constraint algorithms and Markov chains by using a Markovian transition table as a constraint rule. Conventionally, it is accepted that stochastic processes such as Markov chains effectively capture overall musical tendencies (general distribution of events in time, ambit, etc.), at the cost of neglecting the local scope of musical information due to their inherent probabilistic nature. On the other hand, constraint algorithms offer a more detailed control through the definition of rules that constrain the local scope (individual notes, rhythmic figures, time position, harmonic intervals, etc.) [25]. Therefore, it is expected that a combination of these two methods could potentially yield more flexibility for developing sets of rules seeking a compositional balance between these two scopes.

Fig. 1. A schematic representation of the flow using constraint algorithms. A domain can consist of any symbolic information (such as pitches, durations, text, etc.). The three subsequent modules exemplify a chain of three rules. The constraint engine enforces these rules in possible combinations of elements, and finally, the solution -or solutions- is output.

[4] In the field of Artificial Intelligence, the set of solutions among which the desired solution resides is called *search space* or *state space* [23].

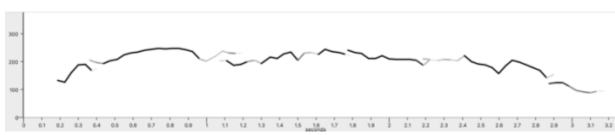
[5] Previous investigations on steered Markovian generation have been carried out for example Pachet and Roy [24]. My implementation essentially follows the same logic, namely, using deterministic control constraints, and defining the Markovian property as a cost function to optimizein my case, the weight for a heuristic rule-. My interest, however, was to implement this process in Max, using the CAC libraries 'PWconstraints' and 'Cluster-Engine' recently ported to Max by Julien Vincenot and Örjan Sandred.

For the composition of the piece, I relied extensively on a computer-assisted composition tool named Cluster-Engine created by Örjan Sandred [26], a constraint engine specialized for music composition that is part of an external Max[6] library named MOZ'lib[7]. I have focused on the experimental feature of Cluster-Engine known as *Multidomains*, which allows the generation of constraint relations between any musical information existing within a polyphonic texture that can be encoded symbolically as numbers or strings of characters, for example, instrumental articulations and techniques, lyrics or any other type of score annotations[8].

4 Musical material and form

The musical material -in the form of rhythmic and melodic gestures- is mainly derived from the sonification of a speech fragment[9] from Donald Trump. The opening phrase of Trump's inaugural speech as president of the United States, pronounced on Jan. 20, 2016[10], *"The time for empty talk is over"* was analyzed using the Spear[11] software focusing on the duration and energy of its fundamental frequency[12], which was later translated into symbolic information as representations of a melodic contour.

(a)



(b)

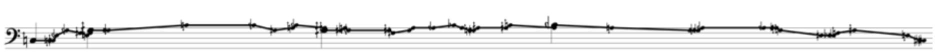


Fig. 2. Above, graph (a) represents the fundamental frequency contour of the phrase "the time for empty talk is over", pronounced by Donald Trump. Below, in (b) is shown the phrase's melo-rhythmic profile.

6 https://cycling74.com/

7 MOZ'lib can be downloaded here: https://github.com/JulienVincenot/MOZLib

8 Other libraries that address this issue are for example Abjad and Straheela [27, 28].

9 Some historical composers who have developed methodologies related to the translation of speech and prosody information into musical notation are Johnathan Harvey [29] and Peter Ablinger [30]. Currently, I have been interested in the work of Fabio Cifariello Ciardi [31, 32].

10 The video of it can be seen here: https://www.youtube.com/watch?v=XZn8tFbISpo (originally retrieved on January 30, 2023)

11 https://www.klingbeil.com/spear/

12 The fundamental frequency or F0, is usually defined as the lowest frequency of a periodic waveform. In the case of speech, it usually equals the perceived melodic contour of a spoken phrase.

Formally, the piece is constructed as a linear flow of chained musical gestures derived from this sonification that moves forward as the dialogical interaction between the instrumental part and the electronics, driven by sound events' chronometric density and *expansions and compression of the musical parameters of frequency and intensity*, allowing for a build-up of energy leading to a formal climax and a posterior recession. The chosen *archetypical* musical form for this piece is strongly connected with the Aristotelian concept of *narrative* or *curse*[13] (fig. 3), which emphasizes that stories should have a narrative structure with a logical chain of cause and effect where earlier events are related to current events and will be perceived as connected to future developments [34].

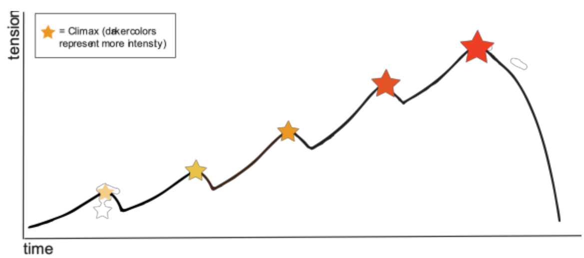


Fig. 3. A schematic representation of the Aristotelian narrative.

The piece begins with breath sounds presented in the electronics, a sound that proposes a type of *inward-outward* gesture and an interruption of the flow towards the end (halted inhaling, silence, and irregular exhaling). These gestures are important and will keep developing throughout the piece, as seeds for processes of expansion and compression and interruptions of the flow. Later in the piece, this category of inward-outward translates into sounds produced with *vibrato*, *tremoli* between harmonics, using notes which are either half-stopped or fully stopped in the left hand and *vertical bowing tremolo* (fig. 4). The gesture of interruption becomes evident within the narrative as certain musical flows are abruptly cut and juxtaposed with contrastive elements. It is possible to see especially in fig. 5, when the appoggiaturas start breaking the flow in increasingly shorter periods, and in the final section from m. 96. The inward-outward gesture becomes the seed for the posterior process of expansion of melodic profiles as *upward-downward*, metaphorically related to the equivalent rhetoric processes of *additio* (or *auxesis*) from the Quadripartite Ratio (fig. 7).

[13] In connection with the conceptual background of the piece, some authors have discussed the persuasive power of *Aristotelian* narrative and how it has been applied to various forms of communication, from fiction to product pitches [33].

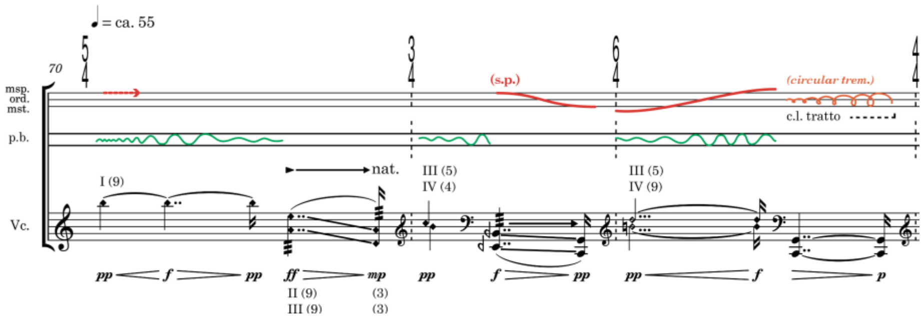


Fig. 4. MM. 70-72.

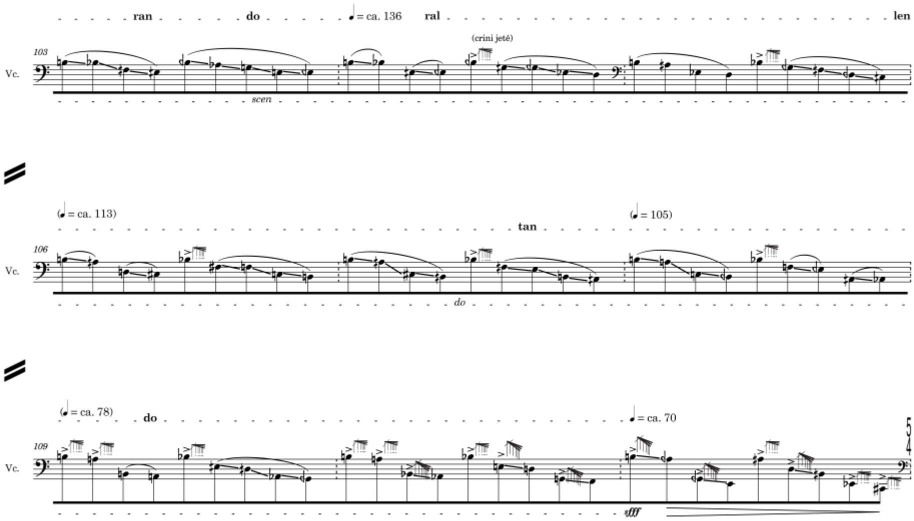


Fig. 5. MM. 102-111.

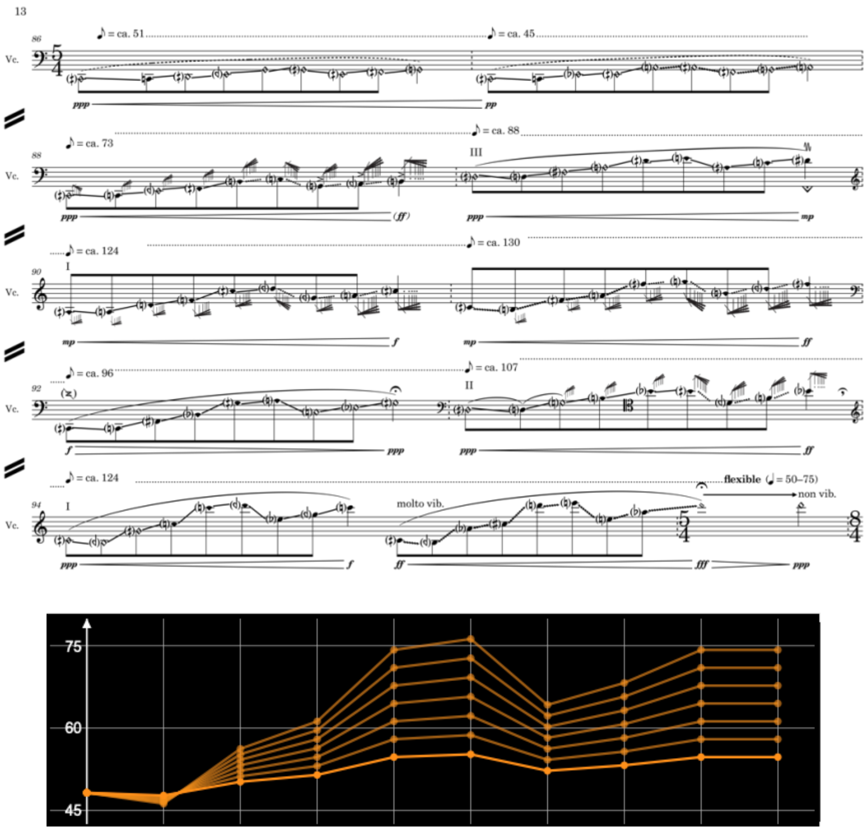


Fig. 6. Measures 86-96. Below can be observed a representation of the same melodic succession and how it expands over the pitch range, as break-point function lines. The x-axis represents MIDI notes, and the y-axis represents time.

After the highest energetic point of the piece in m. 125 (fig. 7), a final fragment breaks the dynamic of the rule-based instrumental sound production unfolding, proposing a shift of the scope of the rules towards durations and pitches. In contrast to the rest of the piece, this last section sounds ironic and somewhat simple, relating a deliberately simple musical narrative resulting from the use of constraint rules for operating permutations of pitch and durations (fig. 8). Moreover, employing a superball to rub the back of the fingerboard imparts a sonic texture reminiscent of nonsensical utterances.

The gesture of interruption of the flow by the lower overpressure C_2, plus the use of silence ends up completely breaking apart the flow towards the end of the piece.

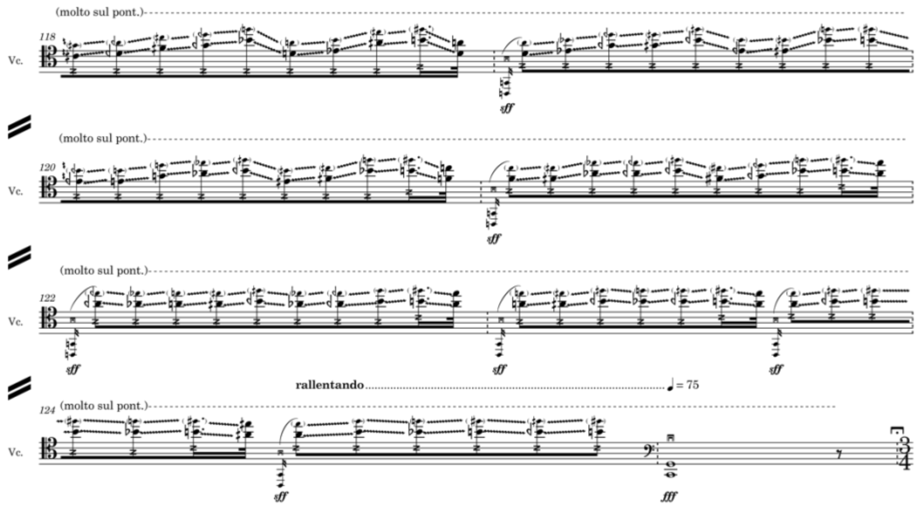


Fig. 7. (a) Climax, m. 118-126.

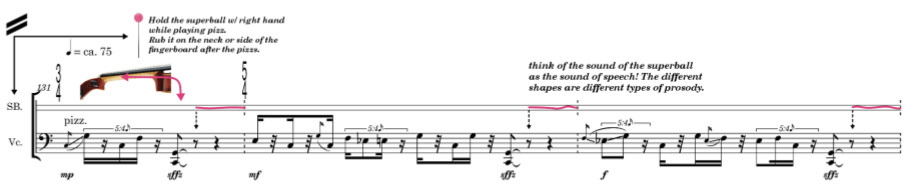


Fig 8. Closing section, m. 131-133.

5 Constraint rules

The permutational and recombinational logic implemented for the structuring of musical parameters inspired by the idea of *permutatio* from the Quadripartite Ratio is enforced by mainly two constraint rules, expressed as LISP code[14]. The first rule involves the repetition of elements within sublists, while the second tackles repetition across consecutive pairs of sublists. In the first example (a), within each sublist, only two elements should repeat, and the rest should be different. In the second example (b), only two elements are repeated in consecutive sublists, and the elements within each sublist are different. Below these rules are expressed as lisp code:

[14] LISP is a programming language considered the father of the programming languages used for artificial intelligence. It is extensively used in computer-assisted composition, in environments such as Open Music, PatchWorks, and more recently Max [35, 36].

```
; rule (a)
( lambda ( a b c d )
  ( = 3
    ( length
      ( remove-duplicates ( list a b c d ))
    )
  )
)

; rule (b)
( lambda ( a b c d e f g h )
  ( progn
    (setq sublist1 ( list a b c d ))
    (setq sublist2 ( list e f g h ))
    ( = 2
      (let ((count 0))
        (dolist (element sublist2)
          (if member element sublist1)
            (incf count)))
    count)
  )
)
```

Rule (a)
(**1 1** 2 3) (1 **2** 4 **2**) (**4** 2 **4** 3) (**2** 3 **2** 1) (**3 3** 1 4) (**4** 1 2 **4**) **etc…**

Rule (b)
(**1 2** 3 4) (**1** 5 7 **2**) (**2** 4 **5** 3) (**3 2** 1 7) (**2** 4 5 **7**) **etc…**

Fig. 9. LISP rules (a) and (b), and examples of results using rules them.

These rules have been applied for the temporal organization of the bowing position, bowing pressure, dynamics, various types of tremolos or vibrato, and instrumental articulation. The candidate elements for this rule consist of a list with all the possible bowing positions -in this case, the choices have been restricted to *ordinario* (ord), *sul ponticello* (sp), and *sul tasto* (st) all the possible trajectories between them. In Figure 3 can be seen how rule (b) constrains the bowing positions. The rule thus enforces that two elements should be common within consecutive sublists of four elements (from bar 1 to 2, the two common elements are 's.t.-s.p'. and 'ord').

Naturally, as more musical parameters become constrained simultaneously, the combinatorial possibilities increase exponentially. In the following example, the constraint engine is used to generate a musical sequence by constraining three musical *voices*[15]: durations, bowing position, and bowing pressure. Here, the three voices are constrained in a cascade fashion: the first voice (V0) follows a permutational rule, and the candidates for the rest of the voices are constrained to certain rhythmic figures. For example, longer bowing trajectories (e.g. “sp-st”) can only occur in whole notes. Static bowing trajectories (e.g. “sp”, “ord”, “sp”) are restricted to quarter and eight notes, and intermediate trajectories occur only in eight and sixteen notes. In addition, voices 1 and 2 are constrained by the rule (a) discussed above.

TABLE II: CASCADE-TYPE ORGANIZATION OF CONSTRAINTS.

Voice/Param.	Candidates			Rules
V0 - Rhythm	𝅗𝅥	♩ ♪	♪ 𝅘𝅥𝅯	Permutations: 𝅗𝅥 ♩ ♪ 𝅘𝅥𝅯
V1 – Bow pos.	(st-sp) (sp-st)	(ord-sp) (ord-st) (sp-ord) (st-ord)	(ord) (sp) (st)	Rule (b)
V3 – Bow press.	(flaut-overp) (overp-flaut)	(flaut-ord) (overp-ord) (ord-flaut) (ord-overp)	(flaut) (ord) (overp)	Rule (b)

6 Electronics

The electronic part is designed as a live-electronics component, originating from the combination of signal processing and sound synthesis with the use of speech sound samples that are played via transducers through the instrument’s body, transforming it into a sort of *speaking* cello. These samples consist mainly of stochastically constrained

[15] *Cluster-Engine considers each musical dimension to be constrained as a musical voice in a polyphonic texture.*

generated text transformed to speech by the software VoxBox by Myifone[16], using a synthetic voice that sounds like Donald Trump.

Stochastic constraints

In order to create the text for the electronic voice, I implemented a generative process that relies on the combination of constraint rules with Markov chains. In practice, the constraint engine uses a Markovian transition table as a constraint rule. This means that the constraint engine checks that every combination of candidates (the candidates are letters or syllables, depending on the order of the Markov chain) is represented as a transition state from the Markov matrix. As Markov chains are based on probabilities for certain transitions to occur, a second rule checks that each transition in the solution matches the probability of appearing as it exists in the Markov matrix. As constraint rules are feasible to be chained, Markov rules can be combined with other constraint rules, for example -in the case of using text-, patterns of rhyme and alliteration[17], or poetic structures.
The example below was created using a 1st-order Markovian transition table of Trump's inaugural speech as a constraint rule in combination with the aforementioned lisp rule (a). Namely, for every four letters, two should be repeated and two different.

The resulting text follows a 1st-order Markov probability matrix of Trump's speech but at the same time follows the aforementioned rule (a):

Peoeceacayaianinvive
Ebyeysesrerbrcrdrudu
Mphpapipepreroeonenb
Eiuinifieferecrcarak

In the following example, I created five shorter words using the same rules discussed above, but I added a rule that constrains the text output to have a determined endinrhyme pattern (a b a b a). The time for searching the possible solutions became considerably longer:

Daldlydyoy (a)

[16] VoxBox by Myifone allows the user to generate speech using AI-based voices trained with voices of celebrities. It can be downloaded here: *https://filme.imyfone.com/voice-recorder/*

[17] It is important to mention that equal typographical beginnings or endings don't imply alliteration or rhyming -and vice versa- in the English language. This issue has been addressed in a different piece, in which this same process was done using the IPA alphabet. In that case, the correspondence between symbols and sounds is univocal. In any case, the result using the Latin alphabet was not so far from what I expected.

Gefeoenean (b)
Puspsosyoy (a)
Olsosasnan (b)
Ewbebyeyoy (a)

To take these ideas further, I expanded the Markov transition table into 2nd and 3rd order, and I started creating sentences with rhyme schemes such as (a b a b a b), as verses from a poem:

Wards overnmen itselves cities (a)
Jobs families alle ones an car (b)
Kingth strulers ourse of dones (a)
Ves of mich unrealittle restar (b)
Lves rebuild first anot to res (a)
Fusing fourished intone an car. (b)

Finally, I implemented combinations of rules in order to create a sonnet with a rhyme scheme abab cdcd efef gg:

Rusten again trulerse ten your in ther,
Rospense future truly mich thangs the an,
Ges contry andscattening its which pover,
Ysted it behindustem flourseasonable jan.
Magnificential mentermies that that nowled,
Repain is neight hearterseasonal trates,
Everts to everts we trationable for sted,
Deprived herent is americh with no longes.

Ge at it bush poverts deple by on today,
Lef buted ind little defend of dolle thou,
Dle face ten to be forgotte rememble decay,
Tries has willions armies that to cou.

Job dolledge alle has neign inner shards,
His movery at it back togethen too mands.

7 Discussion

Formalization strategies

When composing this piece, I was also interested in exploring how computational rule-based methods can be harnessed to include a range of musical parameters exceeding the typical confines of melody, harmony, and rhythm, for example, a wide array of instrumental techniques including bowing position, bow pressure, types of *tremoli* or vibrato. In most cases, these parameters exist as text indications or instructions. In the field of statistics, these are known as *categorical* or *qualitative* variables, and a conventional strategy to formalize categorical variables into mathematical models - which in this case can be equated to computer-assisted composition workflows- is to use logical constraints [38]. Cluster-Engine, in its version for the Max platform, is designed around the Bach[18] external library's environment, and it includes an experimental feature named *Multidomains*, which represents a powerful expansion of its capabilities, as it allows logical constraining of qualitative information as metadata contained in Bach's *slots*, such as score annotations, articulations, and lyrics, among many others. Thus, to formalize and organize these musical domains in the piece, two strategies were taken. The first one was to create logical constraint rules that were expressed mainly as string equality tests. (e.g., bowing positions such as *ord, sp, st, mst*, etc.). For example:

```
( function ( lambda ( a )
 (if ( > ( first a ) 1/8 )
    ( not ( equalp (second a ) 'stacc))
 t )
))
```

Fig 10. Example of a logical constraint rule consisting of a string equality test expressed as follows: If the first value of the list *'a'* is larger than 1/8 (eight-note), the second value of 'a' should be not *'stacc'*.

Secondly, candidates were encoded as integer numbers that were later translated into strings by following a simple chart. This turned out to be a much more efficient way, faster, and lighter in terms of computation resources. For example[19]:

```
( function ( lambda ( a )
 (if ( > ( first a ) 1/8 )
    ( not ( = (second a ) 1))
 t )
))
; 1 = "stacc"
```

[18] https://www.bachproject.net/

[19] This rule is a simplification for the purpose of exemplifying the formalization technique. The actual rules used in the piece involved more complex interactions between musical parameters.

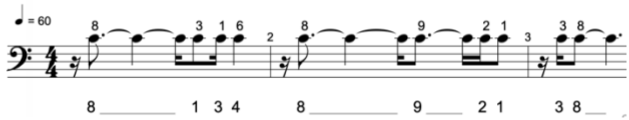



Fig. 11. Example of a logical constraint consisting of a numeric equality test, used to organize the bowing position of the cello over a melodic sequence following the aforementioned **rule (a)** *-only two repeated elements in sublists of length 4-*. Below is shown the output of this rule.

8 Conclusions/Final Reflections

For the composition of this piece, I have implemented a compositional framework that I have termed *musical procedural rhetoric*. The process aims to exemplify how the logic and procedural nature of constrained computation and stochastic methods can be combined to create complex and ultimately expressive musical narratives. In this piece, the sound material, and the constructive method are anchored to a conceptual backdrop, based on the ironic resemblance between an elevator pitch as a content-devoid political speech and an artificial construct implementing a type of 'laboratory' rhetoric, and its potential misuse in political communication as a tool for thought manipulation. The selection and use thought the piece of Donald Trump's speech is not coincidental, as in my view, he exemplifies a type of politician who relies on carefully crafted, rehearsed laboratory-made speeches which most of the time are emptied of meaningful content and are reliant on cliches and effectist phrases. As I see it, the aforementioned speech is a clear example of this. Although Trump may be the most distinctive one and well-known due to his highly mediatized profile and history, he is certainly not the only one and examples of these can be found across the full spectrum of political ideologies. In this sense, an important question -which is still unanswered-revolved around the problem of bringing Trump's voice to the center of the scene, without turning the piece into a parodical cliché, or being labeled with a political stance that observes no nuance, which is not the case. A critical reflection around this problem is still onging.

Finally, the composition of this piece is part of a process of addressing an important personal artistic research inquiry. As a composer mainly reliant on computer-aided composition workflows, I aim to investigate how to create musical narratives that seek a deeper connection between ideas, poiesis, sound results, and reception, aiming to go beyond a merely formalistic approach. However, the audience is ultimately who should judge the aesthetical value of these methodologies.

Acknowledgments

I express my heartfelt gratitude to my supervisors, Daniel Peter Biro and Örjan Sandred, for their generosity and invaluable shared knowledge. I would also like to extend my thanks to Tijs Ham and Julien Vincenot for their comments and suggestions. A special appreciation goes to Lucas Fels for her outstanding performance in the recording of this piece. I am sincerely grateful to the Norwegian Artistic Research Program, the University of Bergen, and the Grieg Academy for their generous funding and support of my artistic research project. Additionally, I would like to thank Norsk Komponisforeningen for the financial support received for this project.

References

[1] K. R. Wallace, Understanding discourse; the speech act and rhetorical action. Baton Rouge: Louisiana State University Press, 1970.

[2] H. Lausberg, Handbook of literary rhetoric : a foundation for literary study (Handbuch der literarischen Rhetorik). Leiden: Brill, 1998.

[3] J. Dubois, F. Edeline, J.-M. Klinkeberg, P. Minguet, F. Pire, and H. Trinon, A General Rhetoric. Baltimore and London: The Johns Hopkins University Press, 1981.

[4] J. W. Hill, "Baroque music : music in Western Europe, 1580-1750," ed. New York: Norton & Co, 2005.

[5] B. Wilson, G. J. Buelow, and P. A. Hoyt, "Rhetoric and music," ed: Oxford University Press, 2001.

[6] G. J. Buelow, "Music, Rhetoric, and the Concept of the Affections: A Selective Bibliography," 1973, vol. 30.

[7] D. Bartel, Musica Poetica: Musical-Rhetorical Figures in German Baroque Music. U of Nebraska Press, 1997.

[8] H. Lenneberg, "Johann Mattheson on Affect and Rhetoric in Music (I)," Journal of Music Theory, vol. 2, no. 1, pp. 47-84, 1958, doi: 10.2307/842930.

[9] K. H. Eschman, "The Rhetoric of Modern Music," The Musical quarterly, vol. VII, no. 2, pp. 157-166, 1921, doi: 10.1093/mq/VII.2.157.

[10] J. Tenney, Meta (+) Hodos: A Phenomenology of Twentieth Century Musical Materials and an Approach to the Study of Form. Frog Peak Music, 1968.

[11] E. Beck, "A Theory of Persuasive Computer Algorithms for Rhetorical Code Studies," Enculturation: A Journal of Rhetoric, Writing, and Culture, pp. 1-29, 2016. [Online]. Available: https://www.enculturation.net/a-theory-of-persuasive-computer-algorithms.

[12] I. Bogost, Persuasive Games : The Expressive Power of Videogames. Cambridge, MA: MIT Press, 2007.

[13] K. Brock, Rhetorical Code Studies: Discovering Arguments In And Around Code. University of Michigan Press, 2019.

[14] J. L. A. Hiller and L. M. Isaacson, "Musical Composition with a High-Speed Digital Computer," Journal of the Audio Engineering Society, vol. 6, no. 3, pp. 154-160, 1958, doi: 10.7551/mitpress/4360.003.0004.

[15] Ö. Sandred, "Approaches to Using Rules as a Composition Method," Contemporary Music Review, vol. 28, no. 2, pp. 149-165, 2009, doi: 10.1080/07494460903322430.

[16] T. Anders, "Compositions Created with Constraint Programming," Oxford University Press, 2018, pp. 0-0. [Online]. Available: https://doi.org/10.1093/oxfordhb/9780190226992.013.5

[17] H. Rosa, Social Acceleration: A New Theory of Modernity. Columbia University Press, 2013.

[18] P. Lorenz-Spreen, B. M. Mønsted, P. Hövel, and S. Lehmann, "Accelerating dynamics of collective attention," Nature Communications, vol. 10, no. 1, pp. 1-9, 2019, doi: 10.1038/s41467-019-09311-w.

[19] C. Clark, "The impact of entrepreneurs' oral 'pitch' presentation skills on business angels' initial screening investment decisions," Venture Capital, vol. 10, no. 3, pp. 257-279, 2008, doi: 10.1080/13691060802151945.

[20] P. Norris, "Political Communications and Democratic Politics," J. Bartle and D. Griffiths Eds. London: Palgrave Macmillan UK, 2001, pp. 163-180.

[21] V. Bakir, "Torture, intelligence and sousveillance in the war on terror : agenda-building struggles," ed. Abingdon, Oxon: Routledge, 2016.

[22] L. Bayer and J. Plucinska. (2018, 2018-04-04) Orbán's media puppetmaster. Politico. Available: https://www.politico.eu/article/viktor-orban-media-empire-hungary-election-antal-rogan-fidesz-propaganda/

[23] D. L. Poole and A. K. Mackworth, Artificial intelligence : foundations of computational agents. Cambridge: Cambridge University Press, 2010.

[24] F. Pachet and P. Roy, "Markov constraints: steerable generation of Markov sequences," Constraints, vol. 16, no. 2, pp. 148-172, 2011/04/01 2011, doi: 10.1007/s10601-010-9101-4.

[25] J. B. Schilingi, "Local and global control in computer-aided composition," Contemporary Music Review, vol. 28, no. 2, pp. 181-191, 2009. DOI 10.1080/07494460903322455.

[26] Ö. Sandred, "PWMC, a Constraint-Solving System for Generating Music Scores," Source: Computer Music Journal, vol. 34, no. 2, pp. 8-24, 2010. [Online]. Available: https://about.jstor.org/terms.

[27] T. Anders and E. R. Miranda, "Constraint programming systems for modeling music theories and composition," ACM Computing Surveys (CSUR), vol. 43, no. 4, pp. 1-38, 2011, doi: 10.1145/1978802.1978809.

[28] T. Baca, J. W. Oberholtzer, J. Trevino, and V. Adán, "Abjad: An open-source software system for formalized score control," in Proceedings of The First International Conference on Technologies for Music Notation and Representation, 2015.

[29] G. J. Bolaños Chamorro, "An Analysis of Jonathan Harvey's Speakings for Orchestra and Electronics," Ricercare, vol. 13, no. 13, pp. 72-107, 2021. DOI: 10.17230/ricercare.2020.13.4.

[30] S. V. Lavrova, "Metaphysics of Sound Objects by Peter Ablinger," (in Russian), Nauchnyy vestnik Moskovskoy konservatorii / Journal of Moscow Conservatory vol. 13, no. 4, pp. 736–751, 2022, doi: https://doi.org/10.26176/mosconsv.2022.51.4.03.

[31] F. C. Ciardi, "Dalla prosodia alla musica strumentale: una sfida compositiva. ," in L'arte orale: Poesia, musica, performance. , L. Cardilli and S. Lombardi Vallauri Eds. Torino: Accademia University Press, 2020.

[32] F. C. Ciardi, "Strategies and tools for the sonification of prosodic data: A composer's perspective," Proceedings of the 26th International Conference on Auditory Display (ICAD 2021), pp. 247-252, 2021, doi: https://doi.org/10.21785/icad2021.041.

[33] O. Sullivan-Tarazi. (2018, 2018-05-04) Narrative insights: notes from Aristotle on storytelling. @Medium. Available: https://medium.com/@odile_sullivan/narrative-insights-what-aristotle-can-teach-us-about-storytelling-239d1b878e74

[34] Aristotle, J. Baxter, P. Atherton, and G. Whalley, Aristotle's Poetics, Montreal,Buffalo: McGill-Queen's University Press, 1997.

[35] I. Duncan, "Scheduling Musical Events in Max/MSP with Scheme For Max," Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University, no. 1, p. 1, 2021.

[36] J. Vincenot, "LISP in Max: Exploratory Computer-Aided Composition in Real-Time," ICMC 2017 proceedings (Shanghai, 2017), 2017. [Online]. Available: http://hdl.handle.net/2027/spo.bbp2372.2017.012.

[37] D. S. Starnes, D. Yates, and D. S. Moore, The practice of statistics. Macmillan, 2010.

[38] M. D Orazio, M. Di Zio, and M. Scanu, "Statistical Matching for Categorical Data: Displaying Uncertainty and Using Logical Constraints," Journal of Official Statistics Stockholm, vol. 22, no. 1, p. 137, 2006.

Automatic Creation of Chordal and Melodic Chromaticism

A. Christou, R. Dannenberg

Yale University | Carnegie Mellon University
antonis.christou@yale.edu | rbd@andrew.cmu.edu

Abstract. Western popular and art music use the diatonic scale as their foundation in both melody and harmony, and restricting to diatonic scales can simplify both analysis and composition. However, even mostly diatonic genres contain some limited use of chromaticism or accidentals. Therefore, we propose that, after creating diatonic music, we can automatically alter selected notes chromatically to enhance the melody and harmony. This paper presents several deep learning models for generating this chromaticism given a purely diatonic input melody and accompanying chord sequence. The altered melody and harmony from the model are generally compelling and conform to musical syntax, making our work a potentially useful tool for both human and computer agents to augment existing compositions or aid in the compositional process.

Keywords. Chromaticism, Machine Learning, Melody, Music Generation and Composition, Trans former

1 Introduction

The concepts of diatonicism and chromaticism are fundamental building blocks in Western music. Diatonicism is generally seen as the most basic harmonic system and refers to the use only of notes from a particular diatonic scale. For example, a song in the key of C major that exclusively uses the notes of that key would be seen as purely diatonic. In contrast, the chromatic scale refers to a division of the octave into twelve equal half steps and does not fit neatly into any one key.

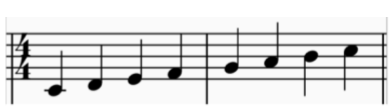



Fig. 1. A C major scale.

While the use of notes and derived chords from diatonic scales is most common, composers from all eras occasionally make alterations to the notes of this diatonic scale and instead use notes and chords from the chromatic scale. This process is called chromaticism and idiomatic use of it is fundamental to the unique sound of many genres of music [1, 2].

Therefore, due to the great importance of using chromaticism effectively in writing compelling music in many genres, developing models that generate or harness it in their compositions is an important but largely unexplored area up until this point. The most common approach has been to create a model that learns from data when to use chromatic and diatonic notes and so freely selects notes not restricted to the diatonic pitches. However, this places the burden on the model of learning the difference between the two sets of pitches and often results in outputs that drift from one key to another or includes a number of "wrong" notes due to the probabilistic nature of these methods.

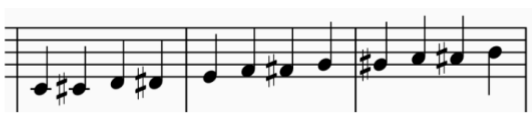



Fig. 2. A chromatic scale.

As a result, some successful automatic composition models are restricted exclusively to the diatonic scale, eliminating the possibility of incorrect chromatic notes but also resulting in simpler sounding outputs. Our goal is to create a tool powered by deep learning that could potentially be used in conjunction with existing diatonic models or human compositions. To that end, we propose a "post-processing" or editing step that inserts chromatic melody notes and chords. We believe that this model breaks new ground in this area of deep learning applied to musical chromaticism and will enable future explorations as well as lead to generally more interesting sounding music generations now and in the future.

2 Related Work

Even before the invention and widespread use of neural methods, composers ranging from Mozart with his games of dice to Xenakis and Cage's aleotoric music have used algorithmic or computational methods to produce musical output [4]. Today, a great deal of work has been produced on the subject of creating models to generate music using a variety of differ ent methods. Markov chains and grammars were used both historically and, in the present, producing notable works such as Hiller and Isaacson's

Illiac Suite, which is one of the earliest scores produced by a computer [6, 7]. More recently, recurrent neural networks have been applied, recognizing the inherently sequential nature of music [5]. Other methods include convolutional neural networks, notably applied by Huang et al. [9] to the problem of completing musical scores in the domain of contrapuntal polyphony. Most recently, the Transformer with its self-attention and ability to encode long term structure [8] in addition to other deep learning methods have been applied to music generation with promising results [9]. Work has also been done in the area of harmonizing melodies with appropriate chord changes [10]. However, the problem of chromatizing existing diatonic melodies and harmonies rather than simply generate them has not been explored previously, despite its great importance in musical style. A variety of models have used the simplifying assumption of generating purely diatonic outputs, ranging from rule-based systems [11, 12], to a deep-learning system [13], opening the door for the creation of a model that could be used to augment other exclusively diatonic compositions. Therefore, we introduce this model as a means of exploring this area and enabling human and computer composers to improve the quality of their material.

3 Data

A primary challenge in creating a model to learn patterns of melodic and chordal chromaticism is finding a data source with sufficiently rich harmonizations. We use the Public Domain Song Anthology book (PDSA), a collection of 347 songs taken from popular American folk songs that usefully for this application are provided with two sets of harmonizations, one slightly more diatonic and one using more advanced and modern jazz harmonies.

As a pre-processing step, we filter the PDSA to use only the songs in simple duple meter (ex. 2/4 and 4/4). Additionally, effort was taken to create a purely diatonic input to train with by eliminating all accidentals in the melody and all chords outside of the basic diatonic ones, leaving us with a chordal vocabulary of the 14 chords found diatonically in the keys of C major and minor for input and the 194 chords found in the PDSA with their chromaticism intact for output. The melody notes are encoded with midi as tuples in the form of [midi note, duration] with 0 being used to mean rest leading up to the highest midi note found in the PDSA, 87 or Eb6. Each song overall is encoded as a stream of these tuples for both melody and the attached chords. We also transpose all songs to the key of C major or minor for easier learning.

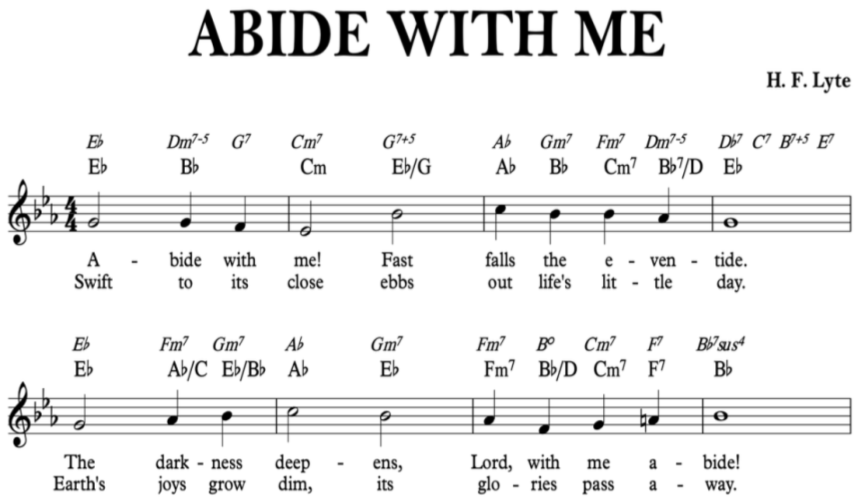



Fig. 3. An example song from the PDSA.

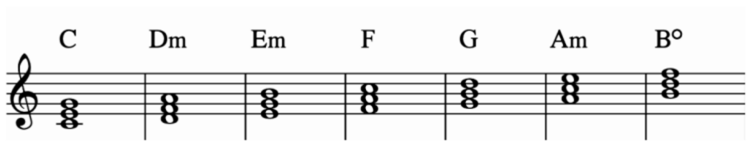


Fig. 4. The chords diatonic to C major.

4 Model

Our model uses both Transformer and LSTM, with the Transformer functioning as the encoder and LSTM as the decoder. This Transformer-LSTM hybrid model has been used successfully in other recent music generation projects [13]. We also tried several other models such as pure LSTM, pure bidirectional LSTM, and pure Transformer. We use a nested model to handle the melodic and chordal chromatization in two passes. For the chords, we encode each chord name as an integer in the vocabulary and pass this input to an encoding layer before going to the Transformer encoder. This representation is then decoded by an LSTM before finally going through a fully connected layer to generate the next predicted chord based upon the output chromatic chord vocabulary. After all chords have been processed, this process is then repeated with the melody notes. Each note encoded as midi is passed through a melodic model with the same architecture to make the binary classification decision of whether or not this melody note should be chromaticized or should remain diatonic. We use the convention of chromatizing notes based upon their most likely alteration in the C major scale (ex. Fs become F#s rather than Gbs, As become Bbs rather than A#, etc) as a simplification. After both models pass on the input, we recombine the newly chromaticized chords and melody with the durations found in the original song and generate a MusicXML file for viewing as a score according to the conventions of Western notation.

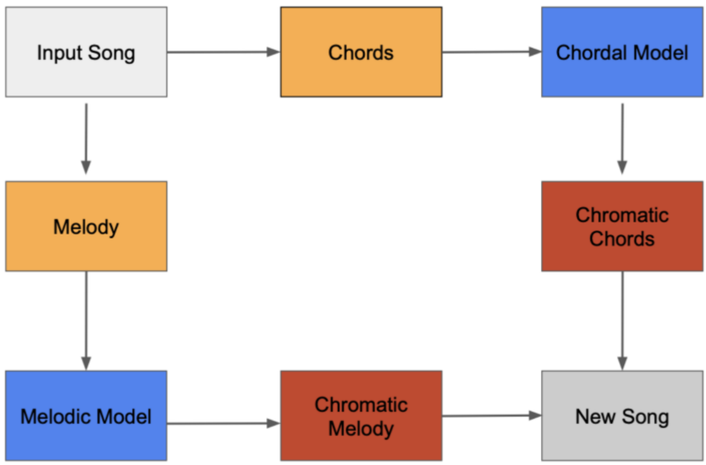


Fig. 5. Generation pipeline for new chromatic song.

This score includes both block voicings of the new chords as well as chord symbols in lead sheet notation, allowing for human musicians to play, alter, and voice the generations in addition to hear them as they are.

5 Experiment and Evaluation

To determine the best model for this problem, we first compared the four architectures above to see which produced the most accurate and idiomatic chromaticism. Melodic accuracy was defined as the percent of correctly made chromaticization decisions, while chordal accuracy was defined as the percent of correctly chosen output chords across the dataset given any particular input chord.

The models were trained using the songs from PDSA with 80% being reserved for training and the remainder for validation. Since the data for melodic chromaticism is already heavily unbalanced, with approximately 90% of the notes being diatonic, we opt to filter out any songs that do not contain any instances of melodic chromaticism for training the melodic model. We also set an optional flag for filtering out minor key songs to make for easier learning. After all filtering, 176 songs were used to train the chordal model and 121 for the melodic. We use Cross Entropy Loss as our loss function and Adam as the optimizer. For the melodic model, as a means of combating the aforementioned unbalanced class distribution, we weight the loss function according to the prevalence of diatonic versus chromatic notes in the original dataset. In all networks, we use a learning rate of 0.001 and a dropout probability of 0.1. The transformer based models used two layers each for their respective encoder and decoders, with two heads in the multi-head attention in the self-attention layer, an embedding dimension of 80, and a feed forward dimension of 400. The LSTM component of the models that include one utilizes a hidden dimension of size 100.

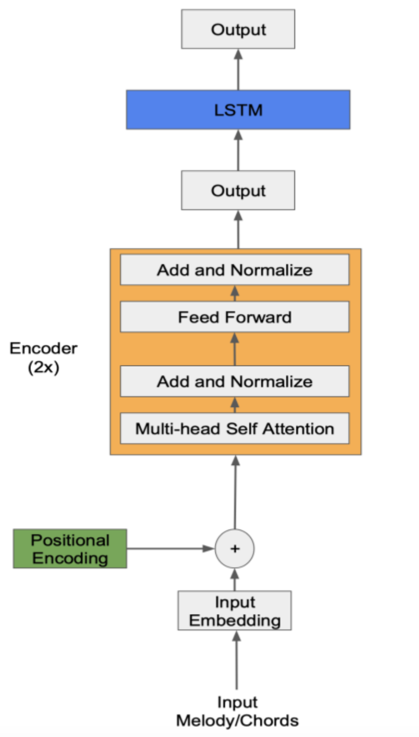


Fig. 6. Transformer-LSTM model used for melody and chordal chromaticization.

The pure LSTM model was augmented with a few hands selected features based upon domain knowledge, such as the duration of the melody note, the duration of the chord, and which chord generated earlier is playing over any given melody note, as these features are important to human musicians when deciding how, if at all, to chromaticize a note or chord. These values are concatenated with those from the embedding layers to form the full input to the model

Model	Melodic Accuracy	Melodic Loss	Chordal Accuracy	Chordal Loss
LSTM	68.5%	0.53	69.5%	1.53
Bidirectional LSTM	79.6%	0.69	72.2%	**1.36**
Transformer	72.1%	0.74	45.8%	3.63
Transformer-LSTM	**80.3%**	**0.36**	**73.16%**	1.47

Fig. 7. Loss and accuracy on validation set for each model type

Overall, the Transformer-LSTM model had the highest accuracy on the validation set. The Trans former model performed poorly as it tended to overfit heavily on the validation set and also struggled to train, likely because of the small dataset. For the LSTM based models, making the model bidirectional improved the accuracy, likely due to the fact that knowing what comes both before and after a given note or chord can help lead to idiomatic and grammatical chordal and melodic patterns.

6 Conclusion

Overall, the outputs from the model can at times be quite coherent and pleasant to listen to.

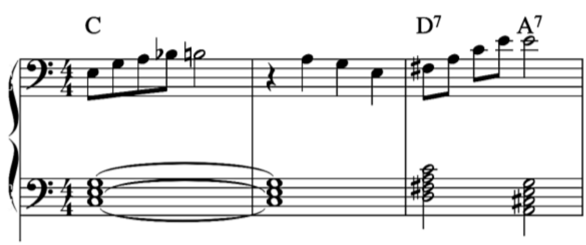


Fig. 8. An output phrase with proper use of chromatic passing tones and chromatic chords according to standard jazz harmony conventions and the PDSA.

The short example phrase above from the validation set correctly learned to use common chromatic melody devices like passing tones, and the generated chord progression uses a number of grammatical and typical stylistic hallmarks of chromatic harmony, such as secondary dominants and seventh chords.

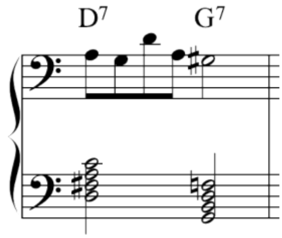



Fig. 9. An example of dissonant, unidiomatic output in the melody.

However, at times some of the generations from the model can be dissonant and not grammatical according to the conventions of the styles represented with the PDSA. In the following example [Fig. 9], the model has outputted a chord and melody sequence that is theoretically plausible in isolation, but the G# in the melody voice directly clashes with the root of the chord, making it an unlikely choice for a human composer to write. Therefore, in future explorations, we would like to experiment more with allowing the model to more consistently select chords and melody notes that do not clash with each other, as well as investigate further into variable length and Seq2Seq based model, especially for melodic chromaticism.
The model in its current form can be used by humans and computer music models as a tool to give new ideas or make existing compositions more exciting. We believe that this model breaks new ground in a mostly unexplored area of music generation and opens the door for further investigation in this field.

References

[1] Meyer, Leonard B. (1956). Emotion and Meaning in Music, p. 217. Chicago: University of Chicago Press. ISBN 9780226521374. Quoted in Brown (1986), p. 1.

[2] Forte, Allen, Tonal Harmony, third edition (S.l.: Holt, Rinehart, and Wilson, 1979): p. 4. ISBN 0-03-020756-8 3.

[3] Cooper, Paul (1975). Perspectives in Music Theory, p. 208. New York: Dodd, Mead, and Company. 4.

[4] Taruskin, Richard. “Music in the Late 20th Century, Chapter 2 ‘Indeterminacy.’ ” The Oxford History of Western Music, Oxford University Press, Oxford Etc, 2010.

[5] Chen, Chun-Chi J., Miikkulainen Risto, 2001, Creating Melodies with Evolving Recurrent Neural Networks, IJCNN-01

[6] Thornton, C. (2009). Hierarchical markov modeling for generative music. In International computer music con ference.

[7] Lejaren A Hiller Jr and Leonard M Isaacson. Musi cal composition with a high speed digital computer. In Audio Engineering Society Convention 9. Audio Engi neering Society, 1957.

[8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention is all you need. In Advances in neural infor mation processing systems (pp. 5998–6008).

[9] Huang, C.-Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, C., . . . Eck, D. (2018). Music transformer. arXiv preprint arXiv:1809.04281 .

[10] S. Rhyu, H. Choi, S. Kim and K. Lee, “Translat ing Melody to Chord: Structured and Flexible Harmo nization of Melody With Transformer” in IEEE Ac cess, vol. 10, pp. 28261-28273, 2022, doi: 10.1109/AC CESS.2022.3155467.

[11] A. Elowsson and A. Friberg, “Algorithmic composi tion of popular music,” in the 12th International Con ference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, 2012, pp. 276–285.

[12] Dai, Ma, Wang, and Dannenberg, “Personalized Pop ular Music Generation Using Imitation and Structure,” Journal of New Music Research, (2023).

[13] Dai, Jin, Gomes, and Dannenberg, “Controllable Deep Melody Generation via Hierarchical Music Structure Representation,” in Proceedings of the 22nd International Society for Music Information Retrieval Conference, Online, Nov 2021, pp. 143-150.

