Real-Time Computational Feedback Systems for Enriched Music Pedagogy

Alberto Acquilino

Computational Acoustic Modeling Laboratory - Music Technology Area
Department of Music Research, Schulich School of Music
McGill University
Montreal, Quebec, Canada

July 2025

A thesis submitted to McGill University in partial
fulfillment of the requirements
of the degree of Doctor of Philosophy

Abstract

This dissertation explores how modern digital technologies can be leveraged to enrich music instrument pedagogy, with a particular focus on trumpet instruction. Beginning with a broad review of existing educational tools and software, the research identifies unrealized potential for delivering real-time performance feedback and supporting student-centered learning.

To address this potential, a modular open-source framework is presented for creating educational applications, enabling the development of browser-based and hybrid mobile solutions through reusable components and services. Grounded in pedagogical best practices, this framework facilitates learning through guided repetition, structured reflection, and multi-sensory engagement. Its capabilities are demonstrated with a case study on trumpet fingering, showcasing how thoughtful app design can address critical technical and auditory skills in music education.

An exploratory study with adult learners was conducted to assess the practical application of the framework. It examined participants' interactions with the software, gathered insights into their perceptions of ease of use and perceived usefulness, and collected feedback for improvements, providing key perspectives on its effectiveness in real-world learning scenarios.

Building on this foundation, the manuscript examines additional technical skills central to music instrument pedagogy, including efficiency in sound production and articulation. Machine learning classifiers are developed and evaluated using an extensive dataset of trumpet tones, enabling the prediction of trumpet sound production efficiency based on timbre quality. Results demonstrate strong alignment with professional instructors' assessments, underscoring the promise of automated feedback in supporting the development of good playing habits. Additionally, a digital signal processing algorithm is introduced to measure the duration of attack transients in wind and bowed string instrument sounds, providing a method for monitoring articulation.

By integrating educational theory, open-source software development, and signal processing, this dissertation offers a flexible and extensible framework for music pedagogy. While primarily focused on trumpet instruction within the Western musical tradition, the presented methods and technologies are designed to be adaptable for other instruments and traditions, opening new possibilities in music education.

Résumé

Cette thèse explore comment les technologies numériques modernes peuvent être exploitées pour enrichir la pédagogie de la pratique instrumentale, avec un accent particulier sur l'enseignement de la trompette. En commençant par un examen approfondi des outils éducatifs et des logiciels existants, cette recherche identifie un potentiel inexploité des technologies numériques afin de fournir un retour en temps réel sur la performance musicale et soutenir un apprentissage autonome de l'élève.

Afin d'exploiter ce potentiel, nous avons utilisé un environnement modulaire et *open-source* dedié à la création d'applications éducatives, et permettant le développement de solutions sur navigateur et mobiles hybrides grâce à des composants et services réutilisables. Fondé sur les meilleures pratiques pédagogiques, ce cadre facilite l'apprentissage par répétition guidée, réflexion structurée et engagement multisensoriel. Les capacités de ce cadre sont illustrées par une étude de cas sur le doigté de la trompette, mettant en évidence comment la conception d'application bien pensé peut révéler des compétences techniques et auditives essentielles pour l'éducation musicale.

Une étude exploratoire auprès d'étudiant.es adultes a été menée afin d'évaluer l'application pratique du cadre. Nous avons examiné les interactions des participant.es avec le logiciel, recueilli leur témoignage concernant la facilité d'utilisation et l'utilité perçue de ce logiciel. Les perspectives des participant.es étaient en effet essentielles pour évaluer l'efficacité du logiciel tant pour des scénarios d'apprentissage réels que pour procéder à des améliorations.

S'appuyant sur cette base, le manuscrit examine des compétences techniques supplémentaires essentielles à la pédagogie de la pratique instrumentale, notamment en ce qui concerne l'efficacité de la production sonore et de l'articulation. Des classificateurs d'apprentissage automatique sont ensuite développés et évalués à l'aide d'un vaste ensemble de données de sons de trompette, permettant de prédire l'efficacité de la production sonore en fonction de la qualité du timbre. Les résultats montrent une forte concordance avec les évaluations des enseignants professionnels, soulignant le potentiel du retour automatisé pour le développement de bonnes habitudes de jeu. De plus, un algorithme de traitement du signal numérique est introduit pour mesurer la durée des transitoires d'attaque dans les sons des instruments à vent et à cordes frottées, fournissant un outil de suivi de l'articulation.

En intégrant la théorie de l'éducation, le développement de logiciels *open-source* et le traitement du signal, cette thèse propose un cadre flexible et extensible pour la pédagogie musicale. Bien que principalement axées sur l'enseignement de la trompette dans la tradition musicale occidentale, les méthodes et technologies présentées sont conçues pour être adaptées à d'autres instruments et traditions, ouvrant ainsi de nouvelles perspectives pour l'éducation musicale.

Acknowledgements

I would like to sincerely thank my supervisor, Gary Scavone, for supervising and guiding my research, accepting me as his student and always making me feel welcome. My deep gratitude also goes to my thesis committee members—Roger Dannenberg, Philippe Depalle, Marcelo Wanderley, Lisa Barg, and Erin Strumpf—for their invaluable insights and guidance.

Special thanks go to my colleagues at the Computational Acoustic Modeling Laboratory, as well as the professors, researchers, and friends within the Music Technology community for their daily support and inspiring discussions. Without these interactions, my diverse and interdisciplinary research journey would not have been as enjoyable and stimulating. In you, I found a true international family that welcomed and embraced me for who I am, and for which I feel deeply grateful and reciprocate sincere affection.

I specifically wish to acknowledge the co-authors of the articles comprising this thesis—Mirko D'Andrea, Ninad Puranik, Ichiro Fujinaga, Gianpaolo Evangelista, Jenny Ji Eun Park, and Keerthi Reddy Kambham—for their fundamental contributions and support. My gratitude extends to the broader community of wonderful individuals whom I've had the privilege to know, enriching my research experience.

A special acknowledgment goes to Mirko, who has shared my passions for music and computational research since before my arrival in Montreal. Together, we developed the initial ideas underpinning this thesis and he has served as both developer and mentor throughout the technological evolution of this project. My sincere thanks also to Emanuela Bussino for her enthusiasm and collaboration in collecting the audio dataset.

I also owe particular thanks to the Music Education area at McGill University—especially Isabelle Cossette, Andrea Creech, Liliana Araujo, and Lisa Lorenzino—for their mentoring, valuable advice, and assistance across various thesis chapters. My appreciation also extends to Bennett Smith and Darryl Cameron for their technical support, Stephen McAdams and Erica Ying Huynh for their methodological guidance, Alessandro Casasola for his essential assistance in developing the interface layout, Sevag Hanssian for implementing the pitch tracking algorithm, and Aleksandra Petrukova for significantly enhancing the software's graphical interface.

A heartfelt thank you goes to Ermanno Ottaviani for the inspiring discussions and valuable exchanges on the educational aspects of this research, and for providing the trumpet reference recordings used in the developed software.

I am profoundly grateful to the anonymous study participants, the experts who generously dedicated their time and expertise to annotate the audio dataset, and the musicians who participated in software testing. In this regard, special thanks to Professor Trevor Dix and the leadership of the McGill Symphonic Band Club and the *Harmonie Nouveaux Horizons de Montréal* for their enthusiasm and willingness to facilitate participant recruitment.

I would also like to express my gratitude to the Center for Interdisciplinary Research in Music Media and Technology for fostering an extraordinary environment of research and collaboration, and for providing crucial support and opportunities to me and the entire Music Technology community.

This research was made possible thanks to the generous support of a Tomlinson Doctoral Fellowship and an *Fonds de Recherche du Québec Nature et Technologies* doctoral training scholarship, enabling me to focus fully on my research without financial concern.

To my dear friends in Montreal and around the world who have shown constant affection—you are countless, and you have no idea how important your presence has been. I was in Montreal alone, yet because of you, I have never felt lonely.

Finally, my heartfelt thanks go to my parents, my siblings, my sister-in-law, my nephews, and my entire family. I know maintaining a long-distance family relationship is challenging—especially for an Italian family—but you cannot imagine how significant it was for me to always feel fully included and loved during our video calls, your visits, and each of my returns to Italy.

I love you deeply.

It is my hope that this research will positively impact and inspire future generations, contributing to the meaningful change I wish to see in the world.

Contributions

This thesis, and the research it describes, is original work by the candidate, except for commonly understood and accepted concepts, or where explicit reference to the work of others is made. The dissertation is formatted as a monograph comprised of seven chapters and five appendices. It includes content from the following journal and conference publications:

1. A. Acquilino and G. Scavone. Current state and future directions of technologies for music instrument pedagogy. *Frontiers in Psychology*, 13, 2022.

2. A. Acquilino, N. Puranik, I. Fujinaga, and G. Scavone. Detecting efficiency in trumpet sound production: Proposed methodology and pedagogical implications. In *Proceedings of the 5th Stockholm Music Acoustic Conference*, pages 72–79, Stockholm, 2023b. KTH Royal Institute of Technology.