Harmonizing Innovation and Ethics: Navigating the AI Revolution in Music Creation and Production

Sam C. Shin

University of California, Riverside, Department of Music
slong028@ucr.edu

Abstract. This article explores the multifaceted implications of artificial intelligence (AI) in the music industry, primarily focusing on its application in music creation, production, and the evolving legal and ethical landscapes. With AI becoming increasingly central in businesses and technology, its role in music has expanded, ranging from AI-driven music generation services to virtual vocalists and AI-assisted music production tools. I delve into these practical applications of AI in music, such as royalty-free background music for content creators and audio mastering, and the more public-facing research-oriented pursuits of industry leaders like Google and OpenAI. I also present examples of how AI has been used in recent commercial music releases, particularly by electronic producers such as Holly Herndon. The emergence of AI-driven virtual vocalists is highlighted as well, as it brings unique challenges in identity and legal rights. I discuss the need for frameworks to empower vocalists, ensuring equitable compensation and control over training models. Similarly, I critically examine the legal and ethical implications of deep learning in music, especially in the context of copyright issues and the economic impact on musicians. I also underscore the importance of transparency in AI applications and the commercial use of the term "AI." Finally, I emphasize the responsibility of AI researchers and developers in understanding the impacts of their creations, advocating for a critical examination of AI's socio-cultural impacts. There is vast potential of AI in music, and I advocate for collaborative efforts among artists, developers, legislators, and other stakeholders to navigate the complexities of AI in music creation and production.

Keywords. copyright-AI, ethics-AI, music-AI, responsible AI

1 Introduction

Since the debut of OpenAI’s ChatGPT, artificial intelligence has become a focal point, with businesses eagerly capitalizing on the heightened interest in AI technologies. The fluid nature of our conception of AI is evident, as encapsulated in axioms such as Tesler’s theorem asserting that “artificial intelligence is whatever has not been done yet” [1]. This evolving understanding is further reflected in phenomena such as the “AI effect," the notion that "when an allegedly uniquely human ability or skill is being automated by means of computer technology, we refer to it as 'AI.' Yet, as soon as this automation is seamlessly and fully successful, we tend to stop referring to it as an 'AI case'" [2].

2 Emerging Applications for AI in Music

In the contemporary landscape, AI utilization spans various domains. Digital culture theorist Lev Manovich outlines a taxonomy of AI applications: "1) *Selecting* content from larger collections," "2) *Targeting* content," "3) *Assistance* in creation/editing of new content," and "4) *Fully autonomous* creation" ([3], emphasis in original). To illustrate a taxonomy for AI-driven music products and services, Table I contains examples of such products and services categorized into music generation, plug-ins, mixing and mastering, source separation, voice models, music-driven productivity or wellness apps, and audio enhancement.

TABLE 1. AI MUSIC PRODUCTS AND SERVICES.

Company	Service/Product	Use
Google	MusicLM	Music generation
Mubert	Mubert	Music generation
Audialab	Emergent Drums 2	Drum machine
Sonic Charge	Synplant 2	Synthesizer
LANDR	Mastering (online/plug-in)	Mastering
RoEx	Mixing and mastering (online)	Mixing/mastering
AudioShake	Instrument Stem Separation	Source separation
LALA.AI	Stem Splitter	Source separation
Holly+	Holly+	Voice model
Voice-Swap	Voice-Swap	Voice models
Brain.fm	Brain.fm	Productivity/wellness
Endel	Endel	Productivity/wellness
insoundz	Revive	Audio enhancement

Many commercial music generation services predominantly market to content creators seeking royalty-free background music. This preference may arise from the simplicity of the music itself, making it likely easier to generate using current models. Conversely, major tech companies like Google, Meta, and OpenAI have introduced music generation models as demos rather than fully realized products: MusicLM is available through Google's AI Test Kitchen, Meta's MusicGen is available as a demo on Hugging Face, while OpenAI's Jukebox is less accessible to novices as users can run it on Google Colab. These demos serve as showcases for the cutting-edge research conducted by these companies, illustrating their commitment to staying abreast of AI developments. This dual perspective highlights both the practical applications of AI in music creation, primarily serving content creators in need of more commonplace background music, and the currently research-oriented pursuits of industry leaders.

AI-driven plugins are primarily oriented toward mixing and mastering tasks, exemplified by tools from iZotope and Sonible. This inclination may stem from the fact that these tools rely on analyzing incoming audio, which has been extensively researched in the field of music information retrieval. However, there are also notable AI-driven instruments, such as Emergent Drums 2, a drum machine, and VOCALOID6, a singing synthesizer. Additionally, effects such as TAIP, a tape saturator, and Neutone, which runs various AI audio processing models, contribute to the expanding repertoire of AI-enhanced music production tools. An increased demand for AI-driven music-making tools is likely to spur the creation of a more diverse range of plugins. However, a potential obstacle lies in the computational overhead associated with running these models, whether locally or remotely. Balancing the need for advanced capabilities with the practical constraints of computational resources will be crucial for the seamless integration and widespread adoption of AI-driven plugins in music production.

3 Examples of AI-Assisted Music Creation

Limited published research has focused on the intersection of artificial intelligence and music creation (see [4] and [5] for examples of existing research). However, insights gleaned from interviews with artists who integrate AI into their work shed light on the role of AI in their music-making process. Vocalist and producer Holly Herndon's 2019 album *PROTO* stands as an early example of AI incorporation, employing a machine learning program, Spawn. Trained on Herndon's voice, Spawn acts as a collaborator throughout the album [6].

The year 2023 witnessed a surge in music integrating AI technologies, potentially reflecting increased accessibility for the average user. Electronic producer Lee Gamble utilized voice models in his album *Models* to explore the potential humanization of AI technologies [7]. Similar to Herndon, Gamble trained one model on his voice and considered the voice models as collaborators in the music-making process. Another electronic producer, patten, asserts to have created the "first LP made entirely from AI-generated sound sources," *Mirage FM*, which, similar to Herndon, he hopes demonstrates the possibilities of AI-assisted creativity [8]. *Mirage FM* is composed of samples taken from Riffusion, a text-to-audio model. Riffusion is based on a fine-tuning of Stable Diffusion, a text-to-image model, on spectrograms [9]. It produces lowfidelity output, however, patten embraces this limitation, considering it an integral part of the aesthetic of the tool. In contrast to the other artists' motivations, electronic producer Oneohtrix Point Never's album *Again* features some tracks composed with the assistance of various AI tools, including Riffusion and OpenAI's Jukebox, to explore the limits and failings of current AI technologies [10].

4 Virtual Vocalists

"Virtual vocalists" deserve special attention because they present a unique set of challenges surrounding the concepts of identity and the legal rights vocalists have to their own voices. Before the sophistication of AI, virtual vocalists were created using either vocal synthesis or samples. Now producers can use AI to create virtual singers using models trained on recordings of real singers. The identity of a virtual vocalist is inherently flexible, offering opportunities for creative expression but also raising concerns about the potential for bias. When delving into these complexities, it becomes evident that frameworks are essential to empower vocalists, granting them control over the use of their voices in AI training models and ensuring equitable compensation for their contributions.

One pivotal event that underscores the intricate challenges in the domain of AI-powered virtual vocalists occurred in August 2022 when FN Meka, a virtual rapper partially powered by AI, was signed, and then subsequently dropped by Capitol Records [11]. The project faced sharp criticism and was accused of perpetuating stereotypes about Black culture and sidelining Black creators during the character and music development process. In a study examining racialized Vocaloids and their reception, Nina Sun Eidsheim asserts that "connections between race and timbre are cultural constructs and practices" that are formed through a "process by which timbre, visual imagery, text, and listening are tied together into a narrative about race" [12]. Consequently, developers of virtual vocalists and artists who utilize their tools must remain acutely aware of their biases to ensure that these products are created and used with respect and sensitivity. This incident also illustrates the ethical questions surrounding developers' ability to configure and profit from identities that differ from their own, particularly those belonging to marginalized groups.

In addition to issues surrounding identity, vocalists' rights to their own voices becomes a paramount concern. Organizations like the American Society of Composers and Publishers (ASCAP) are actively lobbying for "six guiding principles for AI" with the aspiration that Congress will adopt them as legislation to safeguard the rights of musicians and songwriters [13]. Without federal protections, state laws governing the use of another's likeness seem to be the best option for artists whose voices are digitally replicated [14]. The future may also witness legislative developments governing how artists are compensated when their voices are used to train AI models. Notably, YouTube is collaborating with record labels to monetize music that incorporates AI, while companies specializing in licensing AI voices trained on human singers are emerging. Artists like Grimes and Holly Herndon are adopting alternative approaches by making models of their voices available for other musicians to use, albeit with certain stipulations [15], [16]. Additionally, Herndon's involvement with Spawning, a company focused on tools that permit artists to opt out of public training datasets, underscores the importance of giving creators agency in this evolving landscape [17].

Traditionally, we think of artists as unique individuals who create distinct works. But the immensely popular Vocaloid character Hatsune Miku challenges this preconception. While certain aspects of her identity are defined (young, female, Japanese), she "does not exist as a set personality with a specific narrative back story… [so] she can become… emergent, dynamic, and malleable in the hands of musicians and fans alike" [18]. Virtual vocalists with fluid identities will certainly be possible with AI-generated vocalists as well. But because AI represents a blend of the material it was trained on and the patterns it discerns within its training data, developers must confront the complex ethical questions associated with AI-generated work and take responsibility for the tools they create.

The world of virtual vocalists stands at the intersection of identity, rights, and artistic responsibility. As the boundaries of creativity and technology continue to blur, it becomes imperative that we navigate these challenges with sensitivity, ethics, and a commitment to preserving the integrity of artistic expression.

5 Legal and Ethical Implications of Deep Learning

Deep learning is accompanied by a plethora of legal and ethical implications, further complicated by the novelty of the technology. A primary concern is copyright issues, as delineated in a report prepared by the Congressional Research Service [19]. Presently, the U.S. Copyright Office restricts copyright to works authored by a human being, even if the work is generated based on a text prompt. However, if the AI-generated material undergoes arrangement, modification, or combination with human-authored materials, it may be eligible for copyright protection.

The attribution of authorship in AI-generated work also raises significant questions. Determining whether the author is the human who supplied the prompt, the AI program itself, or the developer of the program is an ongoing challenge. At present, there is no clear rule governing this aspect. Some services, such as OpenAI, address this ambiguity by assigning copyright to the user, thereby circumventing the complexities surrounding authorship and copyright ownership.

Generative AI introduces the potential for copyright infringement during both the training and output phases. Notably, organizations like the Authors Guild, The New York Times, and Getty Images contend that training AI programs on copyrighted materials constitutes a violation of copyright law. In contrast, tech companies and their advocates argue that such activities fall within the realm of fair use [20]. As of now, multiple lawsuits on this matter are still pending, leaving both perspectives untested in a legal context. While some AI services decline prompts involving copyrighted characters or the imitation of a particular artist's style, others may permit users to generate such works, potentially exposing both the AI services and their users to legal action.

Training datasets pose a significant challenge, with potential economic ramifications for musicians already contending with existing economic models, such as streaming services, and now facing competition from AI-generated music for business opportunities. The complex nature of training on copyrighted works is exemplified by Ed Newton-Rex, a researcher in AI audio, who resigned from his role as VP of Audio at Stability AI due to concerns about the company's stance that training on copyrighted materials falls under fair use [21]. Acknowledging that "much of the efficacy – and hence much of the value – of machine learners depends on the datasets on which they are trained," it becomes imperative for compensation models to adapt in tandem with technological advancements for artists to sustain their livelihoods [22]. Without legislative changes, the risk persists that the "cultural capital of individual musicians and communities" will continue to be "exploited by capitalist firms for private interests" [23].

Transparency surrounding the datasets used to train AI models and the role of AI in commercial products and services is another concern. While companies may justify the secrecy around datasets as necessary to safeguard proprietary models, it is plausible that the reluctance to disclose stems from models being trained on copyrighted works, exposing companies to potential legal challenges, as discussed earlier. Stability AI is one example of a company that discloses its dataset; their Stable Audio model was trained using audio files supplied through an agreement with AudioSparx, a stock music company [24].

The commercial value of the term "AI" has led to it being "often used interchangeably with, or instead of, the specific kind of machine learning that companies and labs are doing" and obfuscates the part AI plays in a product or service [25]. This ambiguity is exemplified in a study on the AI-powered mastering service LANDR, where researchers posit that the platform likely "uses ML for part of the process, for instance in analyzing the sound of an uploaded audio track, and then select[s] from a matrix of preset possibilities for processing" [26]. This suggests that the machine learning component may not be tailored to each track, but the exact process remains undisclosed, as LANDR does not provide insights into the inner workings of its mastering program. In contrast, Yamaha, although somewhat concealed within the research and development section of their website, offers an article dedicated to explaining how AI is incorporated into their Vocaloid product, providing a more transparent approach to showcasing the integration of AI in their offerings [27].

6 Researcher Accountability

It is imperative for AI researchers and developers to embrace accountability for the technology they create, taking the lead in critically examining its impacts and working towards mitigating negative consequences through research, supporting regulation, and other measures. Researchers have themselves critiqued their respective fields, such as human-computer interaction ("while the field [of human-computer interaction] examines and generates… systems… its political dimension—that is, the varying balances of power between who produces the technology, who designs it, who uses it, and what socio-cultural impact it may have—is rarely addressed" [28]) and music information retrieval ("MIR [music information retrieval] is not yet engaging in the 'ethical turn' that other technology research fields are undergoing" [23]).

An important recommendation is for researchers to collaborate more closely with artists, not only to enhance the development of better models and services but also to understand and address artists' concerns. In an article focusing on "research applying machine learning to music modelling and generation," the researchers emphasize that "rarely does such work explicitly question and analyse its usefulness for and impact on real-world practitioners, and then build on those outcomes to inform the development and application of machine learning" [4]. Another researcher supports this assessment while also challenging the fatalistic assumption surrounding AI deployment: "Also problematic and unchallenged is the axiomatic assumption that 'AI will help musicians make music'. To the best of my knowledge, no investigations have been conducted to pinpoint what specific help musicians need, prove that such help is actually beneficial, or ensure that these tools will contribute to a more just music industry" [23].

A noteworthy project in this realm is Google's DeepMind laboratory's tool, SynthID, designed for watermarking and identifying AI-generated images and audio [29]. This technology holds promise for addressing challenges related to misinformation and enhancing the transparency of AI-generated media. Hopefully, researchers will persist in developing tools to combat the complex issues that accompany AI. A scrutiny of research papers on AI-generated music from major entities like Meta, OpenAI, and Google, reveals a shared acknowledgment among researchers regarding various risks associated with AI deployment. These risks include built-in bias or a lack of diversity [30], [31], [32], [33], [34], the potential for use in misinformation or scams (such as "creating remarkably realistic deep fakes and voice phishing" [32]), misappropriation of training data [33], and ethical concerns surrounding the sourcing of training data and competition for the work of musical artists [34]. Nevertheless, the extent of ongoing efforts to address these challenges remains unclear, and the willingness of researchers to be vocal about the issues, especially when their ideals may conflict with the economic models of their employers, poses an open question.

7 Conclusions

“AI” is an ambiguous term that encompasses a range of meanings. Existing AI-driven tools for music and audio typically engage in selecting content, targeting content, aiding in the creation or editing of content, or generating entirely new content. While there is currently a limited variety of AI-driven products for music production, the sustained global interest in AI technologies is likely to continue fueling their development. Commercial music releases are already incorporating AI into the music-making process. However, more published research on AI-assisted music creation would be beneficial for both developers and musicians. Virtual vocalists are particularly complex, as they can embody the musical identity of an existing artist but can also be given a completely fabricated identity chosen by the creator.

Despite the proliferation of startups and established companies actively engaged in developing AI-driven music products and services, the ethical, legal, and economic implications of this technology remain unsolved. Pressing issues include the copyright status of AI-generated media, whether training models on copyrighted work constitutes a violation of copyright laws, the potential for AI-generated media to infringe upon copyright laws, and how to fairly compensate artists whose music is included in datasets. Transparency from companies is crucial, particularly regarding the sources of datasets and the role of AI within their products or services.

Developers also bear the responsibility to address difficult ethical questions surrounding the systems they develop and to engage with musicians, producers, audio engineers, and others who will be most impacted by advances in machine learning in music. I firmly believe that addressing the challenges posed by AI will require collaboration and mutual understanding among artists, developers, legislators, and other stakeholders. By fostering cooperation, these diverse perspectives can contribute to the development of equitable solutions that navigate the complexities surrounding AI in the domain of music creation and production.

References

[1] E. Arielli, “Techno-animism and the Pygmalion effect,” in *Artificial Aesthetics*, L. Manovich and E. Arielli, Eds. Self-published: Lev Manovich, 2023, p. 15.

[Online]. Available: https://manovich.net/index.php/projects/artificial-aesthetics-book.

[2] L. Manovich, "Seven arguments about AI images and generative media," in *Artificial Aesthetics*, L. Manovich and E. Arielli, Eds. Self-published: Lev Manovich, 2023, p. 3. [Online]. Available: https://manovich.net/index.php/projects/artificial-aesthetics-book.

[3] L. Manovich, *AI Aesthetics*. Moscow: Strelka Press, 2018, p. 5.

[4] B. L. Sturm et al., "Machine learning research that matters for music creation: a case study," *Journal of New Music Research*, vol. 48, no. 1, 2019. https://doi.org/10.1080/09298215.2018.1515233.

[5] O. Ben-Tal, M. T. Harris, and B. L. Sturm, "How music AI is useful: engagements with composers, performers, and audiences," *Leonardo* vol. 54, no. 5, 2021. https://doi.org/10.1162/leon_a_01959.

[6] G. T. Claymore, "Inhuman After All," Stereogum, May 6, 2019. [Online]. Available: https://www.stereogum.com/2041686/holly-herndon-proto-interview/interviews/.

[7] M. McKinney, "An Interview with Lee Gamble," Passion of the Weiss, Oct. 18, 2023. [Online]. Available: https://www.passionweiss.com/2023/10/18/lee-gamble-interview-models/.

[8] T. Waite, "Mirage FM: How patten Created the First LP Made Entirely from AI Sounds," Dazed, June 30, 2023. [Online]. Available: https://www.dazeddigital.com/music/article/60240/1/mirage-fm-patten-damien-roach-first-lp-made-entirely-from-ai-sounds-riffusion.

[9] D. Coldewey, "Try 'Riffusion,' an AI Model that Composes Music by Visualizing It," TechCrunch, Dec. 15, 2022. [Online]. Available: https://techcrunch.com/2022/12/15/try-riffusion-an-ai-model-that-composes-music-by-visualizing-it/.

[10] Petrusich, "Digital Memory: The Emotionally Haunted Electronic Music of Oneohtrix Point Never," The New Yorker, Sept. 23, 2023. [Online]. Available: https://www.newyorker.com/magazine/2023/10/02/the-emotionally-haunted-electronic-music-of-oneohtrix-point-never.

[11] M. Tracy, "A 'Virtual Rapper' Was Fired. Questions About Art and Tech Remain," The New York Times, Sept. 6, 2022. [Online]. Available: https://www.nytimes.com/2022/09/06/arts/music/fn-meka-virtual-ai-rap.html.

[12] N. S. Eidsheim, "Race as zeros and ones: Vocaloid refused, reimagined, and repurposed," in *Listening, Timbre, and Vocality in African American Music*. Durham, NC, USA: Duke Univ. Press, 2019, p. 139.

[13] P. Williams, "Protecting Songwriters in the Age of AI (Guest Column)," Billboard, Sept. 20, 2023. [Online]. Available:

https://www.billboard.com/pro/ascap-artificial-intelligence-principles-paul-williams-guest-column/.

[14] E. David, "Musicians are Eyeing a Legal Shortcut to Fight AI Voice Clones," The Verge, Sept. 21, 2023. [Online]. Available: https://www.theverge.com/2023/9/21/23836337/music-generative-ai-voice-likeness-regulation.

[15] Elf Tech. [Online]. Available: https://elf.tech/connect (accessed Oct. 2, 2023).

[16] H. Herndon, "Holly+," July 13, 2021. [Online]. Available: https://holly.mirror.xyz/54ds2IiOnvthjGFkokFCoaI4EabytH9xjAYy1irHy94.

[17] Spawning, 2023. [Online]. Available: https://spawning.ai/.

[18] S. A. Bell, "The dB in the .db: Vocaloid software as posthuman instrument," *Popular Music and Society*, vol. 39, no. 2, p. 237, 2016. http://dx.doi.org/10.1080/03007766.2015.1049041.

[19] C. T. Zirpoli, "Generative Artificial Intelligence and Copyright Law," Congressional Research Service, updated Sept. 29, 2023. [Online]. Available: https://crsreports.congress.gov/product/pdf/LSB/LSB10922.

[20] M. M. Grynbaum and R. Mac, "The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work," The New York Times, Dec. 27, 2023. [Online]. Available: https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

[21] E. Newton-Rex, "Why I Just Resigned from My Job in Generative AI," *Music Business Worldwide*, Nov. 15, 2023. [Online]. Available: https://www.musicbusinessworldwide.com/why-just-resigned-from-my-job-generative-ai/.

[22] E. Drott, "Copyright, compensation, and commons in the music AI industry," *Creative Industries Journal*, vol. 14, no. 2, pp. 9-10, 2020. https://doi.org/10.1080/17510694.2020.1839702.

[23] F. Morreale, "Where does the buck stop? Ethical and political issues with AI in music creation," *Transactions of the International Society for Music Information Retrieval*, vol. 4, no. 1, pp. 105-113, 2021. [Online]. Available: https://doi.org/10.5334/tismir.86.

[24] "Stable Audio: Fast Timing-Conditioned Latent Audio Diffusion," Stability AI, Sept. 13, 2023. [Online]. Available: https://stability.ai/research/stable-audio-efficient-timing-latent-diffusion.

[25] J. Sterne and E. Razlogova, "Tuning sound for infrastructures: artificial intelligence, automation, and the cultural politics of audio mastering," *Cultural Studies*, vol. 35, nos. 4-5, pp. 751-752, 2021. https://doi.org/10.1080/09502386.2021.1895247.

[26] J. Sterne and E. Razlogova, "Machine learning in context, or learning from LANDR: artificial intelligence and the platformization of music mastering,"

Social Media + Society, vol. 5, no. 2, p. 2, 2019. [Online]. Available: https://doi.org/10.1177/2056305119847525.

[27] "AI Sound Synthesis Technology," Yamaha. [Online]. Available: https://www.yamaha.com/en/about/research/technologies/aisynth/ (accessed Jan. 3, 2024).

[28] B. Caramiaux and M. Donnarumma, "Artificial intelligence in music and performance: a subjective art-research inquiry," in *Handbook of Artificial Intelligence for Music*, E. R. Miranda, Ed., Cham, CH: Springer Nature, 2021, p. 78. https://doi.org/10.1007/978-3-030-72116-9_4.

[29] "SynthID," Google DeepMind. [Online]. Available: https://deepmind.google/technologies/synthid/ (accessed Jan. 3, 2024).

[30] Dhariwal, et al., "Jukebox: a generative model for music," 2020. [Online]. Available: https://doi.org/10.48550/arXiv.2005.00341.

[31] F. Kreuk et al., "AudioGen: textually guided audio generation," presented at ICLR, Kigali, Rwanda, May 1-5, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2209.15352.

[32] R. S. Roman et al., "From discrete tokens to high-fidelity audio using multi-band diffusion," presented at NeurIPS, New Orleans, LA, USA, December 10-16, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2308.02560.

[33] Agostinelli et al., "MusicLM: generating music from text," 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2301.11325.

[34] J. Copet et al., "Simple and controllable music generation," presented at NeurIPS, New Orleans, LA, USA, December 10-16, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2306.05284.

Chord Progression Analysis by Labelled Lambek Calculus

Matteo Bizzarri[1], Satoshi Tojo[2]

[1]Scuola Normale Superiore, Italy
[2]Advanced Institute of Science and Technology, Japan
[1]matteo.bizzarri@sns.it
[2]tojo_satoshi@asia-u.ac.jp

Abstract. Music and language are thought to share a common origin, leading to numerous research endeavors that have sought to analyze music through a compositional approach, grounded in linguistic grammar rules. In this study, we extend and refine this method, demonstrating that the analysis can be systematically represented using rigorous proof theory, akin to how formal logic elucidates natural language semantics. Our approach involves the use of sequent calculus to construct a proof, notably extending Lambek calculus for categorial grammar to a labeled version. The introduction of labels enables each term within a sequent to be interpreted as a chord with a specified key (tonality). This labeling allows us to precisely determine the grammatical validity of a given sequence of chords. Given that music grammar varies with genre and era, our formalism is designed to accommodate flexible addition and reduction of applicable rules. Additionally, we show that this process of analysis is reversed in Tableaux style, to simulate the process of composition. This methodology not only enhances our understanding of music from a linguistic perspective but also provides a versatile framework for adapting to the evolving nature of musical grammar.

Keywords. Music Analysis, Lambek Calculus, Proof Theory, Tableaux, Chord Progression.

1 Introduction

As Charles Darwin mentioned, the origin of human language is said to be one and the same as music [15], which must have been employed for wooing between primordial males and females. Since our natural languages are classified approximately as context-free language (CFL) in Chomsky hierarchy, the syntax of a sentence is structured in tree; thus, if the origin is common music, too, should be structured in a hierarchical tree.

Thus far, Generative Theory of Tonal Music (GTTM) [8], Rohmeier's Generative Synax Model (GSM) [12], and so on, have shown such tree-structured analyses of music, and among such trials Steedman et al. [5] has shown a construction by combinatorial category grammar (CCG). In this paper, we formalize music analysis in a tree style, extending beyond syntactic analysis to include formal semantics. Just as language semantics are represented in logical formalism, we explore interpretation through logic. Our goal is to analyze a music piece, given a sequence of chord names, and determine if a cadential structure can be derived. We treat a sequence of chord names as a syntactic structure, interpreting each chord name as semantics - a pair of a key and a degree. Subsequently, we compose these interpretations in a structural manner, ultimately reducing them to a cadence.

This paper is organised as follows: in Section 2, we set up our formalism. In the following Section 3, we'll show an analysis of "Stella by Starlight". In Section 4, we show that the analysis process can also be viewed as its dual, i.e., Tableaux, Finally, we summarise our contribution and limitations and discuss future issues.

2 Labelled calculus for music analysis

Chord Notation and Lexicon

First, we strictly distinguish the three kinds of chord notations.

- **Berklee Chord Names -** Em7, A7, Cm7, etc… each of which is a set of diatonic notes, and are shown in upright fonts.
- **Key : Degree -** A pair of a key (tonality) and a degree in roman numeric is shown by being connected by colon (:), e.g., C:I, G:iii, a:i, etc. The lowercase in key is a minor, and the lowercase numeric is a minor third, and they are shown in italic fonts.
- **Chord functions - T** (tonic), and **D** (dominant), **S** (subdominant), shown in boldface fonts.

We provide a *Lexicon*, where a Berklee chord is looked up and it can be interpreted in multiple ways, as follows.

F	$\Rightarrow$	$C:\mathrm{IV},\ F:\mathrm{I},\ B\flat:\mathrm{V},\cdots$
G	$\Rightarrow$	$G:\mathrm{I},\ D:\mathrm{IV},\ a:\mathrm{VII},\cdots$
B$\flat$	$\Rightarrow$	$F:\mathrm{IV},\cdots$
C7	$\Rightarrow$	$F:\mathrm{V7},\cdots$
$\vdots$		$\vdots$

Lambek Calculus (LC) is a sequent calculus for Categorial Grammar (CG). Since CG can bind an adjacent word or a category either from the left-hand or the right-hand side, we need to provide rules in two different directions: → and ←[1]. In Kripke semantics of modal logic, □x means the necessity; that is, in all the accessible possible worlds, x must hold, and ◊x means that there exists some possible worlds in which x holds. In we show a set of general rules for labelled Lambek calculus (LLC)[2], apart from the music construction.

Table 1 Labelled Lambek Calculus, where the right-hand side of '⊢' is intuitionistically restricted to only one term. (□) tells that if αRβ deduces β : x then α : □x, and (◊) vice versa. Note that αRβ appears order-free in sequents.

$$\frac{\Gamma \vdash \alpha : x \qquad \Delta, \alpha : y, \Sigma \vdash \beta : z}{\Delta, \alpha : y \leftarrow x, \Gamma, \Sigma \vdash \beta : z} (\leftarrow_L) \qquad \frac{\Gamma, \alpha : x \vdash \alpha : y}{\Gamma \vdash \alpha : y \leftarrow x} (\leftarrow_R)$$

$$\frac{\Gamma \vdash \alpha : x \qquad \Delta, \alpha : y, \Sigma \vdash \beta : z}{\Delta, \Gamma, \alpha : x \rightarrow y, \Sigma \vdash \beta : z} (\rightarrow_L) \qquad \frac{\alpha : x, \Gamma \vdash \alpha : y}{\Gamma \vdash \alpha : x \rightarrow y} (\rightarrow_R)$$

$$\frac{\Gamma \vdash \alpha : x \qquad \alpha : x \vdash \beta : y}{\Gamma \vdash \beta : y} (\text{Cut}) \qquad \frac{\Gamma \vdash \alpha : \Box x}{\alpha R \beta,\ \Gamma \vdash \beta : x} (\Box)$$

$$\frac{\alpha R \beta,\ \Gamma \vdash \beta : x}{\Gamma \vdash \alpha : \Diamond x} (\Diamond)$$

LLC rules for cadences

Hereafter, we use α, β, … for keys, and x, y, … for degrees. Now, we employ LLC for cadence rules. We regard αRβ is a key modulation from α to β, that is, key β is accessible from α if they are related. For example, if R represents a shift to the parallel key, chord α : x becomes β : y in the parallel key. R works in a different way, according to shifts to parallel, relative, dominant, and sub-dominant keys.

$$\alpha: x, \alpha R \beta \Rightarrow \beta: R(x)$$

together with other initial sequents for chord functions:

$$\alpha: I \vdash \alpha: \boldsymbol{T}, \boldsymbol{\alpha} : V \vdash \alpha: \boldsymbol{D}, \alpha: IV \vdash \alpha: \boldsymbol{S} \ldots$$

[1] In the original LC, /' and \' are used for two directions, but /' is confusing in the context of chord representation (/' is used for *doppel*-dominant or to denote a bass in a chord). Therefore, we employ arrows instead.

[2] The original LC includes a construction by dot (·), however, we omit them here since they are not mentioned in this paper.

We introduce Cadence Rules below, instantiating the labelled Lambek rules in Table

$$\frac{\Gamma \vdash \alpha : \mathbf{D} \qquad \Delta, \alpha : \mathrm{I} \vdash \alpha : \mathbf{T}}{\Delta, \Gamma, \alpha : \mathbf{D} \to \mathrm{I} \vdash \alpha : \mathbf{T}} \ (\to_L)$$

$$\frac{\Gamma \vdash \alpha : \mathbf{S} \qquad \Delta, \alpha : \mathrm{V} \vdash \alpha : \mathbf{D}}{\Delta, \Gamma, \alpha : \mathbf{S} \to \mathrm{V} \vdash \alpha : \mathbf{D}} \ (\to_L)$$

$$\frac{\alpha : \mathrm{I} \vdash \alpha : \mathbf{T} \qquad \Gamma, \alpha : \mathrm{I} \vdash \alpha : \mathbf{T}}{\Gamma, \alpha : \mathbf{T} \leftarrow \mathrm{I}, \alpha : \mathrm{I} \vdash \alpha : \mathbf{T}} \ (\leftarrow_L)$$

From the given initial sequents and instantiated rules, the right-hand side of '⊢' is always a chord function; it means that a sequence of chords is interpreted to bear a function. For an easy example, given an input chord sequence: G–C–D–G, we can obtain the following construction, consulting the lexicon.

$$\frac{\dfrac{\mathrm{G}\Rightarrow G{:}\mathrm{I}}{G : \mathrm{I} \vdash G : \mathbf{T}} \qquad \dfrac{\dfrac{\dfrac{\mathrm{C}\Rightarrow G{:}\mathrm{IV}}{G : \mathrm{IV} \vdash G : \mathbf{S}} \quad \dfrac{\mathrm{D}\Rightarrow G{:}\mathrm{V}}{G : \mathrm{V} \vdash G : \mathbf{D}}}{G : \mathbf{S} \to \mathrm{V}, G : \mathrm{V} \vdash G : \mathbf{D}}\ (\to_L) \qquad \dfrac{\mathrm{G}\Rightarrow G{:}\mathrm{I}}{G : \mathrm{I} \vdash G : \mathbf{T}}}{G : \mathbf{D} \to \mathrm{I}, G : \mathrm{I} \vdash G : \mathbf{T}}\ (\to_L)}{G : \mathbf{T} \leftarrow \mathrm{I}, G : \mathrm{I} \vdash G : \mathbf{T}}\ (\leftarrow_L)$$

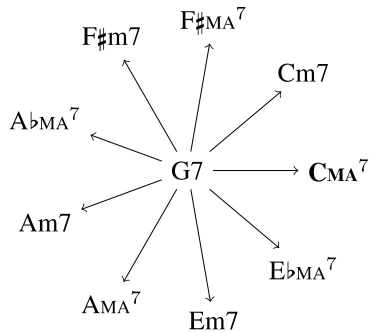
However when the chord C or D occupy longer time duration, we should regard they are local modulations as C:I or D:I. In which case, the latter part of a valid proof tree is constructed, as follows.

$$\frac{\dfrac{\dfrac{\mathrm{C}\Rightarrow C{:}\mathrm{I} \qquad C : \mathrm{I},\ {}_{C}R_{G} \Rightarrow G : \mathrm{IV}}{G : \mathrm{IV} \vdash G : \mathbf{S}} \qquad \dfrac{\mathrm{D}\Rightarrow D{:}\mathrm{I} \qquad D : \mathrm{I},\ {}_{D}R_{G} \Rightarrow G : \mathrm{V}}{G : \mathrm{V} \vdash G : \mathbf{D}}}{G : \mathbf{S} \to \mathrm{V}, G : \mathrm{G} \vdash G : \mathbf{D}} \qquad \dfrac{\mathrm{G}\Rightarrow G{:}\mathrm{I}}{G : \mathrm{I} \vdash G : \mathbf{T}}}{G : \mathbf{D} \to \mathrm{I}, G : \mathrm{I} \vdash G : \mathbf{T}}$$

3 Analysis of "Stella By Starlight"

The main reason for opting for the Labelled Sequent Calculus is the challenging nature of analyzing songs like 'Stella by Starlight' using the instruments introduced in [1]. This challenge stems from the absence of consequent cadences, i.e., that the cadences do not adhere to a strict order, as conventional proof theory rules need for finding a solution. Instead, it is evident that a cadence can resolve one, two, or more measures away from a given point.

Another reason is that encountering a dominant doesn't guarantee a strict resolution to its tonic; instead, it can lead to multiple potential resolutions:



Labelled sequents are helpful to take trace of the current tonality, something that Proof Theory alone couldn't do. To do something like that we will try to use a modified version of labelled sequents, introduced firstly in [4].

As an example, we take "Stella by Starlight" (Figure 1), which is very useful be- cause it contains numerous problematic cadences. For the sake of readability, the har- monic analysis is divided into multiple parts instead of being presented as one very long derivation. To better understand the example, we will describe each part of it.

A. These chords are extracted from measures 5–7, illustrating a classic cadence on E♭.

B. This represents a perfect cadence on B♭MA7, which can be easily deduced.

C. It is the combination of **A** and **B** denoted by $^{(*1)}$ and $^{(*2)}$. In the left section (mm. 3-4), a cadence occurs, but it is not resolved in the subsequent chords. In- stead, resolution happens on the B♭ in measure 9. This brings closure, resolving everything on the tonic of B♭. This kind of cadence is not an exception in music; it is something really common. As a metaphor, we can say that just as in speech where we might start discussing a focus and then change, only to finally come back to the initial topic, in music, cadences like these are frequent. The LLC here fits perfectly, helping us to understand when a certain cadence closes.

D. It is perfect cadence on Dm.

E. This is the union of all the preceding derivations, providing closure to each cadence. The unresolved cadence from the first two measures finds resolution in measure 11, the final measure of this structure, concluding all preceding derivations.

F. From here, we will not write every cadence in all the sequents, as the sequences are now easily closed. This is, in fact, a perfect cadence that resolves from the seventh flat degree.

G. This represents a minor cadence resolving on a seventh chord. Although it may seem a bit complex, the underlying concept is straightforward. The tree indicates an accessible relationship between D minor and D major, simplifying the solution.

H. Follow the same procedure as in **G**.

I. Derive a perfect cadence from the last chord in **H**.

L. A perfect cadence that passes from the seventh degree instead of the fifth.

M. N. O. These are all perfect cadences that are chained because the final chord of **M** is the first of **N** and the final chord of **N** is the first chord of the last cadence of the song in **O**.

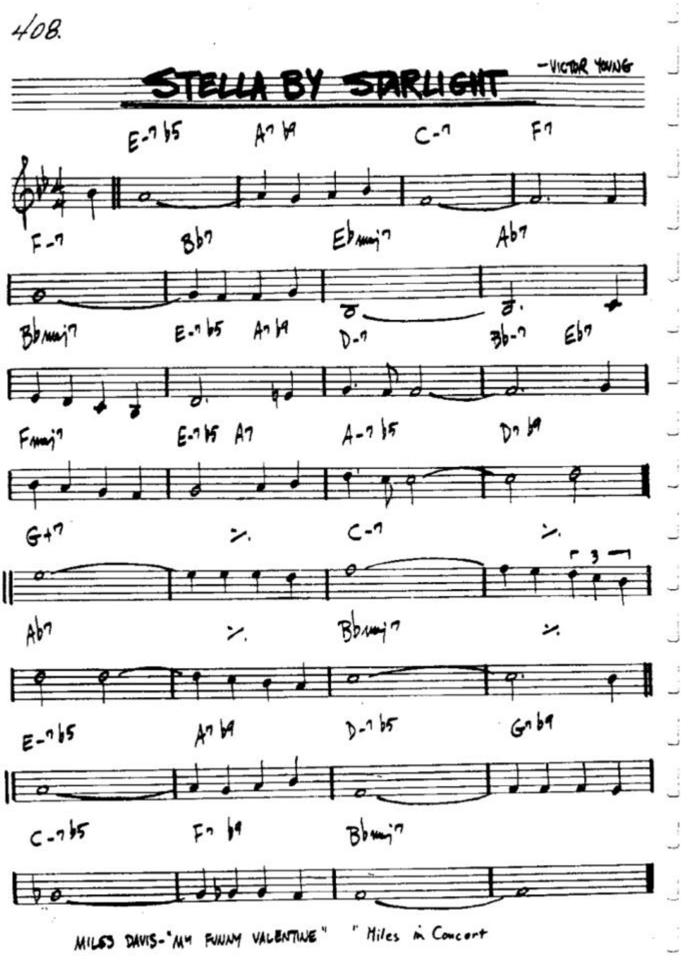



Fig. 1. Stella by Starlight, written by Victor Young.

A

$$\dfrac{\dfrac{\dfrac{\text{Fm7} \Rightarrow \text{E}\flat : \mathbf{II}}{\text{E}\flat : II \vdash \text{E}\flat : \mathbf{S}} \qquad \dfrac{\text{B}\flat 7 \Rightarrow \text{E}\flat : V}{\text{E}\flat : V \vdash \text{E}\flat : \mathbf{D}}}{\text{E}\flat : \mathbf{S} \to V, \text{E}\flat : V \vdash \text{E}\flat : \mathbf{D}}\,(\to_L) \qquad \dfrac{E\flat\,\text{MA}^7 \Rightarrow E\flat : \mathbf{I}}{\text{E}\flat : I \vdash \text{E}\flat : \mathbf{T}}}{^{(*1)}\text{E}\flat : \mathbf{D} \to I, \text{E}\flat : I \vdash \text{E}\flat : \mathbf{T}}\,(\to_L)$$

B

$$\dfrac{\dfrac{\text{A}\flat 7 \Rightarrow \text{B}\flat : VII\flat}{\text{B}\flat : VII\flat \vdash \text{B}\flat : \mathbf{D}} \qquad \dfrac{\text{B}\flat\,\text{MA}^7 \Rightarrow \text{B}\flat : V}{\text{B}\flat : I \vdash \text{B}\flat : \mathbf{T}}}{^{(*2)}\text{B}\flat : \mathbf{D} \to I, \text{B}\flat : I \vdash \text{B}\flat : \mathbf{T}}\,(\to_L)$$

C

$$\dfrac{\dfrac{\dfrac{\text{Cm7} \Rightarrow \text{B}\flat : II}{\text{B}\flat : II \vdash \text{B}\flat : \mathbf{S}} \qquad \dfrac{\text{F7} \Rightarrow \text{B}\flat : V}{\text{B}\flat : V \vdash \text{B}\flat : \mathbf{D}}}{\text{B}\flat : \mathbf{S} \to V, \text{B}\flat : V \vdash \text{B}\flat : \mathbf{D}}\,(\to_L) \qquad (*1) \qquad (*2)}{^{(*3)}\ \text{B}\flat : V, \text{E}\flat : \mathbf{D} \to I, \text{E}\flat : I, \text{B}\flat : \mathbf{D} \to I, \text{B}\flat : I \vdash \text{E}\flat : \mathbf{T}, \text{B}\flat : \mathbf{T}}\,(\to_L)$$

D

$$\dfrac{\dfrac{\dfrac{\text{Em7}^{(\flat 5)} \Rightarrow \text{d}: II}{\text{d}: II \vdash \text{d}: \mathbf{S}} \qquad \dfrac{\text{A7} \Rightarrow \text{d}: V}{\text{d}: V \vdash \text{d}: \mathbf{D}}}{\text{d}: \mathbf{S} \to V, \text{d}: V \vdash \text{d}: \mathbf{D}}\,(\to_L) \qquad \dfrac{\text{Dm} \Rightarrow \text{d}: I}{\text{d:I} \vdash \text{d}:\mathbf{T}}}{^{*(4)}\text{d}: \mathbf{D} \to I, \text{d}: I \vdash \text{d}: \mathbf{T}}\,(\to_L)$$

E

$$\dfrac{\dfrac{\dfrac{\text{Em7}^{\flat 5} \Rightarrow \text{d}: II}{\text{d}: II \vdash \text{d}: \mathbf{S}} \qquad \dfrac{\text{A7} \Rightarrow \text{d}: V}{\text{d}: V \vdash \text{d}: D}}{\text{d}: \mathbf{S} \to V, \text{d}: V \vdash \text{d}: \mathbf{D}}\,(\to_L) \qquad (*3) \qquad (*4)}{\text{d}: V, \text{B}\flat : V, \text{E}\flat : \mathbf{D} \to I, \text{E}\flat : I, \text{B}\flat : \mathbf{D} \to I, \text{B}\flat : I, \text{d}: \mathbf{D} \to I, \text{d}: I \vdash \text{E}\flat : \mathbf{T}, \text{B}\flat : \mathbf{T}, \text{d}: \mathbf{T}}\,(\to_L)$$

F

$$\dfrac{\dfrac{\dfrac{\mathrm{B}\flat m7 \Rightarrow \mathrm{F}{:}\ iv}{\mathrm{F}{:}\ iv \vdash \mathrm{F}{:}\ \mathbf{S}} \quad \dfrac{\mathrm{E}\flat 7 \Rightarrow \mathrm{F}{:}\ VII\flat}{\mathrm{F}{:}VII\flat \vdash \mathrm{F}{:}\ \mathbf{D}}}{\mathrm{F}{:}\ \mathbf{S} \to VII\flat, \mathrm{F}{:}\ VII\flat \vdash F : \mathbf{D}}(\to_L) \quad \dfrac{\mathrm{F_{MA}}^7 \Rightarrow \mathrm{F}{:}\ I}{\mathrm{F}{:}\ I \vdash \mathrm{F}{:}\ \mathbf{T}}}{\mathrm{F}{:}\ \mathbf{D} \to I, \mathrm{F}{:}\ I \vdash \mathrm{F}{:}\ \mathbf{T}}(\to_L)$$

G

$$\dfrac{\dfrac{\dfrac{\mathrm{Em7}^{(\flat 5)} \Rightarrow \mathrm{d}{:}\ II \quad \mathrm{d}{:}\ II,\ {}_d\mathrm{R}_D \Rightarrow \mathrm{D{:}II}}{\mathrm{D}{:}II \vdash \mathrm{D}{:}\ \mathbf{S}} \quad \dfrac{A7 \Rightarrow \mathrm{D}{:}\ V}{\mathrm{D}{:}\ V \vdash \mathrm{D}{:}\ \mathbf{D}}}{\mathrm{D}{:}\ \mathbf{S} \to V, \mathrm{D}{:}\ V \vdash \mathrm{D}{:}\ \mathbf{D}}(\to_L) \quad \dfrac{\mathrm{D7} \Rightarrow \mathrm{D}{:}\ I}{\mathrm{D}{:}\ I \vdash \mathrm{D}{:}\ \mathbf{T}}}{{}^{*(4)}\mathrm{D}{:}\ \mathbf{D} \to I, \mathrm{D}{:}\ I \vdash \mathrm{D}{:}\ \mathbf{T}}(\to_L)$$

H

$$\dfrac{\dfrac{\dfrac{\mathrm{Am7}^{(\flat 5)} \Rightarrow \mathrm{g}{:}\ II \quad \mathrm{g}{:}\ II,\ {}_g\mathrm{R}_{G^{(\sharp 5)}} \Rightarrow \mathrm{G}^{(\sharp 5)} : II}{\mathrm{G}^{(\sharp 5)} : II \vdash \mathrm{G}^{(\sharp 5)} : \mathbf{S}} \quad \dfrac{D7 \Rightarrow \mathrm{G}^{(\sharp 5)} : V}{\mathrm{G}^{(\sharp 5)} : V \vdash \mathrm{G}^{(\sharp 5)} : \mathbf{D}}}{\mathrm{G}^{(\sharp 5)} : S \to V, \mathrm{G}^{(\sharp 5)} : V \vdash \mathrm{G}{:}\ \mathbf{D}}(\to_L) \quad \dfrac{\mathrm{G}^{(\sharp 5)} \Rightarrow \mathrm{G}^{(\sharp 5)} : I}{\mathrm{G}^{(\sharp 5)} : I \vdash \mathrm{G}^{(\sharp 5)} : \mathbf{T}}}{\mathrm{G}^{(\sharp 5)} : \mathbf{D} \to I, \mathrm{G}^{(\sharp 5)} : I \vdash \mathrm{G}^{(\sharp 5)} : \mathbf{T}}(\to_L)$$

I

$$\dfrac{\dfrac{\mathrm{G}^{(\sharp 5)} \Rightarrow \mathrm{G}^{(\sharp 5)} : I \quad \mathrm{G}^{(\sharp 5)} : I,{}_{G^{(\sharp 5)}}\mathrm{R}_c \Rightarrow \mathrm{c}{:}\ V}{\mathrm{c}{:}\ V \vdash \mathrm{c}{:}\ \mathbf{D}} \quad \dfrac{\mathrm{Cm7} \Rightarrow \mathrm{c}{:}I}{\mathrm{c{:}I} \vdash \mathrm{c}{:}\mathbf{T}}}{\mathrm{c}{:}\ \mathbf{D} \to I, \mathrm{c}{:}\ I \vdash \mathrm{c}{:}\ \mathbf{T}}(\to_L)$$

L

$$\dfrac{\dfrac{\dfrac{\mathrm{c} \Rightarrow \mathrm{c}{:}\ I \quad \mathrm{c}{:}\ I,{}_c\mathrm{R}_{B\flat} \Rightarrow \mathrm{B}\flat : ii}{\mathrm{B}\flat{:}\ ii \vdash \mathrm{B}\flat{:}\ \mathbf{S}} \quad \dfrac{\mathrm{A}\flat 7 \Rightarrow \mathrm{B}\flat{:}\ VII\flat}{\mathrm{B}\flat{:}VII\flat \vdash \mathrm{B}\flat{:}\ \mathbf{D}}}{\mathrm{B}\flat{:}\ \mathbf{S} \to VII\flat, \mathrm{B}\flat{:}\ VII\flat \vdash \mathrm{B}\flat{:}\mathbf{D}}(\to_L) \quad \dfrac{\mathrm{B}\flat\mathrm{MA}^7 \Rightarrow \mathrm{B}\flat{:}\ I}{\mathrm{B}\flat{:}\ I \vdash \mathrm{B}\flat{:}\ \mathbf{T}}}{\mathrm{B}\flat{:}\ \mathbf{D} \to I, \mathrm{B}\flat{:}\ I \vdash \mathrm{B}\flat{:}\ \mathbf{T}}(\to_L)$$

M

$$\dfrac{\dfrac{\dfrac{\mathrm{Em7}^{(\flat 5)} \Rightarrow \mathrm{d}^{(\flat 5)} : II}{\mathrm{d}^{(\flat 5)} : II \vdash \mathrm{d}^{(\flat 5)} : \mathbf{S}} \qquad \dfrac{\mathrm{A7} \Rightarrow \mathrm{d}^{(\flat 5)} : V}{\mathrm{d}^{(\flat 5)} : V \vdash \mathrm{d}^{(\flat 5)} : \mathbf{D}}}{\mathrm{d}^{(\flat 5)} : \mathbf{S} \to V, \mathrm{d}^{(\flat 5)} : V \vdash \mathrm{d}^{(\flat 5)} : \mathbf{D}} (\to_L) \qquad \dfrac{\mathrm{Dm7}^{(\flat 5)} \Rightarrow \mathrm{d}^{(\flat 5)} : I}{\mathrm{d}^{(\flat 5)} : \mathrm{I} \vdash \mathrm{d}^{(\flat 5)} : \mathbf{T}}}{*^{(5)}\mathrm{d}^{(\flat 5)} : \mathbf{D} \to I, \mathrm{d}^{(\flat 5)} : I \vdash \mathrm{d}^{(\flat 5)} : \mathbf{T}} (\to_L)$$

N

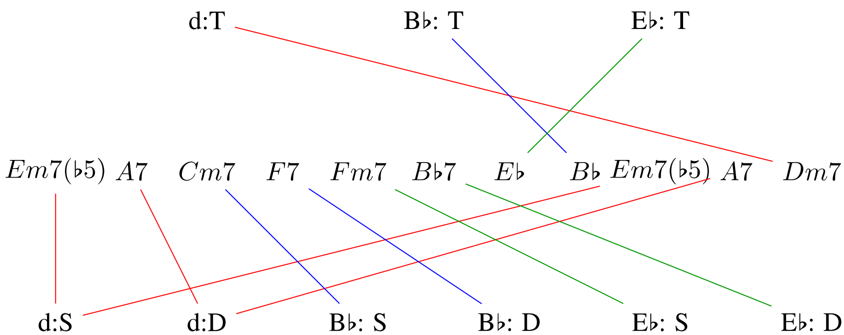
$$\dfrac{\dfrac{\dfrac{*^{(5)}\mathrm{d}^{(\flat 5)} : V \to I, \mathrm{d}^{(\flat 5)} : I \vdash \mathrm{d}^{(\flat 5)} : \mathbf{T} \qquad \mathrm{d}^{(\flat 5)} : I, {}_{d^{(\flat 5)}}\mathrm{R}_{c^{(\flat 5)}} \Rightarrow \mathrm{c}^{(\flat 5)} : II}{\mathrm{c}^{(\flat 5)} : II \vdash \mathrm{c}^{(\flat 5)} : \mathbf{S}} \qquad \dfrac{\mathrm{G7} \Rightarrow \mathrm{c}^{(\flat 5)} : V}{\mathrm{c}^{(\flat 5)} : V \vdash \mathrm{c}^{(\flat 5)} : \mathbf{D}}}{\mathrm{c}^{(\flat 5)} : \mathbf{S} \to V, \mathrm{c}^{(\flat 5)} : V \vdash \mathrm{c}^{(\flat 5)} : \mathbf{D}} (\to_L) \qquad \dfrac{\mathrm{Cm7}^{(\flat 5)} \Rightarrow \mathrm{c}^{(\flat 5)} : I}{\mathrm{c}^{(\flat 5)} : \mathrm{I} \vdash \mathrm{c}^{(\flat 5)} : \mathbf{T}}}{*^{(6)}\mathrm{c}^{(\flat 5)} : \mathbf{D} \to I, \mathrm{c}^{(\flat 5)} : I \vdash \mathrm{c}^{(\flat 5)} : \mathbf{T}} (\to_L)$$

O

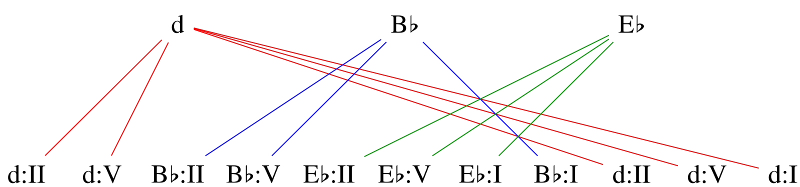
$$\dfrac{\dfrac{\dfrac{*^{(6)}\mathrm{c}^{(\flat 5)} : \mathbf{D} \to I, \mathrm{c}^{(\flat 5)} : I \vdash \mathrm{c}^{(\flat 5)} : \mathbf{T} \qquad \mathrm{c}^{(\flat 5)} : I, {}_{c^{(\flat 5)}}\mathrm{R}_{B\flat} \Rightarrow \mathrm{B}\flat : II}{\mathrm{B}\flat : II \vdash \mathrm{B}\flat : \mathbf{S}} \qquad \dfrac{F7 \Rightarrow \mathrm{B}\flat : V}{\mathrm{B}\flat : V \vdash \mathrm{B}\flat : \mathbf{D}}}{\mathrm{B}\flat : \mathbf{S} \to V, \mathrm{B}\flat : V \vdash \mathrm{B}\flat : \mathbf{D}} (\to_L) \qquad \dfrac{\mathrm{B}\flat \Rightarrow \mathrm{G} : I}{\mathrm{B}\flat : I \vdash \mathrm{B}\flat : \mathbf{T}}}{\mathrm{B}\flat : \mathbf{D} \to I, \mathrm{B}\flat : I \vdash \mathrm{B}\flat : \mathbf{T}} (\to_L)$$

Tableaux for Labelled sequents

The idea of Tableux for LLS is similar to the one presented in [1], but it will be a little more intricate due to the non-linearity of harmony. For example, the first 11 chords of “Stella by Starlight” can be notated in this manner:



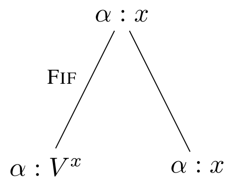
In the upper part, we have connected all the Tonic chords to their respective counterparts. In the lower part, we have established connections for the other functions, including Subdominant and Dominant chords. This comprehensive representation allows us to visualize the harmonic relationships more thoroughly and understand the interplay between different chord functions. It can alternatively be expressed in the form of chord functions:



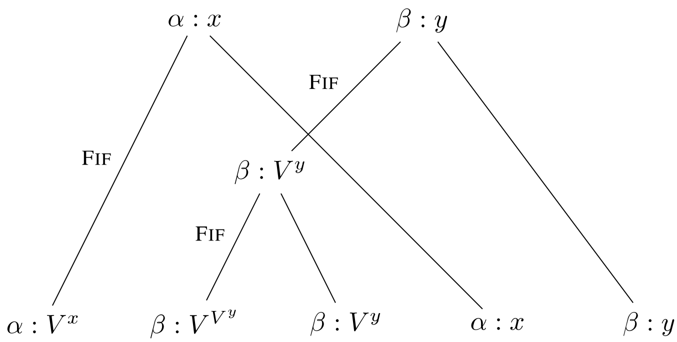
The reasons to use Tableaux are mainly twofold: one can be for a better representation of the entire structure of the harmonic sequence; it is less expressive than sequents, but it can be useful as a map. The other reason is that it can be helpful for composition; it is sufficient to create the dual of the rules introduced for LLC.