3. A. Acquilino, N. Puranik, I. Fujinaga, and G. Scavone. A dataset and baseline for automated assessment of timbre quality in trumpet sound. In *Proceedings of the 24th International Society of Music Information Retrieval Conference*, pages 684–691, 2023a.

4. A. Acquilino, M. d'Andrea, K. K. Reddy, J. J. Park, and G. Scavone. Open-source mobile apps for music education: A case study on trumpet fingering. *Submitted to International Journal of Music Education*, 2025.

5. G. Evangelista and A. Acquilino. An adaptive wavelet-based algorithm for assessing the quality of the attack transients in non-percussive instruments. In *Proceedings of the 28th International Conference on Digital Audio Effects (DAFx25)*, Ancona, Italy, 2025. A3Lab, Università Politecnica delle Marche.

A detailed account of the individual contributions of each author to these publications is provided at the end of each corresponding chapter.

Contents

List of Figures

List of Tables

List of Symbols

A_{thr}	Threshold level for attack onset detection
α	Fractional threshold parameter for attack offset detection
$a_{n,m}$	Wavelet expansion coefficients
BW_N^{tooth}	Bandwidth of each tooth of the harmonic comb at scale level N (Hz)
$b_{N,k}$	Scaling expansion coefficients
$e_f(t)$	Amplitude envelope of the noisy fluctuations associated with the attack transient
$e_h(t)$	Amplitude envelope of the harmonic content of the sound
f_s	Sampling rate of the signal (Hz)
$L^2(\mathbb{R})$	Space of square-integrable functions on real axis
$\ell^2(\mathbb{Z})$	Space of square-summable sequences indexed by integers
μ	Sample mean
μ_{dW}	Mean attack duration estimated using the wavelet-based method
μ_{dT}	Mean attack duration estimated using the input envelope thresholding method
μ_{dTeff}	Mean attack duration estimated using the weakest-effor method
n	Scale index in wavelet transform, $n \in \mathbb{N}^*$
N	Highest scale index (cutoff scale) in wavelet decomposition
P	Time period of discrete-time periodic or pseudo-periodic signal
r	Maximum value of the envelope of the harmonic component
$s(t)$	Signal as a function of time
$s_f(t)$	Scale-truncated wavelet expansion (fluctuating component)

$s_h(t)$	Scaling residue (harmonic trend component)
$s_{in}(t)$	Input signal for attack onset detection
t_{on}	Attack onset time (start of the attack phase)
t_{off}	Attack offset time (end of the attack phase)
t_{dW}	Wavelet-based estimation of the attack duration time
t_{nd}	Noise ducking time (instant when harmonic components overtake initial attack noise)
σ	Standard deviation of a general parameter
σ_{dW}	Standard deviation of the wavelet-based estimation of the attack duration time
σ_{dT}	Standard deviation of the input envelope thresholding method's estimation of the attack duration time
σ_{dTeff}	Standard deviation of the weakest-effor method's estimation of the attack duration time
$\Phi(f)$	Discrete-Time Fourier Transform (DTFT) of the discrete-time scaling function
$\phi(t)$	Scaling function (low-pass function)
$\phi_{N,k}(t)$	Scaling function indexed by scale N and translation k
$\psi(t)$	Mother wavelet
$\psi_{n,m}(t)$	Wavelet basis functions indexed by scale n and translation m

Chapter 1

Introduction and motivation

1.1 Background

Musical training and performance foster brain plasticity, reshaping the brain's structure and function through rich sensorimotor experiences associated with cognitive benefits (Dalla Bella, 2016). Mastering a musical instrument requires intricate coordination of sensory-motor skills, balancing physical execution with emotional expression. Musicians primarily rely on auditory feedback, which can be challenging especially for beginners who may lack a well-developed auditory image of the correct sound technique, making self-correction difficult during individual practice (Steenstrup, 2023). This lack of clarity in practice goals often leads to improper technique development, student frustration, and high dropout rates before attaining the benefits mentioned above (Cremaschi et al., 2015).

To address these challenges, technology can provide additional feedback mechanisms to support the learning process. Visual cues, in particular, offer immediate, actionable feedback on specific aspects of performance, which are likely to drive more effective learning (Welch et al., 2005). While traditional tools such as metronomes and tuners aid in tempo and pitch accuracy, many other musical skills that performers need to express their musical ideas remain underserved by existing technological solutions.

Various technological approaches could contribute to addressing this gap. This research focuses on the potential of web-based interactive technologies as a means to provide real-time aural and visual feedback. Web technologies offer several advantages, including real-time execution within a web browser, which eliminates the need for software installation, ensures consistent performance across different devices, and maintains low-cost, widespread accessibility (Bouras et al., 2015). Recent advancements have also significantly enhanced their computational capabilities, allowing them to handle resource-intensive tasks that previously required dedicated native applications (Perkel, 2024).

By leveraging these benefits, this study aims to develop software applications capable of analyzing audio data from musicians' performances and providing real-time feedback on sound production. These tools are designed to facilitate the acquisition of technical skills in a structured, efficient, and pedagogically sound manner. The research specifically focuses on training within the Western classical music tradition, which follows the structured pedagogical approach taught in conservatories. While the project primarily examines the trumpet as a case study, its principles can be readily extended to other musical instruments, such as winds and bowed strings. By equipping beginners with targeted feedback during individual practice, this approach seeks to mitigate the risk of developing improper playing habits while fostering a more effective and engaging learning experience.

1.1.1 Outline of the Thesis

This section provides an overview of the thesis structure. While the thesis is not formatted as a manuscript-based thesis, each chapter is largely derived from manuscripts that have been submitted or are intended for submission to scientific journals. The specific research questions and methods are detailed within the respective chapters. The thesis is organized as follows:

- **Chapter 2: Current state and future directions of technologies for music instrument pedagogy**
  This chapter provides a review of technologies for music education, identifying common trends and themes. While not exhaustive, it highlights key developments, discusses limitations, and suggests unrealized potential and future directions.

- **Chapter 3: Open-source mobile apps for music education: A case study on trumpet fingering**
  Building on the observations from the previous chapter, this research presents an open-source application designed to integrate pedagogical best practices into music education technology. The app employs a student-centered approach, facilitating guided repetition, structured reflection, and multi-sensory engagement to support technical and auditory skill development. A case study on trumpet fingering showcases its potential in fostering individualized learning and autonomy. The chapter also introduces a modular open-source framework that simplifies the development of educational tools, making them more accessible to educators and developers.

- **Chapter 4: An exploratory study on perceived usefulness in music education technologies**

Building upon the technological framework introduced in Chapter 3, this study explores the usability and potential of the proposed open-source mobile application for trumpet education. Through an exploratory study with adult learners, it examines user interactions with the software, assessing ease of use, perceived usefulness, ability to support structured practice, impact on motivation and areas for improvement. By integrating both quantitative and qualitative analyses, the findings provide empirical insights into the tool's impact on individual practice routines and its potential for broader adoption in formal music education.

- **Chapter 5: Detecting efficiency in trumpet sound production: Dataset, baseline and pedagogical implications**
  This chapter examines the role of timbre quality in brass instrument pedagogy and its connection to efficiency in sound production. A Random Forest classifier is trained on a dataset labeled by experts, achieving an accuracy level comparable to human evaluators. The model is then integrated into the educational interface presented in Chapter 3, extending its capabilities to provide real-time feedback on sound production efficiency.

- **Chapter 6: Adaptive wavelet-based algorithm for measuring attack transients in music sounds**
  This chapter presents a digital signal processing algorithm for estimating the transient duration of a note's attack. The study is motivated by its pedagogical potential, as attack clarity is a key technical skill for articulation in wind and bowed string instruments. The proposed method is evaluated on annotated datasets to assess its reliability and consistency with expert judgments.

Finally, Chapter 7 summarizes the key findings of the thesis, presenting conclusive remarks and outlining future directions for this research.

Chapter 2

Current state and future directions of technologies for music instrument pedagogy

Technological advances over the past 50 years or so have resulted in the development of a succession of hardware and software systems intended to improve the quality and effectiveness of Western music instrument pedagogy during classroom instruction or individual study. These systems have aimed to provide evaluation or visualization of single or combined technical aspects by analyzing performance data collected in real time or offline. The number of such educational technologies shows an ever-increasing trend over time, aided by the wide diffusion and availability of mobile devices. However, we believe there are unrealized opportunities for modern technologies to help music students in their technical development and assist them during their practice sessions in between visits to their teachers.

The ubiquity of PCs and mobile devices with built-in microphones, speakers and cameras has inspired the development of media technologies in support of music pedagogy. They offer an attractive potential for implementing audio signal processing algorithms addressing different technical skills of the performer, providing real-time feedback, collecting data over time and applying statistical models. Despite this potential, most available software for music instrument pedagogy remains very limited in functionality.

This chapter provides a survey of music edTech software available, together with the methods of use, addressed technical skills, commonalities and limitations. Results show that most current software is based on the metronome and tuner, with only a few systems that have limited abilities to follow a performance in real-time and compare it to a given score to monitor correctness of notes, intonation and rhythm. The survey also highlights a high and under-exploited potential regarding the monitoring of other more specific technical skills, which are more instrument-dependent, but no less important, such as the control of dynamic range and clarity of the attack.