Rules

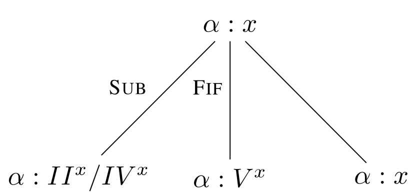
The primary challenge in reversing the rules for Tableaux lies in our intention to employ them for generating new harmonic structures. Consequently, we must adapt them to suit our specific needs. It is important to emphasize here that the rules are easily changeable, given the nature of music, which does not rigidly adhere to set rules. Let's start with a simple rule: from every chord is always possible to add a dominant, keeping in mind that we want a tonal structure. Let's call it *Fif*.



where V^x indicates the dominante of the degree *x*. Differently from the one presented in [1] here the dominants can be chained:

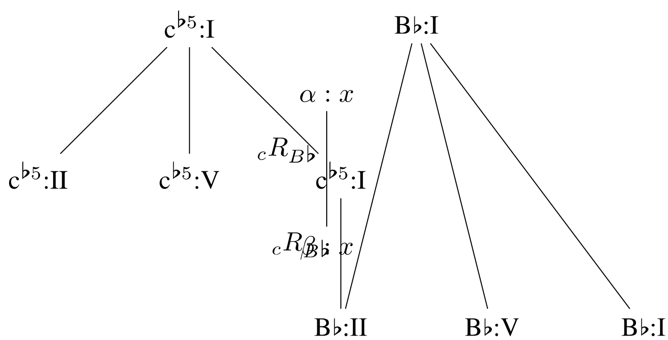


Here, we encounter three instances of the rule *Fif*, and they are intertwined. Obviously we can chain that to every α : S (α : II or α : IV), let's call it *Sub*.



Furthermore, cadences can introduce variations in the function of specific chords. For instance, a C7 chord can serve as the dominant of FMA7, yet it may also function as the first degree in a cadence that resolves to G7. An illustration of this concept can be found in the final part of Stella By Starlight, where the harmonic progression takes an unexpected turn. In general, we can write that:

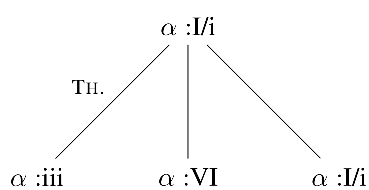
This implies that when there is an accessibility relation αRβ, we have the flexibility to change the interpretation of the chord. For instance, the final part of "Stella By Starlight," it can be seen as:



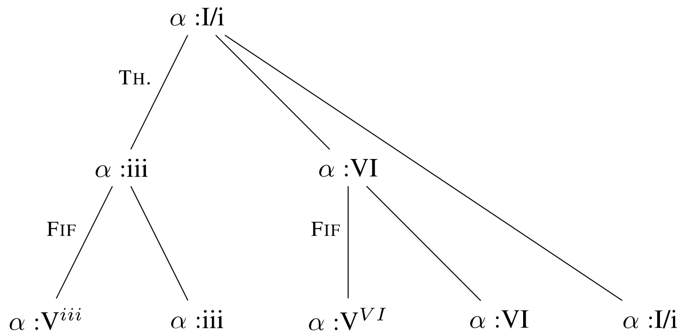
New rules

We didn't want to be complete in the list of the rules, in fact the flexibility of this system proves to be useful when incorporating new rules. For instance, if we wish to introduce a rule allowing the tonic (and only the tonic) to split into its third minor and sixth major, it can be easily formalized in the following manner, calling it *Th.*:

Obviously, these rules can be combined into new harmonic structures:



The significant aspect of this work is that such structures are inherently rational; that is, they adhere to well-established rules or rules determined by the composer. This implies that, on one hand, the structure can be very strict, while on the other hand, it can also be highly malleable.



4 Conclusions

Based upon the traditional view that the origin of language is common with music, we have regarded a sequence of chord names as a given surface sentence, and have shown a structural proof system with proposed sequent rules to decide the grammaticality. Thus far, several works have shown syntactic tree structures for the surface sequence of chords, however, in our work we have further developed the method so as to give its semantics, interpreting each chord and deriving chord functions.

Our new methodology is the labelled sequent calculus; Gentzen's sequent calculus is a strict method to show the adequacy of hierarchical composition, and thus it has been employed as a basis for Lambek calculus (LC) that is a sequent version of categorial grammar (CG). As the labelled sequent calculus has been proposed as a proof system of modal logic, we have applied the notion of label to specify the tonality or the region where each chord resides. Since a Berklee chord name can be interpreted in multiple keys, we could clarify the key with such labels, and the shift to related keys is elegantly shown by the accessibility between possible regions.

With this method, we have analyzed an actual piece of 'Stella by Starlight' and exemplified how our system works. The example is emblematic due to its unique structure; indeed, harmony is not a linear process, necessitating a non-linear analysis easily solved through labeled sequent calculus. Furthermore, we have delineated the ways to use this system in composition, leveraging the dual of sequents, i.e., Tableaux. This novel approach, initially introduced in [1], has been expanded here through various examples and thanks to labels. Labels, in fact, provided easier readability, as well as the creation of more complex harmonic structures.

However, music pieces change their styles according to era and genre, where occur many variations on the authentic theory, to tolerate dissonant chords or non-harmonic notes. Our system is malleable in this sense since we can employ different sets of sequent rules, although at the same time the adequacy of the set of rules should be verified from an objective view. Since some extra rules are rarely utilized, we should also consider the probability in the application of each rule. As for our logic, we have only applied labels for key-shift, but we should embrace the original meanings of labels, addressing the necessity (□) and the possibility (◊), and this is our immediate future work. Another potential task is to automate this analysis system and perhaps create an automatic composer system based on the Tableaux.

Acknowledgment

This work is supported by JSPS 21H03572 and 20H04302.

References

[1] Bizzarri, M. Music and Logic: a connection between two worlds, in Proc. of The 16th International Symposium on Computer Music Multidisciplinary Research (CMMR), (2023).

[2] Chew, E. Slicing It All Ways: Mathematical Models for Tonal Induction, Approxi- mation, and Segmentation Using the Spiral Array. INFORMS Journal on Comput- ing. 18 pp. 305 (2006,8).

[3] Douthett, J. & Steinbach, P. Parsimonious Graphs: A Study in Parsimony, Con- textual Transformations, and Modes of Limited Transposition. Journal of Music Theory. 42, 241-263 (1998).

[4] Gabbay, Dov M. Labelled deductive systems. New York: Oxford University Press (1996).

[5] Granroth-Wilding, M. & Steedman, M. Statistical Parsing for Harmonic Analysis of Jazz Chord Sequences. ICMC 2012: Non-Cochlear Sound - Proceedings of the International Computer Music Conference 2012. pp. 478-485 (2012,1).

[6] Gerhard Gentzen, Untersuchungen u¨ber das logische Schließen. I. Math Z 39. pp. 176–210 (1935).

[7] Krumhansl, Carol L., E. Tracing the dynamic changes in perceived tonal organiza- tion in a spatial representation of musical keys.. Psychological Review. 89, 334-368 (1982).

[8] Lerdahl, F. and Jackendoff, R. A Generative Theory of Tonal Music, The MIT Press (1981).

[9] Messiaen, O. Technique de mon langage musical. (Alphonse Leduc,1944)
[10] Polth, M. The Individual Tone and Musical Context in Albert Simon's Tonfeldthe- orie. Music Theory Online. 24 (2018).
[11] Rohrmeier, M. Towards a generative syntax of tonal harmony. Journal of Mathe- matics and Music. 5, 35-53 (2011).
[12] Rohrmeier, M. & Moss, F. A Formal Model of Extended Tonal Harmony. Proceedings of the 22nd International Society for Music Information Retrieval Conference. pp. 569-578 (2021).
[13] Tojo, S. Modal Logic for Tonal Music. Perception, Representations, Image, Sound, Music: 14th International Symposium, CMMR 2019, Marseille, France, October 14–18, 2019, Revised Selected Papers. pp. 113-128 (2019).
[14] Troelstra, A. & Schwichtenberg, H. Basic Proof Theory. (Cambridge University Press,2000).
[15] Wallin, N. L., Merker, B., and Brown, S. The Origins of Music, The MIT Press (2000).

Examining Liszt's Compositional Language through Schenkerian and Neo-Riemannian Theories in *O lieb, so lang du lieben kannst*

Nikita Mamedov

North America International School, Shanghai, China
nais-music@north-americais.cn; nmamedovmusic@aol.com

Abstract. The study focused on Franz Liszt's compositional language in *O lieb, so lang du lieben kannst*. Integrating Schenkerian and Neo-Riemannian analyses can help develop a unified analytical technique to examine the composer's approach to tonal reasoning and harmonic strategy seen in the song. The compositional innovations seen in Liszt's artistry in *O Lieb* require multiple outlooks that Schenker's and Neo-Riemannian theories can provide. The study dissected Liszt's application of vital key regions and harmonic areas and how these intertwine with the singer's melody.

Keywords. Franz Liszt, Neo-Riemannian theory, Schenkerian theory, Vocal music

1 Introduction

Franz Liszt, one of the leading composers in the Romantic era, transformed the musical world by developing an innovative and unique compositional style filled with virtuosity, technicality, and emotional depth. His novel compositional approaches deserve an in-depth academic and analytical observation, and his pianistic evolution translates into pioneering art forms. Liszt's distinct musical language is multifaceted, revolutionizing what music symbolizes to performers and listeners. Liszt's artistic developments contributed to the musicological knowledge, allowing for a thorough evaluation of the compositional realms and boundaries the composer pushed to ensure his music was stylistically captivating. Liszt's prominent aspect in his compositional language was the art of musical virtuosity, defining his compositions' complexity, technique, and expressiveness while necessitating instrumental strength and excellence on behalf of the performers. Liszt's compositional language had emotional depth, engraving the Romantic era's passion, drama, and zest [1]. Liszt pushed the boundaries of tonal harmonies in his compositions, developing a harmonic language that allowed him to withdraw from traditional composing methods while focusing on unconventional progressions and voice leading. Liszt's music served as a transition into twentieth-century music, defining the meaning of harmonic exploration through chromaticism and dissonance [2].

Liszt's piano and orchestral works are often in the spotlight of music history, showcasing the composer's musical character and compositional output [3]. Integrating academic and performative analysis in Liszt's music should span lesser-studied works, including his vocal music. Liszt's vocal oeuvre deserves recognition due to the composer's ability to uniquely interpret poetry within his compositions while focusing on integrating pianistic lyricism with literature. Liszt's vocal music is significant to his compositional language, offering valuable academic and performative insight into how the composer interpreted vocal music and its place in Romantic vocal genres compared to the music literature of other composers. Liszt's vocal contributions to music history extend the value of his compositional style and offer a diverse comparison to the compositional strategies he employed throughout his career. Therefore, Liszt's compositional explorations in the vocal genres should not go unnoticed and deserve an academic analysis emphasizing harmonic language, musical depictions of poetic elements, and a compositional approach to illuminating literature through music.

This study focuses on Liszt's lieder titled *O lieb, so lang du lieben kannst*, set to text by an acclaimed German poet, Ferdinand Freiligrath [4]. Liszt composed the song for soprano with piano accompaniment, which later appeared as a solo piano composition titled *Liebestraum* No. 3 [5]. The purpose of this research is two-fold. The first is to analyze the musical interrelations and developments between music and poetics. The second is to illustrate Liszt's construal of Freiligrath's text within the music through an analytical construal. The current study argues that Liszt judiciously maneuvers through harmonic regions, employing particular keys to define symbolic and poetic elements of the text. The study draws on Neo-Riemannian and Schenkerian frameworks to interpret the analytical aspects of Liszt's music and conclude the relations between tonal areas and poetic symbolism. The study likewise outlines Liszt's purposeful use of keys and harmonic regions and the interaction such pianistic harmonies develop with the singer's part. The study helps one understand Liszt's interpretation of literary themes and his use of pianistic techniques to depict poetic transformations within Freiligrath's work. Freiligrath's text and the poetic thematics disseminate through Liszt's music as the composer strives to utilize his harmonic language to correspond to a musical setting for an expressive depth to the poetic meaning.

2 Theoretical Framework

The current study utilized Schenkerian and Neo-Riemannian theories as theoretical frameworks to complement the music analysis and define Liszt's choice of compositional techniques. Schenkerian and Neo-Riemannian approaches offer a network of analytical methods to examine harmonies' tonal coherence and chromatic nature and functions, focusing on tonal relationships between the chords within the progression and its tonic [6]. Both analyses help understand the tonal spectrum of Liszt's harmonies while focusing on the musical consistency and structure [7]. The Schenkerian and Neo-Riemannian theories help expand the understanding behind Liszt's choice of harmonic regions and explain their development within the lieder's form. The Schenkerian and Neo-Riemannian analyses offer a perception of Liszt's creative processes, defining the composer's method to structure his work and ensuring that the vocalist's and the piano accompanist's musicianship adequately infer the poet's meaning. The Schenkerian and Neo-Riemannian theories offer a unique outlook on how musical elements interact within Liszt's compositional technique. In this study, Schenkerian analysis helps understand the tonal function of voices within harmonic progressions [8]. In contrast, Neo-Riemannian analysis helps define the triadic behaviors in tonal regions from the perspective of key areas [9].

Applying Schenker's theories to *O lieb, so lang du lieben kannst* allows for a deeper comprehension of Liszt's musical language and complexity within motivic segments that unite to characterize the lieder. The analysis helps uncover the fundamental structures within the music while elaborating the musical purposes behind Liszt's choices. The Schenkerian analysis helps trace voice leading and understand how melodies behave and connect in the composition [10]. The theory likewise offers an analytically centric meaning to what Liszt resembled through his compositions. The Schenkerian approach can illustrate Liszt's ability to intertwine motives in his compositions while focusing on tonal transformations and structural evolvements in the composer's quest to adhere to musical coherence. The Schenkerian analysis examines the tonal architecture that Liszt developed in his work, tracing the composer's compositional techniques and shedding light on novel elements that are prominent to one's poetic comprehension of the song's text. The Schenkerian approach to viewing Liszt's music offers a formulated theory that explains the composer's compositional choices using graphic representation and analytical approaches to define tonality and how Liszt implements it within his music [11]. Schenkerian theory is central to tonal research, helping draw parallels between musical structure, tonal hierarchy, and performative interpretation [12].

Neo-Riemannian theory helps understand the lieder's chordal transformations that evolve as Liszt develops his work's melodic and harmonic elements. The Neo-Riemannian framework helps define the relations among Liszt's harmonies and musically reason their placement in the work. The Neo-Riemannian approach focuses on explaining chromatic and modulatory processes in Liszt's music that are unexplained through Schenkerian analysis. The theory examines harmonic relationships between tonal areas resulting from modulation and chromatic voice leadings [13]. The Neo-Riemannian theory helps explain ways Liszt pushes the tonal boundaries to develop an innovative stylistic musical language. The theory likewise focuses on unexpected progressions that align with symbolic transformations occurring in the song's text. The Neo-Riemannian theory complements the Schenkerian analysis by adding knowledge in a quest to understand the composer's harmonic language and unique compositional voice while focusing on unconventional tonal progressions [14]. Neo-Riemannian theory answers how harmonies interact and how melodic elements within chords transform in musical compositions [15].

The research contends for two research areas within the *O lieb, so lang du lieben kannst*. The first is to define the musical connections that exist between Liszt's music and Freiligrath's poetry. The second is to draw upon Liszt's interpretations of the literature to understand how the composer unites poetic symbolism with his compositional language. The Schenkerian and Neo-Riemannian approaches help define the analytical interpretation and explain Liszt's compositional choices. The literary interpretation explores the poetic depictions, symbolic literary elements, and their place within Liszt's music through framework-based analyses [16]. All figures associated with musical examples appear at the end of the study in the study's Appendix.

3 Literature Review

The scholarly literature on Liszt, his life, and his music offers analyses based on the composer's biography, performance practice perspectives, and the historical context of his musical significance. The select literature review offers sources on Liszt's musical career and analytical sources on the composer's music and compositional language. Sources integrating Schenkerian and Neo-Riemannian theories are of priority to this study, as both approaches form the analytical framework and context to understand *O Lieb* and Liszt's compositional intent. Comprehensive analyses and examinations of Liszt's works are essential to evaluate the existing scholarly literature and define the academic gaps in theoretical research on Liszt's music. Applying Schenkerian and Neo-Riemannian analysis methods to Liszt's works can help expand the knowledge of the composer's artistic and stylistic innovations and musical intentions. The theoretical outlook on the interplay between structure and thematic content in Liszt's vocal repertoire helps construe an academic interpretation of works such as *O Lieb*.

Liszt's music highlights academic and performative research in performance practice, musicology, and music analysis. Walker's work on Liszt is the most comprehensive, up-to-date source, including three volumes on the composer's life, career, and works [17]–[19]. Walker details Liszt's compositional career, emphasizing Liszt's stylistic evolution and keyboard virtuosity. Walker outlined Liszt's influence on the musical culture in Europe and offered a perspective of why the composer is influential in music history. Walker outlines Liszt's contributions to Romantic movements in Western classical music and the influence Liszt had on composers after him.

Gooley's 2004 book titled *The Virtuoso Liszt* offered an outlook on how Liszt's music and stylistic language fit within the European cultural outlook [20]. Liszt's virtuoso career has paved the way for many musical innovations. Contemporary performance practices and the influence of Liszt's pianistic technical abilities impacted contemporary music performance. Liszt's performing style aligns with his compositional language, showcasing similarities in musical innovations that he expanded on in music composition and piano performance. The source analyzes Liszt's place in music history and the innovations he brought into the classical music world.

Hilmes's book – *Franz Liszt: Musician, Celebrity, Superstar* – offers a glimpse of the contributions that Liszt brought to European society through his music [21]. Hilmes covers other areas of Liszt's life outside of his musical composition, such as religion and the influence of other artists from the musical world and their impact on Liszt. Hilmes focuses on Liszt's life as the influence and drive for his music, offering a historical narrative and framework for an innovative musical career trajectory. Hilmes focuses on Liszt's personality and how it affected his music-making and compositional language.

Theoretical sources tend to focus on Liszt's melodic and harmonic output. Wells's study examines rhythm and metrics and how both analytical components influence Liszt's compositional language [22]. The study proposed a new analytical method derived from Lewin's generalized interval system. The repertoire under analysis is Liszt's Transcendental Etude No. 8 in C minor, *Wilde Jagd*. The study decomposes and expands intervallic structures to understand how interval relations function within Liszt's music. The application of various theoretical analyses yielded results that helped define Liszt's compositional approach. The significance of this study exists in developing a new music theory tool to help systematize Liszt's compositional language.

Moortele's study focuses on the form and structure of Liszt's Symphonic Poem No. 10, *Hamlet* – another underrepresented and academically understudied work in Liszt's repertoire [23]. The study discussed the interactions between the programmed meaning and the musical form. The musical form and the deformations within it are components that Liszt utilizes to develop harmonies, homogeneous musical language, and a musical portrayal of Shakespearean characters. The analytical complexity in Liszt's musical meaning grows various musicological and performative interpretations. Liszt's manipulation of form and how musical form interacts with the musical character make this work unique.

Vitalino's dissertation, titled *Franz Liszt's Song Revisions: A Schenkerian Taxonomy*, discusses Liszt's musical landscape in his songs, mainly focusing on underexplored and unexplored thematic ideas of his music [24]. The Schenkerian framework used in the study helps utilize a comparative analysis to understand the theoretical similarities and differences in Liszt's musical thoughts in multiple versions of his compositions. The study uses Schenkerian analysis to categorize the compositional changes and conduct and understand the discourse between the musical meaning of multiple revisions that Liszt went through in his career. Schenkerian research is beneficial in examining the composer's artistic process and how Liszt's musical modifications fit within his compositional language.

Kuczynski's study focuses on the intersection of performance and analysis using Schenkerian theory as a framework [25]. The study discussed Schenker's ideas in understanding how music analysis can construe musical interpretation and how Schenkerian theory can relate to musical uniqueness presented through performance. Various compositions serve as examples in the dissertation. One of such works is Liszt's Transcendental Etude No. 8 in C minor, *Wilde Jagd*. Kuczynski attempted to integrate and unify multiple components of musical performance with the analytical understanding of melody, harmony, and motivic structure.

Damschroder's study utilizes the Schenkerian analysis to examine the structural layering of select Liszt's works [26]. The study focused on the extension of Schenkerian theory and its usefulness in examining Liszt's music to evaluate his tonality, musical ideas, and compositional language. Liszt's harmony contains musical complexities that require theoretical frameworks for thorough understanding and evaluation. The study explores the Schenkerian analysis and its adaptable nature to fit various repertoires. Damschroder provides a structural foundation for Liszt's musical principles and defines the tonal character of Liszt's compositional language.

Chapkanov focused on harmonic analysis and utilized Neo-Riemannian theory to understand Liszt's tonal approaches' functionality and chord relation [27]. The author aims to understand Liszt's compositional language and the voice leading the composer selects in chords filled with chromaticism. The author focuses on chordal functions within a series of chords and the strength the harmonic progressions have on Liszt's overall harmonies. The author seeks to understand Liszt's chromatic approaches and how the Neo-Riemannian theory is a foundation for understanding the composer's musical intentions.

The Neo-Riemannian theory helps understand the chromatic transformations that Liszt selects and graph the chordal progressions in the context of Liszt's harmonies and their tonal resolutions. Liszt's transformative musical language requires more scholarly research regarding Liszt's vocal compositions. The composer's inventive harmonies and innovative use of enharmonic shifts and chromaticism continue to exist in his repertoire [28]. The study examines Liszt's harmonic progressions and the meaning the text contains inside Liszt's compositional language. The study also seeks to understand how Liszt approaches text symbolism and how the composer integrates text with music within his compositional language. The study found that Liszt emphasizes select harmonies and tonal relationships to highlight literary characteristics found in the lyrics.

4 Schenkerian Analysis

Liszt initializes the song with an introductory two-measure segment before the vocalist comes in with the main melody. The initial phrase spans mm. 1-8, beginning and ending at the home key of A♭ major. The piano accompaniment follows an arpeggiated structure as Liszt explores ^1, ^3, ^5, and sometimes ^7 in a broken arpeggiated pattern. Liszt utilizes the I - V/vi - VI - V/V - V - I progression in the first phrase with two sets of four-note patterns in the accompaniment's right hand, creating a unique vocal interplay. The behavior of every fourth note in the four-note design develops a melodic outline with an abrupt halt at A♮, raised ^1 in A♭ major, at the midpoint of the phrase. Figure 1 presents the Schenkerian graph in the piano accompaniment, focusing on the voice leading in the right hand. It is worth noting that Liszt implements standard voice leading in the first, second, and third notes of each four-note pattern while emphasizing changes in the fourth note of each such structure.

Liszt's repetition of mm. 3-5 and mm. 9-11 symbolize musical prominence as the composer proceeds to reinstate a segment and follows it with a variation. The vocalist's line contains 18 notes in the first phrase, appearing in mm. 2-8, and 17 notes in the second phase, appearing in mm. 8-14. Liszt directly repeats the first 9 and last 7 notes in each phrase. The melodic alteration occurs when the composer transforms the [C-F] segment in the first phrase to note F in the second phrase, composed one octave higher. Figure 2 presents the Schenkerian graph notating such a change. The variation allows Liszt to reinstate the main melody while introducing a note in the higher register early in the song. The highest pitch in the vocalist's line is note A, which occurs in m. 46 and m. 56 at the vocal apex of the lieder, merely a major third above note F that Liszt offers in m. 12.

The apogee of the vocalist's section is vital to the music theory analysis because Liszt chooses to introduce a high pitch early in the song. The lieder's first section ends in m. 28, with the second highest note in the vocalist's range occurring at mm. 18-20, note E, and m. 25, an enharmonic note of F♭. A descent to note G follows the F♭ in m. 25, as the musical phrase gradually ends with a piano transition introducing the B-section. Figure 3 shows the Schenkerian graph of the vocalist's line in mm. 15-27, defining the coherence that Liszt builds through the melody. The F in m. 12 is part of the B♭^{7} harmony. Liszt repeats note E in mm. 18-20 on the C major and A minor harmonies before changing the singer's line. The enharmonic F♭ in m. 25 is on the D♭ minor harmony with a non-chord tone of B♭. Another interpretation of m. 25 harmony is to understand the region as B♭ diminished chord.

Liszt's compositional language employed a unique type of chromaticism not seen in

other Romantic composers. Liszt uses hidden chromaticism and tonal shifts in mm. 29-40. The pianist's right hand descends from B to G♯ through A♯ and A♮ in mm. 29-31. Liszt expands the chromatic movement in mm. 35-39 with a B-A♯-A-A♭-G descend, reinstating the bass note G in mm. 37-39 for a total of 3 times, utilizing CM, FM^9, and G^7 harmonies before a resolution to CM in m. 40. Vocalist's line in mm. 29-40 bifurcates into two phrases – mm. 29-34 and mm. 34-40. While both phrases initialize identically, melodic transformations lead to different tonal cadences. In mm. 29-34, Liszt begins and ends on B major, a distant key from the tonal area of A♭ major that predominated a large portion of tonality in the A-section. On the contrary, the harmonic objective in mm. 34-40 is to culminate towards C major. Figure 4 offers a Schenkerian graph representing mm. 29-40 and illustrating the cumulative nature of melody and harmony that Liszt presents in the score. In the vocalist's line, mm. 33-34 ends the phrase with a G♯-A♯-B-D♯-C♯-C♯-B-G♯-F♯ segment, defining the dominant tonic relation between F♯^{7} and BM harmonies. In mm. 39-40, the [B-C-E-E-D-C] segment outlines the G^7 and CM tonal areas as Liszt pursues his harmonic exploration.
Liszt continues to outline the hidden chromaticism starting from m. 40. The [E-D♯-D-C♯] segment in mm. 40-42 extends to [E-D♯-D-C♯-C] in mm. 46-48 with a repetitive G♯ that moves into the enharmonic A♭. Liszt experiments with multiple tonal regions in mm. 46-48 moving from EM to Fm through G♯^{7}, E♯dim^7, and C♯M^7 while building up the musical climax and extending the progression while delaying the cadence, as seen in Figure 5. The melody peaks at m. 56 with the note A – the highest pitch of the vocalist's line – gradually descending into the subsequent theme. The hidden chromaticism and the repetitive notes in the vocalist's part signify the energetic build-up and musical tension that requires a cadential resolution, which never occurs as Liszt proceeds into mm. 50-51 that include seventh harmonies in the pianist's accompaniment and a recitative for the vocalist.
The Schenkerian theory is prominent in understanding Liszt's compositional techniques to culminate *O Lieb*. Using chromatic ascends and descends in the piano without using the voice emphasizes Liszt's approach to pianistic details at the song's end. The inner voices are of particular interest in the pianist's right and left hands, generating motivic patterns that structure chromatism within tonal progressions in search of cadences, ending on diminished seventh chords. The [E♭-E-F-G♭] ascending pattern emerges in mm. 94-95 in the right hand's inner voice, while the [C-B♭-A-C♭] pattern descends in the left hand's inner voice, supported by the [D♭, C♭, A♭, F], a fully-diminished seventh chord. A similar occurrence exists in mm. 96-97 in the right hand, with the inner voice ascending via a [D♭-D-E-F♭-F-G♭] with the left hand descending via [B♭-A♭-G-A], finalized by the [C-A-G♭-E♭], a fully-diminished seventh chord. Figure 6 showcases the Schenkerian graph of mm. 93-102, outlining the chromatic interplay between the inner voices of the pianist's music.

5 Neo-Riemannian Analysis

Liszt was not unfamiliar with harmonic exploration and tonal manipulations in his works. The Neo-Riemannian analysis centers on three primary transformations: (1) *parallel*, appearing in the study as *P*; (2) *relative*, appearing in the study as *R*; (3) *leading tone*, appearing in the study as *L*. Integrating the three primary Neo-Riemannian transformations allowed for a thorough examination of Liszt's tonal regions and their comparison to the Schenkerian understanding of theory in *O Lieb*. Liszt's progressions push the musical boundaries through distant key relations intertwined with chromaticism and chordal complexity. *O Lieb* is in the key of A♭ major, starting and ending with A♭ as the tonic. The A♭ is not the sole prominent harmony in the work since other key areas appear in the music without any preparation. Liszt utilizes A♭ major in mm. 1-28, generating a small piano transition in mm. 27-28 while using an E♭^{7} harmony, illustrating the V, refusing the expected A♭ major key in m. 29 that would return to the tonic, and employing the BM instead. The triadic skeleton of E♭^{7}, the [E♭-G-B♭], shares two notes with the upcoming BM harmony, [B-D♯-F♯] – the B and with E♭ and D♯ being enharmonic. Such a compositional approach allows Liszt to swiftly move from A♭M, a key with four flats, towards BM, a key with five sharps, both being distant keys in relation to each other, separated by an augmented second. The *LR* transformation forwards A♭ major to E♭ major with *A♭M(L) = Cm*; *Cm(R) = E♭M*. The *LPLP* transformation occurs between E♭ major and B major, with *E♭M(L) = Gm*; *Gm(P) = GM*; *GM(L) = Bm*; *Bm(P) = BM*.

Liszt does not spend much time on B major, defining the musical emphasis within the text at the start of m. 29. Liszt continues onto the C major – another primary key in *O Lieb*, transiting to C major through a V-I harmonic movement in mm. 39-40, defined by G^7 to CM harmonies, incorporating an *RL* transformation between GM and CM. The key change does not result in the key signature change since Liszt sustains five sharps in the key signature. The composer never formally returns to B major as the newly established tonic, modulating through various harmonies as the song progresses. One of such harmonies occurs at m. 44, where Liszt designates E major as the new primary tonal area, as seen in E major bases in m. 49-50. The next tonal area that Liszt explores is F minor in the piano transition in m. 54 and F major in m. 56, which contains the highest pitch in the vocalist's line. Liszt makes it a commonality in *O Lieb* to navigate his harmonies swiftly, not spending much time on any particular key. The key areas of A♭M, BM, CM, EM, and FM are prominent upon which Liszt builds his progressions. The Neo-Riemannian key exchange among central regions, as utilized by Liszt in *O Lieb*, shows a utilized chordal alternation, generating a Neo-Riemannian network of keys throughout the musical passage.

The Neo-Riemannian perspective centers not on integrating Liszt's harmonies into a sole tonic. The Neo-Riemannian transformations extend the Schenkerian understanding to view *O Lieb* as a set of transformative harmonies that function without referencing the tonic. The tonal area of A♭ major in m. 14 presents an example of transformative harmonies Liszt adventures through, moving towards E♭ major in m. 22. The simplified I-to-V relation carries a strong transformative character, experiencing multiple key areas, explained by the Neo-Riemannian theory. A♭ major in m. 14, D♭ minor in m. 15, and A♭ major in m. 16 are defined through an *RLP* transformation between A♭ major and D♭ minor and the *PLR* transformation, its reverse, between D♭ minor and A♭ major. A more tonal approach is taken in the next segment, with A♭ major in m. 16 moving to F minor in m. 17 towards C major in m. 18, producing an *R* in mm. 16-17 and *PLR* in mm. 17-18. The proceeding transformation between mm. 18-20 includes C major, A minor, and E major. The final Neo-Riemannian component includes mm. 20-22, incorporating E major, C♯ minor that is enharmonic to D♭ minor, and E♭ major, utilizing *R* in mm. 20-21 between E major and C♯ minor and *RPRPRPR* between D♭ minor and E♭ major in mm. 21-22. The transformation of C♯ minor to D♭ minor in m. 21 is of analytical significance due to Liszt's desire to return to E♭, hinting at the half cadence in relation to the A♭ tonal area initiated in m. 14. The two outside harmonies of A♭ major and E♭ major resemble a tonic-dominant relation, containing harmonic complexity within its borders.

A prominent harmonic cycle occurs in mm. 40-44, imitating the previous occurrence, initializing with C major and ending with E major, the two chords sharing an *LP* transformation. The C major in m. 40 leads to F minor in m. 41 and returns to C major in m. 42, creating an *RLP* Neo-Riemannian transformation, defined through an I-iv-I harmonic relation. The F major in m. 42 transfers to A minor in m. 43, resolving to E major in m. 44, defined through an *L* transformation between F major and A minor in mm. 42-43 and an *RLP* transformation between A minor and E major in mm. 43-44. Such a chordal movement generates a more extended Neo-Riemannian structure within a simplified form of C major and E major key regions of m. 40 and m. 44. The harmonic transformations in mm. 50-52 solely depend on the vocalist's G♯ and A♭, as Liszt transforms E major to F minor, sustaining the third scale degree of one chordal structure in m. 50 and utilizing its enharmonic equivalent as the third scale degree of the second chordal structure in m. 52. The progression outlines, driven by E major and F minor, contain a series of seventh chords, moving in a chromatic pattern, as seen in the bass of the pianist's left hand from E to C. The chords under analysis are EM, G♯^{7}, Edim^7, C♯^{7}, and F minor.

6 Conclusion

The unity of Schenkerian and Neo-Riemannian analyses helped develop a holistic understanding of Liszt's harmonic goals in *O Lieb*. Both theories help uncover the underlying tonal structures within the song's voice-leading complexities. The combination of Schenker's and Neo-Riemannian theories helps enhance the analytical side of *O Lieb*, assist in understanding Liszt's compositional language in the vocal repertoire, and uncover the unit between the music and text. This combination of analytical approaches and their construal within the pianism and vocalism of *O Lieb* provide a multifaceted view of Liszt's compositional methods and reveal the song's depth and harmonic complexity.

References

[1] C. M. Morrow, *Franz Liszt's Life and Music: A Dramatic-Monodrama Piano Recital*. PhD Dissertation, New York University, 1993.

[2] J. Chang, *The Bridge to Modernism: Franz Liszt and the Late Piano Music*. PhD Dissertation, Indiana University, 2021.

[3] J. F. Penrose, "The Piano Transcriptions of Franz Liszt," *The American Scholar*, vol. 64, no. 2, pp. 272-276, Spring 1995. https://www.jstor.org/stable/41212325.

[4] F. Liszt, *Favorite Songs by Franz Liszt*. New York: G. Schirmer, 1888.

[5] D. Aldach, *Liszt's O lieb' so lang du lieben kannst: The Song That Inspired Liebesträume and the Composer's Programmatic Literature*. Southern Illinois University, 2012.

[6] H. Cho, "Music Analysis as Poetry," *Perspectives of New Music*, vol. 53, no. 1, pp. 143-187, Winter 2015. https://muse.jhu.edu/article/778165.

[7] C. Schachter, *The Art of Tonal Analysis: Twelve Lessons in Schenkerian Theory*. Oxford University Press, 2018.

[8] J. Koslovsky, "The Early Schenkerians and the "Concept of Tonality," *Gamut: Online Journal of the Music Theory Society of the Mid-Atlantic*, vol. 7, no.1, pp. 151-185, May 2014.

[9] E. Gollin and A. Rehding, *The Oxford Handbook of Neo-Riemannian Music Theories*. Oxford University Press, 2014.

[10] R. Cohn, "Schenker's Theory, Schenkerian Theory: Pure Unity or Constructive Conflict?" *Indiana Theory Review*, vol. 13, no. 1, pp. 1-19, Spring 1992. https://www.jstor.org/stable/24044686.

[11] D. Temperley, "Composition, Perception, and Schenkerian Theory," *Music Theory Spectrum*, vol. 33, no. 2, pp. 146-168, October 2011. https://doi.org/10.1525/mts.2011.33.2.146.

[12] A. Forte, “Schenker’s Conception of Musical Structure,” *Journal of Music Theory*, vol. 3, no. 1, pp. 1-30, April 1959. http://www.jstor.com/stable/842996.

[13] R. Cohn, “Introduction to Neo-Riemannian Theory: A Survey and a Historical Perspective,” *Journal of Music Theory*, vol. 42, no. 2, pp. 167-180, Autumn 1998. https://www.jstor.org/stable/843871.

[14] D. Harrison, “Review of Alexander Rehding, Hugo Riemann and the Birth of Modern Musical Thought,” *Music Theory Online*, vol. 11, no. 2, June 2005. https://mtosmt.org/issues/mto.05.11.2/mto.05.11.2.harrison.html.

[15] M. Siciliano, “Two Neo-Riemannian Analyses,” *College Music Symposium*, vol. 45, pp. 81-107, 2005. https://www.jstor.org/stable/40374521.

[16] J. Cormac, “Liszt, Language, and Identity,” *19th-Century Music*, vol. 36, no. 3, pp. 231-247, March 2013. https://doi.org/10.1525/ncm.2013.36.3.231.

[17] A. Walker, *Franz Liszt: The Virtuoso Years, 1811-1847*. Cornell University Press, 1988.

[18] A. Walker, *Franz Liszt: The Weimar Years, 1848-1861*. Cornell University Press, 1993.

[19] A. Walker, *Franz Liszt: The Final Years, 1861-1886*. Cornell University Press, 1997.

[20] D. Gooley, *The Virtuoso Liszt: New Perspectives in Music History and Criticism*. Cambridge University Press. 2009.

[21] O. Hilmes, *Franz Liszt: Musician, Celebrity, Superstar*. Yale University Press, 2016.

[22] R. L. Wells, “A Generalized Intervallic Approach to Metric Conflict in Liszt,” *Music Theory Online*, vol. 23, no. 4, December 2017. https://mtosmt.org/issues/mto.17.23.4/mto.17.23.4.wells.html.

[23] S. V. Moortele, “Form, Program, and Deformation in Liszt’s Hamlet,” *Dutch Journal of Music Theory*, vol. 11, no. 2, pp. 71-82, 2006.

[24] M. Vitalino, *Franz Liszt's Song Revisions: A Schenkerian Taxonomy*. PhD Dissertation, University of California, Santa Barbara, 2014.

[25] A. J. Kuczynski, *From Diminution to Autonomous Themes in Schenkerian Theory: Bridging the Divide between Tonally Organic and Modern Thematic Music*. PhD Dissertation, University of Rochester: Eastman School of Music, 2014.

[26] D. A. Damschroder, *The Structural Foundations of “The Music of the Future”: A Schenkerian Study of Liszt's Weimar Repertoire*. PhD Dissertation, Yale University, 1981.

[27] B. Chapkanov, *Harmony and Tonality in the Late Piano Music of Franz Liszt – Functional and Transformational Analytical Perspectives*. PhD Dissertation, City University of London, 2022.

[28] R. Bass, H. Savage, and G. Patricia, “Harmonic Text-Painting in Franz Liszt’s Lieder,” *Gamut: Online Journal of the Music Theory Society of the Mid-Atlantic*, vol. 6, no. 1, October 2013, https://trace.tennessee.edu/gamut/vol6/iss1/2.

Appendix

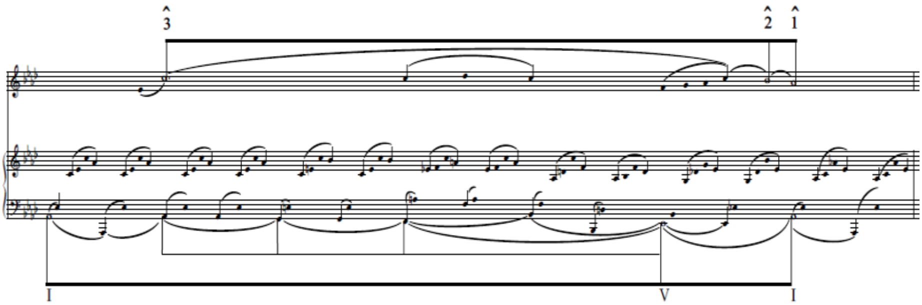



Fig. 1. Schenkerian analysis of Liszt's *O Lieb*, mm. 1-8.

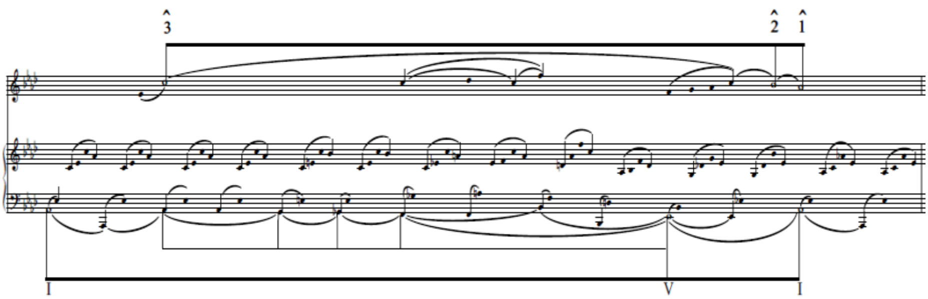


Fig. 2. Schenkerian analysis of Liszt's *O Lieb*, mm. 9-14.

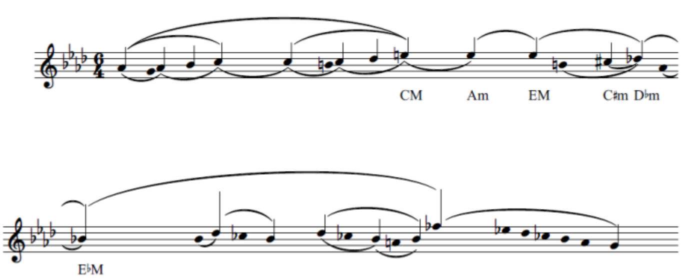


Fig. 3. Schenkerian analysis of Liszt's *O Lieb*, vocalist's melody, mm. 15-27.

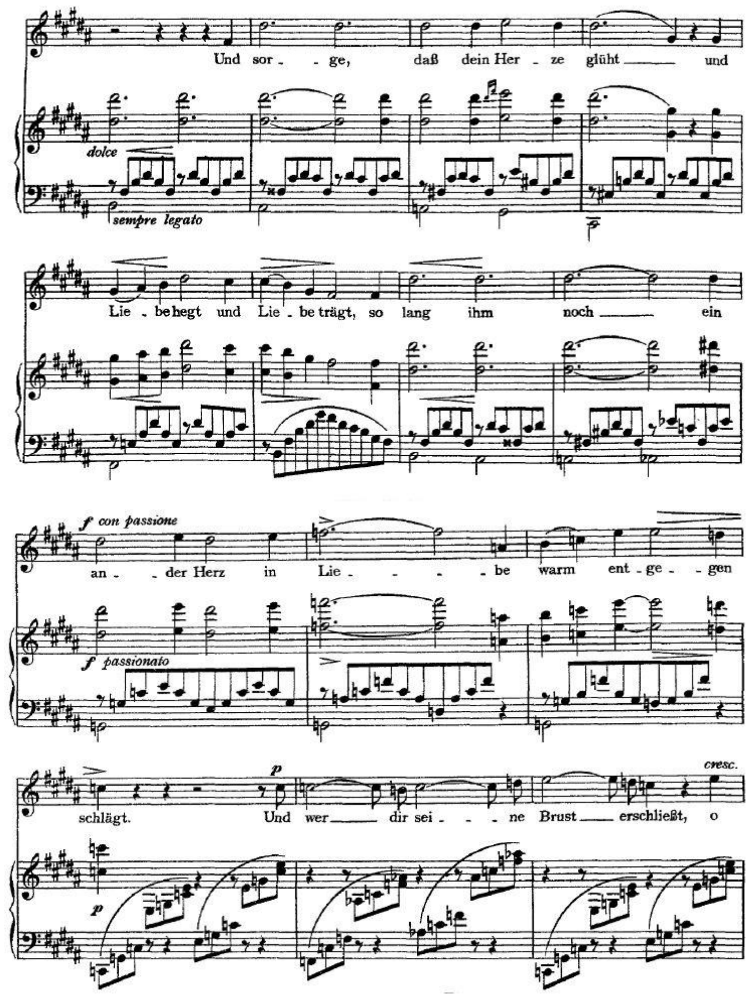



Fig. 4. Liszt's *O Lieb*, mm. 29-40.