The chapter ends with a discussion of possible directions for future development of technologies to support the practice of music students at different levels, with some consideration for the corresponding signal processing methods that can be utilized or that need advancement.

By helping students to more efficiently achieve a high level of proficiency of their instruments with assistive technologies, we hope to minimize stress and afford better enjoyment of the music performance experience for all.

This chapter is based on the following research article:

A. Acquilino and G. Scavone. Current state and future directions of technologies for music instrument pedagogy. *Frontiers in Psychology*, 13, 2022.

2.1 Introduction

Musical pedagogy for learning traditional Western musical instruments is currently most often delivered in one-on-one or group contexts through a master-apprentice model, typically one time per week for 30 to 60 minutes per session (Hanken, 2017). Between those meetings, students practice on their own the assigned exercises and, attempting to apply the suggestions received in class, try to reach the learning goals set out by the instructor. It is a common problem that students either misunderstand or do not correctly remember the details of a performance technique (Welch, 1985), which can lead to frustration, slower development and potentially termination of music studies.

Evidence from a wide variety of motor control tasks shows that real-time visual feedback can accelerate the learning progress (Shea and Wulf, 1999) and can help learners to identify, become aware of, or modify specific bodily actions (Welch et al., 2005). These findings suggest the development of technological tools based on audiovisual feedback to help music students address the aforementioned problems. Indeed, improvements in the effectiveness of learning classical music through aural and visual feedback has been demonstrated in different study applications (Ferguson, 2006; Leong and Cheng, 2014; Pardue and McPherson, 2019; Malandrino et al., 2019).

Among the oldest assistive technologies available for musical practice is the tuning fork, invented in 1711 by John Shore in London (Feldmann, 1997). Presenting a resonance frequency almost constant under any weather condition, this tool was used as a reference for tuning musical instruments. About a century later, the metronome was devised, providing a periodic "tick" sound at a desired tempo, typically in beats per minute, that can be set by the user. Patented for musical purposes in 1815 by Johann Maelzel, the metronome was proposed as a tool for composers, to indicate in a simple and objective way the speed of execution of their scores, and for music students, to develop a proper observance of time (Parker, 1825). More recently, based on the tuning fork principle, electronic tuners have become widely available and inexpensive, providing feedback on a player's intonation with respect to a particular tuning system, though they are often only used at the beginning of a practice session to make sure an instrument is correctly tuned.

The rapid spread of digital technologies with ever greater computational capabilities has made possible the continuous development of increasingly refined musical educational software. The metronome and tuner have been transformed from dedicated hardware devices to software that use the integrated components of PCs and mobile technologies. Furthermore, new functionalities and methods of interaction have been added that create greater engagement between the musician and the system.

A comprehensive survey on software for musicians and music teachers was provided by Axford (2015), although this field is constantly evolving and characterized by a high launch and dropout rate, making the list partially outdated after a few years. Despite there being a large number of software developed for music pedagogy in recent years, these systems appear to be underused due to interface inefficiency, technological complexity and lack of institutional support (Kenny and McDaniel, 2011; Gall, 2013; Fautley, 2013). One might expect this situation to stem from a reluctant and conservative philosophy of thinking toward technology in music education (Creech and Gaunt, 2012; Gaunt, 2017). However, Waddell and Williamon (2019) found evidence of a generally positive attitude toward current and future technology use among teachers, amateur, students and professional musicians. This also points to a general problem in perceived or actual effectivity of current software technologies for music pedagogy.

Musicians appear to be interested in integrating new technological tools into their practice routines and the ubiquity of mobile devices offers a convenient platform through which such tools can be made available. In this context, the present study provides a survey of existing technologies in the field of music education. By analyzing how they are structured, classifying them and discussing their pedagogical potential, we attempt to show their strengths and weaknesses, with the objective of providing an explanation regarding the gap between the wide availability of edTech music software and its relative under use in music education. We then discuss promising directions for future technologies in this field.

Section 2.2 outlines how educational technologies have been researched to assist music students. Section 2.3 presents a collection of the most common and innovative technologies in support of music education, proposing different classifications. Finally, in Section 2.4 we discuss their pedagogical potential in music classrooms, highlighting strengths and weaknesses, in order to illustrate future directions in the development of educational technologies in this field.

2.2 Review of musical instrument educational technologies

The present study is focused on technologies that support music students in their development in learning to play a musical instrument. Such technologies are more applicable to the learning of standard technical skills (e.g., control of dynamics, articulation, intonation) rather than musical expression, which can be more subjective. Thus, it is expected that these tools will be more beneficial for beginning, rather than advanced, students as they work to develop basic functional skills on a given instrument. From a technological standpoint, we believe that tools designed to evaluate music learning from sound signals, rather than video or special purpose sensors, hold the most promise for widespread acceptance.

There have been several past academic research projects aimed at developing tools to assist with music instrument learning. A project to support piano instruction for beginner students was pursued during the 1980s and early 1990s, with reported achievements in polyphonic score following, page turning, analysis, feedback and the application of Instructional Design theory (Dannenberg et al., 1993). Another study examined the effect on improving harmonic intonation skills, specifically the ability to play justly tuned major thirds on a reference tone, using Coda Music Technology's Intonation Trainer software program (Swift, 2003). This technology is based on the concept that musical instruments with variable pitch (e.g., strings, woodwinds, brass) can adjust their pitch as they are played. Players of these instruments are therefore released from the equal tempered intonation system and it becomes important for them to develop the ability to play chords with improved harmonic ratios (compared to equal-tempered tuning), and thus reduce beating effects. However, the idea did not find widespread adoption at a time when accessibility to a computer workstation and recording equipment was still limited to music students.

The Interactive Music Tuition System (IMUTUS) was a European project that ran from 2002–2005 with the goal of developing an open platform for training beginner students on the recorder (Tambouratzis et al., 2002; Raptis et al., 2005; Schoonderwaldt et al., 2005). It focused on score matching pitch and note onsets, with a user interface that "graded" students on their overall performance, indicated locations in the score where mistakes were made and provided some basic description on each error.

Another project was focused on the evaluation of saxophone performance using a system to track the fundamental playing frequency and perceived loudness level for specifically prescribed exercises consisting of long tones of both fixed and varying dynamic level (Robine and Lagrange, 2006; Robine et al., 2007; Percival et al., 2007; Percival, 2008). The use of such exercises helped avoid problems in distinguishing between technical errors and deliberate expressive decisions by performers, whereby they may intentionally nuance their playing to achieve expressive effects. The results of the analysis were reported to users via a simple computer interface, with additional features to allow comparison of results between other students in a class.

In the field of music information retrieval, a research project investigated the possibility of using machine learning algorithms to differentiate between good and poor quality trumpet notes (Knight et al., 2011). Each of the notes were analyzed and rated individually in a monophonic and unaccompanied context. Although the results of this study were not conclusive, the widespread application of artificial intelligence methods in nearly all computing contexts offers opportunities for the development of tools to provide useful feedback to students learning to play music instruments.

A more recent European Commission project, Technology Enhanced Learning of Music Instruments (TELMI, 2016–2019) included the design and implementation of new interaction paradigms for music learning and training based on state-of-the-art technologies (Ortega et al., 2017; Kholykhalova et al., 2017; Giraldo et al., 2019; Perez-Carrillo, 2019). The project focused primarily on violin performance, with the development of a prototype tool called SkyNote that can provide real-time feedback on pitch and intonation, dynamics, tone quality, and rhythm. When combined with a motion-tracking system, SkyNote can also monitor specific aspects of bowing technique including bow tilt, speed, weight, contact point, inclination, and direction. A recent project reported the use of an interactive robot for recorder tutoring (Bagga et al., 2019).

A limited number of technologies have been commercially developed to assist with general music learning, such as software systems for music theory, ear and rhythm training, music notation and music instrument practice. In Section 2.3, we provide an overarching overview of these software, analyzing classifications between them and examining their functions. In Section 2.4, the potential limitations of such software and the possible future directions from the perspective of optimal technology enhanced music learning are discussed.

2.3 Review of current educational technologies

In this section, a list of computer software and mobile apps, chosen among the most popular for number of downloads and the most innovative systems created for music pedagogy, is analyzed and described. The software selected in alphabetical order are (refer to Appendix

A for a list of URL references): Anytune Pro+, Amazing Slow Downer, EarMaster, Estill Voiceprint Plus, forScore, GNU Solfege, Guitar Pro, GuitarToolkit, GuitarTuna, KORG cortosia, Knock Box Metronome, Modacity: Pro Music Practice, liveBPM - Beat Detector, Piascore, QuantiForce, Rec'n'Share, Rhythm Teacher, Rhythm Trainer, Riyaz, RTFactory Rudiments, SkyNote, SmartMusic, Tempo, The Metronome by Soundbrenner, TonalEnergy, tonestro, Visual Note, Yousician. These edTech systems offer functionality normally applicable to all categories of musical instruments, with some exceptions for technologies dedicated to plucked strings (i.e., Guitar Pro, GuitarToolkit, GuitarTuna, Visual Note, Yousician), to percussion (i.e., liveBPM, Knock Box Metronome, RTFactory Rudiments) or winds and bowed strings (i.e., KORG cortosia, QuantiForce, tonestro). Some of the systems provide flexibility in terms of expected proficiency level, allowing the learning goals and exercise levels to be modified as the student progresses.