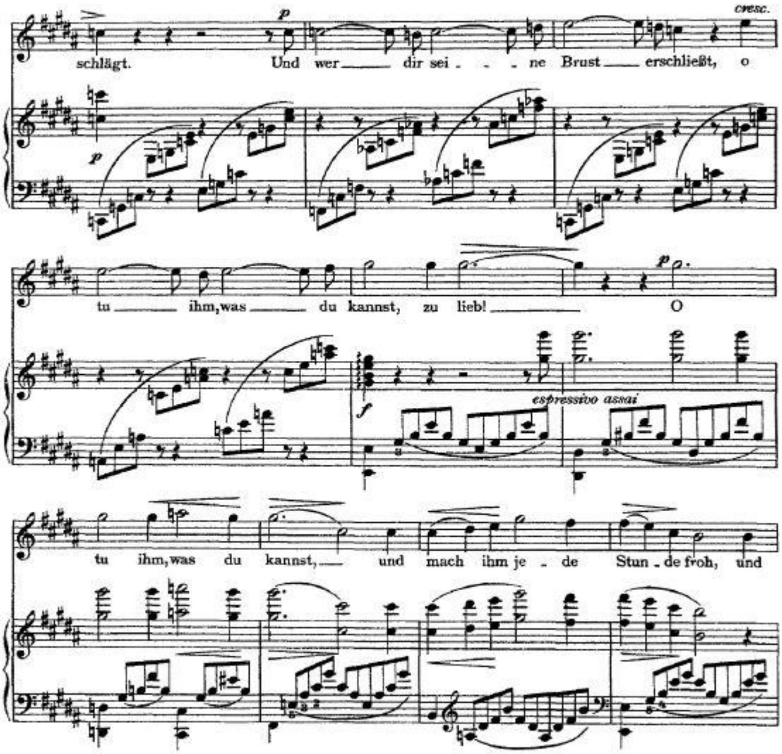



Fig. 5. Liszt's *O Lieb*, mm. 40-49.

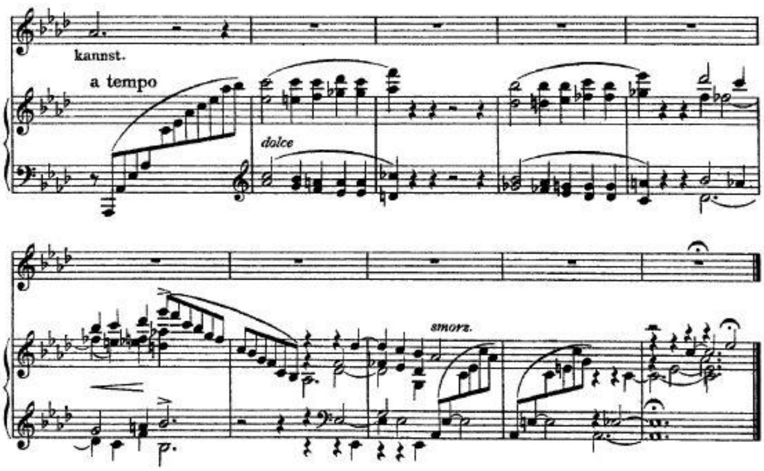



Fig. 6. Liszt's *O Lieb*, mm. 93-102.

Evaluating Music Repertoire Used in the Music Analysis Textbook to Teach AP Music Theory

Nikita Mamedov and Jihong Cai

North America International School, Shanghai, China
nais-music@north-americais.cn; nmamedovmusic@aol.com;

Abstract. Music theory – an indispensable facet of music education and a broader component of comprehensive pedagogy for K-12 students – intertwines music's analytical and creative dimensions. The AP Music Theory course, equivalent to a university-level course, is offered to high school juniors and seniors, serving as the initial foray into advanced music theory studies, provides a balanced and comprehensive representation of the subject content. Among the several eminent textbooks utilized in standard AP Music Theory courses, *Music in Theory and Practice* is the one this study analyzed, delving into an in-depth analysis of the repertoire referenced in the text, with a particular focus on the composers, styles, and genres (Benward & Saker, 2009). The study examined these variables in relation to the book's overarching design and structural framework. The study elucidated the implications of repertoire diversification from the observations and comparative analyses. The study likewise proposed a set of cogent recommendations to enhance music theory's pedagogical efficacy in repertoire and its appearance throughout the text.

Keywords. AP Music Theory, music analysis, music education, music repertoire.

1 Introduction

Music education is significant in the holistic development of artistry and creativity in K-12 education. The value of music pedagogy helps integrate students' performative and academic pursuits as they combine systematized learning paths with musical creativity, developing artistic links through the study of instruments, history, and theory [1]. Music education consolidates various musical abilities and helps students develop a comprehensive and well-balanced approach to musicianship. Music coursework is manifold, encompassing various musical disciplines, including the most common studies in the K-12 curriculum of music appreciation, history, theory, technology, and instrumental classes, equipping students with all-rounded knowledge. Music coursework bifurcates into performative and academic branches, integrating different facets of music into the curriculum and instruction. Both performative and academic coursework disperse students' knowledge of the musical repertoire. The historical interpretation and awareness of the repertoire are vital for all-inclusive and robust music education programs engraved into lifelong and transformative music education processes [2].

Music theory is an academic subject allowing music programs to utilize music analysis to develop students' analytical skills. The course focuses on integrating the performative comprehension of music and teaching students how music works. Music analysis is a fundamental approach in music education, necessary to help students develop and maximize their musicianship levels. Music theory is taught at all levels and is the most common academic course in the music curriculum. The course should offer a neutral approach to music repertoire, emphasizing works that help best outline the analytical concepts being taught [3]. Music theory and analysis studies offer the most academically rigorous learning content, influencing the systematic side of schools' music curricula. AP Music Theory is the highest music theory course offered in the high school curriculum, is university-equivalent, and delivers students an opportunity to study music theory at the higher education level.

The standardized nature of the AP Music Theory coursework helps unify the curriculum to provide a consistent and thorough understanding of analysis concepts, ensuring that all students – regardless of their prior knowledge or skill level – can retain the content and prepare for the proceeding examination. Focusing on music theory rudiments introduces the analytical techniques students should know to define, comprehend, and analyze music. The AP Music Theory learning content is practical and applicable to what students play and listen to daily. The music theory textbooks' repertoire should continue to develop and expand as new music enters analytical discussions in music education circles [4]. Music theory would lack meaning without appropriate applicability, necessitating the instructors to apply the music theory components to appropriate musical selections continually. The curriculum's musical repertoire is pertinent to how music theory affects the learning process and students' motivation. The musical repertoire is integral to ensuring that the systematized learning offered by the AP Music Theory course is meaningful to curriculum and instruction. Students are more motivated to study music they relate to and works that pertain to their current societal interests [5]. A broad musical repertoire helps develop music theory as a source of knowledge while developing a rich teaching tradition for music theory teachers [6].

2 Research Emphases & Framework

Various textbooks exist to help students cover the AP Music Theory coursework and prepare for the AP exam. The study evaluated the musical selections found in the *Music in Theory and Practice* textbook, written by Bruce Benward and Marilyn Saker, designed for high school and university coursework in music theory [7]. The study examined the textbook's repertoire, emphasizing composers, styles, and genres. Musical examples are vital to students' applicability, and the diversification of musical repertoire that aligns with music theory concepts is necessary for a meaningful comprehension of the subject content [8]. The study focused on two components related to music theory pedagogy in alignment with repertoire diversification and students' motivation. The first is whether the musical repertoire in the textbook is diversified to include and apply to various genres and styles across music history. The focus is on repertoire diversification and how much it applies to music theory pedagogy [9]. The second is whether the current repertoire is meaningful and applicable to students' learning depending on students' musical interests and performative pursuits engraved in students' motivation. The study attempted to answer whether the *Music in Theory and Practice* textbook reflected the current changes in music theory pedagogy and which improvements to repertoire selections can help facilitate a more robust music curriculum that helps develop students' knowledge of musical repertoire through music theory.

Recent changes in music theory pedagogy philosophy align with repertoire and how music theory sources utilize musical examples in music analysis [10]. Studying a variety of musical compositions in music theory classes can enhance students' understanding of music theory and composition and how both subjects align within the music education network. Students can better understand different musical forms, structures, and techniques by focusing on a larger sample of musical examples. The repertoire diversification is vital to students' success in music theory for two reasons.

First is the practical component that encircles music repertoire and students' accessibility to musical compositions. Students are more likely to study music that is more meaningful to what they sing, play, and listen to in their music courses [11]. Musical examples that resonate with students' current experiences can offer a more productive learning experience and help students find individual perspectives in pursuing music theory. More meaningful music produces more robust teaching opportunities for music theory instructors [12]. Repertoire to which students relate helps them find meaning in their analytical objectives. Appropriate repertoire likewise makes the learning more suitable when covering the concepts of harmony, melody, and counterpoint.

Second is the more inclusive perspective of music composition and its relation to music history. Promoting a variety of musical examples helps students learn about works that influence a wide range of musical styles [13]. Inclusive music composition in music education is a critical approach to extending traditional music boundaries with a focus on acknowledging rich musical heritage and many musical traditions, all necessary to understand the historical retrospective of how composers influence music history. Incorporating more musical examples in theory texts would help instructors expand the boundaries of music composition and the recognition of the subject within musicological studies.

Recent studies on music theory pedagogy identified three ways of repertoire diversification that book authors can utilize to expand their musical examples. The first approach is implementing more non-Western music, paving the way for alternate music analyses. Using world music and national music from non-Western countries and societies is one method of how curricularists can increase repertoire expansion across textbooks [14]. Employing examples from popular music repertoire is another approach to expanding the musical repertoire and offering students opportunities to analyze modern twenty-first-century non-classical works. Pragmatic pertinence of musical works and compositional techniques with appropriate curricular material is vital to students' retention of musical theories and their applicability in future studies of music analysis.

The second approach is to implement non-tonal music in analysis courses. Post-tonal music helps students understand non-tonal relationships and focus on post-tonal musical techniques and trends [15]. Such repertoire helps students understand the music of composers outside of the tonal spectrum while focusing on their style and historical significance. Post-tonal music allows instructors to develop students' creativity without emphasis on the tonal center and expand students' musical expertise and vocabulary in avant-garde styles. Post-tonal repertoire likewise offers innovative teaching approaches that instructors can utilize to help students retain repertoire and music theory concepts in the classroom [16].

The third approach is to implement underrepresented composers and genres, ensuring students have an equal opportunity to study works not commonly found in general music curricula [17]. Underrepresented repertoire allows students to appreciate music that has yet to flourish in the historical narrative while learning about alternative musical influences. Including novel works would add to musical innovations that music theory outlines, offering students novel perspectives and ideas in the analysis-centric course.

The repertoire diversification and its analysis highlight the prominence of student-centered teaching in music theory education [18]. Students are more likely to focus on learning that they find meaningful [19]. Students are more likely to be aware of the course content and find motivation to pursue music theory with a focus on repertoire with which they relate. Diversifying musical repertoire will enable students to cover different musical compositions across historical epochs. A diversified repertoire is more meaningful to students' learning and more impactful on students' music theory education. Diversified musical repertoire, used as examples in music theory classrooms, would help music instructors generate a more rounded music theory curriculum. The study evaluated the repertoire used in the *Music in Theory and Practice* textbook, concentrating on the book's strengths and weaknesses. The study likewise offered recommendations on closing out the missing gaps in the pursuit of repertoire diversification in the music theory textbook.

3 Literature Review

The research literature prominent to the study involves operating music repertoire in music theory pedagogy and its cross-relation with contextual educational issues, AP Music Theory teaching practices, and music theory and motivation. The utilization of music repertoire drives the curriculum and instruction in music theory education. The AP curriculum is the most extensive program of music study, offering the most substantial opportunity to include various musical compositions in the text. Curricular success in implementing and developing musical works within music theory education is beneficial for students and their motivation capacities to pursue academic music studies.

Music theory pedagogy is a medium between traditional music education and curricular expansion, which is possible through music theory studies [20]. The emphasis on obstacles and educational complexities that intertwine the world of curriculum and instruction and music theory learning is vital to ensure robust instruction in analysis-centric music. The focus on the digital approach to understanding and teaching music theory through an abstract rather than standardized approach helped conclude that an extended and developed curriculum is necessary in music theory to achieve a transformative educational process. Transformative music education is at the core of contemporary musical issues, necessitating resolution through diverse musical repertoire through texts and learning materials.

Music theory's interdisciplinary nature allows the subject to integrate with mathematics, creating prominent connections between both studies concerning music analysis and its meaning [21]. Connections between music theory and mathematics offer novel perspectives on musical conceptualization, allowing for the employment of objective references to define artistic subjectivity. Interpreting computational approaches in music theory and the place such approaches hold in the music theory curriculum is essential to the academic nature of music theory and the content it strives to teach. Mathematical applications and interdisciplinary relations between music and mathematics help students improve their music theory expertise. Various educational components in music theory and mathematics multidisciplinary fields can increase students' analytical expertise.

Another contemporary music education topic relevant to expanding musical repertoire develops out of focus on understanding the links between diversity and creativity in music theory pedagogy [22]. The systematized academic side of music theory influences the performative artistic component of music analysis seen through music repertoire and what kind of music the theoretical perspective aids in explaining. The focus is on expanding the current pedagogical practices and expertise and understanding the creative spectrum students can achieve through the systematized study of music theory. Concentrating on teaching practices can help enhance the music theory classroom and make learning more meaningful within the subject's objectives and course goals.

The discussions on the place of music theory in the music education curriculum allow a scholarly comparison of students' analytical expertise at various levels of education [23]. Music theory training is necessary, and more course development is vital for students' holistic understanding of music and its meaning. Many students graduate from various institutions - both at high school and university levels - showcasing inefficient theoretical preparation, making it more challenging to interpret music at the following levels of their education and careers. Music theory pedagogy is vital for all-inclusive musical development and successful academic comprehension of music making. Theory and analysis pedagogy focuses on the familiarity, experience, and rendition of music, which would increase in a student-centered learning environment by expanding the musical repertoire in academic music courses.

The focus on music theory pedagogy philosophy is prevalent in current educational matters and helps locate ways strategic music analysis can enhance the music theory curriculum [24]. The emphasis on methodological issues in the music theory curriculum and interactive learning approaches can help increase the learning and student proficiency in the subject. The discussions on behalf of the technical complexities of music theory content and the necessity to outline music theory knowledge as an academic subject are vital in comparing the learning results to performative music curricula. The concentration on fostering the theoretical understanding of music analysis and its representation in the subject matter leads to the necessity of a broader range of musical repertoire to appear in the textbooks.

The extension of music theory pedagogy processes and its history improves the contemporary learning methods and issues music theory instructors face in modern classrooms [25]. The historical comprehension of music theory teaching helps instructors understand the challenges students face when retaining music theory concepts. Music theory evolution and its pathway are prominent in the analysis-centric representation of performative music components in analysis and theory. Music theory aims to explain how music works and interpret it from cultural and artistic perspectives. Music theory can generate unexplored characteristics of its instructional capacity through expanded musical repertoire. The fundamental theory concepts within the historical context emphasize the pedagogical implementation of the learning content in the music theory classroom.

The music theory curriculum can lead toward a multicultural music education, ensuring that analysis and its purpose align with various musical disciplines [26]. Music theory is an applicable subject, standardized and mandated in high school and university learning. Music theory is likewise interpretational since much of the curriculum focuses on Western art music. Developing repertoire in Western and non-Western musical directions aligns music theory with other prominent music subjects, such as musicology and ethnomusicology. Questioning, examining, and evaluating various pathways for the music theory curriculum helps conceptualize and realize the course's goals from a broader performative music perspective. Such a procedure implicates the necessity to integrate musical repertoire into a comprehensive understanding of the relationship between theory and practice and how theory impacts the development of compositional, creative, and interpretational analyses.

Music theory development spans beyond music studies, including integrating academic music with cognitive sciences [27]. The interdisciplinary studies of music theory and cognition enhance and enrich music education and offer new perspectives and methodologies for teaching and learning music. Pedagogical practices in music theory often extend into performative and academic music curricula. The learning structures music theory as an influential course that strongly impacts the direction of music education. Alternative collaborative opportunities for music theory are essential to understanding its academic nature. Empirical assessment and implications for cognitive development in curriculum and instruction can serve various interdisciplinary methodologies where music theory can thrive.

Scholarly literature on AP Music Theory helps align educational theory and practice to enhance the instructional approaches to students' academic results [28]. Learning gaps between high school music theory and the AP Music Theory coursework develop varied expectations and students' educational outcomes. Integrating learning content and its extent affects students' preparation for the final AP examination. The pacing strategies instructors should implement when teaching the AP curriculum are vital to explore when determining the strategic planning and movement among various course modules. Motivation and academic results are crucial strategies to ensure students have a positive and applicable learning experience.

Examining musical experiences from a student-centered perspective helps shed light on learners' perceptions and awareness of the course content [29]. The focus on evaluating skills and learning patterns in the course and how students' musical backgrounds affected their learning path are essential for examining student-centered music theory experiences. Self-efficacy on behalf of students' perspective is prominent in academic music classes, helping seek answers to what the AP Music Theory instructors can do for students regarding instructional design to improve overall motivation and music theory learning performance. Students' achievement and learning experiences are at the core of the educational framework to examine students' stances and abilities to tackle advanced university-level learning content at the high school level.

The aural skills components in the AP Music Theory classroom and students' abilities to meet expectations in the performative elements of the analysis-centric course are vital in discussing AP-centric repertoire selection [30]. Evaluating music improvisation, sight-singing, and sight-reading learning from the current musical repertoire and focusing on approaches to instructing aural skills help students integrate the retained content with music analysis. Various practical techniques can help students improve their expertise in aural skills while preparing to enhance their overall performative capacities in the music theory curriculum. Employing a more comprehensive range of musical selections in the aural skills section can increase interest levels for students studying music theory at the AP level.

Melodic dictations are other components of the music theory curriculum and hold much prominence in the AP studies of academic music [31]. The AP coursework emphasizes melody and its theoretical and performative components. Evaluating teaching approaches in interpreting and instructing melody and melodic dictation can help improve students' expertise in the AP course. The focus on teachers' perception of what students need to improve their understanding of theoretical and performative components related to the study of melody can help increase academic productivity and results in the AP Music Theory course. The psychological and motivational barriers in curriculum and instruction challenge students and their learning processes and require practical answers.

Motivation is a prominent component in music theory education research due to the subject’s systematized learning and lack of performative opportunities compared to performative coursework. Students invest time and commitment to studying music theory, and the lack of results would demotivate their studies and course perceptions [32]. Motivational capacities help offset instructional challenges when studying AP Music Theory and preparing for the course’s external examination. Various motivational theories can help measure students’ motivational capacities in the classroom. Self-determination theory is one such framework that allows understanding students’ perspectives towards the coursework and its content, focusing on students’ competence and interest levels. Scholarly literature on various motivators that impact students’ learning can help increase students’ academic results and help them comprehend the motivational spectrum within the AP course.

The bifurcation of motivation into intrinsic and extrinsic can help researchers understand students' interests and their lack when pursuing music theory learning [33]. Studying musical repertoire can help increase motivation and how students perceive the music analysis content. The various musical narratives help identify the levels of autonomy that students need for adequate motivational pushes to pursue music studies. The focus on pedagogical characteristics that can help increase motivational capacities and students' musical results in the classroom can help offset students' motivational challenges in the music classroom. Understanding students' perceptions of the textbook's musical repertoire can positively affect their learning habits.

Suitable repertoire selections and fitting musical examples in the textbook help introduce music theory concepts in the classroom and can impact students' motivation to study theory [34]. The advocacy for expanding repertoire that scopes outside of the Western classical canon is vital for a well-rounded music theory education. Concentrating on the teachers' perception and ensuring that students can familiarize themselves with different works from various styles and epochs are likewise prominent in students' success in the AP Music Theory classroom. The repertoire choices instructors make in music theory classes can affect motivation and students' perception of music theory and its representation.

Musical selections that schools introduce in music classes affect students and the learning paths that the music repertoire develops [35]. No systematized curriculum mandates particular music repertoires to appear in the AP Music Theory curriculum, meaning that textbook authors of the AP teaching material can choose music that best outlines an example of a particular music theory concept. The ability to diversify the musical repertoire can lead to a more robust music curriculum that focuses on various styles and musical forms. Musical diversification can help develop students' cultural heritage and appreciation for different musical societies with a focus on a comprehensive outlook of varied musical works within music history.

4 AP Music Theory Curriculum

AP Music Theory is an introductory university course that familiarizes students with the rudiments of music theory. The course is practical and applies to music performance, where students can utilize theory concepts while focusing on performative elements of ear training, sight reading, sight singing, dictations, and interval and chord recognitions. Students learn about musical concepts, covering the music theory concepts of pitch, rhythm, form, analysis, and counterpoint. Students focus on musical and analysis notation, learning to understand theoretical techniques that help define and evaluate music. The AP Music Theory course consists of eight modules, often completed in one-to-two semesters at the high school level. The AP Music Theory exam contains multiple choice and free response sections. Mastering the course allows students to evaluate their knowledge of AP Music Theory through the final exam. The cumulative nature of the final exam is not the ultimate knowledge point in the music theory curriculum. Instead, the AP Music Theory course is an initial course that helps with knowledge of more advanced music theory topics. Students can pursue further music theory courses at the university level. Students can likewise continue learning other more content-specific theory courses focusing on Schenkerian analysis, Neo-Riemannian theory, and atonal theory.

The eight units in the AP Music Theory curriculum combine to create the college experience for students learning about music analysis. The first unit includes content on music fundamentals, as students work through pitches, notation, rhythms, meter, intervals, and scales. The first unit offers the fundamental building blocks students should understand that apply to analysis in future units. The second unit develops music theory content focusing on different scalar forms and key signatures, emphasizing melody and its analysis. Students learn about timbre and texture. Students learn about various melodic features and how to analyze them. Students dive deeper into meter, rhythm, and syncopation, setting the stage for more advanced rhythmic analysis. The third unit in the AP Music Theory course focuses on harmonic understanding of music analysis. Students learn about triads and seventh chords and pitch relations within chordal structures. Students learn about different musical inversions, chord types, and qualities of seventh chords, setting up for the proceeding unit on harmony and analysis. The fourth unit helps students understand harmonic analysis by exploring voice leading, contrapuntal writing, four-part harmony, and harmonic progressions. Students learn to voice-lead various chordal structures in relation to tonic. Students also learn about cadences and how tonally-minded composers resolve cadences in their repertoire. The fifth unit developed more knowledge on voice leading and harmonic analysis. Students learn about harmonic functions and purposes that chords have within progression. Students also learn about music analysis and part-writing and how various voices and tones are observed and resolved according to theoretical rules. The sixth unit moves towards more advanced concepts in harmony and voice, leading with students covering motivic development, musical transformations, analysis-centric embellishments, and sequences. Students can analyze full-scale compositions and interpret and comprehend the context of music analysis and its applicable nature to music performance. The seventh module focuses on understanding the relations between musical keys and scale degrees within the context of Roman numeral analysis. Students likewise learn about modulation, tonicization, and its use in practical music theory. The eighth module introduces modes, phrasal relationships, and musical form, learning to understand music from alternative theoretical perspectives.

Music in Theory and Practice by Benward and Saker contains 17 chapters, covering the eight modules required to prepare learners for the AP Music Theory exam. The first chapter focuses on notation, covering the topics of musical staff, notes, rhythms, articulations, and dynamics. The book likewise includes a historical interpretation of how notation is prominent and evolving in music theory and develops opportunities in music performance and performance practice. The second chapter covers scales, keys, key signatures, and modes, focusing on melodic combinations of notes and how such components help set up the music theory curriculum to instruct analysis in further chapters. The chapter covers diatonic scales, tonalities, tonal regions, and relations between notes and scales. The third chapter covers intervals and transpositions, introducing students to a harmonic understanding of pitch-to-pitch relations and ways melodic and harmonic comprehension of music translates into theory and its practical perspectives. The fourth chapter discussed chords, including triads and seventh chords, as well as their theoretical positions. The chapter likewise introduced students to Roman numeral analysis, symbols, and symbolic representation used in music theory to analyze scores. The fifth chapter introduced structural elements in music, focusing on cadences, phrases, and nonharmonic tones. The chapter is the extension of Roman numeral analysis, helping students understand the performative characteristics of what music theory analyzes. The sixth chapter focuses on melodic organization, covering the concepts of motive, sequence, and phrases. The melodic distinction is prominent in music theory pedagogy to ensure students can analyze musical lines and separate them into motivic segments. The seventh chapter focused on texture and musical reduction, emphasizing the applicability of more advanced concepts that students can learn in the future, such as Schenkerian analysis. The eighth chapter focused on counterpoint and its influence on the music theory curriculum. The ninth chapter discussed voice leading and four-part harmony writing. Students learn about applications of voice leading to music analysis, and attempt to work with voice ranges while adhering to voice spacing rules. The tenth chapter focuses on harmonic progressions and their relation to Roman numeral analysis. Students learn about chord relations and their purpose within a musical harmony. The eleventh, twelfth, thirteenth, and fourteenth chapters focus on seventh chords and their place in music analysis. Particular emphasis exists on dominant seventh chords, leading-tone seventh chords, nondominant seventh chords, and secondary dominant chords. These chapters outline the advanced analytical techniques, preparing students to analyze any form of music. The fifteenth chapter focuses on modulation, developing students' understanding of common-chord, chromatic, and phrasal ways of changing keys. Students also learn about modulatory construction and macro analysis. The sixteenth and seventeenth chapters focused on musical form and its relation to Roman numeral analysis, completing the fundamental knowledge of music theory concepts.

5 Research Method

This study examined the choices of repertoire cited in Benward's and Saker's *Music in Theory and Practice*. This research aimed to offer an apprehension of the authors' musical repertoire choices in the context of introductory music theory. The study utilized statistical methods to present data and evidence that provided grounds for recommendations and conclusions. The study analyzed the repertoires concerning the book's structure and organization, providing insight into the musical and curricular design the authors implement, which they believe will deepen students' music appreciation, prepare students for further music studies, and potentially enroll in sequential courses to study more advanced music theory. A discussion on the reasons behind these choices and analysis into the potential benefits and issues such collection brings followed that shed light on the curricular implications of the repertoire selections.

The following terminology helped explain the definition and variables used in the study. The textbook consists of chapters and exercises that utilize musical examples. The study focused on the repertoire distribution, leading to the materials discussed in the book and practice problems receiving separate discussions. The study employed the word *assignment*, which coincided with the term the author used in the book's content. The term *text* represented the entire chapter without the assignment sections. The symbol (*T*) referred solely to text, the symbol (*A*) referred to assignments, and (*A/T*) defined the combination of the chapter's text and assignment. The study likewise analyzed the composer's birth and death dates for music used in the textbook. The focus was on composers' dates rather than when the pieces were composed, allowing for a thorough evaluation of the stylistic nature of musical works the textbook authors selected. In the rare cases when some composers did not have an exact record of their birth or death date, estimates found in music history literature were applied to the data. This study categorized all music's stylistic features into classical, popular, and *N/A* when the composer is unknown in the textbook. When a piece had multiple composers, the study listed all composers in the composer category without treating the musical selection for numerous occurrences. The dates for such cases included the range of the earliest birth to the latest death. If a piece appeared multiple times in the textbook, the study listed such work once per appearance in the table. The study helped evaluate the data based on the composer's name and dates, page and chapter of occurrences, their stylistic categorization, and the functionality in the text.

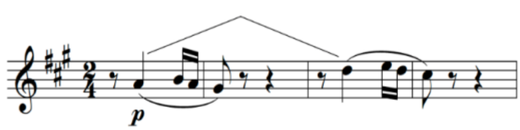



Fig. 1. Data collected for Beethoven's String Quartet in C♯ Minor, op. 131, movement IV, mm. 1–4.

Figure 1 is an example of music cited on page 14 of the textbook. The composer is Ludwig van Beethoven, and the piece's name is String Quartet in C♯ Minor, op. 131. Beethoven lived from 1770 to 1827. The following work appeared in the first chapter of the textbook titled Notation. This piece appears in the text and not the assignment section. It is worth noting that if the textbook mentioned the music's name but did not use the sheet music as an example, such a work did not count towards the study's collection. The textbook's authors utilized 328 pieces throughout the book as part of analysis examples, of which 327 appeared in the text and 1 in the book's appendix.

The repertoire data helped with evaluations and observations of general statistical information the textbook authors used to demonstrate the musical materials. The textbook is 371 pages long, including 327 music citations as theoretical examples. Several works appeared multiple times; the most frequent work appeared four times throughout the textbook. One composition appeared in the appendix in *Summary of Part-Writing Practices*. Johann Sebastian Bach's compositions appear extensively in the text, constituting around one-third of all musical examples. A substantial number of composers only appeared in the textbook once. Popular music appeared infrequently in the textbook in comparison to classical repertoire.

This study presented the total number of pieces, number of composers, distribution of music citations among composers, the data of which all composers lived, distribution of pieces among stylistic categorizations, and density of musical citations in text. This study discussed the musical and educational consequences, such as if students receive a well-balanced representation of all musical repertoire. For example, Bach's works and the authors' emphasis on his compositions attributed to the composer's prominent contribution to musical literature and music theory. This study helped analyze the justifications and potential risks of the collection and representation and other questions for discussion that naturally follow based on the observations.

The study examined the links between the repertoire and other factors, such as the textbook's structure and organization of variables related to musical repertoire. The study focused on six binary comparisons to the repertoire. These are composers versus chapter in *T* and *A/T*, composer versus *A/T*, composer's dates versus chapter, style versus chapter in *T* and *A/T*, style versus *A/T*, and chapter versus *A/T*.

The study emphasized the distinctions between *T* and *A/T* made in comparing composers, chapters, and styles, representing the structural functionality of the piece. The assignments are formative assessments, and the texts are descriptive materials of the music theory concepts. Therefore, the music utilized by the textbook authors in the *T* is for discussing the theoretical application or to see examples of the learning components in music theory practice. At the same time, the *A* serves as a tool for reinforcement, where students apply the learned material in the *T* to analyze music on themselves. The appearance of music examples in the text is denser when compared to musical appearances in the assignments.

6 Findings

According to the musical example index in the book's appendix, the textbook identified and adjusted 238 compositions using the collection criteria above. Among these, 167 pieces (50.91%) appeared in assignments, and 160 (48.78%) appeared in the text. Classical music appeared 279 times, which is 85.06% of all occurrences. Popular music appeared 20 times, which is 6.10% of all pieces. The textbook contained 29 pieces that do not have a composer or, if composers were not specified, identified as *N/A*, which is 8.84% of all music. This research analyzed music in the text (*T*) and assignments (*A*) so that a perspective that aligns the choices of repertoire with their functions in the book. The textbook's 166 pieces appeared in the assignments section at the end of each chapter, while 161 pieces appeared in the text for explanatory purposes. Wolfgang Amadeus Mozart's Fantasia in C Minor, K. 475, is cited in Appendix A as a *Summary of Part-Writing Practices*. The 328 pieces of music in the textbook belong to 84 composers from the 15th century until the 20th century. An additional 29 pieces do not have a composer, which is 8.84% of all musical works the authors used in the textbook. The compositions of Johann Sebastian Bach appeared 105 times, followed by 23 appearances for music by Franz Joseph Haydn. None of the other composers appeared more than twenty times throughout the textbook. There are 11 (10+1) composers whose music contributed to more than 1.00% of the overall collection. In addition to Bach and Haydn, and the non-attributed ones, the music of Wolfgang Amadeus Mozart, Ludwig van Beethoven, Frederic Chopin, Franz Schubert, Robert Schumann, Claude Debussy, Felix Mendelssohn, George F. Handle appeared 19 times (5.79%), 15 times (4.57%), 10 times (3.05%), 9 times (2.74%), 8 times (2.44%), 7 times (2.13%), 6 times (1.83%), and 5 times (1.52%) respectively. The above composers constitute 71.95% of all music cited in the book, with 236 pieces in total. There are 60 composers whose music appeared once in the entire book. The 60 composers constituted 71.43% of all musicians cited in the book, and their music comprised 18.29% of all collections.

On average, the density of citations in the textbook is 0.88 pieces per page. The calculations do not include the appendix. Chapter 6 on melodic organization and Chapter 7 on texture and textual reduction are the most densely cited chapters, where they have 2.27 and 2.28 music per page on average, respectively. Other densely citing chapters include chapter 12 on the leading-tone seventh chords, chapter 13 on nondominant seventh chords, and chapter 15 on modulation, which has 1.19, 1.06, and 1.14 pieces per page, respectively. It is worth noting that Chapter 8 utilized no examples of musical works to discuss counterpoint and contrapuntal compositional techniques.

Figure 2 shows the Gantt Graph, illustrating the date range of composers' birth and death dates. The representability of the musical examples showed the textbook's usage of composers across musical timelines. The following Gantt chart demonstrates the composers' lives. The graph included all identified composers except Bernart de Ventadorn, born in 1130 and died in 1190. Ventadorn's exclusion from the list occurred due to his dates being the outliers, generating a 250-year gap between the death of Ventadorn and the birth of the following composer, Josquin des Prez, who was born in

1440 and died in 1521. The data likewise demonstrated a gap in pre-Renaissance music, as only one piece appeared as an example of Medieval music, and no works appeared as examples of antique music. The data showcased that more diverse composers appeared starting from the second half of the Baroque period.

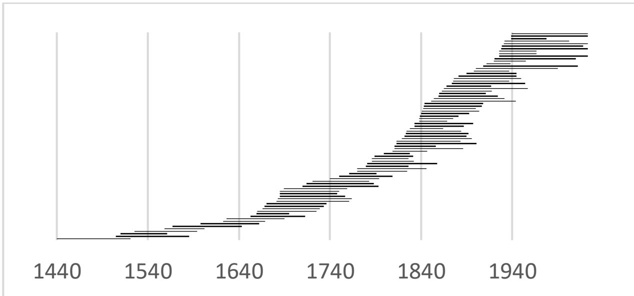


Fig. 2. Gantt chart representing dates when composers used in the textbook lived.

The estimation of the trend of composers who appeared in the textbook was approximated via the curve by the left endpoints, forming a shape for an exponential function, meaning that the textbook cites more composers from later musical epochs than pre-Baroque music. For example, in 1640, 4 composers were alive; in 1740, 10 were alive; in 1840, 18 were alive. Such data indicated a lack of appearances on behalf of the pre-Renaissance composers. For example, 10 composers were born before 1640, around 10% of all composers appearing in the textbook.

Very little musical material from contemporary and modern composers appeared in the book. For example, no music appeared from composers born after 1940, meaning that the youngest composer appearing in the textbook would be over 60 years old today. Another interpretation is that more emphasis is necessary on the more recent music repertoire in the textbook. The inability to close out repertoire gaps would lead to issues with the repertoire being relevant to music theory students in the modern twenty-first-century classroom.

The study concluded that there are 15 composers whose music appears more than twice in the textbook. These 15 composers contributed 248 pieces, 75.61% of all the music in the theory textbook. The study showed that 22 composers only had their music appear in assignments. Some composers had more than 50% of their music appear in the assignments. Most notably, 77.78% of Schubert's music appeared in the assignments, and 68.97% of all music composed anonymously appeared in assignments. Additionally, 66.67% of J. S. Bach's music appears in assignments, which includes 70 out of 105 compositions. Two-thirds of Scarlatti's music, 52.17% of Mendelssohn's music, and one out of two pieces from C. P. E. Bach, Brahms, Copland, Corelli, Sousa, and Walther's music are in assignments. On the other hand, all three examples from Purcell's compositions appeared in the text. Additionally, 38 composers have their sole appearance in the book's text. Another 15 composers had their music appear in the text for more than 50% of all the music cited. The only two composers with music cited more than 10 times in text are Haydn, with 11 occurrences (47.83%), and J. S. Bach, with 35 occurrences (33.33%).

Appendices 1 and 2 summarized the distribution of the most frequently cited composers and the distribution of their citations in each chapter. J. S. Bach has contributed with 105 of all pieces. Among these, 14 occurrences happened in Chapter 9; 12 occurrences happened in Chapter 5; 11 occurrences happened in Chapter 15; 10 occurrences happened in Chapter 7 and Chapter 14 each.

For all the anonymously composed works, 6 compositions appeared in Chapter 15, and 5 in Chapter 6 and Chapter 10 each. Haydn is the second most cited composer in the textbook. Haydn's 12 out of 23 pieces appeared in Chapter 6, of which half appeared in text. Haydn has the highest percentage of music cited in one chapter. For Haydn, 6 out of 11 appearances, 54.55%, appeared in the text of Chapter 6. Mendelsohn is the only other composer whose music appears more than half the time in a single chapter. Mendelsohn's music appeared 6 times, 3 of which were in Chapter 6. Specifically for text, 2 out of 3 appearances occurred in the text of Chapter 6. Schubert's music appeared 9 times in total, of which 2 appearances were in the text. Both of the occurrences were in Chapter 2. Chapter 8 contained no musical examples. Chapter 9 had 15 musical examples, with 14 belonging to J. S. Bach, whose music likewise appeared 12 times out of 20 in Chapter 5.

Figure 3 shows the link between style and chapter. Chapters 6 and 7 had the most citations, with 59 and 41 pieces each. These two chapters have the most classical music and popular music appearances. Based on the data, 30.47% of all classical music and 45.00% of all popular music appear in Chapters 6 and 7 combined. Classical music appeared more than other genres throughout the textbook. Among its 20 appearances, popular music appeared 5 times in Chapter 6 and 4 times in Chapter 7. The study likewise showed that there were 6 occurrences of anonymous music in Chapter 15 and 5 appearances of anonymous music in Chapter 6 and Chapter 10 each.

	Classical	Popular	N/A	Total
Ch 1	8	0	1	9
Ch 2	17	1	2	20
Ch 3	6	0	0	6
Ch 4	11	1	0	12
Ch 5	20	0	0	20
Ch 6	49	5	5	59
Ch 7	36	4	1	41
Ch 8	0	0	0	0
Ch 9	14	1	0	15
Ch 10	14	2	5	21
Ch 11	16	2	1	19
Ch 12	18	1	0	19
Ch 13	12	1	4	17
Ch 14	22	1	2	25
Ch 15	19	0	6	25
Ch 16	8	0	1	9
Ch 17	8	1	1	10
Appendix	1	0	0	1
Total	279	20	29	328

Fig. 3. Stylistic categorization in each chapter.

Among all 328 compositions the textbook utilized, 166 appeared in the assignments, and 161 appeared in the text, as per Figure 4. The ratio of music used in assignments and texts indicated that the authors made substantial efforts to use musical examples to showcase the theory content while reinforcing the concepts in the assignment sections. Among 279 classical pieces in the textbook, 1 appeared in the appendix, 135 appeared in the text, and 143 appeared in the assignment sections. The authors included 17 out of 20 popular repertoires in the text for popular music. In contrast, only 3 appeared in assignments, indicating that the authors attempted to use a more diversified repertoire to illustrate musical examples of the theory they discussed.

	Classical	Popular	N/A	Total
Appendix	1	0	0	1
Assignment	143	3	20	166
Text	135	17	9	161
Total	279	20	29	328
	85.06%	6.10%	8.84%	

Fig. 4. Stylistic categorization in *Appendix*, *Assignment*, and *Text*.

There was an even ratio distribution between music examples in assignments and texts, as per Figure 5. Similar musical appearances occurred in most chapters in both assignment and text sections. For instance, in Chapter 1, only 2 musical examples appeared in the text, and seven were in the assignment section. In Chapter 4, 3 musical examples appeared in the text section, and 9 musical examples appeared in the assignment section. There were five chapters with 50% more music in the text and four with 50% less music in the text. Most notably, twice as much music appeared in the assignment section than in the text section in Chapter 3. Also, the fourth chapter contained three times more music appearing in the assignment section than in the text section. Twice as much music appeared in the text section than in the assignment section in Chapter 9.

Chapter Number	A	T	
Ch 1	7	2	28.57%
Ch 2	7	13	185.71%
Ch 3	4	2	50.00%
Ch 4	9	3	33.33%
Ch 5	10	10	100.00%
Ch 6	23	36	156.52%
Ch 7	18	23	127.78%
Ch 8	0	0	N/A
Ch 9	5	10	200.00%
Ch 10	10	11	110.00%
Ch 11	11	8	72.73%
Ch 12	11	8	72.73%
Ch 13	10	7	70.00%
Ch 14	16	9	56.25%
Ch 15	17	8	47.06%
Ch 16	4	5	125.00%
Ch 17	4	6	150.00%

Fig. 5. Number of music in assignment and text per chapter.

The textbook contained 370 pages, counting only the content in the assignment and text sections. The textbook's 327 pieces of music appeared in these 370 pages. On average, 0.88 pieces of music appear per page. Chapters 6 and 7 have the highest density in terms of music citations, with 2.27 and 2.28 music per page on average. These are the only two chapters where the density is above 2.00. Chapters 12, 13, and 15 have a density of 1.19, 1.06, and 1.14, respectively.

Chapter Number	Density (Piece per Page)
Ch 1	0.35
Ch 2	0.71
Ch 3	0.33
Ch 4	0.55
Ch 5	0.91
Ch 6	2.27
Ch 7	2.28
Ch 8	0.00
Ch 9	0.68
Ch 10	0.95
Ch 11	0.95
Ch 12	1.19
Ch 13	1.06
Ch 14	0.89
Ch 15	1.14
Ch 16	0.56
Ch 17	0.50

Fig. 6. Density of music citations per chapter.

7 Implications

Diversifying musical repertoire in music theory textbooks can positively affect music education in administrative, pedagogical, and research circles. The implications of diversification of repertoire in music theory are multifaceted, encompassing a variety of pathways to how novel musical works can influence music theory learning and the results such learning brings to students' musical educational upbringing. The implications offer an interpretative context of how Western and non-Western musical works can set up musical systems necessary for students to pursue learning through critical thinking and exploratory curriculum while exploring the theory's academic side.

The diversified music theory composition selection has a societal implication for music theorists, scholars, and curricularists. Such an implication involves a broadened music repertoire, an extensive analytical repertory, and the ability to challenge the common framework of centering music analysis on European musical tradition. While the music of Bach, Beethoven, and Mozart is prevalent in musical circles, including diverse musical traditions will foster a more equitable learning environment and offer a more robust development for music appreciation, opening new curricular lines into other composers' lives and careers. Such a societal implication provides a more comprehensive development for supplementary musical expression and additional artistic perspective, focusing on novel musical characteristics represented by composers from different stylistic periods. Scholarly implication likewise allows music theory researchers and educators to explore the cultural heritage of new music with a focus on novel musical narratives. A deepened understanding of how various musical compositions align with the artistic narrative presented by music theory helps develop unknown cultural identities while carving a way into novel curriculum and instruction methodologies that will enhance students' interests in music that often do not exist in modern music theory books.

Traditional music theory pedagogies necessitate supplementary motivational incentives. Lack of motivation on students' behalf leads to a lack of interest in music theory on behalf of students. Studying a more comprehensive range of musical styles and utilizing musical examples of ways to explain analytical phenomena in music theory books would encourage students to pursue academic music studies and help curriculum makers develop new frameworks upon which to build music theory knowledge. Expanding repertoire and including various musical compositions help students broaden musical horizons and expand artistry and creativity through interpreting newly-included music. Introducing different music to AP Music Theory students would help them with musical literacy, leading to a more comprehensive and holistic music education system in which students can thrive and develop. Diversification of musical repertoire involves expanding musical boundaries, leading to the knowledge of more musical styles and novel, experimental, and analytical techniques designed to include different musical repertoires. The exploration of new musical concepts through novel musical compositions would enhance students' musical expertise and increase productivity and knowledge of music theory elements and their applicability. Expanding musical repertoire in AP Music Theory books helps curriculum makers focus on extending the traditional learning modes while expanding the curriculum to include classical, popular, film, and traditional musical genres.

Diversification of repertoire in music theory books likewise leads to implications for studying music education. Cultural consciousness and musicological knowledge are at the center of repertoire awareness. Differentiated musical repertoire would foster a more well-rounded music education system, enabling students to engage with different musical genres and appreciate multiple music traditions outside Western European music. Musical diversification would help with interdisciplinary connections in music education with the subject's ability to integrate with other artistic and non-artistic studies. Enabling students to engage with different music and understand the meaning of music theory through various repertoires extends musical traditions beyond music education into the spectra of music history and appreciation. Interdisciplinary studies in music theory with other subjects would help students promote musical ideas, develop a more holistic understanding of music's role in society, and offer a comprehensive insight into how composers and their musical styles relate to past and current music history.

One vital implication of diversifying repertoire lies in its potential impact on performance practice and the interpretative nature of music education. Music is a subjective study with no systematized learning paths to musical compositions and what should and should not exist in the curriculum. Including a broader musical narrative in the history, book would help enhance the artistic understanding of what music represents and how different composers express their artistry through unique compositional language while focusing on musical meaning. Musical styles are prominent due to their appearance in the musical sources. History, theory, and performance practice studies shed light on musical composers' place in the musical world. Performative subjects help practicalize the academic nature of music, allowing students to perform the works. Musical diversification would assist with the recognition of different composers and their works, presenting performers with more choices and theorists with more opportunities to understand the analytical process.

8 Conclusion

The implications of a diversified musical repertoire would lead to an increased understanding of musical aesthetics. Expanding musical repertoire in theory books would help interpret musical authenticity without an objective reference to categorizing music. Musical works must appear in academic sources, offer a varied reflection of musical society, and concentrate on music theory's broadened perspectives. The subject of music theory should open up musical repertoire and offer new insights into the usefulness and applicability of musical examples. Instructors disseminate such knowledge within the context of a motivational music classroom. Expanding musical repertoire would give students a more profound interpretation and understanding of musical societies and how music theory reflects them through an academic nature. The AP Music Theory curriculum has an opportunity to develop a more comprehensive outlook on music repertoire and what it represents in music theory pedagogy. The AP Music Theory course can promote cultural expansion by integrating musical examples from historical epochs. It allows music students to enhance their musical knowledge through rich historical contexts related to musical composers from different epochs and eras. The teaching material catalyzes ongoing educational reforms regarding the musical selections that students cover.

References

[1] E. Varner, "Holistic Development and Music Education: Research for Educators and Community Stakeholders," *National Association for Music Education*, vol. 32, no. 2, pp. 1-21, September 2018. https://doi.org/10.1177/1048371318798829.

[2] J. Southcott, "Egalitarian Music Education in the Nineteenth Century: Joseph Mainzer and Singing for the Million," *Journal of Historical Research in Music Education*, vol. 42, no. 1, pp. 29-45, October 2020. https://www.jstor.org/stable/27139235.

[3] T. de Clerq, "A Music Theory Curriculum for the 99%," *Engaging Students: Essays in Music Pedagogy*, vol. 7, pp. 1-8, June 2019. https://doi.org/10.18061/es.v7i0.7359.

[4] L. VanHandel, *The Routledge Companion to Music Theory Pedagogy*. New York: Routledge, 2020.

[5] D. M. Rolandson, "Motivation in Music: A Comparison of Popular Music Course Students and Traditional Large Ensemble Participants in High School," *Contributions to Music Education*, vol. 45, pp. 105-126, 2020.

[6] M. K. Kramer, and E. G. Floyd, "Required Choral Repertoire in State Music Education Performance Assessment Events," *Contributions to Music Education*, vol. 44, pp. 39-53. https://www.jstor.org/stable/26724259.

[7] B. Benward, and M. Saker, *Music in Theory and Practice*. Boston: McGraw Hill Higher Education, 2009.

[8] M. Hatfield, *Choral Repertoire: Promising New Directions for Music Theory Teaching*. 2023.

[9] C. S. Palfy, and G. Eric, "The Hidden Curriculum in the Music Theory Classroom," *Journal of Music Theory Pedagogy*, vol. 32, no. 5, pp. 79-110, 2018. https://digitalcollections.lipscomb.edu/jmtp/vol32/iss1/5.

[10] D. J. Jenkins, "Music Theory Pedagogy and Public Music Theory," in *The Routledge Companion to Music Theory Pedagogy*. Routledge, 2020.

[11] J. M. Renwick, and G. E. McPherson, "Interest and choice: student-selected repertoire and its effect on practising behaviour," *British Journal of Music Education*, vol. 19, no. 2, pp. 173-188, July 2002.

[12] G. Barton and S. Riddle, "Culturally responsive and meaningful music education: Multimodality, meaning-making, and communication in diverse learning contexts," *Society for Education, Music, and Psychology Research*, vol. 44, no. 2, pp. 345-362, May 2021. https://doi.org/10.1177/1321103X211009323.

[13] B. Juan-Monera, I. Nadal-Garcia, and B. Lopez-Casanova, "Systematic Review of Inclusive Musical Practices in Non-Formal Educational Contexts," *Education Sciences*, vol. 13, no. 1, 2023, pp. 1-43. https://doi.org/10.3390/educsci13010005.

[14] M. Reid, "Nonwestern Music and Decolonial Pedagogy in the Music Theory Classroom," *Journal of Music Theory Pedagogy*, vol. 36, no. 4, pp. 103-120, 2022. https://digitalcollections.lipscomb.edu/jmtp/vol36/iss1/4.

[15] M. A. Roig-Francoli, "A Pedagogical and Psychological Challenge: Teaching Post-Tonal Music to Twenty-First-Century Students," *Indiana Theory Review*, vol. 33, no. 1-2, Summer 2017, pp. 36-68. https://doi.org/10.2979/inditheorevi.33.1-2.02.

[16] A. Guler, "Thinking with Atonal Music in Visual Arts Education," *Journal of Qualitative Research in Education*, vol. 33, p. 1-12, 2023. https://doi.org/10.14689/enad.33.1660.

[17] C. D. Meals, "Composer Diversity in State Music Lists: An Exploratory Analysis," *National Association for Music Education*, December 2023. https://doi.org/10.1177/00224294231218272.

[18] J. P. Cumberledge, and M. L. Willians, "Representation in Music: College Students' Perceptions of Ensemble Repertoire," *Research Studies in Music Education*, vol. 45, no. 2, pp. 344-361, 2023. https://doi.org/10.1177/1321103X211066844.