As mentioned in Section 1, an inclusive list of software in support of music education is provided by Axford (2015) in a 250+ page book published in 2015 that is now partially outdated, given the high birth and death rate of these technologies. For this reason, we prefer to avoid the replication of a similar updated work, but to focus on the classification of the pedagogical aspects addressed. Thus, we have chosen to present a comprehensive list of software across the range of provided functionality and adopted hardware components. Within each category, we select the most popular – in terms of number of downloads – or innovative systems reported in publications.

2.3.1 Classification based on functionalities

Table 2.1 provides a list of the computer software and mobile apps considered in this study. The categories adopted for the classification are described below.

Digital score rendering

All software applications in this category provide a score in Western diastematic notation. The musician can add annotations as on a paper score (i.e., forScore, Piascore), play by turning the page through a specific functionality (e.g., foot switchers, touch pad, wink detection on camera), write in musical notation directly by playing the instrument (i.e., Guitar Pro), or following the score on a rolling window. Such software can also keep track of how much time the user spends on each exercise, allowing statistical calculations on the distribution of study time. While applications in this category do not directly assist with pedagogy, they provide a useful and popular functionality in music performance, especially as digital versions of music scores become prevalent. This type of software contains pedagogical potential especially

Table 2.1: List of software in support of music instrument learning classified according to the provided macro-functionalities.

Software	**Digital Score**	**Metronome**	**Tuner**	**Rhythmic Refinement**	**Sound**	**Fingering Display**	**Performance Feedback**	**Statistics**	**External Hardware**
Piascore	✓	✓	✓						
forScore	✓	✓	✓						
Modacity		✓	✓					✓	
GuitarToolkit		✓	✓			✓			
Guitar Pro	✓	✓				✓	✓		
GuitarTuna		✓	✓			✓			
Visual Note		✓	✓			✓	✓		✓
SmartMusic	✓	✓	✓			✓	✓	✓	
tonestro	✓	✓	✓			✓	✓	✓	
Yousician	✓	✓	✓			✓	✓	✓	
TonalEnergy		✓	✓		✓				
KORG cortosia		✓	✓		✓				
QuantiForce					✓				✓
Riyaz - Learn to Sing		✓	✓		✓		✓	✓	
Vocal Pitch Monitor		✓	✓		✓				
Estill Voiceprint Plus			✓		✓				
SkyNote	✓	✓	✓		✓		✓		✓
GNU Solfege		✓		✓					
EarMaster	✓	✓	✓	✓			✓	✓	
liveBPM				✓			✓	✓	
Tempo		✓	✓	✓					
Anytune Pro+				✓					
Amazing Slow Downer				✓					
Rec’n’Share				✓					
Rhythm Trainer				✓					
Knock Box Metronome		✓		✓					
RTFactory Rudiments	✓			✓				✓	
Soundbrenner		✓		✓			✓		✓

when embedded in larger-scale systems that include algorithms to analyze the performer's sound in parallel and provide visualization or feedback on musical skills. A popular program in this category includes forScore, which offers the possibility to read PDF scores, organize music through metadata, build set lists, annotate, rewrite lyrics, add music notation, share, download and edit the scores, as well as providing metronome, tuner and MIDI keyboard functionalities.

Metronome and basic rhythm functionalities

This category includes software systems that provide metronome functionality. This can be implemented according to its standard application by marking every beat, playing rhythmic structures of more complex subdivisions (i.e., Soundbrenner, TonalEnergy), detecting the metronomic tempo through tapping (i.e., KORG cortosia, Soundbrenner, TonalEnergy), illuminating the correct fingering in time (i.e., Visual Note), or verifying in real time the rhythmic accuracy of a musical performance on a given score (i.e., EarMaster, Riyaz, SkyNote, SmartMusic, tonestro, Yousician).

Tuner functionalities

Technologies included in this category provide tuner functionality. It can be implemented to facilitate the intonation of strings (i.e., GuitarToolkit, GuitarTuna, TonalEnergy, Visual Note, Yousician), as a chromatic tuner (i.e., EarMaster, forScore, GuitarToolkit, Modacity, Piascore, SmartMusic, TonalEnergy, Visual Note, Vocal Pitch Monitor), to tune on tuning systems other than equal temperament (i.e., Riyaz, TonalEnergy), to tune drums (i.e., Tempo), or to check the accuracy of the pitch of a musical performance on a given score (i.e., EarMaster, Riyaz, SkyNote, SmartMusic, tonestro, Yousician). For example, tonestro "listens" to a student playing along with a given (or purchased) score and provides feedback when pitches or rhythms are incorrectly executed.

Systems that assist with advanced rhythmic refinement skills

Software in this category offer exercises to improve rhythmic skills, such as rhythmic solfeggio tapping with the finger or clapping (i.e., EarMaster, GNU Solfege, Rhythm Trainer), identifying the metronomic value through sound analysis in real time (i.e., liveBPM) and offline (i.e., Rec'n'Share), setting tempo changes and rhythm patterns with increasing speed at any given number of beats (i.e., RTFactory Rudiments, Tempo), changing the tempo of an audio track (i.e., Amazing Slow Downer, Anytune Pro+, Rec'n'Share), setting cycles in which the metronome plays intermittently to check if the tempo is maintained during the absence of the beats (i.e., Knock Box Metronome), or providing rhythmic pulses on wearable hardware (i.e., Soundbrenner).

Systems that assist in the technique and control of sound production

This category includes features that provide an analysis or visualization of sound characteristics and technical aspects other than pitch, such as vibrato (i.e., Riyaz, Vocal Pitch Monitor), sound spectrum (i.e., Estill Voiceprint Plus, TonalEnergy), articulation and timbral characteristics (i.e., KORG cortosia, SkyNote), bow and brass mouthpiece pressure (i.e., QuantiForce), or posture and bow control (i.e., SkyNote). An interesting application in this category, KORG cortosia, was developed through a collaboration between KORG Inc. and Pompeu Fabra University (Bandiera et al., 2016). It provides an evaluation of what is defined as sound "goodness" by rating in real time five elements: pitch stability, dynamic stability, timbre stability, timbre richness, and attack clarity.

Fingering display

All software applications in this category provide correct fingering to play a specific note or chord. It can be displayed in the form of a chord library (i.e., GuitarToolkit, GuitarTuna), on a rolling score window in real time (i.e., Guitar Pro, Yousician), offline (i.e., SmartMusic, tonestro), or by illuminating the keys via a purchased external hardware component (i.e., Visual Note). A popular software in this category includes Yousician, which illustrates the appropriate fingering on a scrolling window in real time with the performance of a song. For plucked string instruments, it shows which string should be plucked, the corresponding fret number to press, and different colors recommend which finger to use for playing the note. In case there are different alternative fingerings for playing the same note or the same chord, Yousician suggests the most convenient solution to perform the specific song more easily.

Systems providing feedback on music performances

This category includes functionalities that display, monitor and/or assess the correctness of a music performance. The implementation of these functionalities is coupled with algorithms that check the accuracy of rhythm and pitch (i.e., Guitar Pro, Riyaz, SmartMusic, Tonestro, Yousician), timbre and articulation (i.e., SkyNote) for a given score to provide an overall grade of the performance. This type of software is generally applied to the overall evaluation of pieces from the repertoire of performance and musical expression. However, alternative applications can be found dedicated to individual technical aspects, such as monitoring tempo (e.g., LiveBPM, Soundbrenner) and indicating fingering (e.g., Visual Note) in real time.

Systems applying statistical models to keep track of the user's proficiency

Software in this category collect data on performances, displaying or analyzing them according to specific parameters, and store and process the results over time by applying statistical models to illustrate the progress of the musician (i.e., EarMaster, Riyaz, RTFactory Rudiments, SmartMusic, tonestro, Yousician). For example, EarMaster provides a window interface where users can visualize their achieved results and the time spent on each exercise, to help them monitor their progress and analyze strengths and weaknesses. The statistics functionality is also used to provide a visualization of a specific parameter over a short period of time for a single performance (i.e., liveBPM).

Systems requiring external hardware

This category highlights technologies that rely on dedicated hardware components, instead of using the built-in sensors of PCs and smartphones. They can include cameras to provide indications about posture and bow tilting angles through motion capture techniques (i.e., SkyNote), wearable devices (i.e., the Soundbrenner metronome smartwatches), force transducers (i.e., QuantiForce), or LED lighting systems (i.e., the LED keyboard adapter for guitar proposed by Visual Note).

2.3.2 Classification based on hardware components

In Table 2.1, a set of macro-functionalities for technology enhanced music learning is represented. An alternative classification consists in subdividing the aforementioned software according to the hardware components used:

- **Graphic display**: Many software systems use a graphic display to illustrate sheet music, show fingerings, provide light pulses as metronome indication, and generally explain the software functionalities. Some systems also use touch displays, for example, to add annotations or determine rhythmic information by finger tapping.

- **Microphone**: Systems that record audio signals for further processing and display make use of microphones in order to extract specific sound information, such as the fundamental frequency, onset detection, spectral descriptors for timbral information retrieval, articulation, vibrato, and loudness metering.

- **Speaker**: Some systems output audio signals through speakers, such as metronome ticks, edited audio tracks or melodic and harmonic accompaniment.

- **Camera**: Visual information can be collected using a camera in order to provide indications about posture and bow tilting angles through motion capture techniques or detect specific cues, such as winks, to turn page.

- **Other hardware components**: The software systems previously mentioned in the external hardware category all make use of non-standard hardware components not provided on PCs or mobile devices.

This further classification clearly indicates how the development of this type of software has tried to exploit the use of built-in sensors normally installed in PCs or mobile phones. Although software programming and calibration difficulties may be introduced, this choice is largely justified by marketing constraints. Indeed, systems highlighted in the rightmost column of Table 2.1, which rely on external hardware components, generally result in a significantly higher overall cost (software and hardware combined), often exceeding by more than one order of magnitude the cost of systems relying solely on built-in components.

Music pedagogy software systems that support audio and video recording of performances for subsequent analysis by students or teachers (e.g., Modacity) are not considered in Table 2.1. These systems allow students to externally identify weaknesses that need improvement and develop their own critical sense. Although this technology is still under-used, it offers very promising pedagogical potential for students of music (Fautley, 2013). However, this study intends to consider systems whose support and feedback are provided by the technology itself through the implementation of dedicated algorithms (and not as subjective judgments provided by the user).

Other categories of functionalities useful for music learning not included in Table 2.1 are the inclusion of videos on educational courses and masterclasses (e.g., Pickup Music[1], Riyaz, tonestro, TrueFire[1], Youtube[1]) or the availability of a platform to receive individual private lessons via video with professional teachers (e.g., Play with a Pro[1], Riyaz, tonestro). However, in this case the technology is used just as a communication platform to carry out live or recorded music lessons with a human teacher. This category is beyond the scope of this study, which intends to analyze an exclusive relationship with technology that the student can turn to and rely on during practice sessions in between visits to their instructors. Since music lessons for beginners typically take place once a week, we believe that the individual practice sessions between lessons contain a high learning potential which, when exploited effectively, can improve and speed up the overall learning experience.

Although the list of software examined is far from exhaustive, the described classifications give an idea of the state of the art on how software supporting music pedagogy are structured

[1]Refer to Table A.1 in Appendix A for URL references

and what types of algorithms and technologies they implement. Section 2.4 discusses the classifications provided, identifies their possible limitations and proposes future directions of technologies for music instrument pedagogy.

2.4 Discussion

The software survey and classification demonstrates the extent to which the metronome and tuner have been widely adopted by nearly all current music pedagogy technologies. They are implemented in most of the systems considered in Table 2.1, indicating a high level of perceived usefulness. Initially implemented on dedicated hardware devices, the metronome and tuner functions were integrated into PC software or mobile apps, using their built-in components. Despite the huge technological advancement of the Digital Revolution, the functionalities of the metronome and tuner are clearly considered essential in music learning contexts.

We believe that the widespread use of metronome and tuner stems mainly from the fact that they are focused on teaching or assisting with an abstract technical concept. The metronome provides an audible indication of the tempo the player has to maintain during the performance, while the tuner provides a visualization of the fundamental frequency played, comparing it to a previously selected reference frequency. Such tools help the musician to understand musical concepts that are often difficult for performers to consistently internalize or perceive. By clearly understanding the technical concept and then the musical goal to be pursued through an audiovisual learning approach, students can therefore considerably improve the quality of their practice sessions and internalize more quickly a correct way of playing. Thus, music students develop and improve procedural memory, which allows them to learn movements, habits and skills almost independently of their conscious thought (Squire, 1992). These skills, learned automatically and internalized correctly, guarantee musicians a solid and effective technical background on which to rely during the performance and allow them to improve response and recovery to mistakes during performance (Lam, 2020). In fact, being based on abstract concepts, the metronome and tuner can be effectively applied in flexible ways and without particular limitations in most performance contexts, demonstrating their universality of application.

2.4.1 Current limitations

The widespread use and perceived usefulness of the metronome and tuner in music pedagogy has inspired numerous other musical software, as previously surveyed, which have focused on developing their application on predetermined musical scores drawn from the repertoire of performance and musical expression. Indeed, current developments in many software systems have focused on expanding the metronome and tuner functionalities to provide real-time feedback on pitch and rhythm correctness during the performance of a given musical score (i.e., Guitar Pro, Riyaz, SkyNote, SmartMusic, Tonestro, Yousician). By applying an objective judgment on the accuracy of rhythm and pitch, these software offer an evaluation of the overall musical performance. However, the adoption of this technological method in the field of music education for beginners may present significant limitations to the effectiveness of their pedagogical experience:

- This type of music software, which evaluate the correct pitch and rhythm, can give the false impression that to play well and be a good musician it is sufficient to play the right notes and in time. However, this is obviously not true. A good musician is a performer capable of communicating emotions through sound, drawing on their wealth of technical skills developed and refined over time. While the musician needs to execute the notes and rhythms correctly, artistic expression fundamentally involves often subtle deviations from exact rhythmic or pitch accuracy. The attention of the performance should be mainly linked to the expressiveness and communication of emotions with the audience (which normally varies according to the type of audience, their response, the acoustics of the environment, the type of concert, etc.); the overall quality of the performance is therefore less suitable to be judged by the software, but rather by human sensitivity. In fact, musicians are granted a flexibility of expression within the technical rules to be less rigid and more communicative. This is one of the main differences between a mere MIDI performance and an artistic interpretation.

  Informal experiments with tonestro, for example, have shown that a very inexpressive performance, in which the notated dynamic and articulation marks were ignored, can achieve very high scores. On the other hand, more expressive musical performances with proper attention to notated articulations and dynamics generally earn poorer scores.

- If the software provides an evaluation of the performance by rigidly judging rhythmic and pitch correctness on a note-by-note basis, according to a subtractive method of judgement (i.e., each error lowers the overall judgement score), the musician's attention will be focused on playing correctly each note in order to achieve the highest final score. This can inhibit the expressiveness of the performer, who concentrates on playing note by note in a pedantic manner, breaking up the melody, instead of artistically playing longer and more expressive musical phrases.

  Moreover, all this can cause an incorrect approach to performance, especially for beginners, who have not yet developed a solid personal style of expression. Musicians become more focused on receiving positive feedback from the software, trying to avoid the appearance of red error marks in the display, rather than trying to express their musical ideas by seeking empathetic contact with the audience. This approach to performance, based on trying to avoid mistakes instead of proposing musical ideas and communicating emotions, can even generate tensions in musicians that ultimately affect their wellbeing.

- Some of the reviewed software follows student progress through statistical analysis of their score. Implementing statistical models applied to collected data to generate a learning curve over time is an effective way to identify strengths and weaknesses for targeted practice. However, this indication is not pedagogically relevant if the software expects the musician to sound like a robot.

Despite these potential limitations in the pedagogical experience for beginners, such software offers powerful playful and entertaining aspects for music players, which greatly encourages user motivation. In particular, the aspect of playing along with backing tracks leads the musician to imagine playing together with others, bringing a deeper involvement in the experience, although the feedback component still continues to present the aforementioned drawbacks.

Another barrier to the adoption of technology within music courses might be represented by ineffective and overly-complicated interfaces. For example, KORG cortosia is one of the few software systems that intends to address different technical aspects beyond rhythm and pitch: pitch stability, dynamic stability, timbre stability, timbre richness, and attack clarity. Although the idea of tackling different technical skills within a single app is compelling, it is severely limited in terms of the interface. The KORG cortosia software shows a five-axis view, each associated with the five different skills considered, and provides an overall numerical score averaged over those five parameters. It is therefore complicated to isolate one parameter at a time, and it is difficult for a student to focus on and manage five at once. For example, a student may need to study pitch stability while playing a crescendo or diminuendo, without the overall numerical score being affected due to changes in dynamics.

Furthermore, even if the functionality of isolating one parameter at a time were easily accessible, a numerical score may not be the most pedagogically effective way to provide feedback. For example, wind instrumentalists need to develop different types of attack or articulation, using different pronunciations, to fulfill equally varied musical needs. It is therefore difficult to implement an algorithm that gives a consistent judgment of attack clarity for all types of attacks. A generic numerical score on this technical skill may not give the students a clear understanding of what they are doing wrong and how to fix the problem. This type of feedback easily risks confusing the students further. A visualization of sound initiation, on the other hand, is much more effective from a teaching point of view, because it allows musicians to associate an image with the execution of a technical skill, and once they understand how the interaction between their body and the musical instrument affects the image, the student has the opportunity to understand how to self-correct and improve. Moreover, a visualization provides flexible feedback that can be adapted to give useful information about different types of a technical skill. For example, a wind musician may associate different images with different types of attack and, by seeking out those images during practice sessions, gain greater clarity on how to manage and master the various articulations.

Other examples of software with possible interface problems are the timbral indications of Estill Voiceprint Plus and TonalEnergy. These systems illustrate the evolution of the audio spectrum over time or the height of harmonic peaks in real time in order to provide indications of the timbral quality of the sound. The sound spectrum and its relative harmonic distribution contain important information about the correctness of the sound produced. An unnatural or strained sound may indicate the presence of muscular rigidity in a wind performer and inefficiency in playing (Thompson, 2003; Jacobs and Nelson, 2006). However, being able to extract this information by referring only to the spectrogram and its harmonic distribution is a difficult or almost impossible task for a music student.