[19] S. J. Priniski, C. A. Hecht, and J. M. Harackiewicz, "Making Learning Personally Meaningful: A New Framework for Relevance Research," *The Journal of Experimental Education*, vol. 86, no. 1, pp. 11-29, 2018. https://doi.org/10.1080/00220973.2017.1380589.

[20] J. Gutierrez, "An Enactive Approach to Learning Music Theory? Obstacles and Openings," *Frontiers in Education*, vol. 4, pp. 1-48, November 2019. https://doi.org/10.3389/feduc.2019.00133.

[21] C. Barroso, C. M. Ganley, S. A. Hart, N. Rogers, and J. P. Clendinning, "The relative importance of math- and music-related cognitive and affective factors in predicting undergraduate music theory achievement," *Applied Cognitive Psychology*, vol. 33, no. 5, pp. 771-783, January 2019. https://doi.org/10.1002/acp.3518.

[22] M. Hoag, "Integration, Diversity, and Creativity in Current Music Theory Pedagogy Research," *College Music Symposium*, vol. 56, pp. 1-30, 2016. https://www.jstor.org/stable/26574445.

[23] J. S. Snodgrass, "Current Status of Music Theory Teaching," *College Music Symposium*, vol. 56, pp. 1-10, 2016. https://www.jstor.org/stable/26574444.

[24] J. Day-O'Connell, "Putting the Theory Back in 'Music Theory'," *Engaging Students: Essays in Music Pedagogy*, vol. 7, pp. 1-14, 2019. https://doi.org/10.18061/es.v7i0.7368.

[25] B. Utne-Reitan, "Music Theory Pedagogy in the Nineteenth Century: Comparing Traditions of Three European Conservatories," *Journal of Music Theory*, vol. 66, no. 1, pp. 63-91, April 2022. https://doi.org/10.1215/00222909-9534139.

[26] Y. Kang, "Defending Music Theory in a Multicultural Curriculum," *College Music Symposium*, vol. 46, pp. 45-63, 2006. http://www.jstor.org/stable/40374439.

[27] S. Gates, "Developing Musical Imagery: Contributions from Pedagogy and Cognitive Science," *Music Theory Online*, vol. 27, no. 2, pp. 1-25, June 2021. https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.gates.html.

[28] N. O. Buonviri, "Successful AP Music Theory Instruction: A Case Study," *Update: Applications of Research in Music Education*, vol. 36, no. 2, p. 53-61, February 2018. https://eric.ed.gov/?id=EJ1165736.

[29] N. Patton, *Student Musical Experiences and Self-Efficacy in AP Music Theory*. PhD Dissertation, Temple University, 2023.

[30] A. Grey, "Improving Students' Aural Skills on the AP Music Theory Exam," *Music Educators Journal*, vol. 107, no. 3, p. 47-53, March 2021. https://eric.ed.gov/?id=EJ1290967.

[31] A. S. Paney, and N. O. Buonviri, "Teaching Melodic Dictation in Advanced Placement Music Theory," *Journal of Research in Music Education*, vol. 61, no. 4, pp. 396-414, 2014. https://doi.org/10.1177/0022429413508411.

[32] P. D. MacIntyre, and B. Schnare, "Self-Determination Theory and Motivation for Music," *Psychology of Music*, vol. 46, no. 5, pp. 699-715, 2017. https://doi.org/10.1177/0305735617721637.

[33] R. H. Woody, “Music Education Students' Intrinsic and Extrinsic Motivation: A Quantitative Analysis of Personal Narratives,” *Psychology of Music*, vol. 49, no. 5, pp. 1321-1343, August 2020. https://doi.org/10.1177/0305735620944224.

[34] J. Check, “Repertoire Choices in the Classroom,” *College Music Symposium*, vol. 52, pp. 1-4, 2012. https://www.jstor.org/stable/26564877.

[35] S. G. Nielsen, A. Jordhus-Lier, and S. Karlsen, “Selecting Repertoire for Music Teaching: Findings from Norwegian Schools of Music and Arts,” *Research Studies in Music Education*, vol. 45, no. 1, pp. 94-111, 2023. https://doi.org/10.1177/1321103X221099436.

Appendices 1 and 2

In A/T Only	N/A	J.S. Bach	Beethoven	Chopin	Debussy	Handel	Haydn	Mendelssohn	Mozart	Schubert	Schumann	All Others		
1	1	4	2	0	0	0	0	0	0	0	0	2	9	44.44%
2	2	2	0	1	1	0	2	0	2	0	0	10	20	50.00%
3	0	3	0	0	0	1	0	0	1	0	0	1	6	50.00%
4	0	5	0	0	0	0	0	0	1	0	0	6	12	50.00%
5	0	12	0	0	0	1	1	0	0	0	0	6	20	60.00%
6	5	1	3	2	0	0	12	2	2	4	0	28	59	47.46%
7	1	10	3	3	3	0	2	3	3	2	1	10	41	24.39%
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
9	0	14	0	0	0	0	0	0	0	0	0	1	15	93.33%
10	5	9	2	1	0	1	0	0	0	0	0	3	21	42.86%
11	1	7	1	1	0	0	2	0	1	0	0	6	19	36.84%
12	0	8	2	0	1	0	1	0	3	0	0	4	19	42.11%
13	4	6	0	0	1	1	0	0	1	0	3	1	17	35.29%
14	2	10	2	1	1	0	0	0	1	0	0	8	25	40.00%
15	6	11	0	0	0	0	1	0	1	3	0	1	23	47.83%
16	1	3	0	0	0	1	1	0	0	0	2	3	11	27.27%
17	1	0	0	1	0	0	1	1	2	0	2	2	10	20.00%
Appendix	0	0	0	0	0	0	0	0	1	0	0	0	1	100.00%
	29	105	15	10	7	5	23	6	19	9	8	92		
	20.69%	13.33%	20.00%	30.00%	42.86%	20.00%	52.17%	50.00%	15.79%	44.44%	25.00%	30.43%		

Appendix 1. Citations of main contributors (more than 1%) in each chapter.

In T Only	N/A	J.S. Bach	Beethoven	Chopin	Debussy	Handel	Haydn	Mendelssohn	Mozart	Schubert	Schumann	All Others		
1	0	0	2	0	0	0	0	0	0	0	0	0	2	100.00%
2	2	1	0	1	1	0	1	0	1	0	0	6	13	46.15%
3	0	1	0	0	0	0	0	0	0	0	0	1	2	50.00%
4	0	0	0	0	0	0	0	0	0	0	0	3	3	100.00%
5	0	5	0	0	0	1	1	0	0	0	0	3	10	50.00%
6	3	1	3	1	0	0	6	2	1	2	0	17	36	47.22%
7	1	7	0	2	2	0	1	1	2	0	0	7	23	30.43%
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
9	0	9	0	0	0	0	0	0	0	0	0	1	10	90.00%
10	0	4	2	1	0	1	0	0	0	0	0	3	11	36.36%
11	0	1	0	1	0	0	0	0	1	0	0	5	8	62.50%
12	0	0	2	0	1	0	0	0	2	0	0	3	8	37.50%
13	0	1	0	0	1	1	0	0	1	0	2	1	7	28.57%
14	0	1	1	1	1	0	0	0	0	0	0	5	9	55.56%
15	1	2	0	0	0	0	1	0	1	0	2	1	8	25.00%
16	1	2	0	0	0	0	1	0	0	0	0	1	5	40.00%
17	1	0	0	1	0	0	0	0	1	0	1	2	6	33.33%
Appendix	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
	9	35	10	8	6	3	11	3	10	2	5	59		
	33.33%	25.71%	30.00%	25.00%	33.33%	33.33%	54.55%	66.67%	20.00%	100.00%	40.00%	28.81%		

Appendix 2. Citations of main contributors (more than 1%) in the *T* for each chapter.

Consolidating Visual Genres of Opera and Ballet in Film Music Curriculum in Music Education

Yingshu Wang and Xuanyuan Jin

Department of Music Education
yingshuwang1110@outlook.com, jinxuanyuan@outlook.com

Abstract. In the high school music curriculum, film music plays a significant role in developing an overall understanding of music appreciation. Film music is a new genre of musicology and is expanding in the musicology curriculum. Research shows that teachers can reinforce other musical subjects by including film music in their curriculum (Haghverdi, 2015). This research shows how film music integrates with classical and popular music curricula in high school music education. This research also discusses a novel educational strategy, expanding the film music curriculum to include classical music genres such as opera and ballet. This new approach will help in understanding the visual aspects of ballet and opera as predecessors to film music. The ballet repertoire used in this article for analysis includes Tchaikovsky's *Swan Lake* and *Nutcracker*. The operatic repertoire used in this article for analysis includes Bizet's *Carmen.* Extending the film music curriculum to include the teachings of ballet and opera can also advance students' interests and develop a greater appreciation of music.

Keywords. ballet, film music, music curriculum, music education, opera, visual music

1 Introduction

Film music is a genre of musicology new to the high school music curriculum [1]. Film music focuses on compositions with a purpose to accompany visual arts. Film music can evoke strong emotions, build tension, create suspense, and enhance dramatic moments. Film music aims to add to what the audience sees on screen, accompanying the sounds to the images and actions in the film's plot. The film music genre is becoming more extensive in music history because of music and film technology in the twenty-first century [2]. The history of film music began in the silent era when no sound was to accompany the film [3]. Gradually, composers introduced music to cinema, becoming an industry, creating a visual with music. A film incorporating music and screen action is not the first example of music with visuals. For example, musicians often play in the pit orchestra, accompanying characters on the stage in musicals. The pit orchestra members, including actors, singers, and dancers, play the music

accompanying the show presented on the stage. Today, film music has its independent industry, providing a different platform for composers to express their creative ideas.
Film music plays a significant role in the music curriculum. Film music learning integrates movie plots, musical scores, symbolic meaning in what characters do in the movie, and historical narrative in literature. Learning about film music is essential in music education to develop an all-rounded knowledge in high school music appreciation. Film music appears in the university curriculum but is limited in its appearance in high school music classes. High schools focus on the knowledge of classical and popular music in music education [4]. Integrating film music and what it can include in its study plan can improve students' knowledge of music history and their appreciation for different musical genres. The integration of film music can also increase students' awareness of music technology and its use in other musical styles and cultures [5].
Film music can help teachers reinforce other musical subjects in the school curriculum [6]. For example, film music can help teachers support fundamental music theory concepts for their students. Students can use film scores as part of analysis projects and in more advanced activities that involve identifying the different elements of the film scores. Students can also learn about diverse musical symbols and characteristics in the film music. Many stylistic features are helpful in the music education curriculum when integrated with films and film themes. Film composers determine the characteristics of different film music based on the descriptions and actions of movie characters, all of which can be a part of the film music curriculum [7]. Studying film music can help students understand the different emotions in the films and how film music supports the movies they watch.
This research focuses on expanding the film music curriculum of classical and popular music in high school music education. Expanding the curriculum is an educational strategy to give more interest and emphasis to film music for students interested in growing their appreciation of music. Incorporating film music into the school music curriculum can help capitalize on students' enthusiasm for film music, which is necessary to develop new music ideas in music education [8]. Educators can use film music and its developed curriculum to help students explore more musical styles and show a more substantial musical diversity.

2 The Focus on Film Music Curriculum

High school music education classes focus on music performance more than academics [9]. Schools value music performance because it is a music elective with a stronger appeal to high school students. It gives performing opportunities to students, raises their performance confidence on the stage, teaches artistic skills, and improves students' instrumental proficiency [10]. However, a strong performance curriculum must balance with a music academic curriculum, including music theory and history classes. The lack of academic music classes creates limited learning opportunities for students in analytical knowledge.

A balanced music curriculum is essential for academic access and more robust learning initiatives for music education [11]. Students playing instruments improve artistic and performance skills, but knowledge of theory and history is necessary to understand the performance and interpretation. Music is not required, so students select their music classes based on their interests. Students are more likely to choose courses based on intrinsic motivation [12].

Music performance is motivational because students can play instruments and collaborate in a group. Ensembles, orchestras, and bands all provide students with these learning initiatives. Academic music needs more motivation and contains more formal assessments for students to complete. Motivation in music education is essential to understanding the type of classes that students will select [13]. Motivation drives students' passion, and the loss of motivation decreases interest in music [14]. Music history can be more motivational for students with higher music performance included in the curriculum, which is possible for film music. Classical music history involves the formal study of artistic history and has less interest for students when learning about musical genres, such as ballet and opera. Implementing ballet and opera as visual genres in film music curricula can raise students' motivation and awareness for these genres and teach students how visual classical genres integrate with film music.

Societal trends change in the twenty-first century, and students' interests in music align with what they experience in society. Popular and film music is more prominent among younger students, making it more attractive for high school students to study as part of their academic music curriculum [15]. Students with different musical tastes find learning about specific genres more attractive to their musical education. When comparing music history, popular music and film music are more appealing than classical music, the idea that should drive music history education [16]. With the increase in popular and film music in society, schools should not eliminate classical works and genres. Less interest in studying classical music refers to less motivation in these musical styles for students [17]. However, it does not mean that classical music needs to be eliminated from the curriculum. To create more awareness of classical music and to emphasize its influence in history, the visual genres of ballet and opera can appear in the film music curriculum, expanding the film music history enough to offer balanced music education.

Certain composers and styles in classical music appear in history and have had enormous musical impacts. For example, Tchaikovsky is a renowned composer known for his ballets [18]. While Tchaikovsky holds influence in the ballet world of Western European music, his works of Nutcracker and Swan Lake are under-represented in the high school music history curriculum. Including Tchaikovsky's ballets in film music programs of study in high schools can be a predecessor to teaching about movies based on the ballets or films that use Tchaikovsky's music.

3 Research Objectives

Multiple objectives exist for conducting research in this field. First, this article aims to expand the film music curriculum of classical music and popular music in music education. Film music is a new topic with a history smaller than other substantial music history genres compared to Western Classical and American Popular music. Film music is not an independent art form and evolves with the film industry's expansion [19]. Curriculum growth is essential in music education to incorporate new learning and create a different systematized curriculum path for students in music history classes [20]. This research assists with analyzing how the musicology curriculum can be more engaging for high school students and how film music can enhance students' overall music appreciation and knowledge.

Second, this research evaluates ways that help high school music students find more interest in classical music, such as ballet and opera. The research proposes teaching the genres of ballet and opera in film music. This curricular action will develop more appeal in music learning for students studying and understanding ballet and opera. Both genres are underrepresented in classical music learning, but both genres are essential in classical music with rich and influential history, which students should comprehend [21], [22]. Ballet and opera are visual genres that combine actions on the stage with the music, similar to film music. The ballet visuals include dancing and choreography, while the opera visuals include acting and singing. Both ballet and opera use performed music as accompaniment to the visual. Therefore, to introduce visual elements in film music, ballet, and opera must appear in the film music curriculum to help students comprehend the visual characteristics of different art forms.

Third, this research extends the definition of film music curriculum and compares the meanings of music history and music education for the art of film music. This research shows a new perspective and comprehension of film music and how students in school can interpret it. The ideas in this research can help music teachers create a more interactive academic music curriculum that will cover visual styles and genres that integrate well with film music.

4 Nomenclature

The following section of the research defines three divisions of music history. These are classical music, popular music, and film music. These definitions are for music curriculum to create a distinction between teaching material. The following terms show distinct features in how music history is divided in the high school music curriculum.

Classical music refers to works composed by Western classical composers. Baroque, Classical, Romantic, and Impressionist composers exist in classical music history and combine Western art music. For example, Ludvig van Beethoven, a composer from the classical and romantic genres, born in 1770, is an example of a known artist from the Western arts [23]. Claude Debussy, born in 1862, is a French composer from the Impressionist period [24]. Classical music has a rich history, and schools should promote and foster the academic study of classical music for a balanced education. Classical music has been losing appeal in high school music history learning because of limited interaction between students and classical music [25]. The classical genres of ballet and opera and their potential association with the extension of film music curriculum can help students find identity and new links with classical genres.

Popular music originated in the United States in the late nineteenth century. Popular music spread through the United States as new popular styles continued to appear. Popular music has more listeners than classical music worldwide because it appeals to more individuals. Popular music has been growing in the school curriculum, and students find popular music more appealing than classical music [26]. Popular music has a shorter history, but as it evolved, it was influenced by classical, traditional, and other styles. The knowledge of Western European music is essential in understanding how it influenced popular music.

The birth of jazz in New Orleans and the development of country music, which originated in the 1920s in the southern United States, are examples of popular music styles growing over time [27]. Popular music is favored over the classical music industry and is considered more business-centered [28]. From a music theory perspective, popular music contains fewer analysis elements, making it easier to comprehend. For instance, early jazz musicians were self-taught and mainly performed by ear [29]. Additionally, country music that appeared in the early nineteenth century was suitable for musicians of all levels [30].

Film music has the function to accompany movies. Film music can be original. Film music can also be from classical and popular repertoire. Film music composers need to use sounds to develop the cinematic plot, allowing the audience to feel the same emotions as the characters in the movie. Film music composers create music based on the characters' actions and the environment set in the movie. Film music should complement the movie. Different plots will have different characters that film music needs to express. Strong rhythms and melodic musical lines can highlight tension and peacefulness, depending on what is happening in the movie. Lyrical music in film music repertoire can highlight the characters' emotions in the movie.

The following overview of styles shows boundaries in musicology and definitions in music education to outline stylistic features. Classical, popular, and film music are general divisions of musicology and how curriculum can divide music history. However, musicology is continual, as styles transform depending on musical influence. For example, composers might borrow classical music and orchestrate it into popular music, adding modern and popular elements. A similar process occurs with film music, with some compositions borrowed from classical and popular styles. Music is never static. Stylistic vocabulary exists to define the characteristics of genres. At the same time, categorizing music can be detrimental to music education because students will focus on musicological definitions more than on artistry and authenticity.

5 Review of Research Literature

Various sources appeared in the research literature that discusses film music, its integration into education, and its analysis. Berg discussed the selection of films for music education classes [31]. Berg stated that the development of film follows the growth of society. From the perspective of music education, film music appears spontaneously through technology, television, and other ways. Students can hear and be aware of the music without watching the film. It is also common for students to watch the movie and remember certain parts of the film's music. Film music strongly influences high school students and their awareness of music history and its repertoire knowledge. Teachers should use film music to their advantage and develop knowledge in music history classes. Teachers of film music should ensure that the works they showcase to students meet educational standards.

Webb and Fienberg demonstrated ways for students to develop analytical, musicological, and performance skills to understand better aspects of Australian indigenous musical cultures and the Australian experience [32]. Webb and Fienberg discussed the transformation between music and narrative in engaging students with history and social relationships. By immersing in performance, students used their imagination in history, which gave them a deep connection with music history. Less theoretical music foundations become a struggle for students studying music. Webb and Fienberg demonstrated the connection between ethnomusicology in Indigenous film and school music education.

Jorgensen discussed the "musical connections with myths evident in societies around the world and well established in the anthropological, ethnomusicological, and philosophical literature" [33]. Jorgensen gave the example of *The Lord Of The Rings*; musical scores help highlight the character of the narrative. Music is the role of blurring past and present. In *The Lord of the Rings*, the arts are brought together in unity, fitting the education standard. Students do not only study music but the broad environment, focusing on learning about the artistic emotions, cultures, and stories that music complements. The musicological connections of film and its music are also associated with learning philosophy in society.

Hunter and Frawley explored incorporating an arts-based pedagogical approach that uses film, art, and music to enhance students' overall learning experiences [34]. Arts-based learning can help students sustain interest in the course material and understand the theory of the course material. Arts-based learning also lets students think and analyze at a higher level, giving them more confidence in their analytical and critical thinking abilities. Hunter and Frawley found that students have been positively influenced to incorporate ideas with film, art, and music in their theory and history lectures. With these approaches, students learned the material in exciting ways, which helped students comprehend and recall the learned knowledge.

Jarvis discussed a graphic technique for large-scale film-music analysis that utilized the film's narrative structure as a scaffold for conducting musical inquiry [35]. Jarvis's approach focused on building a temporal, narrative-based scaffolding on music's relationship to film context. Jarvis's work is based on music theory analysis, helping students learn about musical function, relationships between events and compositional styles in film music, and the scores project information over long periods. Jarvis stated that NDS diagrams, Narrative/Dramatic Structure diagrams, can be applied in any multimedia work.

Broesche evaluated Glenn Gloud's analogy between live theatre and film, as well as concert performance and studio recording [36]. By examining the history of Glenn Gloud's studio processes, Broesche stated that Gould's understanding of studio recording differed from the familiar recordings, creating a simulacrum in the live event. Broesche's research can help understand the differences and similarities between film music recordings and music recorded for live theatre.

Neumeyer conducted tests to understand how Schoenberg's *Begleitungsmusik zu einer Lichtspielszene* functions as background music and defines the implications of questions that arise about musical culture and classes [37]. Creating film scores is complex, requiring vital creativity and good movie comprehension. Creating film music leads to cognitive biases and standards of judgment of compositional traditions. Neumeyer found the use of Schoenberg's compositional language and the implemented musical analysis method. Neumeyer also suggested a potential crisis in the links between technical musical criticism tools and the ideology of compositional culture.

6 Analysis: Extending Film Music Curriculum with Music for Visualization

Opera and ballet are both classical music genres that are precursors of modern-day movies. Both genres are similar to film because the music is secondary in these art forms and accompanies the actions on the stage or the screen. Similar to films, operas, and ballets tell the story through visuals and auditory art forms. An opera is a form of theater with dramatic action occurring on the stage, similar to a movie expressing the plot to the audience [38]. Ballet expresses the story through dancing and accompanying music, related to how the art of film works, involving the characters' actions. As substantial classical music genres, opera and ballet influence the film industry.

Opera is a performing art that combines music, drama, and visual arts. The actor uses singing to express emotions and illustrate a story. Opera stage designs, costumes, and synopsis are all essential components of visual arts that the audience enjoys while listening to the music played by the pit orchestra. As a secondary element, operatic music helps the audience perceive and comprehend the synopsis. Ballet involves the technical performance of dance on the stage, relying heavily on physical expressions without involving spoken or sung dialogues. At the same time, genres of different historical periods, ballet music, operatic music, and film music carry the nature of accompanying the visual representation that occurs on the stage or the screen.

Operas and ballets have been adapted into film formats. One known example is the *Carmen* opera, composed by the French composer Georges Bizet. There are several adaptations of this opera. There is a ballet called the *Carmen Suite*, composed by Rodion Shchedrin, a Russian composer, with choreography by Alberto Alonso. There are also two African films based on the opera Carmen called *Karmen Gei*, created in 2001, and another film called *U-Carmen eKhayelitsha*, created in 2005. It is advantageous in music education for opera adaptions to appear in the film music curriculum. Extending the recent film to incorporate the historical birth of Carmen will help music teachers in high school music history classes to integrate the visual aspects of classical and film music genres.

Another example of ballet is Tchaikovsky's *Nutcracker*, a monumental Romantic work. A fantasy adventure film based on the ballet is Hallstrom's and Johnston's *The Nutcracker and the Four Realms*, premiered in 2018. The recent film and Tchakovsky's ballet originate from Hoffmann's book called *The Nutcracker and the Mouse King*. Film music curriculum can support music learning, its elements, and literature influence. The adaptations above retain the elements of opera and ballet, such as the dramatic narrative, the music, and the choreography.

The audience has a separate perspective of visual genres. Ballets, operas, and films share similarities because they serve as entertainment genres. The audience watches and observes the plot and actions on the stage and the screen, with secondary music. Teaching ballets and operas through movies is very helpful for students in the music history curriculum. The adaptation of films can give audience members different perspective of composers, their influences, and the compositional process. By learning about ballets and operas and their adaptations to films, the film music curriculum can give students different knowledge on emotions and cultures that composers express in different forms. Extending the film music curriculum to include classical visual genres is helpful for students to understand different music genres and how composers adapted different art forms throughout music history.

Film music's importance in music history is significant [39]. Film music gives emotional engagement to the film, providing the audience with a fundamental musical base for the film's plot. Film music gives musical support to the characters in the film and offers musical reasoning for the film's main idea and topic. Film music is a cultural sign in music history, expanding as film scores continue to accompany new movies. Film music, as a genre that helps film directors produce a form of entertainment, develops with the expansion of technology. Film music is indirectly influential because of the primary theme in the film's plot and the actions on the screen. The film composer's task is to ensure the music adapts to the movie, making film music an important visual genre influential in music history. As the film industry develops, film music composition continues to grow.

Film music is not designed to be concert music, although film music has been played on the stage without the visual. A traditional music concert also contains aural and visual art forms. However, the aural component, including listening to the music, is primary, as composed to watching the performance. This changes in the opera and ballet music, where the visual aspect includes the opera's and ballet's plots, including various characters, with music helping the audience shape their interpretation. The auditory aspect of ballet and opera requires composers to enhance the plot, allowing the audiences and actors to immerse themselves in the same emotions quickly.

Swan Lake, a ballet by Tchaikovsky that premiered in 1877 holds significant historical meaning in the ballet art form. The *Swan Lake* shows powerful music mixed with strong emotions, supported by the thematic ideas that Tchaikovsky envisioned to illustrate the drama of the ballet's plot. Tchaikovsky's *Swan Lake* music has four acts with different musical styles and moods, depending on the plot. At the beginning and introduction to the ballet's plot in the opening act, Tchaikovsky uses grandeur music, transforming into the smooth lyricism of the swan theme. Tchaikovsky's use of different musical styles is based on the ballet themes associated with the different characters and emotions in the plot. For example, in *The Swan Theme*, Tchaikovsky used a delicate musical style to represent the purity of the swan maidens; the type of music played by the string gives the beauty of the swan maidens. In *The Rothbart Theme*, Tchaikovsky uses the dark musical style to show the evilness of Rothbart, where the music is played by the brass and percussion instruments, indicating the darkness of Rothbart.

Nutcracker, another ballet composed by Tchaikovsky, is based on Hoffmann's book *The Nutcracker and The Mouse King*. In *Nutcracker*, Tchaikovsky divides the ballet into two acts, where each act has a different musical compositional style, depending on the characters, their description, and their actions. For example, in Act I, the stage is full of joyful music, which helps to have a connection with the ballet's plot.

Tchaikovsky used joyful dances and waltzes to show excitement and create a particular atmosphere onstage. In Act II, Tchaikovsky used the ethereal music to express the gorgeousness. For instance, *Dance of the Sugar Plum Fairy* is a delicate ballet piece that creates a magical atmosphere. While the ballet's music is often performed without the visual, the combination of music and dance is ideal for understanding the meaning behind the music and what it provides for Tchaikovsky and his ballet's plot.

Because of similarities in how visuals combine with the music, many ballets have been adapted into movies. For example, the *Black Swan*, directed by Darren Aronofsky, uses the idea from Tchaikovsky's *Swan Lake*. Nina Sayers, the character in Aronofsky's *Black Swan* movie, shows the protagonist and antagonist character traits. The movie *Black Swan* integrates some elements of Tchaikovsky's music with music composed by other composers. The film score for *Black Swan* can teach how ballet music integrates with movies and how ballet music can transform into a film industry production. However, for this to happen, Tchaikovky's Swan Ballet must be introduced and taught before *Black Swan*.

Tchaikovsky is a significant composer in music history. His contributions to the Romantic era of classical music are numerous, which include symphonies, operas, chamber music, and ballets. Ballets, in particular, are examples of Tchaikovky's music that combine visual representation and Romantic music. Teaching Tchaikovsky and his ballet music in film music history can help students understand the importance of Tchaikovsky as a composer and the compositional style that he used to develop ballet music. Tchaikovsky's music has also had a significant impact on other films. His compositions have been used in different movies to add emotion to the film's story. For example, Disney's *Fantasia* in 1940, Stanley Kubrick's *A Clockwork Orange* in 1971, and James Cameron's *Titanic* in 1997.

Integrating ballet and film and resemblances between both art forms is significant in education and art history. Music in films helps people understand the purpose of the visual themes and learn about the purpose of music in the visual art form. Another famous example of a movie that uses classical music is *Brief Encounter*, a British Romantic drama directed by David Lean. The music uses excerpts from Sergei Rachmaninoff's Piano Concerto No. 2 in C minor, which continually reappears throughout the film. No other music is used.

Learning about this film and music helps us understand the connection between Rachmaninoff's themes and how they fit into the movie. Music analysis is different when Rachmaninoff's Piano Concerto No. 2 is played at the concert hall, as opposed to watching the film and understanding the purpose the music has in the film. Using concert music to explore the connections between different art forms and better understand music's role in enhancing visual storytelling will help expand the film music curriculum. Ballet is a historic art form with a rich history and cultural identity. The ballet can be a transformative art form, and understanding ballet music can help understand how film composers structure their works.

7 Conclusions

Expanding the film music curriculum to include ballet and operatic genres will create a more robust learning opportunity for students in music history classes. Integrating film and classical music would help teachers create more exciting ways of teaching music history and leading music education. Integrating classical music genres such as ballet and opera can raise awareness for European classical music and make it appear more in the film music curriculum. Understanding how visual aspects affect music is essential in ballet, opera, and film music. The cultural aspects of ballet and opera music are essential to learn because these genres are predecessors of film music and other visual music genres. Understanding the visual characteristics of different art forms can help students understand how the visuals integrate with sound in combined art forms.

Classical music and its academic side are less motivating for students. However, Western classical music is an important source and foundation of musical knowledge. Ballets and operas are influential genres with a rich history in classical music. Students can benefit from learning these genres and understanding how they integrate visual aspects of musicianship with the performance on the stage. Expanding the film music curriculum and combining classical musical genres allows students to find more interest in music history and appreciation.

Motivation is essential in music history classes, and a motivating classroom can make the class more exciting. When teaching about classical music, teachers can use films to make learning exciting. For example, the 1940 Disney-animated film called *Fantasia* combines classical music with impressive visuals. This film set many compositions by composers, including Beethoven, Tchaikovsky, and Stravinsky. Because there are many visual interpretations of the music, students can find interest in classical music selection. In music classes, students can discuss how the music fits the animation, how the emotions in the film are being conveyed, and how to analyze the techniques of storytelling in the film and the music's impact on it. Using this approach, students can gain a deep understanding and appreciation of music.

This article can set a base for future research in musicology and education. Integrating music education and musicology is an inclusive and interdisciplinary way to align genres and show their benefit in different musical studies. Connecting music and integrating contrasting genres helps students see music history as a transformative study. Students become more aware of different compositions and visual genres, which makes musical repertoire more accessible. Extending the film music curriculum helps build the bridge between classical visual music and modern film music, where students learn to understand different cultures and histories of music.

References

[1] E. Varner, "Holistic Development and Music Education: Research for Educators and Community Stakeholders," *National Association for Music Education*, vol. 32, no. 2, pp. 1-21, September 2018. https://doi.org/10.1177/1048371318798829.

[2] J. Southcott, "Egalitarian Music Education in the Nineteenth Century: Joseph Mainzer and Singing for the Million," *Journal of Historical Research in Music Education*, vol. 42, no. 1, pp. 29-45, October 2020. https://www.jstor.org/stable/27139235.

[3] T. de Clerq, "A Music Theory Curriculum for the 99%," *Engaging Students: Essays in Music Pedagogy*, vol. 7, pp. 1-8, June 2019. https://doi.org/10.18061/es.v7i0.7359.

[4] L. VanHandel, *The Routledge Companion to Music Theory Pedagogy*. New York: Routledge, 2020.

[5] D. M. Rolandson, "Motivation in Music: A Comparison of Popular Music Course Students and Traditional Large Ensemble Participants in High School," *Contributions to Music Education*, vol. 45, pp. 105-126, 2020.

[6] M. K. Kramer, and E. G. Floyd, "Required Choral Repertoire in State Music Education Performance Assessment Events," *Contributions to Music Education*, vol. 44, pp. 39-53. https://www.jstor.org/stable/26724259.

[7] B. Benward, and M. Saker, *Music in Theory and Practice*. Boston: McGraw Hill Higher Education, 2009.

[8] M. Hatfield, *Choral Repertoire: Promising New Directions for Music Theory Teaching*. 2023.

[9] C. S. Palfy, and G. Eric, "The Hidden Curriculum in the Music Theory Classroom," *Journal of Music Theory Pedagogy*, vol. 32, no. 5, pp. 79-110, 2018. https://digitalcollections.lipscomb.edu/jmtp/vol32/iss1/5.

[10] D. J. Jenkins, "Music Theory Pedagogy and Public Music Theory," in *The Routledge Companion to Music Theory Pedagogy*. Routledge, 2020.

[11] J. M. Renwick, and G. E. McPherson, "Interest and choice: student-selected repertoire and its effect on practising behaviour," *British Journal of Music Education*, vol. 19, no. 2, pp. 173-188, July 2002.

[12] G. Barton and S. Riddle, "Culturally responsive and meaningful music education: Multimodality, meaning-making, and communication in diverse learning contexts," *Society for Education, Music, and Psychology Research*, vol. 44, no. 2, pp. 345-362, May 2021. https://doi.org/10.1177/1321103X211009323.

[13] B. Juan-Monera, I. Nadal-Garcia, and B. Lopez-Casanova, "Systematic Review of Inclusive Musical Practices in Non-Formal Educational Contexts," *Education Sciences*, vol. 13, no. 1, 2023, pp. 1-43. https://doi.org/10.3390/educsci13010005.

[14] M. Reid, "Nonwestern Music and Decolonial Pedagogy in the Music Theory Classroom," *Journal of Music Theory Pedagogy*, vol. 36, no. 4, pp. 103-120, 2022. https://digitalcollections.lipscomb.edu/jmtp/vol36/iss1/4.

[15] M. A. Roig-Francoli, “A Pedagogical and Psychological Challenge: Teaching Post-Tonal Music to Twenty-First-Century Students,” *Indiana Theory Review*, vol. 33, no. 1-2, Summer 2017, pp. 36-68. https://doi.org/10.2979/inditheorevi.33.1-2.02.

[16] A. Guler, “Thinking with Atonal Music in Visual Arts Education,” *Journal of Qualitative Research in Education*, vol. 33, p. 1-12, 2023. https://doi.org/10.14689/enad.33.1660.

[17] C. D. Meals, “Composer Diversity in State Music Lists: An Exploratory Analysis,” *National Association for Music Education*, December 2023. https://doi.org/10.1177/00224294231218272.

[18] J. P. Cumberledge, and M. L. Willians, “Representation in Music: College Students' Perceptions of Ensemble Repertoire,” *Research Studies in Music Education*, vol. 45, no. 2, pp. 344-361, 2023. https://doi.org/10.1177/1321103X211066844.

[19] S. J. Priniski, C. A. Hecht, and J. M. Harackiewicz, “Making Learning Personally Meaningful: A New Framework for Relevance Research,” *The Journal of Experimental Education*, vol. 86, no. 1, pp. 11-29, 2018. https://doi.org/10.1080/00220973.2017.1380589.

[20] J. Gutierrez, “An Enactive Approach to Learning Music Theory? Obstacles and Openings,” *Frontiers in Education*, vol. 4, pp. 1-48, November 2019. https://doi.org/10.3389/feduc.2019.00133.

[21] C. Barroso, C. M. Ganley, S. A. Hart, N. Rogers, and J. P. Clendinning, “The relative importance of math- and music-related cognitive and affective factors in predicting undergraduate music theory achievement,” *Applied Cognitive Psychology*, vol. 33, no. 5, pp. 771-783, January 2019. https://doi.org/10.1002/acp.3518.

[22] M. Hoag, “Integration, Diversity, and Creativity in Current Music Theory Pedagogy Research,” *College Music Symposium*, vol. 56, pp. 1-30, 2016. https://www.jstor.org/stable/26574445.

[23] J. S. Snodgrass, “Current Status of Music Theory Teaching,” *College Music Symposium*, vol. 56, pp. 1-10, 2016. https://www.jstor.org/stable/26574444.

[24] J. Day-O’Connell, “Putting the Theory Back in ’Music Theory’,” *Engaging Students: Essays in Music Pedagogy*, vol. 7, pp. 1-14, 2019. https://doi.org/10.18061/es.v7i0.7368.

[25] B. Utne-Reitan, “Music Theory Pedagogy in the Nineteenth Century: Comparing Traditions of Three European Conservatories,” *Journal of Music Theory*, vol. 66, no. 1, pp. 63-91, April 2022. https://doi.org/10.1215/00222909-9534139.

[26] Y. Kang, “Defending Music Theory in a Multicultural Curriculum,” *College Music Symposium*, vol. 46, pp. 45-63, 2006. http://www.jstor.org/stable/40374439.

[27] S. Gates, “Developing Musical Imagery: Contributions from Pedagogy and Cognitive Science,” *Music Theory Online*, vol. 27, no. 2, pp. 1-25, June 2021. https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.gates.html.

[28] N. O. Buonviri, "Successful AP Music Theory Instruction: A Case Study," *Update: Applications of Research in Music Education*, vol. 36, no. 2, p. 53-61, February 2018. https://eric.ed.gov/?id=EJ1165736.

[29] N. Patton, *Student Musical Experiences and Self-Efficacy in AP Music Theory*. PhD Dissertation, Temple University, 2023.

[30] A. Grey, "Improving Students' Aural Skills on the AP Music Theory Exam," *Music Educators Journal*, vol. 107, no. 3, p. 47-53, March 2021. https://eric.ed.gov/?id=EJ1290967.

[31] A. S. Paney, and N. O. Buonviri, "Teaching Melodic Dictation in Advanced Placement Music Theory," *Journal of Research in Music Education*, vol. 61, no. 4, pp. 396-414, 2014. https://doi.org/10.1177/0022429413508411.

[32] P. D. MacIntyre, and B. Schnare, "Self-Determination Theory and Motivation for Music," *Psychology of Music*, vol. 46, no. 5, pp. 699-715, 2017. https://doi.org/10.1177/0305735617721637.

[33] R. H. Woody, "Music Education Students' Intrinsic and Extrinsic Motivation: A Quantitative Analysis of Personal Narratives," *Psychology of Music*, vol. 49, no. 5, pp. 1321-1343, August 2020. https://doi.org/10.1177/0305735620944224.

[34] J. Check, "Repertoire Choices in the Classroom," *College Music Symposium*, vol. 52, pp. 1-4, 2012. https://www.jstor.org/stable/26564877.

[35] S. G. Nielsen, A. Jordhus-Lier, and S. Karlsen, "Selecting Repertoire for Music Teaching: Findings from Norwegian Schools of Music and Arts," *Research Studies in Music Education*, vol. 45, no. 1, pp. 94-111, 2023. https://doi.org/10.1177/1321103X221099436.

Double Meanings in Operatic Social Interactions, Characters, and Symbolisms in *The Magic Flute*, *La Traviata*, and *Carmen*

Lyuming Xu

North America International School, Shanghai, China
xulvming@126.com

Abstract. Opera is connected to society through character behavior and the impact of characters on the plot and music. The research focuses on understanding characters through alternative perspectives and how their actions help interpret their personalities and implications for opera. The paper analyzes the characters' social statuses and power struggles that affect the meaning of emotions and intentions in the plot. The three operas, Mozart's *The Magic Flute*, Verdi's *La Traviata*, and Bizet's *Carmen*, reference the novel character analysis approach that dissects the symbolic contrary perspectives of understanding the operas.

Keywords. Music history, opera analysis, opera society, opera interpretation

1 Introduction

Opera is an artistic and complex art form representing music to the audience. Opera combined with orchestra off the stage and the singing on the stage creates a solid visual and aural genre that can be explored artistically and through research. Opera is an art form that originated in Western Italian musical culture, involving groups of musicians and performers working together, culminating in a grand performance. Opera originated in the late 16th century, first playing in courts as dramatic musical works before expanding its influence all over Europe and beyond [1]. One of opera's most significant characteristics includes sung lyrics that create meaning for the synopsis and the character's actions onstage. The development of scripts for opera characters requires a detailed examination of each character's personality, actions, and behaviors [2]. Character evaluations help opera performers and researchers understand why certain operatic protagonists and antagonists choose specific actions over others. There are many ways society can understand the meaning opera brings to music and how to interpret the synopsis performed by the singers. Each opera reflects different important meanings hidden behind the opera's relation to the culture and society [3].

There are many explanations for character behaviors in operas, depending on the different perspectives one takes to interpret the actions. A deep understanding of characters and their behaviors is essential to visualizing the opera and its meaning in society. Comparing character traits and their personalities can help analyze their behavior and influence on operas. Categorizing opera characters by personality and by their intentions helps analyze operatic meanings.

This research evaluates different perspectives of primary operatic characters and the meaning these characters present on opera's social and cultural aspects. This research does not take a sole stance on each character. Instead, this research shows how character visualization can help interpret their personalities in multiple ways.

This research looks into multiple operas in Western music history. The first example discussed Papageno and the Queen of the Night from *The Magic Flute* by Mozart. The Queen of the Night can be described both as a protagonist and an antagonist. Papageno can be described as both a brave and cowardly character. The second example from *La Traviata* by Verdi shows how Violetta's character is shaped through love dilemmas and marriage, reflecting different perspectives on her actions. The third example from *Carmen* by Bizet shows how the opera's main character is powerful and powerless in her abilities and the lack of abilities to influence the opera's plot.

2 Opera and Society

Historically, the meaning of opera to society has changed through time. Opera's place in European music history changed depending on the genre's importance in different musical cultures [4]. According to Rosselli, opera can be seen as a social occasion and an opportunity to interact socially [5]. An opera is an event focusing on music, art, and entertainment. In Vienna's early 17th and 18th centuries, opera existed for nobles and upper-class citizens. The noble class would have the best seats for the performance. In Italy, there were a few public opera houses where everyone could watch opera, although the cost of attendance was high. The opera performance started very late, and people could attend the opera after finishing work, dinner, or shopping, creating an open and relaxing operatic environment.

"At the New York Academy of Music 1865, the bulk of the audience makes for the foyers when a notoriously poor tenor is announced as an emergency replacement in *Il Trovatore*, but when he suddenly manages a loud high C many of them come back" [5]. Rosselli's comparison of the audience in the past and today shows the difference in opera's meaning between two historical points. On one side, audiences have similar ideas and behaviors about wanting unusual and surprising entertainment. For example, an uneventful opera performance may make the audience feel immersed. A high C is challenging for tenors to sing, but not impossible. The audience is focused on the performance but not the performer, meaning they are interested in the music quality. A strong performance of an aria can change the mood of the audience. Audiences' focus on the stage and singers has been crucial in all historical epochs of opera performance. Conversely, the meaning of opera and its interactions in music history have changed over time. Opera's relaxing and social atmosphere was more evident in the 17th and 18th centuries when compared to today. Throughout history, moving through the chronological timeline, operatic performance became more formal, setting up a stage of more seriousness in opera performance practice. The 17th and 18th centuries pin importance on the operatic environment, where the social element of opera can be shown on and off the stage. In modern professional opera performances, the focus is on the stage, particularly on music, synopsis, and characters. The operatic history and its environment develop two different meanings in how opera was and is perceived in societies.

The transformation between different settings and perceptions of opera in European societies occurred gradually. Vienna was known as a rich people's paradise in the 18th century. Vienna's upper class spread into the old and new nobilities, creating many problems in Vienna with the development of the middle class, which made the society complex. Integration into one society created more equality among Vienna citizens, which led to more opportunities in the opera world. With time, less importance was placed on one's social class. The new wealthy class supported Mozart, and his ideas were encouraged by the intelligentsia from the new nobility. The operas Mozart wrote were preferred by people, which shows Mozart's status in opera and society. Mozart's work also reveals what was happening in society, especially the decay of the monarchy, which is reflected in Mozart's compositional style. Dancing became a part of opera shows, bringing the audience happiness and transforming the operatic environment. The waltz, one example of a dance form, became more appealing to the audience and attracted more people to visit and watch opera performances. There were different dance styles, too, giving society more charm. Steptoe states, "Amateur musicians flourished, and publishers hastened to supply the market not only with chamber works but with keyboard reductions and variations on popular operatic pieces." [6]. This was the time when opera music became more popular. Professional and amateur music enthusiasts could attend performances, which added new audience members to the opera culture. Producing music became much more familiar for everyone, giving more chances for the opera's music development.

3 The Personality Contradictions in Mozart's *The Magic Flute*

In his opera *The Magic Flute*, Mozart hides the double meaning between bravery and cowardliness, as seen in the actions and personality of Papageno. Queen of the Night is another character who simultaneously displays the characteristics of protagonist and antagonist. Analyzing operatic characters helps understand and interpret the plot within their actions. The character Papageno in *The Magic Flute* is a talkative and kind birding man. After Papageno finds that Tamino fainted in Act I Scene I, Papageno talks a lot to Tamino about saving him and introduces himself as a bird catcher. However, he is not speaking the truth. Papageno faces many challenges with Tamino, but Papageno always stays with Tamino and saves Pamina. In the opera's music, Papageno's playing of the ascending scale on the flute creates a beautiful decoration due to the speedy notes, as shown in Figure 1, setting a comical characteristic in the opera, revealing Papageno's conversational personality [7].

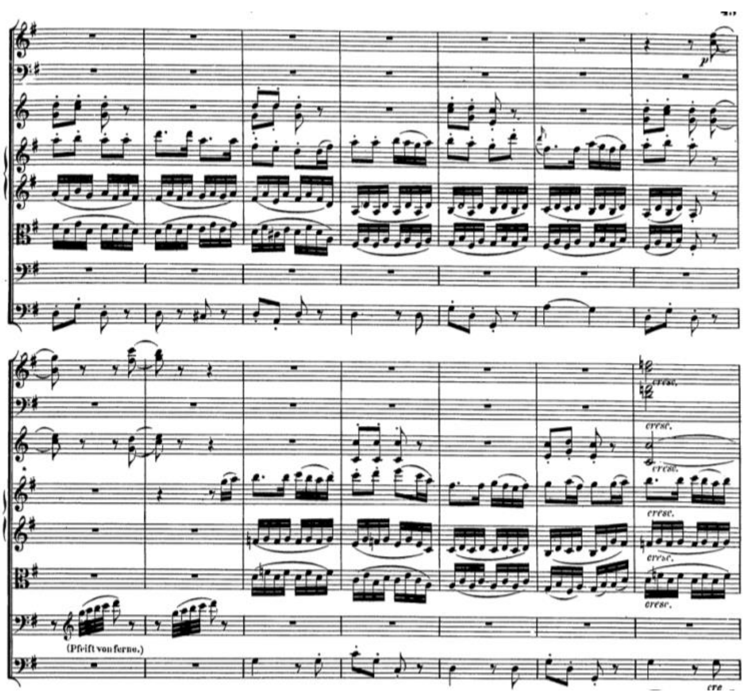



Fig. 1. Measures 6-19 from Aria No. 2 from Act I Scene I from *The Magic Flute*.

Papageno can summon many animals, showing listeners his association with nature and connections with the animal world. At the same time, Papageno's weakness is being challenged. For example, Papageno worries about challenges and fears failing to rescue the princess. At the same time, Papageno shows bravery when facing many challenges. When Papageno sees Monostatos wanting to bother Pamina, Papageno forgets everything he was afraid of before, swaying the silver bell to make Monstatos dance while saving himself and Pamina. Papageno's characteristics resonate due to his innocence. One of the charms of Pagageno's character is that Papageno tries everything

and succeeds.
When compared to Papageno, Tamino's character is more stable and responsible. Tamino is confident, thinking deeply and making many decisions before starting the challenge. In Act I Scene I, Tamino sees Pamina's picture, as the music is slow and has lyrical accompaniment, as shown in Figure 2 [7].

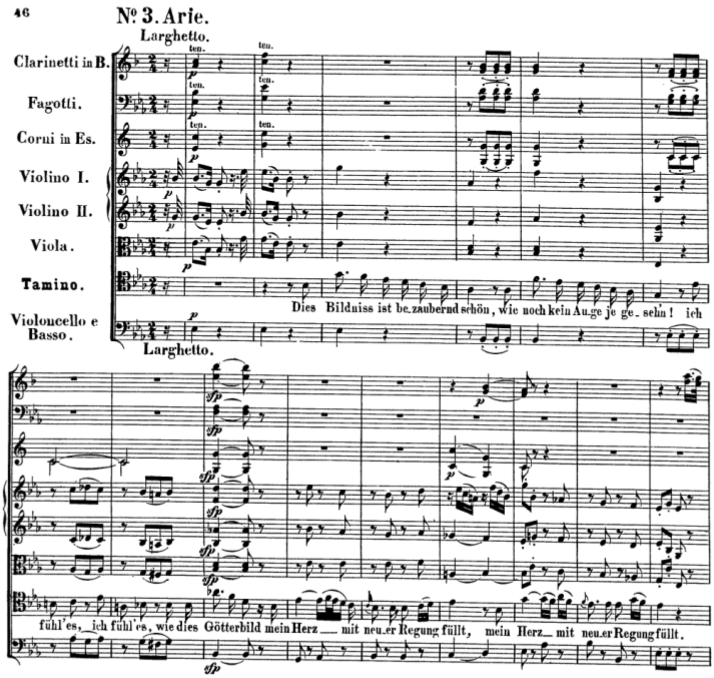



Fig. 2. Measures 1-15 from Aria No. 3 from Act I Scene I from *The Magic Flute*.

Tamino's voice resembles an instrument performed in the music, with a soft and peaceful sound, making the atmosphere very romantic. Tamino's high and bright voice integrates when the strings sustain harmony. Tamino answers by singing, and the dynamics of his voice show a powerful, determined heart.
The Queen of the Night can represent evil. At the request of the Queen, the three ladies intentionally rescue Tamino from the snake. The Queen pre-plans everything and carefully strategizes her actions in the opera. For example, if she lets Papageno tell Tamino a lie, then she can get a reason to punish Papageno. The Queen's purpose is to show how her power is in front of Tamino, thus gaining Tamino's trust. As a reward, Tamino will go to save the Queen's daughter. However, the Queen can also be considered the opera's protagonist. The Queen wants revenge, wishing to reclaim the Circle of the Sun, wanting her powers back from the priest Sarastro. "Sarastro and the

priests insinuate that death and despair is the Queen's nature as well as a punishment meted out to all who associate with her, but the Queen's reference to death and despair in 'Der Holle Rache' seems less a threat than a shriek against fate" [8]. The Queen is helpless and hopeless, wanting to express her unlucky fate and emphasize the emotion because she tries to save Pamina, who does not listen to the Queen. The Queen is suppressed and hopeless, losing everything in the process. The Queen's might and strengths are evident in the opera, but a hidden meaning shows an opposite situation. With the strengths behind her abilities, she also lacks power and control, which decreases her authority and ability to influence actions around her.