These difficulties in analyzing particular technical abilities – such as timbre quality or technical skills considered by Estill Voiceprint Plus, TonalEnergy and KORG cortosia – are further accentuated by the fact that these software systems analyze audio data collected by microphones embedded in PCs or mobile devices. The recorded audio signal therefore depends on the particular model of microphone sensor installed (usually not suitable for recording musical instruments with sufficient quality), on the distance and position of the microphone with respect to the sound source, and on the acoustics of the room. For example, if a trombone player changes orientation or places the smartphone behind the bell in order to better see the display, the feedback provided by the software will be altered compared to when holding the smartphone in front of a stationary bell, making the system unrepeatable and unreliable. In fact, sound dynamics is a determining factor in identifying the timbral properties of an instrument (Fabiani and Friberg, 2011).

Another limit to the creation and production of technologies for music instrument learning involves the cost and complexity of necessary external hardware components. SkyNote, for example, presents excellent goals regarding what we believe can support music pedagogy. However, the project never left the research phase to find a real application in music classrooms, as it requires hardware equipment that is too expensive and sophisticated to be easily obtained and installed by a music student.

In the next subsection, we propose possible directions for technologies in support of music pedagogy that address the limitations mentioned above.

2.4.2 Future directions

Given the issues discussed in the previous section, a sensible direction for the development of new pedagogic software systems is to focus on teaching a specific technical concept in an "exercise-like" context (in comparison to a context in which the player may be inclined to be musically expressive). In this way, the musician learns the technical skill in a universal context and, once internalized, can apply it confidently to any performance without incurring the aforementioned risks and limitations. Considering feedback and visualization on a technical aspect, rather than a performance, allows the system to provide higher accuracy and reliability, given fewer variables involved in software development. By focusing on a specific technical skill, players are expected to play like a robot, in order to train their muscle memory through deliberate practice. Systems designed in this way would have a type of functionality that is similar to the metronome and tuner.

There are other technical aspects besides pitch and rhythm that can be addressed with newer technologies and a development in this direction could open new ways to enrich musical pedagogy. These technical abilities are generally more dependent on the particular musical instrument played, requiring greater specificity of the parameters analyzed and provided by the system. Here lies significant potential that is still under-explored in the field of technology-enhanced music learning.

Skills which are fundamental for the optimal technical control of a musical instrument include for example dynamics, vibrato, articulation, staccato/tonguing, sound resonance, body setting (e.g., efficient embouchure, bow and sticks handling), or legato quality. Some of the software listed in the Table 2.1 pursue this direction, although in some cases their pedagogical potential may face the mentioned limitations. In the following, a selection of addressed technical aspects, are analyzed and discussed.

Dynamics and timbre characteristics

Some software provide a real-time display of the sound spectrum or of the harmonic energy content, through which the musician can search for specific timbral characteristics and dynamics (e.g., TonalEnergy, Estill Voiceprint Plus). However, as discussed in Section 2.4.1, their application in music pedagogy is limited due to feedback interpretation difficulties and because the audio recording conditions of a mobile device microphone in a practice room is not guaranteed to provide sufficient levels of repeatability and accuracy. An attempt at interpretation is provided by SkyNote, but only still in an exploratory research setting.

To provide feedback based on timbral characteristics, we might suggest to use a dedicated external microphone that has a configuration to be installed at the same distance and position from the sound source (e.g., clip-on microphones). In this way, the variable of dynamics, which is crucial for the identification of the timbral properties of a sound (Fabiani and Friberg, 2011), is normalized. With this solution, more robust software algorithms can be developed that rely on the recording characteristics of a single microphone sensor, suitable for recording musical instruments, instead of relying on recordings taken from several microphones, usually optimized for voice calls, embedded in different devices. In this case, the additional cost of having to use an external hardware component is justified by the improved reliability of the overall system. The adoption of such a microphone would open the opportunity to provide feedback and visualizations on the sound dynamics produced and on timbral aspects for which a higher quality recording is required.

Attack clarity

Attack clarity refers to the purity or accuracy of the onset of a sound, especially with respect to achieving the desired fundamental frequency that is not contaminated by noise or undesired frequency components. Attack clarity may involve different characteristics depending on the musical instrument considered. Optimal articulation usually requires a very short time duration between the silence before the attack and the achievement of a fully developed sound, regardless of the particular type of articulation, dynamics, or accent required.

Among the software systems listed, some of them (i.e., TonalEnergy, Vocal Pitch Monitor, Estill Voiceprint Plus) include useful features to provide a visualization of this skill. They in fact provide a display of the evolution of the fundamental frequency or spectrum over time. By looking at the graphs, musicians can partially verify the accuracy of their articulation. However, this functionality could be greatly improved by providing detailed visualization of the attack of the notes produced, using short time windows to analyze the audio signal. Also, it could be very useful to provide feedback (e.g., in milliseconds) on the time duration used to achieve a relatively stationary sound from a timbral point of view.

Advanced sound processing algorithms implemented in SkyNote, within the TELMI project[2], have been developed to identify different types of violin pronunciation (e.g., staccato, martelé, détaché) (Ramirez et al., 2018; Giraldo et al., 2019). However, the project has not yet found use in music pedagogy, as it has remained in the research phase.

[2] http://telmi.upf.edu/

Vibrato

Some systems provide a visualization of the evolution of the fundamental frequency or sound spectrum over time (e.g., Vocal Pitch Monitor, Estill Voiceprint Plus). This provides visual indications of the amplitude, frequency and extent of the vibrato.

Possible applicable extensions to these features could include interactive exercises that assess control of these parameters. For example, a system could specify a sequence of long notes, embedded in a rhythmic context, that the performer has to play at predetermined vibrato patterns (e.g., duines, triplets, quatrains at each beat) within specific frequency and amplitude threshold values. Training on these exercises would allow the musician to learn vibrato control under different conditions and master this skill from a technical standpoint. In this way, when performers later want to expressively interpret a piece of music (e.g., aria, sonata, cantata), they will have the flexibility to produce the type of vibrato they feel is most appropriate for that performance, without being constrained by technical limitations.

Relative tuning

Standard tuners represent useful tools to develop a consistent intonation through intervals, scales, dynamics and articulation. However, players of variable-intonation pitched instruments (e.g., violin, trombone) must adjust their pitch relative to that of others when performing in ensemble music contexts. It is therefore important that students of these musical instruments develop the ability to listen to the sound of others as they play, understand how much it differs from their own pitch, and correct any discrepancies.

We believe that modern technology has the potential to help musicians of these instruments develop this skill, using graphic displays, microphones and headphones. It would help future students better integrate into ensemble music groups, more easily find a common pitch, and generally better control the dynamic balance of their sound.

In summary, this study intends to highlight the scarcity of low-cost technologies that provide visualization and feedback on the technical concepts necessary for a complete learning of a musical instrument, as the metronome and the tuner do. Their development, coupled with data collection and statistical analysis capabilities to monitor the level of the musician in their respective technical skills, would provide significant support to visualize progress over time, identify effective practice routines and method of study, as well as represent important tools for stress management and improving performance wellbeing.

2.5 Conclusions

This study presents a review of the main features provided by the technological tools that have been developed to support music instrument learning and investigates their potential benefits and utility. The widespread success of the metronome and tuner have prompted the subsequent development of numerous software and mobile applications that attempt to go beyond basic rhythm and pitch accuracy. However, their use in applied performance repertoire contexts, where the system makes an evaluation that discourages artistic expression, can present important drawbacks in pedagogical experience especially for beginners, who generally have less technical control and sense of self-evaluation.

There are numerous other facets of learning to master an instrument that are still poorly addressed by current music technologies, such as control of dynamics, attack and release precision and refinement, flexibility with timbre, vibrato, embouchure configuration and variation, finger position and movement, posture and breathing, to name a few. We believe that the development of new technologies that provide visualization or perception of technical concepts related to the learning of a specific musical instrument may find broad use in music practice rooms, if they are relatively cheap and have user-friendly interfaces. Clearly understanding a musical concept to be researched and pursued in individual study sessions through audiovisual systems can consistently help instrumentalists in becoming more efficient with their practice. In addition, such systems would represent objective yardsticks for teachers to verify proposed recommendations and improve lesson effectiveness.

By suggesting these new directions for future assistive technology supporting music pedagogy, we hope to better connect the field of technology development with the music school community so that students can enjoy a more fulfilling artistic experience.

Author Contributions

Alberto Acquilino: Conceptualization, survey, discussion, visualization, writing—original draft, and review and editing. **Gary Scavone**: Supervision, discussion, visualization, and review and editing.

Chapter 3

Open-source mobile apps for music education: A case study on trumpet fingering

The previous chapter provided a literature review of existing educational technologies, highlighting their limitations and the need for simple, targeted educational interfaces that focus on developing specific technical skills through structured exercises, rather than in a performance context (Section 2.4.2).

To address these gaps, this chapter presents a modular open-source framework designed to facilitate the learning of these skills. Its potential is demonstrated through a case study on trumpet fingering, showcasing how thoughtful app design can address technical and auditory challenges in music education.

A modular open-source framework is designed to support the development of musical competencies. The framework is built around modular components that can be assembled in different ways to create educational exercises tailored to various technical skills. Its potential is demonstrated through a case study on trumpet fingering, showcasing how a well-designed app can help learners overcome technical and auditory challenges. While fingering has been targeted by existing educational technologies – primarily for plucked strings and keyboards, and often within a performance context (see Table 2.1) – its application to brass instruments in a structured, exercise-based format remains largely unexplored. The flexibility of the framework allows individual components to be modified and adapted for different instruments and technical skills, extending its applicability beyond the trumpet.

The content of this chapter is based on the following research article:

A. Acquilino, M. d'Andrea, K. K. Reddy, J. J. Park, and G. Scavone. Open-source mobile apps for music education: A case study on trumpet fingering. *Submitted to International Journal of Music Education*, 2025.

3.1 Introduction

As educational technology continues to evolve, its potential to improve learning outcomes in various fields, including music education, has become increasingly apparent. Several studies have explored the role of technology in facilitating both vocal and instrumental training (Webster, 2012; Bauer, 2020; Lã and Fiuza, 2022). These advancements highlight the transformative power of digital tools in enriching traditional music pedagogy.

One of the key benefits of integrating technology into music education is its capacity to support the development of sensorimotor schemes (Leman and Nijs, 2017). These schemes enable musicians to anticipate the sensory outcomes of their actions, ultimately refining fine motor control. Educational technology, when thoughtfully designed, can play a critical role in reinforcing these schemes by providing real-time feedback and structured practice environments, thus facilitating more efficient skill acquisition. By automating technical aspects of music performance, technology allows students to shift their focus towards the creative and expressive dimensions of music-making, enhancing both the learning process and artistic outcome.

Educational technologies have often been designed with a teacher-centered approach, where students are relatively passive recipients of instruction. Studies, such as those by Welch et al. (2005), highlight the use of visualization and real-time feedback systems to enrich vocal pedagogy, providing teachers with tools to deliver more immediate and frequent feedback during lessons. While these systems improve the learning experience by increasing the frequency and quality of feedback, they can limit students' autonomy and opportunities for independent learning.

Recently, the pedagogical discourse in music education has shifted significantly (Lennon and Reed, 2012). Traditional instrumental and vocal instruction approaches focused on demonstrating technical skills and providing extensive verbal instruction, modeling, and feedback, while dedicating relatively little time to active student engagement in playing or singing during lessons (Crocco and Meyer, 2021), thereby limiting opportunities to foster student autonomy and lifelong learning skills (Gaunt, 2008). In contrast, contemporary pedagogical trends increasingly prioritize learner-driven approaches that emphasize the learning processes over musical products. This shift encourages strategies such as self-assessment and autonomous error correction (Harrison and O'Bryan, 2014). New trends promote student independence and incorporate peer learning methods, thus encouraging students' active participation in the educational process.

Despite the willingness of students and professionals to embrace technology in music education (Waddell and Williamon, 2019), there remain gaps in addressing the full spectrum of music technical skills. Widely adopted music educational technologies are based on the metronome and tuner, due to their ability to make abstract musical concepts (i.e., rhythm and pitch) immediately accessible in real time (Acquilino and Scavone, 2022). Without these tools, teaching such concepts would be significantly more challenging. These technologies facilitate a clearer understanding of technical concepts and musical goals through an audiovisual approach, enhancing the quality of students' practice sessions and promoting self-regulated learning.

However, beyond rhythm and intonation, there are other technical musical skills that musicians must develop to effectively express their musical ideas, which have received comparatively less attention in terms of dedicated technological solutions. These skills may be broadly applicable across various categories of musical instruments (e.g., dynamics) or may be more instrument-specific (e.g., articulation). This gap between musical skills documented in the pedagogical literature and available educational technology underscores the opportunity for technology to play a more significant role in music education.

A few studies have developed interactive, multimodal systems that serve as real-time feedback mechanisms, monitoring performance parameters such as pitch, sound quality, and movement across various musical contexts (Romani et al., 2015; Bandiera et al., 2016; Brennan, 2016; Blanco et al., 2021; Acquilino et al., 2023b). However, the widespread adoption of these technologies is often hindered by their reliance on complex PC software, overcomplicated interfaces, and costly hardware setups (Acquilino and Scavone, 2022). These limitations highlight the need for more accessible and user-friendly technological solutions.

In recent years, mobile devices such as smartphones and tablets offer new possibilities for educational approaches and applications in music learning (Nijs and Leman, 2014). Mobile apps provide flexible, portable, and ubiquitous solutions, enabling learners to practice anywhere. These platforms enable learners to receive real-time feedback and tailor content to their skill level through user-friendly interfaces, making them valuable tools for enhancing learning experiences (Paule-Ruiz et al., 2017; Shi, 2023).

Despite these advantages, developing mobile applications requires advanced programming skills that might not be part of the expertise of those working in music education technology research. This presents a significant barrier for academic researchers and educators, who are then compelled to seek external support – which is often difficult to obtain – or to settle for simpler, potentially less effective alternatives. These obstacles hinder the development of mobile solutions that could contribute to bridging the gap in empirical research needed to fully understand the impact of technology on music education, as highlighted by Leman and Nijs (2017).

To address this challenge, the present study describes a hybrid open-source application developed for iOS and Android devices. A hybrid application leverages web technologies for its content, while being distributed through app stores like any other native app. The programming environment used to build the application features modular components that can be easily combined, enabling the creation of customized educational tools without requiring advanced programming expertise. By supporting flexible, adaptable tools, the environment fosters innovation and collaboration in music education.

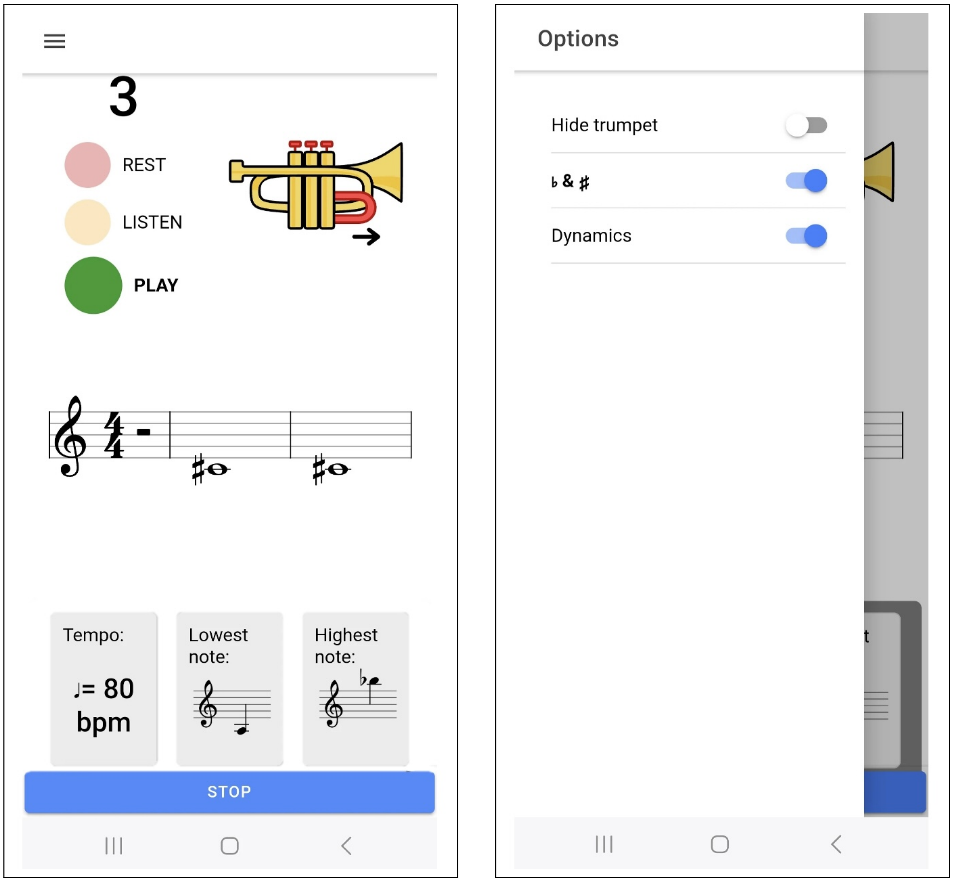
This chapter presents the pedagogical approach that guided the creation of an open-source app, illustrated through a case study on trumpet fingering (Section 3.2). The pedagogical principles underpinning the case study technology and its implications for music education are discussed in Section 3.3. Opportunities for further development and research are highlighted in Section 3.4. The technical details of the application's structure and functionality are provided in Appendix B for those interested in replicating or extending the app[1].