There are different identifications that Queen takes in the opera's plot, and one of her identities is being strong-powered. The Queen in the opera wants to save her daughter Pamina, but she also wants to control and limit her. When Pamina does not listen to her, the Queen is powerless. The Queen in the opera was also perceived as kind and helpful. In one example, the Queen reveals her begging character trait to ask Tamino to rescue her daughter. The Queen lacks the power to do this herself and needs others to help her, showing that she is not evil or dark.

Multiple scale variations with complex chords in different keys with compact or loose tempi make the music fluid. Music tells the emotion that happened in the story, and Brown-Montesano writes, "Gone are the emotional vacillations of the Larghetto; the Allegro expounds renewed confidence and an almost manic delight at the possibility of victory" [8]. The Queen of the Night's first aria is symbolic of the opera's plot, concentrating on the coloratura and showcasing the changes in mood and energy in the music, representing the darkness and its sudden appearance with nervous emotion, where the Queen is ambitious, as her characteristics alter from tenderness to unpowered shouting. The music is sudden and unexpected, represented by the Queen's agitated and powerful voice, reaching high notes, revealing her hopelessness and anger.

4 Violetta's Personal Dilemmas in Verdi's *La Traviata*

The opera *La Traviata* reveals multiple events about the relationship between society levels and love presented in society [9]. Violetta is displayed as a courtesan in this opera, meaning she does not hold much power. This is because the noble class sometimes controls Violetta. In terms of Violetta's character, she falls in love at the beginning of the opera and stays in true love until the very end. In Act I, in the song *Libiamo ne'lieti calici*, the music sounds fluid and romantic. Verdi uses a time signature 3/8 in the dance, implementing counterpoint and legato in the strings section, creating a deep harmonic movement. When Violetta sings "La vita è nel tripudio," translated as "Life is just pleasure," she thinks pessimistically that she fears the day when she is no longer in love. Violeta claims that her life would be finished when there is nothing worthwhile for her to live for without love. Violeta becomes a realist when facing the world, given her past living a poor lifestyle. However, Alfredo sings back, “Quando non s’ami ancora,” translated as “If one hasn’t known love,” encouraging her, showing that he does not believe her love will disappear, telling her to wait for love to make her life more pleasant. During this time, Violetta’s mind breaks away from daily life, and she starts believing in love. Her illness at the beginning of the opera tells that she was hurt deeply and found it hard to live for a long time, which gives a foreshadowing cover in the opera, making Violetta’s life troublesome. At the end of Act I, Violetta deeply loves Alfredo and is prepared to spend her life with him. On the other side, Violetta knows she has little time to live, succumbing to the idea that she has not lived enough to understand or experience the full extent of love.

Alfredo’s father asks Violetta to leave Alfredo, and Violetta knows that she is not rich enough for Alfredo. Her thoughts are more profound; Violetta respects Alfredo’s father’s decision not to allow Alfredo to marry her. Wealthy men had absolute rights in that period because Giorgio told Violetta his daughter’s fiancé did not like a courtesan. Alfredo’s father is more slant to his daughter but not Alfredo. Violetta does not have enough power compared to Giorgio’s daughter, which is why Giorgio’s daughter’s husband is trying to avoid Violetta. In the music of *Pura siccome un Angelo*, the vocal parts between Giorgio and Violetta are loud and energetic. Violetta first shows her anger, refuses to listen to Giorgio, and then feels hopeless. Violetta may also think about Alfredo’s future; Alfredo can find a better soulmate. This is also a way to express love because Violetta understands that getting married to Alfredo will make Giorgio unsatisfied. Violetta sacrifices herself and tries to save Alfredo and his family's lives, which shows her respect and what she can do for Alfredo.

In Act III, Violetta dies as the opera develops towards the climax. Violetta always cares about Alfredo. At the same time, she is deceptive because she does not want to tell Alfredo that she is going to die. Alfredo suggests they live together, even though he knows it is too late. Violetta wants to live with Alfredo but rejects his proposal to live together. Violetta’s heart becomes unstable, and she does not know what to do. Should Violetta listen to her heart or follow the noble class? On the way to death, nothing is more critical than Violetta’s ideas. She is a conscientious person and always states her thoughts. The control of the wealthy class destroyed what her life could have been. On one side, she wants to get married. On the other side, she knows she will die. Violetta’s dilemmas create a decisive inner struggle for a struggling character with a few possible action choices in the opera but many inner emotional struggles regarding her possible actions.

5 Bizet's *Carmen*: Powerful or Powerless?

Carmen is a brave and energetic woman. Carmen acts differently than other female characters in the opera, expressing her opinion and disregarding others' thoughts. Carmen's actions show that she wants to be free, as seen in her dance in the *Habanera* in Act I. *Habanera* starts rhythmically yet gracefully. Initially, the vocal part has a repetitive rhythm, creating mysterious and robust emotions based on its motivic structure. The song *Habanera* holds a different mood compared to other music inside opera. The music of *Habanera* has a fluid sound, freeing it from the musical perspective and revealing Carmen's true personality. Carmen interacts with many men, showing her confidence, acting boldly, and letting the characters in the opera around her realize her allure. Carmen publicly conveys her ideas. This characteristic is appreciated because, in the opera, these women are portrayed with more traditional traits or as workers in a factory. Carmen has more power and bravery compared to other women in the opera. Carmen exaggerates in conveying her ideas about love and finding a husband. The lyrics, "Love is a rebellious bird," deeply state that Carmen thinks there are no rules when facing love, which is how she expresses her individuality and resistance to life.

Carmen shows her decisive and assured personality when Don Jose is sent to lock up Carmen in Act I. Carmen can acquire Don Jose's trust, even as a powerless character. This shows the weaknesses that Jose possesses. Don Jose listens to what others say, falling for Carmen's charm, which is a turning point in how Don Jose becomes immersed in Carmen. Don Jose gradually stops following the rules, revealing the contradiction of personality between Carmen and Don Jose. Carmen tries her best to reach her goals in any way possible. Don Jose's job as a soldier includes listening to what his officer says, meaning that Don Jose should be a good listener and respond to directions. Don Jose cannot control himself and starts losing his power toward Carmen. The subsequent part confirms this when Don Jose returns from jail because he released Carmen. Carmen shows that she does not care whether Don Jose is doing well in prison because she already reached her goal of escaping. In Act III Scene III, in the aria *La fleur que tu m'avais jetée*, the flower symbolizes Jose's love and loyalty. Carmen starts to think Jose is an enormous burden because Jose becomes increasingly immersed in her, but Carmen hates the behavior of Jose's control because she believes love is free and starts to become bored with Jose. This shows Carmen is strong and does not feel guilty for her wrongdoing. Don Jose sacrifices a lot, as indicated by the flowers he gave to Carmen, to show his loyalty to her. Don Jose gives up the rules he must follow as a soldier and betrays his officer.

In Act III, Frasquita and Mercedes turn the cards for Carmen's fate. Carmen seems to believe that card turning is influential. Carmen's fate, as shown on the card, includes death. Destiny does not stop Carmen because she indicates her acceptance of what is happening and does not fear the future. The cards can symbolize the events happening around Carmen, foreshadowing her destiny. At the end of Act IV, Carmen dies, showing her looking for freedom and perceiving her ideas as a higher priority than her life. While Carmen is physically dead, her soul is free. Don Jose can not accept that Carmen has fallen in love with another man. It is not surprising that Don Jose kills Carmen. Don Jose shows the loss of control of the situation after he kills Carmen, which offers a bright comparison between these two main opera characters, who are opposites in their characteristics throughout the plot.

6 Conclusions

The opera reveals the connections between society and culture. The art of opera, combined with musical innovations developed by composers, transforms the characters' meaning and how these characters influence the genre [10]. Opera contains complex connotations. One-sided aspects of character analysis show little to opera analysis. Multiple perspectives of characters and their actions offer more depth into operatic analysis and compositional intentions of those involved in the opera production [11]. The paper focused on detailed analysis and interpretation of primary opera characters, conveying different perspectives and reflecting corresponding backgrounds based on operas' synopses. Contrasting interpretations open varied cultural and societal understandings of how characters shape the opera and how their actions and behavior affect the plot.

References

[1] G. R. Martin, "The Role of Culture in Global Structural Transformation: Opera and the Baroque Crisis in Seventeenth-Century Europe," *International Political Science Review*, vol. 18, no. 2, pp. 153-166, Apr. 1997.

[2] Z. Hongrui, "Thinking and Analysis of Characterization in Opera Performing Art," *Art and Performance Letters*, vol. 3, no. 2, pp. 68-72, 2022.

[3] P. Ther, *Center Stage: Operatic Culture and Nation Building in Nineteenth-Century Central Europe*, Purdue University Press, 2014.

[4] V. Kotnik, "The Adaptability of Opera: When Different Social Agents Come to Common Ground," *International Review of the Aesthetics and Sociology of Music*, vol. 44, no. 2, pp. 303-342, Dec. 2013.

[5] Rosselli, "Opera as a Social Occasion," in R. Parker, *Oxford Illustrated History of Opera*, Oxford University Press, pp. 304-321, 1994.

[6] Steptoe, *The Mozart-Da Ponte Operas*, Oxford University Press, 1988.

[7] W. A. Mozart, *Mozart's Werke, Serie V*, Breitkopf & Hartel, 1879.
[8] Brown-Montesano, *Feminine Vengeance II: (Over)Powered Politics*, Oxford University Press, 2007.
[9] E. Dillon, "Violetta, Historian Verdi, 'Sempre libera' (Violetta), "La traviata", Act I (1853)," *Cambridge Opera Journal*, vo. 28, no. 2, pp. 191-197, 2016.
[10] Williams, and M. Tanner, *On Opera*, Yale University Press, 2016.
[11] Ingraham, J. So, and R. Moodley, *Opera in a Multicultural World: Coloniality, Culture, Performance*, Routledge Research in Music, 2015.

Introducing Conceptual Simplification: How to Simplify Complexity when Analysing, Learning and Memorising Post-Tonal Piano Music

Laura Farré Rozada

Royal Birmingham Conservatoire, Birmingham City University
laura.farrerozada@mail.bcu.ac.uk

Abstract. There is a gap in music performance, education and psychology in terms of memorisation training for post-tonal piano music. Despite the repertoire spanning over 100 years, pedagogues and professionals still lack effective tools for developing this skill. Existing research on this domain is mostly focused on observing practitioners' behaviours during practice, to understand how these prepare for a memorised performance of a selected repertoire. However, a systematic method for effective memorisation is not provided. This paper discusses a new method for analysis, learning and memorisation of post-tonal piano music, named Conceptual Simplification, which was developed, tested and formalised with my PhD thesis. This presents a novel implementation to musical memorisation building on certain areas of mathematics and computer science to improve human memory and musical performance. However, Conceptual Simplification does not require any previous scientific training to be successfully implemented and works for different learning styles and types of complexity. This method could also be adapted to other instrumentalists, singers and conductors; and musical genres; and presents enough flexibility for other practitioners to incorporate additional strategies, adapting it to their needs accordingly. Finally, Conceptual Simplification can also assist in preventing performance anxiety through greater confidence and reducing the potential for injuries that usually result from repeated practice. The method's systematic approach toward engaging conceptual memory and reasoning leads to more confident memorised performances, while needing less repetition during practice.

Keywords. analysis, learning, memorisation, method, post-tonal piano music.

1 Introduction

This paper discusses some of the main outcomes of my doctoral research at the Royal Birmingham Conservatoire, which focused on testing, extending and formalising Conceptual Simplification [1]. This is a new method for memorising post-tonal piano music that I developed over the years with my own experience as a pianist and mathematician.

My interest in researching musical memory started when I was an undergraduate student in Barcelona. Back then, I was simultaneously pursuing a bachelors in piano performance and another in mathematics. Given that I had a very solid schedule, I needed to be very efficient in my practice. Therefore, I started experimenting with using mathematical strategies in my piano performance for practising, but especially for memorising. Later on, in my Master's thesis at the Royal College of Music in London, I organised and further developed this pool of strategies, under a first prototype of Conceptual Simplification [2]. However, beyond my interest and background, it is reasonable to use mathematical strategies for musical memory, because mathematics is one of the best tools for solving problems and for identifying patterns [3]. Moreover, that is what musicians do all the time: they are constantly trying to find patterns in the music and are also trying to solve different kinds of challenges [4]-[6]. Furthermore, given that post-tonal music presents less standard tonal patterns, because composers develop their own composition principles [6]-[7], mathematics can also help in simplifying the score, to better identify such new patterns [1].

2 Research Background

Existing research on musical memory is quite extensive, but it has mostly focused on observing how musicians practise and memorise, to determine whether some strategies are more efficient than others (e.g., [4], [8]-[12]). However, a systematic method on how to effectively memorise has not been proposed (e.g., [4], [8], [9], [13]), particularly for post-tonal music ([6], [10]-[12]). Additionally, most soloists, especially pianists, are required or expected to perform from memory during their studies and at professional level, but memorisation is not a topic frequently taught or discussed at conservatoires [14]. This is partly due to an existing gap in music performance, education and psychology in terms of how memorisation should be trained [15]-[17]. But also, because research findings do not always transfer or are pedagogically implemented at conservatoires ([14], [15], [17]). Consequently, memorisation is still a taboo with which performers struggle with, leaving musicians to find their own ways for achieving this goal (e.g., [10], [11], [13], [15]), which are not always effective under pressure or within tight deadlines ([8], [10], [12], [13]). Also, this is one of the main reasons why post-tonal music is likely to be performed from the score, since regular memorisation strategies (namely, using traditional harmony and standard patterns) are not always applicable ([6], [10]-[12]). Therefore, Conceptual Simplification, which is briefly discussed in this paper, aims to provide a systematic method for analysis, learning and memorisation.

3 Methodology

The doctoral thesis *Conceptual Simplification: an Empirical Investigation of a New Method for Analysis, Learning and Memorisation of Post-Tonal Piano Music* (2023) presents the findings of testing Conceptual Simplification with other pianists and myself, with a substantial body of repertoire [1]. For this purpose, I completed:

- A couple of Self-Case Studies with myself as practitioner.
- Interviews with professional performers specialised in post-tonal music.
- Studies with recruited participants, who mostly consisted of advanced piano students.

The main outcome of this research was a three-stage method, as shall be later detailed. However, first it should be explained the main philosophy behind how Conceptual Simplification operates.

4 Understanding Mathematical Thinking

Given this problem below, there are different ways in which this could be approached:

$$1 + 2 + 3 + \cdots + 100 = ? \qquad (1)$$

In a straightforward manner, this sum could be solved by adding one term after the other, until reaching the result. Despite this eventually leading to solving the problem, it is not the most efficient procedure. Alternatively, a pattern could be sought within this sum, noticing that by successively grouping in pairs the biggest and lowest terms, these always total 101:

$$\begin{gathered} 1 + 100 = 101 \\ 2 + 99 = 101 \\ 3 + 98 = 101 \\ 4 + 97 = 101 \\ \ldots \end{gathered} \qquad (2)$$

Hence, the original problem simplifies into adding 50 times 101, which is the same as multiplying 50 by 101 [18]. However, this is not only an effective strategy for solving the problem faster: the time initially spent analysing the problem permits developing a deeper understanding of it, leading to higher proficiency in solving similar problems in the future. Consequently, the original problem can be generalised for all positive integers n:

$$1 + 2 + 3 + \cdots + n = \frac{n}{2}(n + 1) \quad (3)$$

Furthermore, by understanding the problem's underlying patterns and finding an effective problem-solving strategy, there is no need to memorise the formula as such. This can be forgotten and deduced when needed, by reconstructing the original steps. Accordingly, only one key idea must be remembered: how the terms of this sum are to be paired.

5 Main Issues on Current Memorisation Strategies

It may be puzzling how human memory can relate to mathematics, but this example illustrates the main concern in how musicians approach a new piece: analysis, learning and memorisation are typically regarded as linear problems. On the one hand, less experienced musicians (i.e., novices) tend to memorise through mindless repetition [15], exclusively relying on Sensory Learning Styles: i.e., kinaesthetic memory, aural memory and visual memory [16]. Therefore, memorisation results from repeated practice ([5], [15]). On the other hand, expert musicians follow a more analytical approach that engages conceptual memory and implements problem-solving strategies for achieving certain goals (e.g., [12]). Hence, in contrast with novices, experts effectively use their knowledge of tonal patterns to encode music and memorise by triggering these familiar entities (e.g., [13], [16]).

Nevertheless, when experts deal with post-tonal music's unfamiliar languages, this approach becomes more challenging and time-consuming (e.g., [10]), since composition principles do not necessarily concur, making it harder to develop common codes [6], [7]. Within this context, these practitioners can lose some of their advantage in respect to novices, at being exposed to an unknown framework with less evident ways to proceed [6], [10], [11]. Thus, a frequent strategy used is fitting the music within a tonal framework, if possible, to restore some familiarity (e.g., [10], [12]). However, this process is slow and difficult: patterns are not explicit as in tonal music, but content is reinterpreted according to these, significantly increasing the time investment that an equivalent tonal context would require [10]-[12]. Furthermore, the more challenging a musical work is, the higher the tendency in linearly segmenting it in smaller units than usual. Therefore, practice becomes mostly ruled by a linear understanding of the music, according to the structure identified, leading to a time-consuming approach: the smaller these are, the more time needed (e.g., [10]). This is why Conceptual Simplification reduces complexity differently, allowing to work with bigger chunks [1].

6 Conceptual Simplification's Paradigm

On a larger scale, Conceptual Simplification uses a series of paradigms from computer science for algorithm optimisation. On a smaller scale, the method's pool of strategies is informed by mathematical thinking and based on problem-solving techniques that are frequently used in number theory, geometry and group theory. Discussing these is out of the scope of this paper, but further details can be found in my doctoral thesis cited earlier [1]. Still, it is worth reviewing the reasons for relating human memory to algorithms, since humans and computers think differently and excel at different tasks ([3], [5], [19]).

Computers have limited memory resources and need to complete tasks within a reasonable time [19]. Similarly, musicians have a limited working memory capacity, and finite time for practising conditioned by deadlines ([4], [13]). Therefore, Conceptual Simplification's main principle is the following:

What essential information do I need to memorise to remember this music?

Accordingly, the method identifies and encodes the least amount of information needed to learn and memorise effectively. It also seeks ways of triggering, deducing or reconstructing the content of a passage, through a series of conceptual clues or instructions. That is translating into music the same procedure used for solving the mathematical problem discussed earlier. All this is possible, because Conceptual Simplification proceeds by slicing into layers of complexity a musical score, to scaffold learning and memorisation. Consequently, the practitioner is always comfortable with the amount of difficulty involved, without tackling more information than can be successfully managed or internalised [1].

7 Conceptual Simplification's Overview

In general terms, Conceptual Simplification comprises three main steps:

1) The first step of the method is a *Triage*, which is an initial stage to become acquainted with the music and identify what strategies could be useful for facing the challenges presented.

2) The second step is *Simplifying Layers of Complexity*, which proceeds to slicing complexity by layers, while dealing with bigger chunks that are easy enough to

approach. This process is done preferably mentally using the piano, but it can be written down too, if that's more helpful.

3) Finally, the third step is *Conceptual Encoding*, which is like the reverse process of Simplifying Layers of Complexity. Basically, Conceptual Encoding consists in restoring layers of complexity once a certain modified version of the musical work has been successfully internalised.

8 Triage

As anticipated, the Triage consists in developing a first impression of the music, while identifying its main challenges and how these can be best tackled with the available pool of strategies. This purpose is achieved by implementing several mental and physical strategies as listed below. But, as the literature has consistently highlighted, it is very important to spend some time at the beginning before solving a problem, to understand its nature and logic (e.g., [4]), which is exactly what was done earlier with the mathematical problem (1)-(3). Hence, the proposed strategies for this purpose are:

(1) Score Overview
(2) Listening to recordings
(3) Sight-reading (as opposed to sight-playing)
(4) Decision making on fingerings and hand arrangements
(5) Formal analysis
(6) Assessment of main challenges
(7) Decision making on potential effective strategies

9 Simplifying Layers of Complexity

After becoming acquainted with the piece, the method proceeds to slice complexity, by identifying and removing those elements that are an obstacle for learning or memorisation. The set of information that is removed each time with this procedure is a *layer of complexity*. For this purpose, the simplifying strategies are classified according to four main elements, which are: pitch, harmony, rhythm and context.

(1) **Pitch:** pitch, octaves.
(2) **Harmony:** voicing, chords, hands.
(3) **Rhythm:** rhythm, repetition, tempo.
(4) **Context:** extended techniques, structure, preceding structure.

10 Conceptual Encoding

Once a certain version of the music is successfully learned and memorised, the method proceeds to restore a layer of complexity, to make it slightly more difficult. Hence, Conceptual Encoding consists in encoding and practising the resulting patterns each time, using the following strategies:

(1) **Pitch:** interval conceptualisation.
(2) **Harmony:** chord conceptualisation.
(3) **Rhythm:** *solkattu* vocalisation and clapping, rhythm conceptualisation.
(4) **Context:** pattern, switches and dynamics conceptualisation.

After providing an overview of the three stages of Conceptual Simplification [1], I proceed to discuss its implementation on David Lang's *Cage* (score to be retrieved from [20]).

11 How to Implement Conceptual Simplification?

Implementing Conceptual Simplification to a musical work produces versions of reduced complexity by removing layers of information to enhance understanding. Once that amount of difficulty is assimilated, it can be slightly increased by restoring removed layers [1]. The piece *Cage* by American composer David Lang (b.1957) is a clear example of one extreme of complexity that involves uniformity and self-referencing (score to be retrieved from [20]). Thus, the main difficulty of this piece is switches, which are places that look similar, but resolve differently, and can be a real challenge for memory [13].

As an example, focusing on bars 1-30 of this piece, four elements need to be tackled: uniformity, hand coordination, independent melodies and changes of register. These are respectively identified as four layers of complexity, namely: repetition, rhythm, hands and octaves. Hence, these shall be temporarily removed and restored.
To implement Conceptual Simplification to this passage, the method starts by identifying what information is essential: here, repetition and rhythm have an ornamental role. Therefore, these are removed, obtaining two independent melodies. Then, two further layers of complexity are identified: combining both hands and playing the notes in different octaves. Hence, both hands are examined independently. Also, by removing the octaves, it is easier to see the underlying pitch-sequence of each hand. However, these might have not been spotted, if memorising this piece by repetition or by how it sounds ([5], [15]-[17]). In this latter case, possibly leading to a less confident performance, due to the piece's uniformity and self-referencing ([1],

[10], [13]). Once this is clear and memorised, the layers of hands and rhythm are restored, focusing now on learning how both hands interact within the middle register of the piano.

After this, the rhythm is removed, to learn each hand's pattern in the original octaves. Then, the same process is repeated: the combination of both hands and the rhythm is restored, but this time in the original octaves. Finally, once this is clear, it is simply a matter of adding the repetition to internalise the original excerpt. Consequently, this excerpt can be confidently learned and memorised, despite being misleading due to switches ([10], [13]). Thus, it is just a matter of repeating this same process as many times as needed.

Therefore, implementing Conceptual Simplification does not imply ignoring how the piece sounds, looks or feels in the physical sense: using a conceptual approach allows to organise the rest of senses and memory resources more effectively. This way, memorisation does not exclusively rely on the Sensory Learning Styles [16], but regard these as complementary [1]. Also, repetition is still used as an overlearning strategy: despite using Conceptual Simplification, mental and physical run-throughs will be needed. However, these will be much more effective, probably needing less repetition to achieve the same result.

12 Conclusion

As a summary, the main goals of Conceptual Simplification are:

1) Scaffolding analysis, learning and memorisation into different progressive stages, to prompt always being comfortable with the amount of difficulty to deal with, not taking more information that can be successfully processed at once [21]. Also, even if the final goal is not performing from memory, this method is also very effective for preparing a confident performance from the score.

2) Once confident with a certain amount of difficulty, this is slightly increased by adding a new layer of complexity. Thus, working by layers enhances the understanding of the piece.

3) For difficult pieces, such an approach could lead to more fluent and convincing performances. Particularly, for post-tonal music, this could make a difference to audiences that are less familiar or eager of this repertoire.

4) Internalising music by scaffolding complexity instead of through repetition can help in preventing performance anxiety and injuries ([5], [13]). This procedure could also be adapted to the needs of other instrumentalists, besides pianists.

However, so far, this method has only been tested with a limited number of practitioners and repertoire [1]-[2]. Since musicians have different learning styles, I designed Conceptual Simplification in a way that is flexible enough for other practitioners to include further strategies, according to their needs [1]. But it is not certain that such an approach will work for everyone. Still, Conceptual Simplification:

1) Provides a new method for analysis, learning and memorisation.

2) Simplifies complexity, not necessarily proceeding in a linear way.

3) The method is flexible, and each strategy can be used on its own, or in combination with others, without needing previous expertise on a certain musical genre or composer.

Presents a novel implementation to musical memorisation, building on certain areas of mathematics and computer science to improve human memory and musical performance [1]. However, it does not require any previous scientific training to be successfully implemented and works for different learning styles and types of complexity.

References

[1] Farré Rozada, L. (2023) *Conceptual Simplification: an empirical investigation of a new method for analysis, learning and memorisation of post-tonal piano music*. PhD Thesis. Royal Birmingham Conservatoire, Birmingham.

[2] Farré Rozada, L. (2018) *Memorising George Crumb's Makrokosmos I: exploring new strategies for non-tonal music*. MMus Thesis. Royal College of Music, London.

[3] Sáenz de Cabezón, E. ([2016] 2020) *[Mathematical intelligence] Inteligencia*

matemática. Barcelona: Plataforma Editorial.

[4] Chaffin, R., Imreh, G., Lemieux, A. and Chen, C. (2003) “Seeing the big picture”: piano practice as expert problem solving. *Music Perception*, 20(4), pp. 465–490.

[5] Sloboda, J. A. (1985) *The musical mind: the cognitive psychology of music*. Oxford: Clarendon Press.

[6] Thomas, J. P. (1999) *Interpretative issues in performing contemporary piano music*. PhD Thesis. University of Sheffield, Sheffield.

[7] Auner, J. (2017) *[Music in the twentieth and twenty-first centuries] La música en los siglos XX y XXI*. Madrid: Ediciones Akal.

[8] Chaffin, R. and Imreh, G. (1997) “Pulling teeth and torture”: musical memory and problem solving. *Thinking and Reasoning*, 3(4), pp. 315–336.

[9] Chaffin, R., Lisboa, T., Logan, T. and Begosh, K. (2010) Preparing for memorized cello performance: the role of performance cues. *Psychology of Music*, 38(1), pp. 3–30.

[10] Fonte, V. (2020) *Reconsidering memorisation in the context of non-tonal piano music*. PhD Thesis. Royal College of Music, London.

[11] Soares, A. (2015) *Memorisation of atonal music*. DMus Thesis. Guildhall School of Music and Drama, London.

[12] Tsintzou, T. and Theodorakis, E. (2008) Memorization strategies of atonal music. *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08)*. Thessaloniki, Greece, 3-6 July 2008, pp. 1-10.

[13] Chaffin, R., Imreh, G. and Crawford, M. (2002) *Practicing perfection: memory and piano performance*. New Jersey: Erlbaum.

[14] Jónasson, P. and Lisboa, T. (2016) Shifting the paradigm: contemporary music, curriculum changes and the role of professional musicians as researchers. In: E. K. M. Chong, ed. *Proceedings of the 21st International Seminar of the ISME Commission on the Education of the Professional Musician*. Saint Andrews, UK, 20-23 July 2016. Saint Andrews: University of St. Andrews, pp. 78–92.

[15] Ginsborg, J. (2004) Strategies for memorizing music. In: A. Williamon, ed. *Musical excellence: strategies and techniques to enhance performance*. Oxford: Oxford University Press, pp. 123-141.

[16] Mishra, J. (2005) A theoretical model of musical memorization. *Psychomusicology*, 19(1), pp. 75–89.

[17] Mishra, J. (2010) A century of memorization pedagogy. *Journal of Historical Research in Music Education*, 32, pp. 3–18.

[18] Meavilla, V. (2021) *[Human calculators: biographies, feats, and tricks of the great mental calculators] Calculadoras humanas: biografías, hazañas y trucos de los grandes calculadores mentales*. Córdoba: Guadalmazán.

[19] Cormen, T. H., Leiserson, C. E., Rivest, R. L. and Stein, C. (2009) *Introduction*

to algorithms (third edition). Cambridge: The MIT Press.

[20] Lang, D. (1992) *Memory Pieces*. New York: Red Poppy.

[21] Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. and Wittrock, M. C. (2013) *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. London: Abridged Edition-Pearson.

Telescola (TVSchool): Music Education and Media in Portugal from Dictatorship to Democratization

Baltazar, Ângela Flores[1]

CESEM-IN2PAST/ FCSH-NOVA University
angelafbaltazar@gmail.com

Abstract. *Telescola* (TVSchool) was a television program broadcast (and later recorded on VHS) in Portugal between 1964 and 2003. As a public initiative, its primary goal was to expand access to education through mass communication technologies across the Portuguese territory. This endeavour significantly contributed to reducing illiteracy rates in Portugal and fostered the cultural development of the population. During the Estado Novo dictatorship, this project served as a democratizing tool for mass education, which remained ideologically relevant after the Portuguese Revolution of 1974. This article primarily explores the subjects dedicated to developing musical skills that were provided in this TV program - musical education, musical listening, and choral singing. These disciplines mainly relied on the practice of songs for the development of rhythmic, harmonic, and melodic skills, and to cultivate abstract thinking that would be useful for other scientific disciplines.

Keywords. TVSchool, television, music education, mass communication.

1 Introduction

This paper is part of my PhD research project, which primarily focuses on TV programs that promoted the so-called "classical music" to a broader audience in Portugal during the 1960s and 1970s. The main goal of these programs was to provide access to the classical music genre, that was considered elitist - ideally reaching people who didn't have access to it due to social, geographic and economic reasons -, and television was a great medium to do it. Therefore, to cultivate this kind of appreciation among the Portuguese population, it was important to integrate it into mass education and mass communication dynamics. These TV programs, especially *Telescola* (TVSchool), were part of major mass efforts to educate the Portuguese population in the 1960s, of which cultural stimulus and musical skills development were a small part of it.

[1] Financial co-operation: CESEM UIDB/00693/2020, IN2PAST LA/P/0132/2020 and FCT 2023.05001.BD

Telescola was a television program broadcast by the National Broadcaster (Emissora Nacional) and recorded on VHS in Portugal between 1964 and 2003. As a public initiative, its main goal was to endorse the use of mass communication technology to expand access to education, particularly across various regions of Portuguese territory, aiming to improve literacy levels and intellectual development of the population. The cultural stimulation took a part in these programs, as a way to develop transversal skills that would be important to the performance in other scientific disciplines and to engage the educational activity in the state propaganda.

Even though, the political propaganda bias did not dominate this democratizing action. As we will see later on, the *Telescola* was a project that was flexible and adaptable to various ideological positions. It had a diverse origin that incorporates visions of nationalism, globalization, democratization, and propaganda, which are articulated with the political changes in Portugal, especially in the transitions before and after the Estado Novo [1]. The inquiry into music education classes in *Telescola* is particularly relevant for this kind of analysis because those activities were institutionally designed to pursue an ideal of community through the collective practice of songs, therefore, it is more likely that they convey political ideals.

Accordingly, this paper begins by focusing on the institutional context of *Telescola* and how it operates within this ideological flexibility and will then proceed with an analysis of the subjects of Music Education, Musical Listening, and Choral Singing. This research was primarily based on the analysis of *Telescola* guidelines bulletins from the first phase of transmissions (1965 to 1968), which were distributed throughout various official transmission spots. These bulletins provided insight into how objectives were structured in each lesson, how personnel on the ground were managed, and also included some notes on the national-level project progression. Although there is limited literature on the subject, this data has been cross-referenced with other sources, including interviews with project participants.

2 To educate massively

During the dictatorship of Oliveira Salazar in Portugal, high levels of illiteracy and increasing migration posed significant challenges to developing a skilled workforce. Despite the regime's desire to maintain intellectual control over the population, the government introduced changes in the national education system and policies to enhance labour productivity. The development of the school system during Estado Novo operated under a positivist and critically limited framework, focusing primarily on literacy and mathematics. Nevertheless, culture was an important part of propaganda and the ideological re-coding of national values through national education and mass communication [2]. As noted by Fernando Rosas, educational action was integrated as

a means to reformulate nationalist standards and re-educate towards the "politics of the spirit" [3] within the context of nationalist propaganda and for the benefit of the nation. This ideological reformulation aimed to bring about cultural change by re-educating elites and reproducing them in social hierarchies. Additionally, it sought to initiate mass cultural movements that would be driven by the "spirit of true national interest" [4]. According to Rosas, this re-coding of "national values" is evident in the "education of the masses" [5] of Estado Novo. As he stated: "However, there was also a decisive investment in the formation of the masses, not only in terms of conforming them but also in morally and spiritually educating them, especially in the values of a 'popular culture', national-rural, ethnographic, and corporatist, created by national propaganda and conveyed at the level of urban popular media and in the rural world by F.N.A.T. and the Central Board of People's Houses (J-C-C-P-). This action was complemented by the investment in 'national education' in the massification of primary education, albeit marked by a minimalist and integrative logic" [6].
The mass mobilization of educational resources is something that Valter Lemos also refers to as the main characteristic of the modernization and democracy throughout the 20th century, whose relationship with increasing globalization is essential to understanding the development of tendentially universal systems [7]. Thus, even during a regime that was entirely nationalist at an ideological level, international relations and globalization played a significant role in this mass movement in education. In addition, the renewal of educational structures and the expansion of their reach promoted a democratizing trend.
Therefore, to understand the involvement of educational policies, it is necessary to understand the politics of external cooperation in a post-war period and the gradual opening of Portugal to the international scope. International relations after the World Wars intensified a climate of cooperation and the creation of transnational models seeking recovery solutions, especially at an economic level. Alongside a conservative and nationalist regime, Portugal inevitably absorbed those international models at an economic, political, and cultural level that resulted in the social transformations characterizing the 1960s. As Fernanda Rollo states in this excerpt: "The signs of failure in the first phase of reconstruction and the recognition of the ineffectiveness of national responses led European countries to look for new solutions; the most viable, the one that at the outset seemed to ensure the best conditions for success, was undoubtedly international cooperation which, in a way, made it possible to revive the old forms of cooperation that had been temporarily abandoned and, in some more ambitious minds, the idea of European unification" [8].
That said, the O.E.C.D. cooperation was a part of this international involvement. As a consequence of the Marshall Plan, the Organization for European Economic Cooperation (O.E.E.C.) was founded in 1948 to manage international cooperation resources for economic recovery after the Second World War [9]. Subsequently, in

1961, the Organization for Economic Cooperation and Development (O.E.C.D.) was founded to assume the functions of the former O.E.E.C. [10]. For this essay, amid the complexity of transnational policies in the post-war period, it is important to understand the involvement of the O.E.C.D. in education policies and how they initiated some democratizing dynamics also at a cultural level. As David Lemos stated: "The democratization of access to education and the construction of mass schooling in the so-called "Western countries," that is, the liberal economies that constituted or integrated the O.E.C.D., took place after the war, coinciding with the creation, growth, and action of the organization itself. Considering the important role it undoubtedly played in the public education policies of these countries, the analysis of its action in the case of Portugal cannot fail to be an interesting object of study, since the country has been a member of the organization since its foundation" [11]. In this context, educational policies were operationalized to drive economic development and cultivate a skilled workforce. The cultural functions were intricately linked to the notion that educating the people, intellectually engaging them and disseminating the ideological principles of the regime were essential for enhancing national productivity levels. The first revision of the O.E.C.D. involved countries including Spain, Greece, Italy, Turkey, and Yugoslavia [12].

Due to the high illiteracy rate in Portugal, which was one of the highest in Europe, it was essential to increase access to technical and vocational education. It was necessary to reform the current system and create plans that motivated the training of both children and adults. To achieve the goals of the O.E.C.D, the National Education Board (Junta de Educação Nacional - J.E.N.) was restructured. The J.E.N. was an institution that had been established during the First Portuguese Republic (1910-1926), thus having this democratic ideological basis. Regarding the institutional history of the J.E.N., Fernanda Rollo emphasizes the ideological flexibility of this institution among various political transformations in Portugal—from the First Republic to the Estado Novo and subsequently to the beginning of democracy: "A history in which Science, Culture, and Language shaped a political project, surviving various circumstances, empowered by different institutional formalizations and supported by sometimes fundamentally different ideological formulations, from "Public Instruction" to "National Education," from "scientific culture" to "cultural expansion," between logics of "erudite culture" or "scientism" and "commissioned industrialism," between "pure science" and "applied science"; in short, between diverse, successive, or overlapping policies—rhetorics and agendas at play, in a coexistence or confrontation of powers that reveal, in reality, different facets of one of the main characters of Western civilization" [13].

In 1952, the J.E.N. was renamed the Institute of High Culture (Instituto de Alta Cultura I.A.C.) to mediate the Portuguese "cultural life" [14], and to “contribute to the development and improvement of higher culture and artistic culture, scientific research, cultural relations with foreign countries, and the dissemination of Portuguese language and culture” [15]. The Institute of High Culture was tasked to explore these changes in the national system, and it was through this institute that research began on the optimal utilization of new communication technologies for mass education. Following the first television broadcast in Portugal, the Center for Audiovisual Pedagogy Studies was established within the administrative structure of the Institute of High Culture on December 9, 1963, to “promote the study and expansion of modern teaching processes and educational action through audiovisual means” [16]. These functions included scientific research on technological means, hosting scientific events, devising strategies for pedagogical application, human resources training, creating audiovisual content, and archival organization.

The Institute of Audiovisual Media in Education (I.M.A.V.E.) was then established in 1964 “under the direct authority of the Minister of National Education” [17] to collaborate with the Center for Audiovisual Pedagogy Studies and exercising strategies for the application of audiovisual means in education. I.M.A.V.E. aimed to “promote the use, expansion, and improvement of audiovisual techniques as auxiliary means and for the dissemination of teaching and the cultural enhancement of the population” [18] and to promote educational television, radio programs and other educational and cultural content [19].

Consequently, on December 31, 1964, *Telescola* was established under the I.M.A.V.E. structure, to "serve as a framework for various school radio and television courses" [20]. This educational service was meant to function as a "support or extension of courses taught in official educational institutions", expanding the dissemination of educational action and pedagogical content through audiovisual mediums. This TVSchool model was based on the Italian *Telescuola,* which was created under similar institutional circumstances, and influenced by the O.C.D.E. examination. It was broadcast for the first time in 1958 by the Radio Televisione Italiana (R.A.I.) and included the three years of the official high school program [21].

Telescola planning was divided into two services. The first service corresponded to the television programmes of the "Unified Course", the "Adult Education Course", "Educational Television" and also the broadcasting of "Radio School" [22]. The Unified Course included "the teaching of the subjects of the preparatory cycle of Professional Technical Education and, as for French, of the first High School Education" [23], while the other television programmes were mainly dedicated to adult education and the free dissemination of topics that were not directly linked to the school plan. This educational activity was complemented by "Radio School" (Radio Escolar), which was dedicated to extracurricular activities that were linked to some of the subjects on the television programmes, such as reciting texts on Portuguese, history, etc.

The activities of the second service consisted of investing in and organizing a media library to support the activities of the first service.
The Nacional Broadcaster (Emissora Nacional) recorded the classes given by the teachers, and the monitors operated the sessions *in site*. The broadcasts on the official spots of *Telescola* were interspersed by the monitors with three moments: the "preparation period" for getting the students up to speed before each lesson, the "immediate exploration" for revising the subject after each broadcast and the "mediate exploration" for a general revision and rectification of the subject given throughout the day [24]. Radio School was broadcast from 10.30 a.m. to 3.15 p.m. from Monday to Friday (except Thursdays) and the Unified Course of *Telescola* was broadcast from 3.15 p.m. to 6.30 p.m. from Monday to Friday, and from 3 p.m. to 5 p.m. on Saturdays [25].

3 Music classes on *Telescola*

Although educational action in Portugal was predominantly positivist in the 1960s, music education activities were integrated into the radio and television transmission plans of I.M.A.V.E. Radio School (broadcast) two weekly music activities - Music Education, Musical Listening - and Choral Singing were broadcast in the Unified Course on television, which was supposed to complement the other practical musical activities.
The Music Education subject was broadcast on the Radio School of the Portuguese National Broadcasting (Emissora Nacional) focusing mainly on ear training to develop components of "rhythm, melody, and harmony" [26]. Two different classes were broadcast throughout the week: Series E for 1st and 2nd graders, and Series F for 3rd and 4th graders. The investment in these components is described as a means of self-development in an almost "poetic" articulation as mentioned in the I.M.A.V.E. Bulletins: "1 - Modern musical education is based on the psychological relationships established among the three elements of music: rhythm, melody, and harmony with the three aspects of human nature - physiological life, affective life, and mental life. 2 - Rhythm is ordered movement; it is the expression of life. (...) 3 - Melody has its central point in affective life. (...) 4 - Harmony is the most intellectual element of music. (...) 5 - The song is the synthesis of all these elements. (...)" [27]
The October-November 1965 catalogue suggests beginning the first exercises with a song as a basis for identifying ascending and descending melodic movements and recognizing rhythmic pulsation. To achieve this, the use of a kinescope for the projection of the score was recommended [28]. The monitors had to prepare the classroom in advance of the broadcast and use the moments of "mediate exploration" and "immediate exploration" to clarify some theoretical points related to the lessons.

The monitors were also responsible for suggesting new exercises following this model, as long as they "do not exceed the level of difficulty proposed by the teacher" [29]. The five lessons in this document were based on aural perception and reproduction of melodic and rhythmic material using songs. Although there was a progression of difficulty in the exercises, they were primarily based on ear training, with no deepening regarding music theory. For these lessons, some teaching materials were also suggested for exercise, such as "small sticks", a "xylophone", and a "flute" [33] The Music Education classes followed a consistent format and gradually increased in difficulty level throughout the lessons. The songs in this class were all from the Portuguese popular repertoire, such as "Balão do João", "Senhora do Almortão", "O Malhão" or "Os olhos da Marianita", but they always followed the same format [30], [31], [32].
The classes of Choral Singing and Music Education were closely related as they both focused on learning and practicing songs. In 1966, there was an attempt to merge these two subjects. However, when *Telescola* returned the transmissions after the summer break of that year, the subjects remained separate in the guidelines section, with a plan that divided the 1st and 2nd grades [33]. Choral Singing classes were scheduled for every Saturday at 3:00 p.m. and lasted for 25 minutes. They were only conducted during the broadcast time and did not involve any preparation or exploration moments outside the class. These classes included melodies that were already part of the ideological imagination, such as the National Anthem and the Anthem of the Portuguese Youth [34].
The Music Listening activities were designed for higher grade levels, specifically 3rd and 4th grades. These activities mainly consisted of listening to excerpts from classical music pieces. Due to the less interactive nature of the exercise, this choice of repertoire was possible because there was no need to conform to the technical limitations of the students' musical practice. In the first edition, there was a more elaborate exercise on formal analysis of the sonata form, including recordings by pianists Maria João Pires and José Carlos Picoto. However, this type of exercise never appeared again after the first edition, and performers were no longer mentioned in the bulletins.

4 Conclusion

Musical practice was the only performing art expression to be included in the *Telescola* curriculum, probably because of the theoretical characteristics of musical language, with a structure of fixed rules similar to grammar and syntax. Although classroom devices were used to fulfil the exercises, Music Education, Music Listening and Choral Singing followed models that could be taught remotely with some ease. Even though the exercises were elementary, the reach of the *Telescola* meant the dissemination of music education material, which was distributed not only by official and unofficial spots but also on national television broadcast, by Emissora Nacional. This means that the lessons not only reached those enrolled in the *Telescola*, but also those who had access to a television set.

The songs used in music education classes, and especially in Choral Singing, have an ideological charge that ties in with the propaganda of the Estado Novo. The practice of songs is frequent in these activities, and is prone to collective environments, thus being a platform for a spirit of unity and sometimes propaganda. The choices of Portuguese music repertoire reinforce nationalist ideals, which are even clearer in the choice of the Portuguese youth anthem, which has a proto-fascist character. Even so, the *Telescola* programmes were created within the intermediary of the Institute of High Culture, which internally has a republican ideological past that tends to democratize education, something that was maintained in practice during the period under study. After the Portuguese Revolution of 1974, the *Telescola* project already had all the conditions to adapt to a welfare state and a democratic political environment. The *Telescola* ended the broadcast in 2003, but this system were replicated 2020 during covid-19 exceptional education program - *Study at Home, Estudar em Casa* - , to maintain the elementary school in quarantine.

References

[1] The Estado Novo was a political regime that existed in Portugal between 1933 and 1974, led by Oliveira Salazar and later by Marcello Caetano. It is considered an authoritarian and dictatorial regime, based on a corporatist organization. Lasting approximately 41 years, this was the longest-lasting authoritarian regime in Europe in the 20th century. Fernando Rosas argues, as described in *Salazar e o Poder: A arte de saber durar,* that the durability of this regime is primarily due to its corporatist organization, which allowed for wide state control over the economic, institutional, cultural, and educational structures, as well as a preventive and punitive force to control the population. See: Rosas, Fernando, *Salazar e o Poder: A arte de saber durar*, Tinta da China, Lisboa: 2012.

[2] See: Rosas, Fernando, *Salazar e o Poder: A arte de saber durar*, Tinta da China, p. 332, Lisboa: 2012.

[3] The "politics of the spirit" (política do espírito) was a basic principle of the Estado Novo's national propaganda. This term was coined by António Ferro, developed during his mandate as head of the National Propaganda Society, and is based on the promotion of the approach of the immaterial to elevate the spirit of the population. This principle, on a practical level, resulted in cultural policies that were ideologically aligned with the Estado Novo. See: Santos, Graça dos.

"Política Do Espírito': O Bom Gosto Obrigatório Para Embelezar a Realidade." Media e Jornalismo, vol. 12, 2008, pp. 59–72.
[4] See: Rosas, Fernando, *Salazar e o Poder: A arte de saber durar*, Tinta da China, p. 331, 2012.
[5] Ibid.
[6] Ibid.
[7] In: Lemos, Valter. *A Influência Da OCDE Nas Políticas Públicas de Educação Em Portugal*. Almedina, pp. 33-37, 2014.
[8] In: Rollo, Maria Fernanda. Portugal e o Plano Marshall. Editorial Estampa, p. 65, 1994.
[9] Ibid.
[10] In: Lemos, Valter. *A Influência Da OCDE Nas Políticas Públicas de Educação Em Portugal*. Almedina, p. 42, 2014.
[11] Ibid, p. 6.
[12] See: Silva, Renata Maldonado da. "A Construção Das Políticas de Comunicação e Educação Em Regimes de Excepção: Uma Análise Da Expansão Da Escolarização Por Meio Da Televisão Em Portugal." *Intermeio*, vol. 27, p. 134, 2021.
[13] In: Rollo, Maria Fernanda. *Ciência, Cultura e Língua Em Portugal No Século XX : Da Junta de Educação Nacional Ao Instituto Camões*. Imprensa Nacional-Casa da Moeda, p.14, 2012.
[14] In: Law-Degree 46038, November 16 in *Diário da República*
[15] Ibid.
[16] Law-Degree n.º 45418 in *Diário da República*
[17] Law-Degree n.º 46135 de 31 de Dezembro in *Diário da República*
[18] Ibid.
[19] Ibid.
[20] Law-Degree 46136 in *Diário da República*
[21] See: Silva, Renata Maldonado da. "A Construção Das Políticas de Comunicação e Educação Em Regimes de Excepção: Uma Análise Da Expansão Da Escolarização Por Meio Da Televisão Em Portugal." *Intermeio*, vol. 27, p. 135, 2021.
[22] This information is on the I.M.A.V.E. Bulletin to *Telescola* directions - IMAVE. *Boletim Do Instituto de Meios Audio-Visuais No Ensino*. Ministério da Educação Nacional, p.2, October-November, 1965.
[23] Ibid., p. 12.
[24] Ibid., p. 13.
[25] Ibid., p. 8-11.
[26] Ibid., p. 169.

[27] Ibid, 170.

[28] Ibid, 171.

[29] Ibid, 170-171

[30] Ibid, 171-174

[31] IMAVE. *Boletim Do Instituto de Meios Audio-Visuais No Ensino*. Ministério da Educação Nacional, December, 1965

[32] IMAVE. *Boletim Do Instituto de Meios Audio-Visuais No Ensino*. Ministério da Educação Nacional, October, 1966.

[33] IMAVE. *Boletim Do Instituto de Meios Audio-Visuais No Ensino*. Ministério da Educação Nacional, October, 1966.

[34] The Portuguese Youth (Mocidade Portuguesa) was an organisation of the Estado Novo that had a mandatory character until 1971, to provide education with nationalist, military and Catholic content (Decree-Law 26/611-1936 in *Diário da República*). See: Rosas, Fernando, *Salazar e o Poder: A arte de saber durar*, Tinta da China, p. 334, 2012.

iPad vs. Music Glove Use in The Music Classroom: Differences in Children's Learning Performance, Ease of Use and Concentration

Andrew Danso[1,2], Alicia Lucendo-Noriega[1,2], Joshua S. Bamford[1,2,3], Geoff Luck[1,2], Alessandro Ansani[1,2], Rebekah Rousi[4], Marc R. Thompson[1,2]

[1] Centre of Excellence in Music, Mind, Body and Brain
[2] Department of Music, Arts and Culture Studies, University of Jyväskylä, Finland
[3] Institute of Human Sciences, University of Oxford, UK
[4] School of Marketing and Communication, Communication Studies, University of Vaasa, Finland
andrew.a.dansoadu@jyu.fi

Abstract. This study uses a mixed-methods approach to investigate the learning outcomes, student experiences, and concentrated-behavior patterns of two classes of elementary school music students ($N = 42$). The study examines the academic performance resulting from students using established music technology, the iPad, in comparison to using an experimental technology, a wearable Music Glove. The musical knowledge score improved significantly in both the iPad ($W = 1$, $p < .001$) and Glove ($W = 28.5$, $p = .043$) classes; however, the iPad class improved more than the glove class ($d = 1.83$ vs $d = .48$). Following this, we conducted a Mann-Whitney U test to assess whether the improvement in scores between the two groups from before to after the study was significant. Results show a significant difference in test score improvements during the learning process ($p = <.01$). Full contents of this study can be found in the primary author's doctoral dissertation.

Keywords. Music technology, music education, iPad, education

1 Introduction

In this study, our goal was to better understand the impact that established and experimental music technology has on children's musical academic performance, user experience and concentrated behavior. To do this, we used a mixed-method approach, analyzing both quantitative and qualitative data. The quantitative data came from a (1) student musical knowledge assessment, measuring the student's learning performance, (2) a user experience survey measuring student perceived ease of use of their assigned device and, (3) concentrated-related behavior as measured via qualitative video analysis. The study consisted of two groups of students using either the iPad or the Music Glove, a wearable device that activates musical sound when its sensors are touched. Both devices were used over a 6-week period. The three primary research questions in this study are:

RQ1. What is the difference in musical knowledge before and after using the Music Glove between the two music classes?

RQ2. What are the students' ratings in perceived ease of use before and after using the iPad or the Music Glove?

RQ3. What is the difference in concentration-related behavior patterns of the student's while playing and using the iPad or Music Glove in the two music classes?

1.1 Technology Integration in the Music Classroom

In recent years, music classrooms have seen significant technological integration. These integrations involve using innovative devices to encourage hands-on interaction and touch-based feedback as part of the learning process. The most well-known of these devices is the iPad, a tablet computer made by Apple, which has become widely used in modern classrooms. Studies conducted by [1, 2, 3] have shown that iPad use in classrooms has a positive effect on learning outcomes.

1.2 The Music Glove

The Music Glove is a musical MIDI controller using touch sensors and an electronic unit to produce data. These touch sensors are arranged in multiple rows on the glove, forming a musical scale. To activate these sensors, the user presses them with a finger from their other hand. The touch sensors are positioned from the index finger to the little finger, with each fingertip corresponding to specific notes in the first octave, namely C, D, E, and F. Notably, the semitone E-F is positioned between the ring finger and the little finger. When a user touches the glove's sensors, it generates a data signal that is then transmitted to an external device, such as a MIDI device, a PC, or a computer tablet, for details see, e.g., [4, 5].

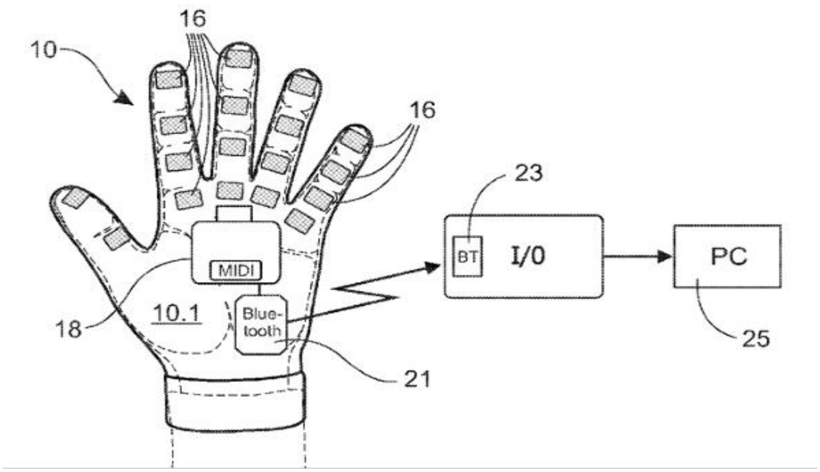


Fig. 1. A diagram of the Music Glove with a musical instrument digital interface (MIDI) and Bluetooth (BT) connected to a personal computer (PC). The numbers in the diagram refer to the hardware that the Music Glove implements: 10. A Glove device. 16. Touch sensors. 18. Central MIDI electronic unit. 21. Bluetooth transmitting MIDI code. 23. Bluetooth receiver. 25. Personal Computer (host device) (U.S. Patent No. 9,905,207, 2018).

1.3 Concentrated Behavior in the Music Classroom

Concentration is said to be achieved during genuine engagement in learning, as the student is cognitively and affectively attuned to acquiring requisite information during lesson time [6]. Concentration is noted to be a core component of educational endeavors, as [7] argue that the amount of time and effort spent in a classroom is wasted if students are not learning, and this happens exclusively within the concentration span of the learners.

1.4 Technology Acceptance in the Music Classroom

The Technology Acceptance Model (TAM), originally developed by [8], helps us understand why people choose to use technology and how they use it. TAM consists of two key factors that are crucial for understanding computer acceptance: perceived usefulness and perceived ease of use [9, 10, 11]. In a survey investigating musicians' uses and attitudes towards using technology in individual learning settings, these two aspects (perceived usefulness and perceived ease of use) were predictors of using technology in music learning [12]. We chose perceived ease of use from the TAM as a theoretical concept to investigate the degree of effort the students anticipated had to exert while interacting with technology in their class.

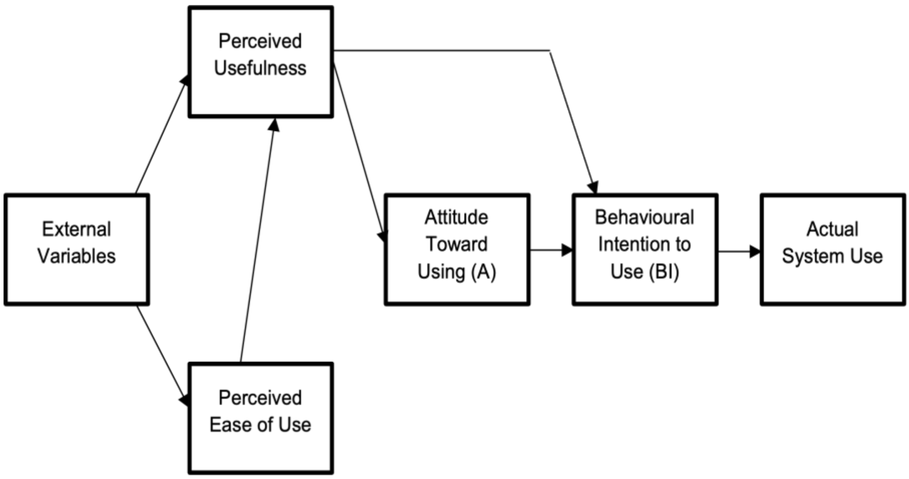


Fig. 2. Original Technology Acceptance Model (TAM) [8].

2 Methodology

We carried out a mixed-method research study in Central Finland, involving two elementary school classes. In one class, students were allocated iPads using a music production app (the Keyboard Touch Instrument app found in GarageBand) for their music lessons. The other class were allocated the Music Glove as their primary device for music lessons.

2.1 Participants

The study involved two classes, each consisting of 21 students, totaling 42 participants. These students, aged 8 to 9 years, were enrolled in regular music classes at Jyväskylän Normaalikoulu in Central Finland. Their average age was 8.3 years (SD = 0.5). To ensure the anonymity of the students, each child was allocated a number from #1 to #21 within their respective classes, allowing consistent identification throughout the data collection processes. Concurrent mixed-method sampling [13] was chosen as it aligned with the needs of both the quantitative and qualitative aspects of the study.

2.2 Apparatus: iPad and Music Glove

The iPad is a tablet with a multitouch screen interface, operating on the iOS platform. It has the capacity to function as a versatile platform for running various applications. Furthermore, we connected the Music Glove to an iPad via USB, with the iPad serving as the host device.

2.3 Familiarization with the Music Glove Device

We conducted two familiarization sessions before giving the students the pre-study knowledge test and the Week 1 user experience survey. These sessions served two purposes: (a) allowing the children to learn and experience the equipment before the actual study began and (b) helping the children and researchers get to know each other. This is crucial when conducting research with children as it builds trust and encourages them to express themselves as they normally would [14].

2.4 Learning Performance: Student Musical Knowledge Assessment

Both classes of students participated in a musical knowledge assessment as the study began (pre-), and at its conclusion (post-). These pre- and post-tests were used to evaluate the students' retention of musical knowledge and their learning progress before and after utilizing their allocated technologies.

2.5 User Experience Survey: Perceived Ease of Use

To assess how the two groups of students rated their ease of use about using their respective technologies during their music class, the students were asked to complete a subjective experience survey before and after using either their allocated iPad or Music Glove. The survey applied a Likert-type scale, which was visually represented in the form of thumb pictures. We conducted a reliability test of the survey during the familiarization sessions. We found no inconsistencies in their responses to the user experience survey after the familiarization session, indicating preliminary validity.

2.6 Video Analysis: Student Concentrated Behavior

To examine variations in students' behavior associated with their concentration levels while using the two music technologies, we carried out a qualitative analysis of video recordings capturing students engaged with their allocated devices in the classroom. The two researchers selected lessons from Weeks 1, 3, and 6 for video analysis, covering the beginning, middle, and end of the study. They used a three-phased coding process to analyze the video recordings. After the three-phased coding process and independent analysis, they agreed on two categories to analyze student concentrated behavior: off-task behavior and on-task behavior.

2.7 Learning Outcomes

As previous literature suggests [15], when introducing technology in the classroom, the pedagogical practices, context, and purpose defined are significant to the potential effect any technology will have on students. Therefore, in the current study, the teacher supervised the integration of both technologies (iPad and the Music Gove) within both classes.

3 Results

3.1 Student Musical Knowledge Assessment

All students in both classes completed a Musical Knowledge Assessment. This test served as a baseline measurement of their understanding of the musical syllabus before they started using the devices. After using the technologies, the same test was given again to see if the use of these technologies had an impact on their musical knowledge. We used Pearson's correlation coefficient to analyze the relationship between the group that used only iPads for music learning (21 students) and the group that used the Music Glove (21 students) for the same purpose. We also compared the test scores at the beginning of the study (Week 1) with the scores at the end (Week 6) of the study. The analysis showed a moderate positive correlation between the initial and final test scores for both groups. The correlation coefficient was .73 for the iPad group and .77 for the Music Glove group.

The results of the knowledge test before and after using the technology are shown in Figure 3. First, we ran two Wilcoxon signed-rank tests to check whether the improvements of the post scores were significant. The MKS improved significantly in both the iPad ($W = 1$, $p < .001$) and Glove ($W = 28.5$, $p = .043$) classes; however, the iPad class improved more than the glove class ($d = 1.83$ vs $d = .48$). Furthermore, we conducted a Mann-Whitney U test to assess whether the improvement in scores between the two groups from before to after the study was significant. The Mann-

Whitney U test showed that there was a significant difference in the change of test scores between the students who used iPads (median change = 8) and those who used the Music Glove (median change = 1). The results were statistically significant (U = 115, p = .004 two-tailed), and the effect size was medium (d = .77).

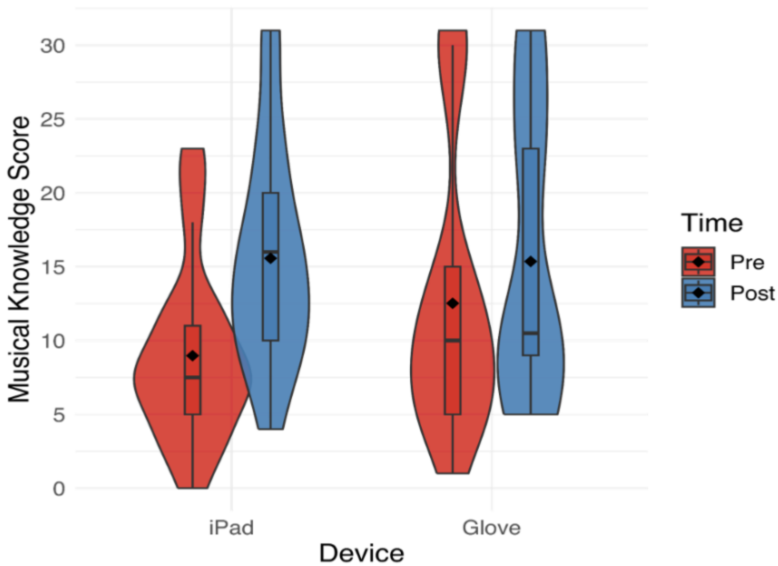


Fig. 3. Violin plots comparing the pre- and post-learning academic test of knowledge results of the iPad and Music Glove class. The violin plot provides a visual representation of the distribution curve, with interquartile ranges (IQRs) and median values displayed using boxplots, denoted by black horizontal lines. Additionally, mean values are depicted as black rhombi. *Note.* Pre-test presents scores before using allocated technology in the class. Post-test presents scores after using allocated technology in the class. The total test score is out of 31.

3.2 Student Perceived Ease of Use

TABLE 1. Descriptive statistics showing the ease-of-use response ratings during weeks 1, 3 and 6, before and after classes for the iPad and Music Glove classes.

	Week 1			Week 3			Week 6		
Factor	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
I think the iPad will be easy to use today	3.84	4.00	1.12	3.50	4.00	1.43	3.76	4.00	1.25
Today I found the iPad easy to use	2.84	3.00	1.07	2.80	3.00	1.28	2.53	2.00	1.55
I think the glove will be easy to use today	4.30	4.50	0.95	3.95	4.00	0.95	3.70	4.00	1.17
Today I found the glove easy to use	2.80	3.00	1.06	2.35	2.00	1.82	2.10	1.50	1.21

To analyze the change in perceived ease of use responses before and after the students used their allocated technologies, we used two Wilcoxon Signed-ranks tests to examine their perceived ease of use survey responses for both groups.

Before using the iPad at Week 1, the median perceived ease of use rating was 4.00. After using the iPad at Week 6, the median rating decreased to 2.00. Analysis showed a significant change ($Z = -2.58$, $p = .009$, $d = .80$). Before using the Music Glove at Week 1, the median perceived ease of use rating was 4.5. After using the Music Glove at Week 6, the median rating decreased to 1.5. Analysis revealed a significant change with a large effect size ($Z = -3.42$, $p = .001$, $d = .80$).

3.3 Concentrated Behavior Observations

Two researchers independently rated and coded the behavior of the selected students at each observation point. To ensure the accuracy of the collected data, we measured interrater reliability using Cohen's kappa (k). The kappa value was found to be $k = .81$, indicating a strong agreement between the researchers in their coding categories.
Tables 2 and 3 present the researchers' analysis of the behavior of two representative students in both the iPad-using and Music Glove-using classes. In the iPad-using class,

students numbered #7 to #12 were chosen for video analysis, while in the Music Glove-using class, students numbered #3 to #10 were selected for behavior analysis through video recordings.

TABLE 2. Total tallies of the researchers' analysis of behavior in the iPad class.

Week	Student	Off-Task Behavior	On-Task Behavior
1	#7	3	0
	#12	2	0
3	#7	0	0
	#12	0	0
6	#7	1	0
	#12	4	0
TOTALS:		10	0

Note. Student indicates the number assigned to students in the class (categorized from #1 to #22). Students numbered #7 and #12 selected for analysis.

TABLE 3. Total tallies of the researchers' analysis of behavior in the Music Glove class.

Week	Student	Off-Task Behavior	On-Task Behavior
1	#3	0	0
	#10	0	0
3	#3	0	0
	#10	0	0
6	#3	0	3
	#10	0	3
TOTALS:		0	6

Note. Student indicates the number assigned to students in the class (categorized from #1 to #22). Students numbered #3 and #10 selected for analysis.

4 Discussion

In this study, it was found that students improved more than the glove class (d = 1.83 vs d = .48) who used the iPad for music learning. As the change in perceived ease of use ratings is analyzed before using both technologies, statistically significant results are reported. The findings from the qualitative video analysis tentatively indicate that concentration-related behavior appeared to be more prominent in the two students who used the Music Glove compared to the two students using only the iPad.

RQ1. What is the difference in musical knowledge before and after using the Music Glove between the two music classes?

As evident from the analysis of post-test results, students who used the iPad displayed greater improvements in their musical knowledge over the 6-week learning period compared to those who used the Music Glove. When we compared the differences in post-test results related to the change between these two groups, we found the difference to be significant, with a medium effect size. These results suggest that the use of the iPad contributed more to enhanced learning compared to the Music Glove. These findings align with a previous study that utilized the same data [4] as well as a recent meta-analysis [3].

RQ2. What are the students' ratings in perceived ease of use before and after using the iPad or the Music Glove?

The results from the perceived ease of ratings indicate that both the iPad and the Music Glove were perceived as more difficult to use after six weeks of use, compared to their initial ratings at Week 1.

RQ3. What is the difference in concentration-related behavior patterns of the student's while playing and using the iPad or Music Glove in the two music classes?

The differences observed in concentration-related behavior between the two students from each group could be attributed to the students' varying levels of familiarity with the iPad and the Music Glove. It's worth noting that all participants in this study had previous experience using the iPad in their music classes before this experiment began. Therefore, their on-task and off-task behaviors might have been influenced or even caused by their familiarity (with the iPad) or lack of (the Music Glove) with their allocated technology.

4.1 Student Ease of Use Ratings, Concentration and Technology Acceptance in the Music Classroom

Curiously, the data shows that the students observed for concentration-related behavior while using only the iPad engaged in more off-task behavior than those using the Music Glove, suggesting that the former group was not as fully concentrated during the 6-week study. In this context, higher perceived ease of use corresponded to lower interaction effort with the familiar technology (i.e., the iPad), leading to lower levels of concentration-related behavior. On the other hand, the increased concentration-related behavior observed in students using the Music Glove likely stemmed from their heightened effort in interacting with the technology. Therefore, applying the concept of perceived ease of use from TAM [8] illustrates that students had strong expectations about how user-friendly both device interfaces appeared before playing music, but their concentration-related behavior reflected the challenges these technologies presented in practice. One might also consider limitations of the original version of TAM [8] in this context due to its over-simplicity. Subsequent developments of the model, including TAM 2 [16], the Unified Theory of Acceptance and Use of Technology (or UTAUT) [17], and the e-commerce-oriented TAM 3 [18] each unpack and add detail to model elements. Additional testing and perhaps further development of the model will be necessary to ascertain the suitability of this general framework to the issues discussed in this paper.

5 Conclusions

In summary, the study revealed several key findings:

- Students who used the iPad showed greater improvement in their learning performance compared to those who used the Music Glove.
- Both groups of students initially rated the perceived ease of use of both technologies highly, but these ratings decreased after 6 weeks of use.
- We observed that both groups of students displayed high levels of concentrated-related behavior while using their respective devices. However, students using the Music Glove exhibited even higher levels of concentration compared to those using the iPad alone.
- When considering the relationship between concentration and perceived ease of use (based on the TAM), we found that both technologies were expected to be easy to use for playing music. However, the iPad required less effort during use compared to the Music Glove. This difference in effort could be attributed to students' familiarity with the iPad in their music education.

References

[1] Wario, R., Ireri, B., and De Wet, L. (2016). An Evaluation Of iPad As A Learning Tool In Higher Education Within A Rural Catchment: A Case Study At A South African University.

[2] Wang, B. T., Teng, C. W., and Chen, H. T. (2015). Using iPad to Facilitate English Vocabulary Learning. International Journal of Information and Education Technology, 5(2), 100–104. https://doi.org/10.7763/ijiet.2015.v5.484

[3] Wang, K., Young, J., Li, Y., and Xiao, Y. (2023). Effects of the iPad use on K-12 students' STEM achievement: A meta-analysis. International Journal of Mobile Learning and Organisation, 1(1), 1–1. https://doi.org/10.1504/ijmlo.2023.10051538

[4] Danso, A., Rousi, R., and Thompson, M. (2021). Novel and experimental music technology use in the music classroom: Learning performance, experience and concentrated behavior. Human Technology, 17(1), 81–112. https://doi.org/10.17011/HT/URN.202106223979

[5] Danso, A. (2023). The Use of Technology in Music-based Interventions for Health and Education. JYU Dissertations.

[6] Dansereau, D. F. (2014). Learning Strategy Research. In Thinking and learning skills: Volume 1: Relating instruction to research (1st ed., pp. 209-239). Routledge.

[7] Bester, G., and Brand, L. (2013). The effect of technology on learner attention and achievement in the classroom (South African Journal of Education, Vol. 33, Issue 2, pp. 1–1).

[8] Davis, F. D. (1985). A technology acceptance model for empirically testing new end-user information systems: Theory and results (Doctoral dissertation) Massachusetts Institute of Technology, Cambridge MA, USA.

[9] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–339.

[10] Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models (202409:38:41; Source: Management Science, Vol. 35, Issue 8, pp. 982–1003).

[11] Venkatesh, V. (2000). Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. Information Systems Research, 11(4), 342–365. https://doi.org/10.1287/isre.11.4.342.11872

[12] Waddell, G., and Williamon, A. (2019). Technology use and attitudes in music learning. Frontiers in ICT, 6. https://doi.org/10.3389/fict.2019.00011

[13] Teddlie, C., and Yu, F. (2007). Mixed Methods Sampling: A Typology With Examples. Journal of Mixed Methods Research, 1(1), 77–100. https://doi.org/10.1177/2345678906292430
[14] Barley, R., and Bath, C. (2014). The importance of familiarisation when doing research with young children. Ethnography and Education, 9(2), 182–195. https://doi.org/10.1080/17457823.2013.841552
[15] OECD (2023), PISA 2022 Assessment and Analytical Framework, PISA, OECD Publishing, Paris. https://doi.org/10.1787/dfe0bf9c-en.
[16] Venkatesh, V., and Davis, F. D. (2000). Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. Management Science, 46(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926
[17] Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View (Quarterly, Vol. 27, Issue 3, pp. 425–478).
[18] Venkatesh, V., and Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Melodic Treatment: Non-functional Chord Substitution in Commercial Music

Tony Moreira

Belmont University - College of Music and Performing Arts
tony.moreira@belmont.edu

Abstract. This paper presents one of the chapters from the applicant's book 'Harmonic Explorations' which explores advanced techniques for chord substitutions in jazz harmony. It will focus on one of the chapters that uses a technique called 'Melodic Treatment'. This technique results in harmonization that feature chords that are essentially non-functional, but which create a modern jazz sound. The technique essentially treats each note as if it existed outside of the context of the song's key. In other words, any given note can be harmonized in any key that contains that note. Imagine that a song is in the key of C and the melody note is a E harmonized by an F maj7 chord (the IV chord in the key) you could substitute an F#min 7 chord - which still harmonizes the melody note beautifully but is now non-functional. Often some sort of organizing principle guides the choice of non-functional chords. For example, you could harmonize every melody note longer than a quarter-note in a given passage with the same sonority (e.g. minor 11th chords), or create an interesting bass line.

Keywords. Jazz Harmony, Non-functional Harmony, Tonal Destabilization, Tonicization, Tritone substitution, Jazz Harmonic Structures.

1 Introduction

The chord substitution technique that I call Melodic Treatment (MT) opens up many interesting harmonization possibilities. The essential nature of the MT technique is to find harmonies that support the melody but are <u>not</u> associated with the prevailing key center. To get a better sense of what this means it is a good idea to review how harmonization usually works and then examine how we might open up other possibilities.

In most tonal and modal musical styles the melody is the primary element. It is considered to be in the foreground and all other musical considerations refer to and support the melody. It is also true generally that the melody and harmony are both usually made from notes determined by the key center. For a song in the key of C major the notes comprising the melody are primarily drawn from the seven notes of the C major scale. Likewise, the harmonies must also use those same seven notes.[1] The effect of this is to limit the possibilities in the harmony. For example, the note E in the melody could be harmonized by any chord that can be created from notes in the C major scale – C maj7, D min9, E min7, F maj7, G13, A min7, and B min7b5(11). This list of seven chords is a small fraction of all possible chords that contain the note E.

In the MT technique we temporarily suspend the rule that the harmony needs to be from the same *key* as the melody. In an abstract sense any specific note could be found in numerous possible keys — e.g. the note E is in the major keys of F, C, G, D, A, E, and B and the natural minor keys Dm, Am, Em, Bm, F#m, C#m, and G#m.[2] The MT technique allows us to borrow chords from any of those keys as long as the borrowed chord contains the note E. Each melodic note then becomes a unique and independent element in this form of reharmonization. Another way to say this is that the melody note becomes connected for a short time to a new key signature according to a degree we arbitrarily assign to it. When this is done the resulting harmonies can become difficult to explain 'functionally' (i.e. using Roman numerals) and are sometimes referred to as non-functional harmonies.

Before we look at a few examples here are a few observations that might be useful. This technique is more effective when applied to long notes. Be careful when using MT substitutions. They are a form of harmonic embellishment which can create some beautiful effects but they also tend to destabilize the tonal center.

In Fig. 6.1 the melody comprises the notes A, B, and C. Functionally these are scale degrees 6, 7 and 8 of the key of C. The original harmony uses chords from the key of C – D min7, G7 and C maj6. For the MT harmony let's consider some other chords that harmonize the note A but are not in the key of C. For example, the melody note A could be the 9th of a G min9 or it could be the #11 of an Eb maj7 (#11) chord – there are many possibilities.

[1] Chromatic notes, i.e. notes that are not in the key, are not entirely avoided but they must he handled with care.

[2] This is considering only relative minor keys. If we include other minor modes like Dorian, Melodic Minor, etc then there are many more possibilities.

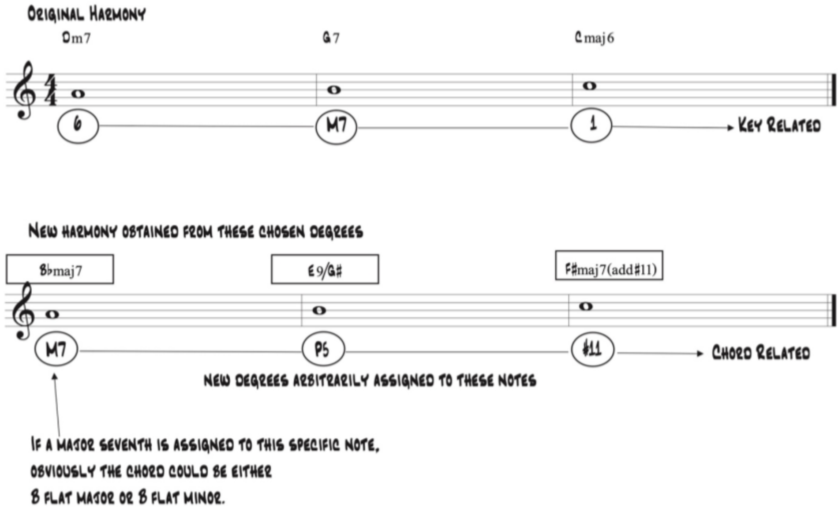


Fig. 1. Melodic Treatment Harmonization.

In Fig. 1 I chose to think of the melody note A as the major 7th of a Bb maj7 chord. The next melody note (B) will also have many possible chords that harmonize it that are not in the key of C. I chose to think of the note B as the 5th of an E9 chord (I put the G# in the bass to provide a nice descending whole-step motion in the bass line that moves in contrary motion with the melody). For the final melody note (C) I used the chord F# maj7(#11) — the C (enharmonic B#) being the #11 of the chord. The effect of these MT substitutions is still harmonious with the original melody but does not support any specific tonal center. In fact, you could spend a whole day trying to find a Roman numeral analysis for these chords and not find anything sensible.

2 More Examples of the MT Technique in Action

We will now consider some examples to demonstrate how the MT technique can result in some really beautiful reharmonizations. In each of the examples below we will use the opening few measures of the tune *A Foggy Day*. The original changes are shown in Fig. 2 in order to give some context. It can be seen that the melody note C is harmonized by an F maj7 chord, the Eb is harmonized by an Ab7 chord, the D is harmonized by a G min7 chord, and the A is harmonized by a C7 (13) chord. All of these chords can be thought of functionally in the key of F major — F maj7 = I, Ab7 = bV substitution for V7/ii, G min7 = ii, and C13 = V.

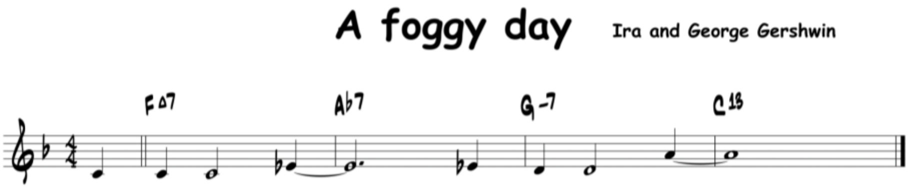



Fig. 2. Excerpt from *A Foggy Day.*

3 The 9th

Fig. 3 shows a reharmonization where each marked note of the melody (the downbeats) is the ninth of a specific chord. In the first bar the melody note is C which would be the 9th of a Bb maj9, Bb min9, or Bb9. Both will sound good but will have a different affect. In the example I chose the Bb maj9 because it has the same chord type as the original F maj7. The Eb note in bar 2 would be the 9th of Db maj9, Db min9, or Db9 – I chose the Db9. For the 3rd measure the note D is the 9th of C maj9, C min9, or C9. Since the original chord is a minor 7th chord, I chose C min9. (Try both the C maj9 and C min9 and choose the one that you like better.) For the last chord the note A would be the 9th of some G chord. In this case we'll use the G min9.

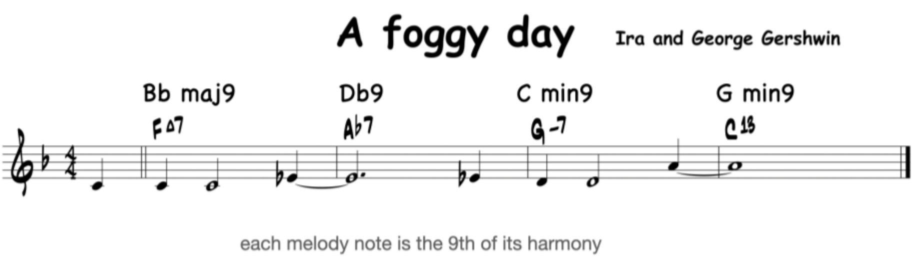



Fig. 3. Excerpt from *A Foggy Day* Reharmonized.

4 The #11th

Fig. 4 reharmonizes the same musical excerpt. But this time each marked note will assume the condition of the #11 degree of a major 7th chord or a dominant 7th chord and from there we will build the chord having that note as a top voice of the chord. Since the #11 is not usually considered a good sound with a minor type chord we will not use any minor 7th type chords. In Fig. 6.4 the melody note C would be the #11 of a Gb maj7(#11) chord, the note Eb is the #11 of A maj7(#11), the note D is the #11 of Ab7(#11), and the note A is the #11 of Eb maj7(#11).

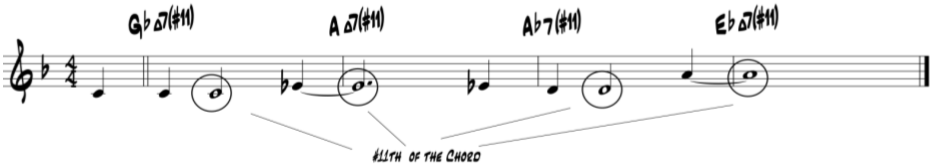


Fig. 4. Excerpt from *A Foggy Day* Reharmonized.

5 The Perfect 4th (11th)

In Fig. 5 we will treat each marked note of the melody as the 4th of a 7th sus chord. (We could use 11th chords instead). The note C is the 4th of G7 sus, Eb is the 4th of Bb7 sus, D is the 4th of A7 sus, and A is the 4th of E7 sus.

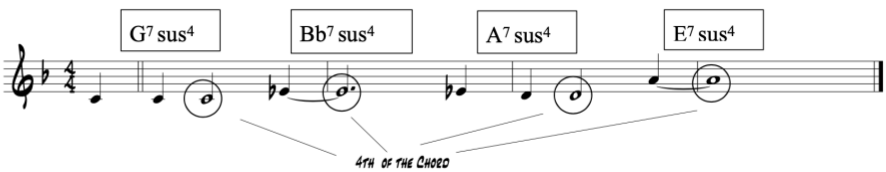


Fig. 5. Excerpt from *A Foggy Day* Reharmonized.

6 The 13th

In Fig. 6 below each marked note will be the 13th of a dominant 7th chord and from there we will build the chord having that note as a top voice of the chord.

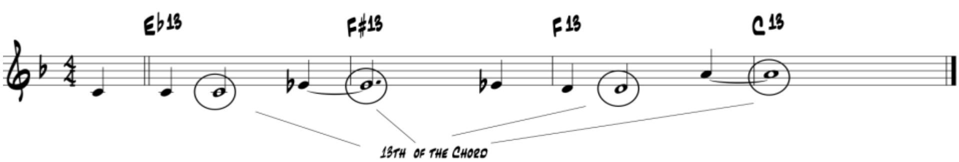


Fig. 6. Excerpt from *A Foggy Day* Reharmonized.

7 The Major 7t

Finally, we will play each marked note of this melody as the seventh of a certain major 7th chord. In these chords it is also possible to include extensions beyond the 7th. For example, the MT in the second measure adds the #11.

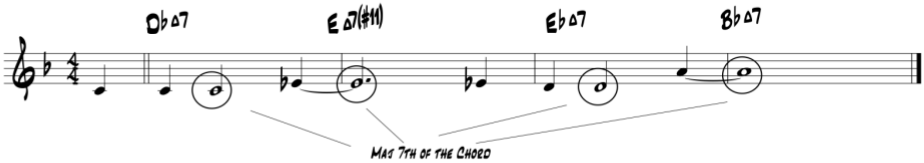


Fig. 7. Excerpt from *A Foggy Day* Reharmonization.

This type of harmonization has an interesting effect and has been very common in jazz harmony for many years. It is important to observe that all of the reharmonizations we have done so far sound excellent but they are very different from each other. This technique opens up a really beautiful and varied palette of harmonies.

8 Melodic Treatment – Random Use

In the following examples we will see how this technique would be applied effectively to the song *It Could Happen to You*. Fig. 6.8 gives the original harmonization for the opening 8 measures.

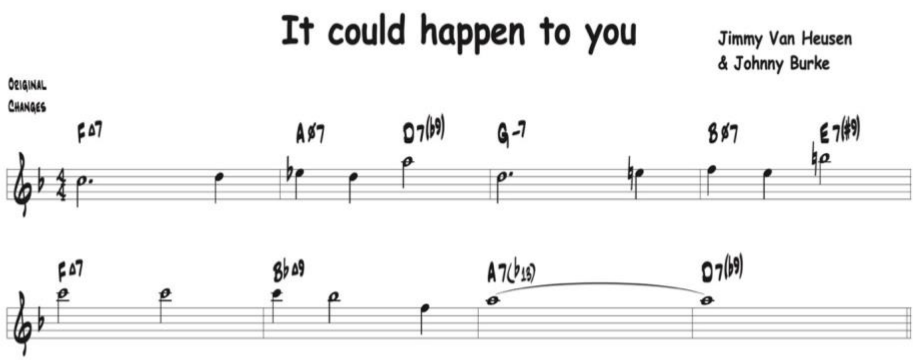


Fig. 8. Excerpt from *It Could Happen to You* with Original Harmony.

In Fig. 9 the chords were chosen to create a bass line that rises by semitone from the Ab in measure 1 to the arrival of the Eb in measure 7. To accomplish this the melody note is not always in the same relationship to the chord as it was in the previous examples. The melody note is the 3rd of the chord for the Ab maj7 and Bb maj7 chords, the 7th for the C maj7, Db maj7(#11), and D min7(11) chords, and the #11 of the Eb maj7(#11) chord.

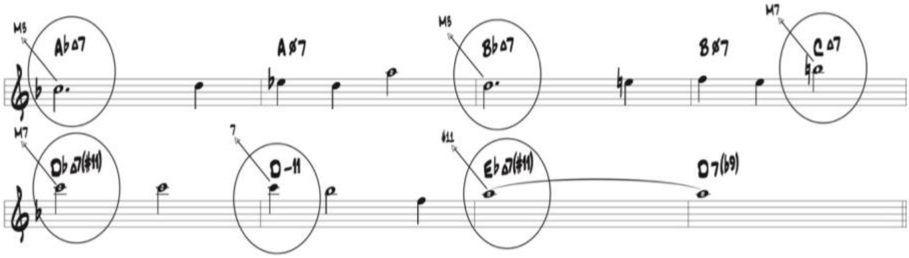


Fig. 9. *It Could Happen to You* with Ascending Bass Line.

In summary:

- if the original chord is a Maj7 or Min7 type chord avoid using a Dominant 7th type chord as the MT substitution;
- if the original chord is a dominant chord it is acceptable to use dominant 7, maj7, or min7 type chords as the MT substitution;
- the MT chord can be for a single chord or a series of substitutions;
- often it is effective to use chords that give an interesting bass line; e.g. by step or minor 3rd;
- using a series of chords that have the same chord type is also a good idea. See Figs. 6.5 and 6.7.

Acknowledgements:

This article represents original work and the development of my jazz harmony technique over many years of experience. Although I did not quote from their works specifically I would like to acknowledge the contributions of the authors Ernst Levy, Barrie Nettles, Mark Levine, and Hugo Riemann. And especially my former mentor in Brazil - Hilton Valente. These scholars helped shape my understanding in important ways.

InLOrk: The Indian Laptop Orchestra

N. Dogra[1], Perry R. Cook[2], and Dharam V. Sharma[3]

[1]Department of Computer Science, University of Reading, UK
[2]Department of Computer Music, Princeton University, USA
[3]Department of Computer Science, Punjabi University, Patiala, India
[1]n.dogra@reading.ac.uk
[2]prc@cs.princeton.edu
[3]dveer@pbi.ac.in

Abstract. The Indian Laptop Orchestra (InLOrk) presents a different way to program Indian music and presents examples of computer-based music-making, through the lens of computer music. InLOrk is a collection of computers designed for the performance of Live Coding and collaboration to generate Indian musical compositions. It is the culmination of a research work on Computer Music at Department of Computer Science, Punjabi University, Patiala, where researcher amalgamates computer science, Indian music and digital music producing techniques into an immersive Laptop ensemble. Members of InLOrk are trained to code new ensemble pieces using the Chuck audio programming language, and to perform with their laptops, computers and other digital devices, which aid in the development of new Indian digital instruments. This paper describes the process of designing this orchestra with Chuck's built-in instruments, such as Mandolin, Flute, Shakers as well as custom built Indian digital percussion instruments, such as Tabla, Geje, Marindanga and Ghatam.

Keywords. Computer music, Laptop Orchestra, Music Technology, Chuck audio programming

1 Introduction

The Princeton Laptop Orchestra (PLOrk) was founded in 2005 by Dan Truman and Perry Cook [1], opening up new avenues for computer music composition, performance, and education. Ge Wang, a key component of PLOrk, established The Stanford Laptop Orchestra (SLOrk) in 2008. This ushered in the "Age of Lork" and stimulated the creation of the world's first assortment of laptop orchestras [2]. The Laptop Orchestra (LOrk) explores new possibilities for computer-mediated musical performances by bringing together traditional musical contexts with cutting edge technology. Additionally, the LOrk explores innovative artistic and pedagogical

possibilities in computer programming, live performance, music composition, and digital instrument design [3].
LOrk ensembles have traditionally focused on Western practices and compositions, and very little work has been done with Indian music. Another research is highlights this western even more [4].
From its inception to the present, Indian music has undergone numerous stages of changes and adjustments. It has both accepted and resisted a great deal of modernity in order to preserve perfect harmony with archaic customs [5]. Particularly when contrasted to Western music traditions, Indian music traditions are highly distinctive; they have produced a wide range of original instruments, music forms, musical practices, and compositions. The Indian subcontinent's diverse musical genres provide a great platform for testing computer music research methodologies. Inspired by all of these initiatives, we set out to create our own laptop orchestra within the framework of Indian music. The project's main goal is to employ Chuck audio programming to create new music in the Indian tradition.

2 Composing with Coding using Chuck Audio Programming

In essence, music is the result of combining instruments and human voice. Melody instruments and percussion instruments are the two categories of musical instruments. If playing traditional instruments and singing live on stage is one aspect of the performance, the other is capturing the music on various storage media and utilizing computers to play it. The following categories apply to computer-based music, or computer music:

1) Use Record/Edit/Sequence tools software tools, also called Digital Audio Workstations (DAWs). Music composers and arrangers often employ one or more of 20 software programs in recording studios, such as FL Studio, Cubase, Logic Pro, ProTools, Reaper, and Audacity.
2) A different technique makes use of Audio Programming Languages, which are specialized computer programming languages. Western artists embrace these languages and use them extensively in the Western music system. Thanks to the growth of digital technology, laptop ensembles in the context of Western music have undergone great development over the past 20 years. The selection of compositions and rhythmic patterns was limited to Western music. It was crucial to select Indian tunes and source materials for the Indian Laptop Orchestra, which is performed live and is programmed using Chuck.

Because of its underlying design, the ChucK programming language is enables the expressive creation, performance, and experimentation of audio and music. With this on-the-fly programmable system, code may be written, edited, compiled, and executed

in real time. With ChucK programming, composers can alter their modules, fine-tune sounds, and swap out composing pieces without having to restart or stop the sound. This makes programming code a real-time expressive tool. It works well for engaging with and managing any kind of digital media, including MIDI, robotics, graphics, and other devices that can communicate with computers. Chuck enables On-the-fly programming with accurate timings and concurrency [6].
Since the author actually sings the voice track live, the programming effort has only concentrated on the instrumental portion. Mandolin, sitar, flute, and other percussion instruments are used in the melody track, while drums, tabla, and other percussion instruments are used in the percussion track.

3 An Experiment on Indian Traditional Composition: *JanaGanaMana*

The engineering process behind creating Indian light music using the audio programming language ChucK, specifically through the composition of Rabindranath Tagore's National Anthem of India, "Jana Gana Mana," with an emphasis on computer music creation by the Indian Laptop Orchestra (InLOrk). The process is divided into three main phases: Music Composition/Design, Technical Design, and Implementation. Let's break down these phases further for clarity: Music Composition/Design Basic Melodic Instrument Tuning: This involves the fine-tuning of instruments to achieve the desired pitch and harmony for the composition. Tuning is critical for ensuring that the melodic instruments are in sync with each other and produce a harmonious sound.Timbre Selection: The timbre, or characteristic sound, of wind, string, and bowing instruments is selected to match the emotional tone and style of the piece. This selection is crucial for creating the intended atmosphere and emotional impact of the music. Percussion Instrument Choice: The composer decides between using Western percussion instruments (like drums, snares, and hi-hats) or Indian classical instruments (like tabla) based on the composition's needs and the desired sound palette.

Technical Design (Music Engineering) Digital Instrument Selection: After the music design phase, the task of music engineering is to select the appropriate digital instrument for each part of the composition. This decision is influenced by the initial music design, ensuring that the chosen digital instrument accurately represents the intended sound of the physical instrument it emulates.

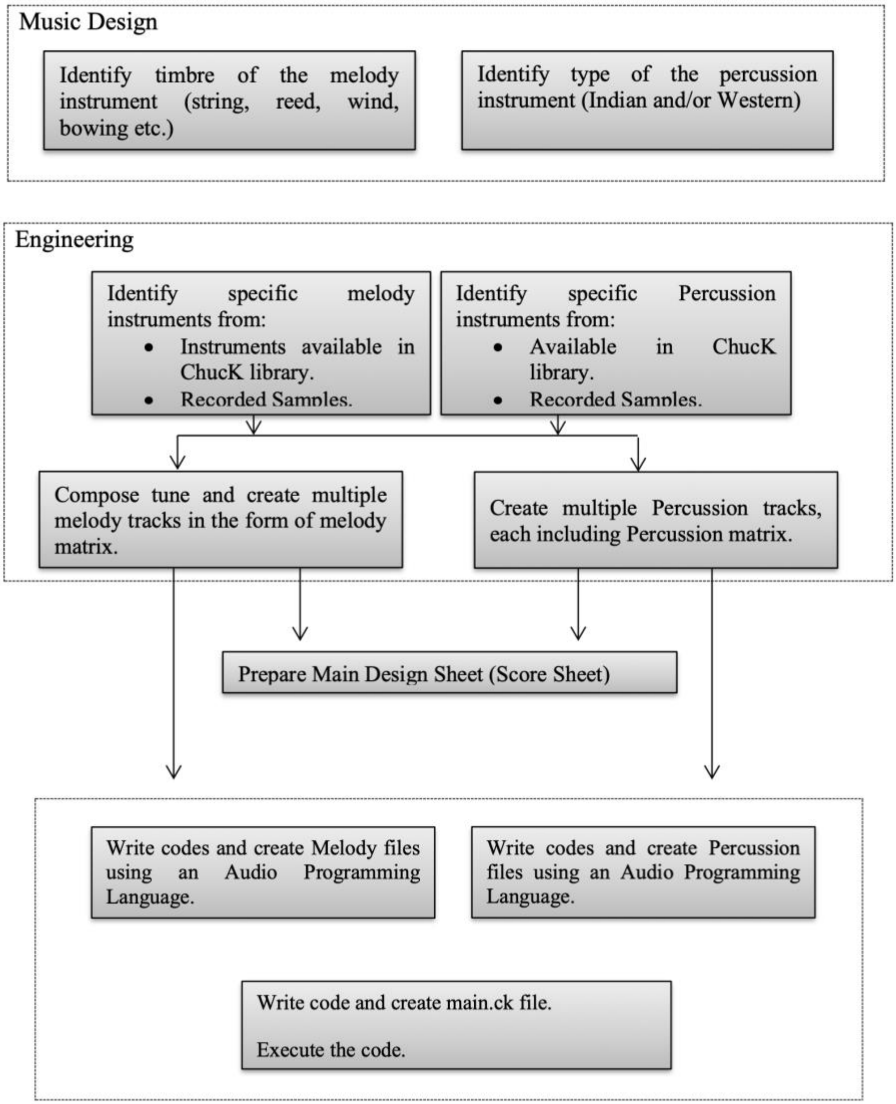


Fig. 1. Music Generation Phase.

Instrument Representation: Whether a string instrument from the music design phase should be represented digitally as a violin, sitar, mandolin, or guitar is decided here, ensuring the digital sound closely matches the chosen physical instrument's timbre.

Implementation

Melody and Percussion Tracks: The implementation phase brings the composed music to life. It involves creating a number of melody tracks, each focused on a different melodic instrument. These tracks are organized using a "melody matrix," which helps in managing the complexity and ensuring that each instrument's role in the composition is clear and distinct. Similarly, percussion tracks are created and organized using a "Percussion Matrix," ensuring that the rhythm section of the composition is cohesive and well-structured.

Additional Insights:
Melody Matrix (Section 3.1): This feature likely involves a systematic approach to organizing and managing different melodic lines within the composition, possibly through software or a conceptual framework that allows the composer and engineers to visualize and manipulate the interaction of these lines.
Percussion Matrix (Section 3.2): Similar to the melody matrix, this system organizes percussion tracks, likely providing a framework for composing, arranging, and synchronizing rhythmic elements within the piece.

3.1 Melody Matrix for the Composition

Chuck refers to the Melody Matrix (see Table 1) as the foundation of the InLOrk Compositions, which are used to create music. Melody Matrix essentially provides the characteristics of each note that is to be performed. The three primary characteristics are: gain, which denotes the note's intensity or volume; note number, which defines the frequency that must be generated; and duration, which determines how long the note must be played.

TABLE 1. MELODY MATRIX.

Note	Duration	Gain	Other attributes
N1	d*	1	Reverb
N2	2d	0.5	
N3	4d	1	Reverb
N4	8d	0.5	
N5	6d	0.5	Echo

*d is the Basic Melodyd is the Basic Melody Unit (BMU) duration. A BMU is a note that has a set frequency. The length of the BMU is determined by the requirements of the composition. Each BMU in the JanaGanaMana melody track is combined into a Basic Melody Sequence Unit (BMSU). The Chuck software selects the element from each matrix row and generates the music. Mandolin and flute, two of Chuck's built-in instruments, are used to compose JanaGanaMana.

The notation part of the song represented as:

Ja	Na	Ga	Na	Ma	Na	Ad	Hi	…	
[61, 8]	[63, 8]	[65, 8]	[65, 8]	[65, 8]	[65, 8]	[65, 8]	[65, 8]	…. .]	@=>int swara[ ][]

3.2 Percussion Matrix for the Composition

The basis for producing the beats' rhythmic patterns is the same principle that the percussion matrix (Table 2) uses in conjunction with the melody matrix. The beat number, duration, and gain are the common characteristics of a percussion matrix. Echo and reverb are optional features that give the note a unique effect. The two components that give the matrix structure are the Basic Percussion Sequence Unit (BPSU) and the Basic Percussion Unit (BPU). Multiple BPSUs can be used in a percussion track.

TABLE 2. PERCUSSION MATRIX.

Beat	Duration	Gain	Other attributes
B1	d*	1	Reverb
B2	2d	0.5	
B3	4d	1	Reverb
B4	8d	0.5	
B5	6d	0.5	Echo

The ChucK library offers a vast array of sounds for western percussion. However, using Indian Percussion Tabla—the beating heart of Indian music—was essential to creating InLOrk. In order to do this, as seen in Figure 2, the seven strokes are captured and saved in the computer in the format of .wav.

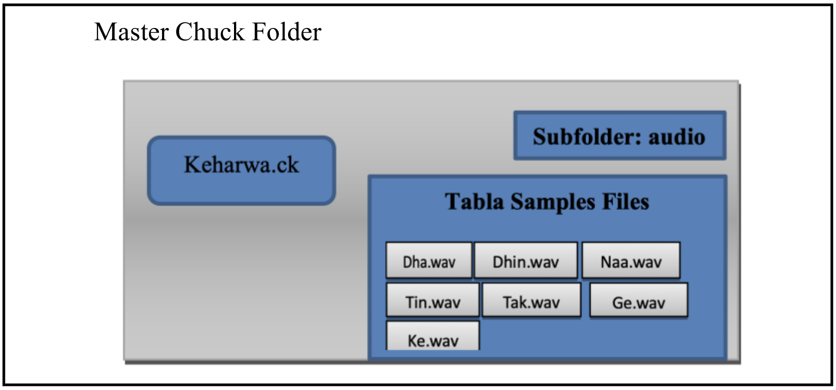


Fig.2. File structure of Percussion Instrument Tabla used for Project.

By using SndBuf to import the necessary samples from the audio folder and by putting together a full file path and file name, three Tabla Patterns are produced. Similar to tabla, the samples of gejje, maridanga, and ghatam from other Indian instruments are recorded in a studio and prepared for the live performance.

Over five Indian musical compositions have undergone successful testing in order to use the aforementioned constructed technology. This work presents one composition from among them. The melody and instrument track used to compose JanaGanaMana are displayed in Table 3 that follows.

TABLE 3. JANAGANAMANA INSTRUMENTAL TRACKS.

Category	Indian Anthem
Language Original Song	Indian National Language - Hindi
Title	*JanaGanaMana*
Total Duration	1 minute 31 seconds
Voice track	Performed Live
Number of Melody Tracks	2
Melody Instruments used	Mandolin and Flute
Available in Chuck Library	Mandolin and Flute
Numbers of Percussion Instruments used	Tabla, Western Drum, Shakers, Mridanga, Gejje, and Ghatam
Available in Chuck Library	Shakers and Drums
Samples Recorded	Tabla, Mridanga, Gejje, and Ghatam

4 First InLork Performance

Members of the orchestra convened at the music studio of Punjabi University to execute a novel form of composition in which laptops and computers are utilized as musical instruments. This piece is the first attempt to build a single music world using MIDI (controlled by Chuck) and a group live performance (3 persons and laptops) (see Figures 3 and 4).

Fig. 3. First Indian Laptop Orchestra (InLOrk) Performance.

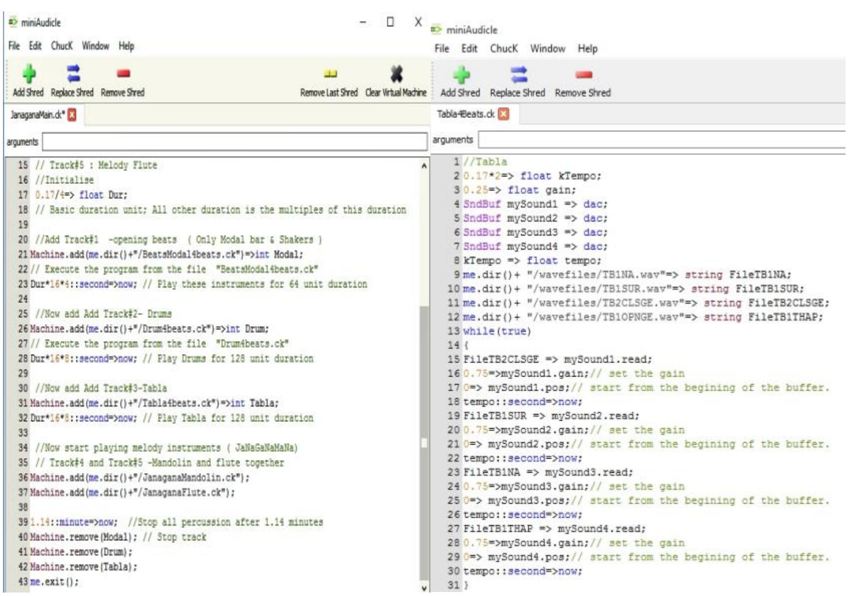


Fig. 4. Coding for InLOrk.

The YouTube video [7] for the InLOrk Final Performance is titled "Indian Laptop Orchestra (InLOrk) - where Technology meets Music." In addition to the enjoyment that comes with creating music and writing creative code, InLOrk aims to monitor the context, culture, and societal implications of the fast-evolving field of computer music. The group actively looks to work with researchers and artists in India, with the goal of contributing to the local computer music scene.

5 Conclusion and Future Work

The efforts described here have added significantly to the Laptop Orchestra field. Nonetheless, one of the most innovative aspects is using the advanced technology to create live performances of Indian compositions created using computer music programming. A variety of rhythmic patterns were created and exported as .Wav files using recorded Tabla, Gejje, Mridanga, and Ghatam samples. These files can be utilized as loops for studio recording as well as live performances. Western music has served as the primary inspiration for a large portion of computer music research. This study thoroughly explores using programs and resources created within the Indian music setting.

Additionally, live coding is used to construct the first-ever Laptop Orchestra performance in India. To the best of our knowledge, InLOrk is the first work that transcribes and performs traditional Indian music using code written using programming abilities. The corpus of work showcased here is a true artistic endeavor that draws on programming expertise and experience from a variety of musical genres. It has empowered musicians, enriched audiences, and fostered creativity through the use of the ChucK audio programming language

Many different musical instruments that are connected to certain Indian states and cultures can be heard in Indian music. Indian classical instruments including the tabla, gejje, mridanga, ghatam, piano, flute, sitar, and mandolin have traditionally been the focus of the InLOrk system for music compositions, audio programming, and live performance. Implementing additional Indian melody and percussion instruments is the next stage of this effort. Furthermore, just five of the more than 300 tabla strokes used in Indian music are used for the final performance. The rest need to be investigated.

ACKNOWLEDGEMENT

We express our gratitude to the University Grant Commission (UGC), India, for their financial support towards the Major Research Project titled "Development of Indian Laptop Orchestra for Digital Artists," under File No.: F1-17.1/2015. Additionally, we extend our thanks to Sabrang International Ltd. UK for their assistance in recording Indian percussion patterns in the studio. We deeply appreciate the online teaching of Chuck programming by Dr. Perry Cook, Dr. Ajay Kapur, and Dr. G. Wang. Special thanks are due to Chidambara Kalamanji for their invaluable assistance in composing Indian music compositions using Chuck audio programming.

References

[1] Trueman, D., Cook, P., Smallwood, S. and Wang, G., *PLOrk: The Princeton Laptop Orchestra, Year 1, in Proceedings of the 2006 International Computer Music Conference. 2006*: New Orleans, Louisiana.

[2] The Symposium on Laptop Orchestras and Ensembles (SLEO), Louisana State University, 2012.
Archived at: https://ccrma.stanford.edu/~prc/SLEO_2012.pdf

[3] Wang, G., Trueman, D., Smallwood, S., and Cook, P., *The laptop orchestra as classroom.* Computer Music Journal, 2008. **32**(1): p. 26-37.

[4] Kalamanji, C., *Generation of Indian light music using audio programming.* International Journal of Advanced Computing and Electronics Technology, 2005. **2**(1): p. 12-23.

[5] Bhattacharjee, A., *Hindustani raga representation and identification: a transition probability-based approach*. International Journal of Mind and Brain Cognition, 2011. **2**(1) pp. 66-91.

[6] Wang, G. and P. Cook, *Chuck: a programming language for on-the-fly, real time synthesis and multimedia, in Proceedings of the 12th annual ACM International conference on Mutimedia.* 2004. ACM: New York, NY, USA.

[7] YouTube Video: “Indian Laptop Orchestra (InLOrk) - Where Technology meets Music”, 2020.
https://youtu.be/SOW2WHoo0zI?si=xBaIBWpNcGJ8Z2eV

Musician's Dystonia: new Pevention and Mitigation Treatments

Joy Grifoni[1], Valeria Crispiatico[2], Anna Castagna[3], Rosa Maria Converti[4], Marina Ramella[5], Angelo Quartarone[6], Teresa L'Abbate[7], Karolina Armonaite[8], Luca Paulon[9], Francescaroberta Panuccio[10], Franca Tecchio[11]

[1, 7, 8] Uninettuno International University, Roma, Italy
[1, 7, 8, 9, 11] Laboratory of Electrophysiology for Translational neuroScience LET'S, Institute of Cognitive Sciences and Technologies ISTC, Consiglio Nazionale delle Ricerche CNR, Roma, Italy
[2] Milano Bicocca University, Milan, Italy
[3, 4, 5] IRCCS Fondazione Don Carlo Gnocchi, Via Capecelatro 66, 20148 Milan, Italy
[6] RCCS Centro Neurolesi "Bonino Pulejo", 98124 Messina, Italy.
[9] Luca Paulon, Engineer Freelance, 00159 Rome, Italy
[10] Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy
[1] j.grifoni@students.uninettunouniversity.net
[2] valeriacrispiatico@gmail.com
[3] acastagna@dongnocchi.it
[4] rmconverti@dongnocchi.it
[5] mramella@dongnocchi.it
[7] t.labbate@students.uninettunouniversity.net
[8] karolina.armonaite@uninettunouniversity.net
[9] luca.paulon@gmail.com
[10] francesca.panucciovg@gmail.com
[11] francamatilde.tecchio@cnr.it

Abstract. This prospective article focuses on the most cited scientific evidence regarding the origin and therapeutic treatments for task-specific musician's dystonia (MD), on which new and increasingly effective therapeutic and prevention approaches can be designed for the future. To this end, we have selected data cited at least dozens of times in Scopus, regarding the characteristics and mechanisms underlying MD and the treatments currently available. Pathophysiological mechanisms demonstrate impaired sensorimotor integration at various levels of brain circuits, involving the limbic system and sensorimotor cortices. The alteration of sensorimotor feedback is accompanied by an impairment of intra-cerebral synchrony and long-term memory plasticity. Task-specific focal dystonia occurs dozens of times more frequently in musicians than in other professions, highlighting that external environmental, social, and emotional factors are crucial dimensions in the origin of MD. This evidence opens the way to the development of multimodal psycho-physical therapeutic strategies to rebalance the alterations of the neural circuits typical of MD through an appropriate personalization of psychotherapy combined with physical rehabilitation activities.

Keywords: task-specific focal musician's dystonia (MD), sensory-motor integration, psychic trauma, multi-sensory multimodal rehabilitation.

1 Definition, etiology, epidemiology

Definition

Task-specific dystonia is a movement disorder characterized by a painless loss of dexterity specific to a particular motor skill [1]. In musicians, the disorder emerges as painless loss of finger motor coordination or embouchure exclusively when playing the musical instrument [2], giving rise to hand or embouchure MD [3]. MD of the hand affects musicians who play strings and plucked strings such as violinists, cellists, guitarists, pianists and more rarely wind and brass players. MD of the embouchure affects brass and woodwind players, involving the perioral, lingual and facial musculature [4].

Aetiology

The causes of MD are still not yet completely clarified, but it is believed to arise from a combination of genetic predisposition and environmental and psychic factors, affecting the brain's motor control system. As for the genetic hypothesis, it is partly supported by evidence that up to 25% of patients with MD have another affected family member with dystonia. Furthermore, a recent genome analysis found an association with the arylsulfatase G gene (ARSG) in both musician's hand dystonia and writer's cramp, but a specific causal mutation within this gene has not yet been identified [3,14]. Possible predisposing risk factors for MD include a positive family history of dystonia, a history of musculoskeletal injury, nerve entrapment or overuse syndrome, and obsessive personality traits [15].

Epidemiology

MD is the most common form of focal task-specific dystonia, with a prevalence of 1:100 compared to 1:6600 for idiopathic dystonia [15]. The prevalence of MD varies depending on the instrument played, with musicians playing piano, violin, guitar, and brass instruments being about 85% [16], and embouchure MD accounting for 13-14% of MD [17,18], and cervical dystonia involving 1-2% of musicians with MD.

How to measure MD

It is essential to define standard scales to evaluate the severity of MD symptoms, as a basis for diagnosis and monitoring and to quantify the effects of changes induced through therapeutic approaches. As key reference we ground on the work [19], which developed a wide overview of the rating scales used in MD. The vast majority of studies use subjective-reported or clinician-reported scales, while objective scales started to be introduced (see Table 1).

2 Current treatments

Among approaches tested over the years to address MD, the most common is *botulinum neurotoxin (BoNT) injections* [20–22]. This commonly used treatment is effective for about 12-16 weeks with the effect dose dependent. Cautiousness is indicated, as BoNT is operator dependent, it depends on guides (electromyography/ultrasound), doses, expertise. It is not possible to consider BoNT as a recommended treatment, but some good results are published in fingers flexors dystonia [23]. However, in the case of embouchure dystonia, most authors report a worsening of musical performance with the use of botulinum toxin [24–26]. Furthermore, a recent investigation [27] revealed that, compared to the control group, the thickness and force strength of injected flexor digitorum superficialis (FDS) and profundus (FDP) muscles were decreased by about 10 -12 % with respect to the non-injected body side, with the extent of weakness and atrophy significantly predicted by the total amount of BoNT injected during the entire treatment period, and the time since the last injection not predicting the amount of strength and muscle mass recovery after cessation of treatment.

Non-pharmacological approaches were developed mainly involving various neuromuscular programs [22,28,29]. most studies were applied on a very small sample of musician with MD, with follow-up assessment absent or very short and often based on subjective nominal or ordinal scale; moreover, other study designs didn't provide any type of randomization to placebo or alternative treatments or blindness.

Sensory tricks are known to be able to provide temporary relief of dystonic symptoms [30–32] usually involving the alteration of tactile or proprioceptive feedback. Some musicians experience a marked improvement in fine-motor control playing with a latex glove or holding an object between their fingers, thus modifying the somatosensory input information [18]. Usually, the effects are very rapid in changing the motor pattern, but they last for a limited time.

Sensory motor retuning (SMR) treatment has been introduced in a seminal study [9,10,33] introducing the immobilization with splints of one or more non-dystonic fingers. The dystonic finger/s then performed repetitive exercises in coordination with the non-splinted ones. The splint maintains the fingers in their characteristic rest position experienced on the instrument during normal performance.. In these studies, positive effects were seen in terms of dexterity, movement speed and accuracy. Sakai and colleagues [34] developed a motor control retraining technique in MD pianists, named "*slow-down exercise*" (*SDE*). During the exercise program, patients undergo basic movement training at decreased speed, making sure that the dystonic patterns would not occur. The pianists would increase metronome speed every 2 weeks as long as they could maintain a normal movement pattern. Berque and colleagues [35] developed a standardized protocol combining SMR plus SDE intervention over a 12-month period. Butler and colleagues [36] introduced a variation including mirror box (six sessions plus daily exercises). Moreover, Ramella et al. [37] have developed the "modified graded motor imagery" (mMGI) treatment, a 4-weeks home – based comprehensive program designed on the three common steps of the GMI. Ackerman & Altenmuller [38] developed an off-instruments progressive retraining protocol for musicians (MusAARP) following an "add-on" approach, starting from individual intrinsic hand muscle actions to more functional hand and arm actions on the instrument.

3 MD mechanisms

Neurophysiological counterpart

Impairment in sensorimotor integration, crucial in the development of task-specific dystonia, is evidenced by findings from animal studies [39], behavioral research [40], and clinical reports utilizing neurophysiological and neuroimaging techniques in humans [41–50] (Quartarone & Hallett, 2013, Mov. Dis.). A seminal animal study [39] demonstrated that repetitive finger compression led to degradation of hand representation and impaired motor control, while a less repetitive strategy preserved sensory integrity and motor function. This suggests that the repetitiveness of motor patterns influences both somatosensory representation and motor execution. Following the SMR model by Byl and colleagues [36], treatment effects were studied on somatosensory hand representation. Pre-treatment, finger relations differed between affected and unaffected hands, but post-treatment, contralateral finger representations resembled those of the less affected side, and were ordered more according to homuncular principles. These physiological changes correlated with behavioral outcomes [2]. Building on extensive research in non-invasive studies of neural network communication in humans [52–58], we have developed a conceptual framework suggesting that neural networks, comprised of nodes and connectors, are governed by a unified functional principle known as the objective-dependent feedback-synchrony-plasticity circuit [59,60]. According to this model, alterations in feedback observed in task-specific dystonia are likely to induce changes in synchronization and plasticity. Regarding synchronization, investigations employing transcranial magnetic stimulation (TMS) protocols revealed the absence of local inhibitory mechanisms during movement initiation and maintenance phases in individuals with task-specific hand dystonia, leading to disrupted cortical inhibition-excitation dynamics [61]. This disturbance affects cortical surround inhibition critical for skill acquisition and relies on inter-hemispheric balances [62,63]. Notably, patients with task-specific hand dystonia exhibit disrupted cortical surround inhibition even during motor preparation [61] (Quartarone et al 2005, Mov. Dis.).

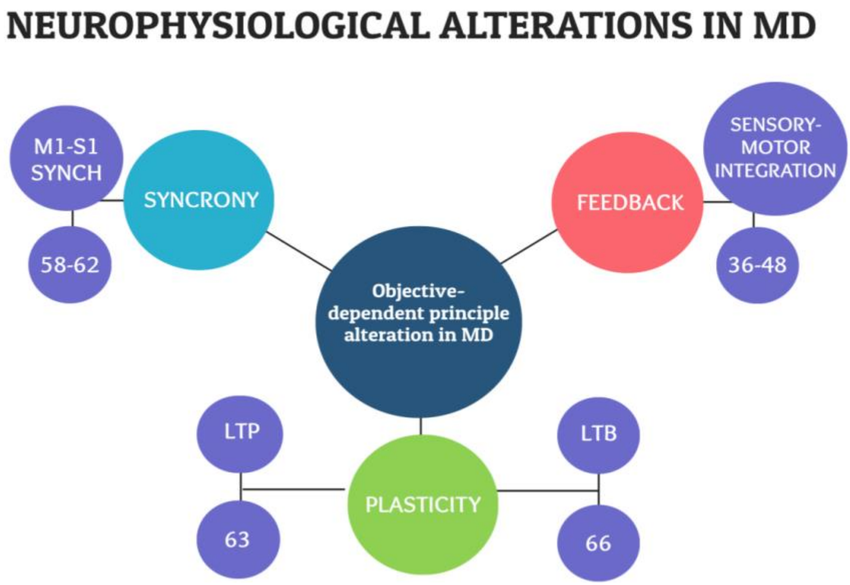


Fig. 1. Neurophysiological alterations in MD. Outline of the alterations secondary to the impairment of sensorimotor integration typical of MD, as supported by the cited literature. LTP: long-term potentiation, LTD: long-term depression, M1: primary motor cortex, S1: primary somatosensory cortex.

Altered excitation and inhibition balances in local motor cortex circuits, reflecting disrupted basal ganglia input, were evidenced by reduced excitability of cortical inhibitory circuits in patients with task-specific dystonia [64]. While no differences were observed in spectral properties between patients and controls in primary motor or somatosensory hand representations, these areas demonstrated reduced functional coupling during movement, along with excessive synchronization in patients compared to controls in the ongoing alpha frequency around 10 Hz [65]. Concerning plasticity, utilizing established non-invasive models of plasticity in humans [66], it has been observed that the motor system exhibits abnormal increases in cortico-spinal excitability and attenuated intracortical inhibitory reinforcement following central associative and peripheral stimulation [67]. Moreover, evidence confirms alterations in long-term potentiation and depression in MD [68–70]. Aligned with the feedback-synchrony-plasticity model, exemplary cases illustrate the intricate relationship between instrument, action, central nervous system, and periphery in MD [1]. For instance, significant modifications in required actions for an illustrator transitioning to a demanding new job had profound psychological effects, whereas a classical pianist experiencing relevant changes in sensorimotor feedback and slight adjustments in motor command faced comparatively less challenging psychological changes when accepting a new role in a musical.

Psycho-social counterpart

While the neurophysiological profile described above is typical of task specific dystonia of the hand, in musicians the occupational dimension decisively modifies the impact of the symptomatology [71] consistent with epidemiological data that indicate MD as the most incident task-specific dystonia [15].

It is immediate to consider that for a musician whose center of professional life is making music, the body is even more immediately at the center of the relationship with the rest of the world. For musicians, health tends to acquire two key dimensions in a dynamic balance, one that concerns the parts of the body directly involved in playing the instrument (especially the hand), and another that focuses on maintaining the health of the whole person [72]. These two dimensions must be integrated in daily attention and care, deepening the awareness that learning through experience, aimed at mutual help and communication, is the key to protect and recover the well-being of the whole body [73].

In the context of the increasingly clear nature of the human being determined within the experience of lifelong learning or trauma (and its profound impact on personal identity), it is possible to speculate that psychological support therapy with EMDR technique (elective for the treatment of trauma) may have particularly encouraging results, especially in its innovation protocol enriched by music and sound stimulation EMDR+ ® [74, 76] where one of the three main dimensions that index the development of a population is the length of the course of study, the learning approaches for musicians become critical in the training years and all along life. In fact, musicians typically start their training at a young age, with the need to develop movements automated through extensive repetitions, gradually increasing complexity, with the need to balance stereotyped production and emotional participation to the story they like to tell. Given the psycho-emotional pressure intrinsic with the public engagement of a musician's activity, affecting the respiratory, circulatory, sensorimotor, and cognitive systems[77–78], the organization of individual and orchestral training is a critical element determining future evolution managing challenges in adapting to the demands of their profession. Indeed, trauma can have far-reaching consequences, extending beyond psychological distress to impact sensorimotor functionality in

musicians. These repercussions are evident in the disruption of motor coordination, manifested as hesitations, involuntary movements, or complete motor blocks, which impede the expression of artistic intentions.
Furthermore, trauma-induced negations can exert influence on cognitive processes, including memory, attention, and decision-making, exacerbating the challenges faced by musicians during performance. Such cognitive impairments may lead to errors or inconsistencies in musical execution, undermining the quality of their performances. In light of these challenges, addressing trauma and its associated negations is paramount for preserving both physical and mental well-being in musicians. By acknowledging and confronting these negations, musicians can embark on a psychotherapeutic path towards reclaiming agency over their bodies and artistic expression. This process fosters resilience and facilitates a more comprehensive approach to musical performance and self-care for MD.

4 Conclusions

We reported about most stable achievements in MD, where the epidemiology underlines how much professional experience implies the highest occurrence of task specific dystonia in musicians with respect to other workers.
Modern rehabilitation approaches aim to restore a physiological position that supports musicians' natural gestures, mainly via botulinum neurotoxin (BoNT) injections and sensory-motor retuning and its variants, emphasizing body awareness and adjusting positions for relaxation in a dynamical training programming.
The consolidated neurophysiological knowledge on the alteration of sensorimotor integration in MD, which coherently with the f*eedback-syncrony-plasticity* functioning principle that governs neuronal networks, is accompanied by altered intra-cortical synchronizations and plasticity, has recently been complemented by the growing awareness of the impact on symptoms of socio-emotional-cognitive experience.
In agreement, new therapeutic strategies are being proposed, which aim to alleviate symptoms that profoundly affect the quality of life of musicians and which often interrupt their career through multidisciplinary approaches with the involvement of interdisciplinary teams including neurology, physiotherapy, occupational therapy, psychology competences integrated into personalized precision approaches.

Author Contributions: Conceptualization, J.G., F.T; methodology, J.G., F.T., V.C.; software, n.a.; formal analysis, n.a.; investigation, n.a.; resources, F.T., L.P.; data curation, n.a.; writing—original draft preparation, J.G. ,V.C., F.T. A.Q. ; writing—review and editing, A.C., RM.C., M.R, T.L. and K.A.; visualization, F.T.; supervision, F.T.; project administration, F.T.,

L.P.; funding acquisition, F.T. All authors have read and agreed to the published version of the manuscript.

Declaration of interests: none

References

[1] Sadnicka, A.; Kornysheva, K.; Rothwell, J.C.; Edwards, M.J. A Unifying Motor Control Framework for Task-Specific Dystonia Europe PMC Funders Group. *Nat Rev Neurol* **2018**, *14*, 116–124. doi:10.1038/nrneurol.2017.146.

[2] Candia, V.; Wienbruch, C.; Elbert, T.; Rockstroh, B.; Ray, W. Effective Behavioral Treatment of Focal Hand Dystonia in Musicians Alters Somatosensory Cortical Organization. *Proc Natl Acad Sci U S A* **2003**, *100*, 7942–7946, doi:10.1073/PNAS.1231193100.

[3] Stahl, C.M.; Frucht, S.J. Focal Task Specific Dystonia: A Review and Update. *J Neurol 264*. doi:10.1007/s00415-016-8373-z.

[4] Frucht, S.J.; Fahn, S.; Greene, P.E.; O'Brien, C.; Gelb, M.; Truong, D.D.; Welsh, J.; Factor, S.; Ford, B. The Natural History of Embouchure Dystonia. *Mov Disord* **2001**, *16*, 899–906, doi:10.1002/MDS.1167.

[5] Fahn S. Assessment of the Primary Dystonias. In: Munsat TL, Ed. Quantification of Neurologic Deficit. . *Oxford: Butterworths* **1989**, 241–270.

[6] Tubiana, R.; Chamagne, P. [Occupational Arm Ailments in Musicians]. *Bull Acad Natl Med* **1993**, *177*, 203–212; discussion 212-6.

[7] Spector, J.; of, A.B.-M.P.; 2005, undefined A New Method for Quantification of Musician's Dystonia: The Frequency of Abnormal Movements Scale. *ingentaconnect.comJT Spector, AG BrandfonbrenerMedical Problems of Performing Artists, 2005•ingentaconnect.com*.

[8] Pesenti, A.; Priori, A.; Scarlato, G.; Barbieri, S. Transient Improvement Induced by Motor Fatigue in Focal Occupational Dystonia: The Handgrip Test. *Mov Disord* **2001**, *16*, 1143–1147. doi:10.1002/MDS.1208.

[9] Candia, V.; Elbert, T.; Altenmüller, E.; Rau, H.; Schäfer, T.; Taub, E. Constraint-Induced Movement Therapy for Focal Hand Dystonia in Musicians. *Lancet* **1999**, *353*, 42. doi:10.1016/S0140-6736(05)74865-0.

[10] Candia, V.; Schäfer, T.; Taub, E.; Rau, H.; Altenmuüller, E.; Rockstroh, B.; Elbert, T. Sensory Motor Retuning: A Behavioral Treatment for Focal Hand Dystonia of Pianists and Guitarists. *Arch Phys Med Rehabil* **2002**, *83*, 1342–1348. doi:10.1053/APMR.2002.35094.

[11] Ramella, M.; Converti, R.M.; Giacobbi, G.; Castagna, A.; Saibene, E.; Borgnis, F.; Baglio, F. The Technical Ability and Performing Scale (TAPS): A Newly

Developed Patient-Reported Functional Rating Scale for Musician's Focal Dystonia. *Parkinsonism Relat Disord* **2022**, *99*, 79–83. doi:10.1016/J.PARKRELDIS.2022.05.015.

[12] Taub, E.; Uswatte, G.; Pidikiti, R. Constraint-Induced Movement Therapy: A New Family of Techniques with Broad Application to Physical Rehabilitation-A Clinical Review. **1999**, *36*.

[13] Jabush, H.C.; Schneider, U.; Altenmüller, E. Delta9-Tetrahydrocannabinol Improves Motor Control in a Patient with Musician's Dystonia. *Mov Disord* **2004**. *19*, 990–991, doi:10.1002/MDS.20214.

[14] Nibbeling, E.; Schaake, S.; Tijssen, M.A.; Weissbach, A.; Groen, J.L.; Altenmüller, E.; Verbeek, D.S.; Lohmann, K. Accumulation of Rare Variants in the Arylsulfatase G (ARSG) Gene in Task-Specific Dystonia. *J Neurol* **2015**, *262*, 1340–1343, doi:10.1007/S00415-015-7718-3.

[15] Rozanski, V.E.; Rehfuess, E.; Bötzel, K.; Nowak, D. Task-Specific Dystonia in Professional Musicians. A Systematic Review of the Importance of Intensive Playing as a Risk Factor. *Dtsch Arztebl Int* **2015**, *112*. doi:10.3238/ARZTEBL.2015.0871.

[16] Aránguiz, R.; Chana-Cuevas, P.; Alburquerque, D.; Curinao, X. Distonía Del Músico: Fenomenología y Desencadenantes Vinculados a La Ejecución Musical PALABRAS CLAVE. **2015**, *30*, 270–275, doi:10.1016/j.nrl.2013.12.024.

[17] Butler, K.; Rosenkranz, K. Focal Hand Dystonia Affecting Musicians. Part II: An Overview Of Current Rehabilitative Treatment Techniques. *The British Journal of Hand Therapy* **2006**, *11*, 79–87. doi:10.1177/175899830601100302.

[18] Jabusch, H.C.; Altenmüller, E. Epidemiology, Phenomenology, and Therapy of Musician's Cramp. *Music, Motor Control and the Brain* **2006**. doi:10.1093/ACPROF:OSO/9780199298723.003.0017.

[19] Peterson, D.A.; Berque, P.; Jabusch, H.-C.; Altenmüller, E.; Frucht, S.J. VIEWS & REVIEWS Rating Scales for Musician's Dystonia. **2013**.

[20] Jankovic, J.; Schwartz, K.; Donovan, D.T. Botulinum Toxin Treatment of Cranial-Cervical Dystonia, Spasmodic Dysphonia, Other Focal Dystonias and Hemifacial Spasm. *Neurosurgery, and Psychiatry* **1990**, *53*, 633–639. doi:10.1136/jnnp.53.8.633.

[21] Cohen, L.G.; Hallett, M.; Geller, B.D.; Hochberg, F. Treatment of Focal Dystonias of the Hand with Botulinum Toxin Injections. *J Neurol Neurosurg Psychiatry* **1989**, *52*, 355–363. doi:10.1136/JNNP.52.3.355.

[22] Gupta, N.; Pandey, S. Treatment of Focal Hand Dystonia: Current Status. *Neurol Sci* **2021**, *42*, 3561–3584. doi:10.1007/S10072-021-05432-7.

[23] Gupta, N.; Pandey, S. Treatment of Focal Hand Dystonia: Current Status. *Neurol Sci* **2021**, *42*, 3561–3584. doi:10.1007/S10072-021-05432-7.

[24] Lee, A.; Al-Sarea, J.; Altenmüller, E. Nonlinear Changes in Botulinum Toxin Treatment of Task-Specific Dystonia during Long-Term Treatment. *Toxins (Basel)* **2021**, *13*. doi:10.3390/TOXINS13060371/S1.

[25] Hallett, M. Mechanism of Action of Botulinum Neurotoxin: Unexpected Consequences. *Toxicon* **2018**, *147*, 73–76. doi:10.1016/J.TOXICON.2017.08.011.

[26] Weise, D.; Weise, C.M.; Naumann, M. Central Effects of Botulinum Neurotoxin-Evidence from Human Studies. *Toxins (Basel)* **2019**, *11*. doi:10.3390/TOXINS11010021.

[27] Ioannou, C.I.; Hodde-Chriske, F.L.; Altenmüller, E. Long-Term Muscular Atrophy and Weakness Following Cessation of Botulinum Toxin Type A Injections in the Flexor Digitorum Muscle of Musicians with Focal Hand Dystonia. *Toxins (Basel)* **2023**, *15*. doi:10.3390/TOXINS15040296.

[28] Enke, A.M.; Poskey, G.A. Neuromuscular Re-Education Programs for Musicians with Focal Hand Dystonia: A Systematic Review. *Med Probl Perform Art* **2018**, *33*, 137–145. doi:10.21091/MPPA.2018.2014.

[29] Jabusch, H.C.; Zschucke, D.; Schmidt, A.; Schuele, S.; Altenmüller, E. Focal Dystonia in Musicians: Treatment Strategies and Long-Term Outcome in 144 Patients. *Mov Disord* **2005**, *20*, 1623–1626. doi:10.1002/MDS.20631.

[30] Kaji, R.; Rothwell, J.C.; Katayama, M.; Ikeda, T.; Kubori, T.; Kohara, N.; Mezaki, T.; Shibasaki, H.; Kimura, J. Tonic Vibration Reflex and Muscle Afferent Block in Writer's Cramp. *Ann Neurol* **1995**, *38*, 155–162. doi:10.1002/ANA.410380206.

[31] Loyola, D.P.; Camargos, S.; Maia, D.; Cardoso, F. Sensory Tricks in Focal Dystonia and Hemifacial Spasm. *Eur J Neurol* **2013**, *20*, 704–707. doi:10.1111/ENE.12054.

[32] Dagostino, S.; Ercoli, T.; Gigante, A.F.; Pellicciari, R.; Fadda, L.; Defazio, G. Sensory Trick in Upper Limb Dystonia. *Parkinsonism Relat Disord* **2019**, *63*, 221–223. doi:10.1016/J.PARKRELDIS.2019.01.006.

[33] Zeuner, K.E.; Shill, H.A.; Sohn, Y.H.; Molloy, F.M.; Thornton, B.C.; Dambrosia, J.M.; Hallett, M. Motor Training as Treatment in Focal Hand Dystonia. *Mov Disord* **2005**, *20*, 335–341. doi:10.1002/MDS.20314.

[34] Sakai, N.; Liu, M.C.; Su, F.C.; Bishop, A.T.; An, K.N. Hand Span and Digital Motion on the Keyboard: Concerns of Overuse Syndrome in Musicians. *J Hand Surg Am* **2006**, *31*, 830–835. doi:10.1016/J.JHSA.2006.02.009.

[35] Berque, P.; Gray, H.; Harkness, C.; McFadyen, A. A Combination of Constraint-Induced Therapy and Motor Control Retraining in the Treatment of Focal Hand Dystonia in Musicians. *Med Probl Perform Art* **2010**, *25*, 149–161. doi:10.21091/mppa.2010.4032.

[36] Butler, K.; Sadnicka, A.; Freeman, J.; Meppelink, A.M.; Pareés, I.; Marsden, J.; Edwards, M.J. Sensory–Motor Rehabilitation Therapy for Task-Specific Focal Hand Dystonia: A Feasibility Study, **2018**, *23*, 53–63. doi:10.1177/1758998318764219.

[37] Ramella, M.; Borgnis, F.; Giacobbi, G.; Castagna, A.; Baglio, F.; Cortesi, M.; Converti, R.M. Modified Graded Motor Imagery for Musicians' Focal Dystonia: A Case Series. *Med Probl Perform Art* **2021**, *36*, 10–17. doi:10.21091/MPPA.2021.1002.

[38] Ackermann, B.; Altenmüller, E. The Development and Use of an Anatomy-Based Retraining Program (MusAARP) to Assess and Treat Focal Hand Dystonia in Musicians-A Pilot Study. *J Hand Ther* **2021**, *34*, 309–314. doi:10.1016/J.JHT.2021.05.007.

[39] Byl, N.N.; Merzenich, M.M.; Cheung, S.; Bedenbaugh, P.; Nagarajan, S.S.; Jenkins, W.M. A Primate Model for Studying Focal Dystonia and Repetitive Strain Injury: Effects on the Primary Somatosensory Cortex. *Phys Ther* **1997**, *77*, 269–284. doi:10.1093/ptj/77.3.269.

[40] Tinazzi, M.; Fiorio, M.; Fiaschi, A.; Rothwell, J.C.; Bhatia, K.P. Sensory Functions in Dystonia: Insights from Behavioral Studies. *Movement Disorders* **2009**, *24*, 1427–1436. doi:10.1002/MDS.22490.

[41] Berardelli, A.; Rothwell, J.C.; Hallett, M.; Thompson, P.D.; Manfredi, M.; Marsden, C.D. The Pathophysiology of Primary Dystonia. *Brain* **1998**, *121*, 1195–1212. doi:10.1093/BRAIN/121.7.1195.

[42] Abbruzzese, G.; Berardelli, A. Sensorimotor Integration in Movement Disorders. **2003**, *18*, 231–240. doi:10.1002/mds.10327.

[43] Quartarone, A.; Rizzo, V.; Bagnato, S.; Morgante, F.; Sant'Angelo, A.; Romano, M.; Crupi, D.; Girlanda, P.; Rothwell, J.C.; Siebner, H.R. Homeostatic-like Plasticity of the Primary Motor Hand Area Is Impaired in Focal Hand Dystonia. *Brain* **2005**, *128*, 1943–1950. doi:10.1093/BRAIN/AWH527.

[44] Quartarone, A.; Sant'angelo, A.; Battaglia, F.; Bagnato, S.; Rizzo, V.; Morgante, F.; Rothwell, J.C.; Siebner, H.R.; Girlanda, P. Neurobiology of Disease Enhanced Long-Term Potentiation-Like Plasticity of the Trigeminal Blink Reflex Circuit in Blepharospasm. **2006**. doi:10.1523/JNEUROSCI.3948-05.2006.

[45] Meunier, S.; Hallett, M. Endophenotyping. *Neurology* **2005**, *65*, 792–793. doi:10.1212/01.WNL.0000177919.02950.4A.

[46] Defazio, G.; Berardelli, A.; Hallett, M. Do Primary Adult-Onset Focal Dystonias Share Aetiological Factors? *Brain* **2007**, *130*, 1183–1193. doi:10.1093/BRAIN/AWL355.

[47] Breakefield, X.O.; Blood, A.J.; Li, Y.; Hallett, M.; Hanson, P.I.; Standaert, D.G. The Pathophysiological Basis of Dystonias. *Nature Reviews Neuroscience 2008 9:3* **2008**, *9*, 222–234.doi:10.1038/nrn2337.

[48] Torres-Russotto, D.; Perlmutter, J.S. Focal Dystonias of the Hand and Upper Extremity. *J Hand Surg Am* **2008**, *33*, 1657–1658. doi:10.1016/J.JHSA.2008.09.001.

[49] Lin, P.T.; Hallett, M. The Pathophysiology of Focal Hand Dystonia. *J Hand Ther* **2009**, *22*, 109–114. doi:10.1016/J.JHT.2008.10.008.

[50] Hallett, M. Neurophysiology of Dystonia: The Role of Inhibition. *Neurobiol Dis* **2011**, *42*, 177–184. doi:10.1016/J.NBD.2010.08.025.

[51] Desrochers; Brunfeldt; Sidiropoulos; Kagerer Sensorimotor Control in Dystonia. *Brain Sci* **2019**, *9*, 79. doi:10.3390/brainsci9040079.

[52] Rossi, S.; Pasqualetti, P.; Tecchio, F.; Sabato, A.; Rossini, P.M. Modulation of Corticospinal Output to Human Hand Muscles Following Deprivation of Sensory Feedback. *Neuroimage* **1998**, *8*, 163–175. doi:10.1006/nimg.1998.0352.

[53] Tecchio, F.; Zappasodi, F.; Pasqualetti, P.; Rossini, P.M. Neural Connectivity in Hand Sensorimotor Brain Areas: An Evaluation by Evoked Field Morphology. *Hum Brain Mapp* **2005**, *24*, 99–108. doi:10.1002/hbm.20073.

[54] Tecchio, F.; Zappasodi, F.; Melgari, J.M.; Porcaro, C.; Cassetta, E.; Rossini, P.M. Sensory-Motor Interaction in Primary Hand Cortical Areas: A Magnetoencephalography Assessment. *Neuroscience* **2006**, *141*, 533–542. doi:10.1016/j.neuroscience.2006.03.059.

[55] Tecchio, F.; Pasqualetti, P.; Zappasodi, F.; Tombini, M.; Lupoi, D.; Vernieri, F.; Rossini, P.M. Outcome Prediction in Acute Monohemispheric Stroke via Magnetoencephalography. *J Neurol* **2007**, *254*, 296–305. doi:10.1007/s00415-006-0355-0.

[56] Tecchio, F.; Zappasodi, F.; Porcaro, C.; Barbati, G.; Assenza, G.; Salustri, C.; Maria Rossini, P. High-Gamma Band Activity of Primary Hand Cortical Areas: A Sensorimotor Feedback Efficiency Index. *Neuroimage* **2008**, *40*, 256–264. doi:10.1016/j.neuroimage.2007.11.038.

[57] Tecchio, F.; Zito, G.; Zappasodi, F.; Dell'Acqua, M.L.; Landi, D.; Nardo, D.; Lupoi, D.; Rossini, P.M.; Filippi, M.M. Intra-Cortical Connectivity in Multiple Sclerosis: A Neurophysiological Approach. *Brain* **2008**, *131*, 1783–1792. doi:10.1093/brain/awn087.

[58] Zito, G.; Luders, E.; Tomasevic, L.; Lupoi, D.; Toga, A.W.; Thompson, P.M.; Rossini, P.M.; Filippi, M.M.; Tecchio, F. Inter-Hemispheric Functional Connectivity Changes with Corpus Callosum Morphology in Multiple Sclerosis. *Neuroscience* **2014**, *266*, 47–55. doi:10.1016/j.neuroscience.2014.01.039.

[59] Tecchio, F.; Bertoli, M.; Gianni, E.; L'Abbate, T.; Paulon, L.; Zappasodi, F. To Be Is To Become. Fractal Neurodynamics of the Body-Brain Control System. *Front Physiol* **2020**, *11*. doi:10.3389/fphys.2020.609768.

[60] Persichilli, G.; Grifoni, J.; Pagani, M.; Bertoli, M.; Gianni, E.; L'Abbate, T.; Cerniglia, L.; Bevacqua, G.; Paulon, L.; Tecchio, F. Sensorimotor Interaction Against Trauma. *Front Neurosci* **2022**, *16*. doi:10.3389/FNINS.2022.913410.

[61] Beck, S.; Richardson, S.P.; Shamim, E.A.; Dang, N.; Schubert, M.; Hallett, M. Short Intracortical and Surround Inhibition Are Selectively Reduced during Movement Initiation in Focal Hand Dystonia. *J Neurosci* **2008**, *28*, 10363–10369. doi:10.1523/JNEUROSCI.3564-08.2008.

[62] Pagliara, M.R.; Cecconi, F.; Pasqualetti, P.; Bertoli, M.; Armonaite, K.; Gianni, E.; Grifoni, J.; L'Abbate, T.; Marinozzi, F.; Conti, L.; et al. On the Homology of the Dominant and Non-Dominant Corticospinal Tracts: A Novel Neurophysiological Assessment. *Brain Sciences 2023, Vol. 13, Page 278* **2023**, *13*, 278. doi:10.3390/BRAINSCI13020278.

[63] Bertoli, M.; Tataranni, A.; Porziani, S.; Pasqualetti, P.; Gianni, E.; Grifoni, J.; L'Abbate, T.; Armonaite, K.; Conti, L.; Cancelli, A.; et al. Effects on Corticospinal Tract Homology of Faremus Personalized Neuromodulation Relieving Fatigue in Multiple Sclerosis: A Proof-of-Concept Study. *Brain Sciences 2023, Vol. 13, Page 574* **2023**, *13*, 574. doi:10.3390/BRAINSCI13040574.

[64] Ridding, M.C.; Sheean, G.; Rothwell, J.C.; Inzelberg, R.; Kujirai, T. Changes in the Balance between Motor Cortical Excitation and Inhibition in Focal, Task Specific Dystonia. *J Neurol Neurosurg Psychiatry* **1995**, *59*, 493–498. doi:10.1136/JNNP.59.5.493.

[65] Melgari, J.M.; Zappasodi, F.; Porcaro, C.; Tomasevic, L.; Cassetta, E.; Rossini, P.M.; Tecchio, F. Movement-Induced Uncoupling of Primary Sensory and Motor Areas in Focal Task-Specific Hand Dystonia. *Neuroscience* **2013**, *250*, 434–445. doi:10.1016/j.neuroscience.2013.07.027.

[66] Stefan, K.; Kunesch, E.; Cohen, L.G.; Benecke, R.; Classen, J. Induction of Plasticity in the Human Motor Cortex by Paired Associative Stimulation. *Brain* **2000**, *123 Pt 3*, 572–584. doi:10.1093/BRAIN/123.3.572.

[67] Quartarone, A.; Bagnato, S.; Rizzo, V.; Siebner, H.R.; Dattola, V.; Scalfari, A.; Morgante, F.; Battaglia, F.; Romano, M.; Girlanda, P. Abnormal Associative Plasticity of the Human Motor Cortex in Writer's Cramp. *Brain* **2003**, *126*, 2586–2596. doi:10.1093/BRAIN/AWG273.

[68] Quartarone, A.; Siebner, H.R.; Rothwell, J.C. Task-Specific Hand Dystonia: Can Too Much Plasticity Be Bad for You? *Trends Neurosci* **2006**, *29*, 192–199. doi:10.1016/J.TINS.2006.02.007.

[69] Elbert, T.; Candia, V.; Altenmüller, E.; Rau, H.; Sterr, A.; Rockstroh, B.; Pantev, C.; Taub, E. Alteration of Digital Representations in Somatosensory Cortex in Focal Hand Dystonia. *Neuroreport* **1998**, *9*, 3571–3575. doi:10.1097/00001756-199811160-00006.

[70] Sadnicka, A.; Hamada, M. Plasticity and Dystonia: A Hypothesis Shrouded in Variability. *Exp Brain Res* **2020**, *238*, 1611–1617. doi:10.1007/S00221-020-05773-3.

[71] Rosenkranz, K.; Williamon, A.; Butler, K.; Cordivari, C.; Lees, A.J.; Rothwell, J.C. Pathophysiological Differences between Musician's Dystonia and Writer's Cramp. *Brain* **2005**, *128*, 918–931. doi:10.1093/BRAIN/AWH402.

[72] Schoeb, V.; Zosso, A. "You Cannot Perform Music without Taking Care of Your Body": A Qualitative Study on Musicians' Representation of Body and Health. *Med Probl Perform Art* **2012**, *27*, 129–136. doi:10.21091/mppa.2012.3024.

[73] Dressler, D.; Altenmueller, E.; Bhidayasiri, R.; Bohlega, S.; Chana, P.; Chung, T.M.; Frucht, S.; Garcia-Ruiz, P.J.; Kaelin, A.; Kaji, R.; et al. Strategies for Treatment of Dystonia. *J Neural Transm (Vienna)* **2016**, *123*, 251–258. doi:10.1007/S00702-015-1453-X.

[74] Grifoni, J., Pagani, M., Persichilli, G., Bertoli, M., Bevacqua, M. G., L'Abbate, T., ... & Tecchio, F. 2**023**. Auditory Personalization of EMDR Treatment to Relieve Trauma Effects: A Feasibility Study [EMDR+ ®]. Brain Sciences, 13(7).

[75] Grifoni, J., Crispiatico, V., Castagna, A., Quartarone, A., Converti, R. M., Ramella, M., ... & Tecchio, F. (2024). Musician's dystonia: an opinion on novel treatment strategies. *Frontiers in Neuroscience,* 18, 1393767.

[76] Ioannou, C.I.; Altenmüller, E. Psychological Characteristics in Musician′s Dystonia: A New Diagnostic Classification. Neuropsychologia 2014, 61, 80–88. doi:10.1016/J.NEUROPSYCHOLOGIA.2014.05.014.

[77] Feener, R. S. (2004). EMDR: Eye movement desensitization and reprocessing, a new method in the treatment of performance anxiety for singers. The Florida State University.

[78] Ioannou, C.I.; Furuya, S.; Altenmüller, E. The Impact of Stress on Motor Performance in Skilled Musicians Suffering from Focal Dystonia: Physiological and Psychological Characteristics. Neuropsychologia **2016**, 85, 226–236. doi:10.1016/J.NEUROPSYCHOLOGIA.2016.03.029.

[79] Ioannou, C.I.; Furuya, S.; Altenmüller, E. Objective Evaluation of Performance Stress in Musicians With Focal Hand Dystonia: A Case Series. J Mot Behav **2016**, 48, 562–572. doi:10.1080/00222895.2016.1161590.

POSTER PRESENTATION

Exploring the Depths of Interpretation to uncover creative Dimensions in Musical Performance

Axelle Bruyne

KULeuven
axellebruyne@hotmail.fr

Abstract. This study delves into the intricate processes involved in the development of musical interpretation, particularly focusing on the creative aspects inherent in the work for solo violin "*When Lilacs last in the dooryard bloom'd*" by composer Roberto Ventimiglia. Building upon existing research in music psychology and cognitive sciences, this investigation proposes three hypotheses. First, it suggests that musical interpretation extends beyond mere technical proficiency, encompassing nuanced and unique expressions that vary among performers. Second, it suggests the presence of distinct stages, strategies, and creative processes specific to expressive musical work, which have yet to be fully explored. Drawing from cognitive theories of creativity, such as idea incubation and the use of metaphors and analogies, this study aims to shed light on mental activities during instrumental work that occur beyond formal rehearsals. Lastly, the study proposes the development of a methodology to systematically capture and analyze the creative processes employed by the performer during rehearsal sessions. By identifying and understanding these creative elements, this research seeks to contribute to a deeper comprehension of how a musician can craft original and compelling interpretations, bridging the gap between technical proficiency and artistic expression in musical performance.

Keywords: creative processes, implicit learning processes, expressive gestures, musical creativity, temporal dynamics, music performance, artistic vision.

This book presents a collection of selected papers that present the current variety of all aspect of music research, development and education, at a high level. The respective chapters address a diverse range of theoretical, empirical and practical aspects underpinning the music science and teaching and learning, as well as their pedagogical implications. The book meets the growing demand of practitioners, researchers, scientists, educators and students for a comprehensive introduction to key topics in these fields. The volume focuses on easy-to-understand examples and a guide to additional literature.

Michele Della Ventura, editor

New Music Concepts and Inspired Education

Revised Selected Papers

ISBN: 978-88-944350-9-2

www.studiomusicatreviso.it