3.2 Technology Case Study Description

A straightforward educational interface has been developed as a case study to aid trumpet players in learning correct fingering, specifically in associating which valves to press and what sound to internalize to play a note displayed on the musical staff. A screenshot of the main interface is shown in Figure 3.1a. It requires the user to input exercise parameters such as a metronome value ranging from 40 to 180 beats per minute (bpm), the lowest note, and the highest note within the typical trumpet range [E3, B♭5]. Once these input values are entered, pressing a *Start* button generates a three-measure score in 4/4 time: the first measure is a rest, and the second and third measures display the same whole note randomly selected within the specified input range. In the top right corner, an image of a trumpet indicates which valves to press and whether to extend the third valve slide to produce the indicated note.

The exercise consists of **Measure 1**: resting during the first measure to allow the musician time to observe the selected note and prepare the corresponding fingering; **Measure 2**:

[1]Released under the Affero GPL license, the app repository is accessible on GitHub: https://github.com/albertoacquilino/music-education-interface-ionic



(a) (b)

Figure 3.1: (a) Main interface of the educational software. (b) Options menu visualization.

listening to a pre-recorded sample of that note played by a professional musician; **Measure 3**: playing the tone.

The software then randomly selects another note within the specified note range and presents the user with another cycle of three measures until the user taps a *Stop* button.

In the top left corner, indications of the corresponding beat (i.e., 1, 2, 3, 4) are provided along with a metronome sound and the required action for each measure: REST, LISTEN, PLAY. These actions are visually reinforced by a traffic light system displayed prominently on the interface: REST is associated with a red light, LISTEN with a yellow light, and PLAY with a green light. The corresponding light is highlighted in real-time during each action to provide the user with an intuitive visual cue.

An additional toggle menu of options allows the user to remove the fingering image, including sharps and flats in the note selection (otherwise only notes from the C major natural scale are selected), and add dynamic indications to the exercise. In the latter case, a random dynamic marking (i.e., *piano*, *mezzo-forte*, *forte*) is displayed below the score, requiring the player to vary their playing dynamics; the reference sound provided during the second measure is also played softer or louder according to the corresponding dynamic indication. A visualization of the options screen is shown in Figure 3.1b.

This software has been developed as a mobile application for Android and iOS devices and is also accessible through a web browser. It has been made available for free download on Google Play[2] , the iOS App Store[3] , and via a web link[4] .

The exercise is grounded in a fundamental learning model framework, which provides the flexibility and motor adjustment necessary for facilitating the prosthesis-dialogue modes of instrumental interaction (Leman and Nijs, 2017). The execution of this exercise aims to adhere to the elements of a cognitive architecture for sensorimotor processing:

- Measure 1 - *The motor command*: During the first measure, which is a rest, the learner can observe the note selected by the software and prepare the correct corresponding fingering on the instrument as indicated in the image on the interface. When done repeatedly, this action trains the neural patterns that control finger movements performing a musical action (Engel et al., 2012).

- Measure 2 - *The expected sensory outcome of the motor command*: In the second measure, the student listens to the sound of the proposed note recorded by a professional musician. This immediate listening before execution helps develop an auditory-motor loop. An fMRI study by Gebel et al. (2013) involved trained pianists and trumpet players using an MRI-compatible trumpet model without auditory feedback. Unlike pianists, trumpet players showed a specific co-activation increase in the left primary sensorimotor cortex, particularly in the regions associated with lip movements, the trunk, the right cerebellar hemisphere, and the left primary auditory cortex. As noted by Steenstrup et al. (2021), this indicates that auditory activity influences motor activity in musicians.

- Measure 3 - *The real sensory outcome*: In this measure, the learner plays the note and perceives the actual outcome of the musical action. The learner experiences the intended sensory feedback, facilitated by the technology in a manner that is both transparent and mentally unobtrusive.

This iterative process occurs flexibly via a smartphone or tablet interface, allowing the learner to practice in their preferred environment.

The following section discusses the pedagogical approaches and principles underlying this case study technology in more detail.

[2]https://play.google.com/store/apps/details?id=com.meiteam.trumpet
[3]https://apps.apple.com/us/app/mei-trumpet/id6452315691
[4]https://mei-trumpet.web.app/home/exercise

3.3 Discussion

Tailoring the learning experience to individual needs is essential for effective skill acquisition and progression. A common strategy for learning a musical instrument is to follow practice methods that present exercises with increasing levels of difficulty. However, this standardized, one-size-fits-all approach may not suit all learners (Steenstrup, 2023).

Digital technologies offer efficient customization in learning process allowing to gradually build their skills and start from their current abilities and incrementally increasing the complexity of tasks. This approach is grounded in several key pedagogical principles, which are outlined below.

3.3.1 Foundational Pedagogical Principles

Starting from what is known

Current pedagogical insights emphasize the importance of constructing new knowledge based on what learners already know (Maton, 2009). The ease of playing specific note ranges can vary depending on individual factors such as physicality or equipment. Generally, beginners are able to perform within a certain range of pitches. Customizing exercises to the learner's technical level, such as allowing them to set their own range of notes to practice, ensures a comfortable and effective starting point. This flexibility supports a more tailored and accessible learning experience, allowing both learners and instructors to design exercises that match the learner's current abilities.

Scaffolding from known to unknown

An effective educational approach involves gradually increasing complexity by setting progressively challenging goals that build on previously acquired knowledge. This ensures that each new step is both intriguing and achievable (Booth, 2009). According to Steenstrup (2017), breaking down large tasks into smaller, manageable parts can help guide learners toward mastering complex musical skills. For example, learners might begin by focusing on a limited range of notes they are comfortable playing, then progressively expand that range to include additional notes.

As in motor learning, reducing feedback over time encourages self-assessment and independent error estimation, making learners more self-reliant and fostering deeper engagement in the learning process (Guadagnoli and Kohl, 2001). This gradual scaffolding approach ensures that learners are consistently progressing at a pace suited to their abilities, making the learning process manageable and motivating.

Errorless practice

The ability to gradually extend the pitch range, with or without alterations, and to increase the tempo (bpm) according to the player's skill level aligns with the principles of errorless practice. This technique is widely recognized for its effectiveness in building confidence while minimizing the frustration associated with frequent mistakes. Errorless practice is particularly beneficial for beginners, facilitating smoother and faster initial learning by creating a supportive environment with fewer opportunities for error (Allingham and Wöllner, 2023). This method helps students maintain motivation and fosters a positive learning experience from the start.

Introduction to combined practice

Combined practice in music involves utilizing various learning methods simultaneously to enhance the overall practice experience. This approach, which includes, for example, physical practice, mental imagery, and singing, has proven to be as effective as extensive physical practice in improving performance and pitch accuracy while also significantly enhancing musical expression (Steenstrup et al., 2021). Alternating different practice methods reduces the total time spent on physical practice, thereby lowering the risk of injuries associated with overuse (Hagglund and Jacobs, 1996). Furthermore, combined practice improves retention more effectively than traditional blocked practice (Granda Vera et al., 2008).

The proposed educational software can introduce learners to combined practice by requiring engagement in both physical practice and auditory imagery. While the current interface does not explicitly involve singing, creative learners and instructors can incorporate it into their routines. For example, students might sing the note in the third measure instead of playing it, or they might sing the note along with the professional recording in the listening measure, and then play the note in the third measure. This sequence would strengthen the connection between mental and physical practice.

A different structure of the software could easily integrate singing by providing specific indications and feedback, further adhering to combined practice principles. By embedding these strategies into the technology, learners are trained not only in accurate fingering but also in developing a healthy and efficient practice strategy for expressive musical performance and retention.

3.3.2 Implementation in Practice

Support correct fingering technique

The exercise presented in this case study can help students develop proper fingering. For instance, the top-right image in Figure 3.1a illustrates which valves to press for each set of three measures to produce the corresponding note and whether to extend the third valve slide (e.g., for C♯4 or D4). When a note can be played with multiple valve combinations (e.g., pressing the first and second valves or only the third valve), one fingering among the most common is chosen.

Additionally, in the specific case of a brass instrument, it is important to note that with the same valve position (which determines a specific geometry of the air column), different notes in the corresponding harmonic series can be played. This can potentially cause confusion among beginners, as they might not know the exact pitch of the note they need to perform among the notes playable with the given valve position. The reference sound played by the software during the second measure of the exercise thus is crucial for clearly indicating the correct pitch to the learner.

Various software solutions in academic literature and commercial products have been proposed to indicate fingering on a screen or projector (Yin et al., 2005), robot tutors (Bagga et al., 2019), and augmented instruments (Li et al., 2022). Most of these educational technologies are primarily aimed at musical performance, providing an overall assessment – often with real-time feedback – of whether the performer has played the correct notes in time. However, this performance-oriented approach, rather than focusing on the acquisition of a specific technical skill, may not be pedagogically effective, as discussed by (Acquilino and Scavone, 2022).

The proposed software design facilitates the acquisition of the fingering technique by associating the visual representation of notes on the staff with their corresponding pitch and fingering position. This association is made in small, manageable chunks, allowing learners to focus on one note at a time, outside the context of performance. The exercise's loop structure ensures sufficient time for internalizing the motor and auditory imagery information involved in the learning process. As with athletes, this guided repetition of single tones facilitates the kinesthetic learning of fingering and sound production in general. While this exercise involves repetition, it focuses on very small units (a single note). These units are not limited to *block practice* but can be flexibly adapted to *interleaved* or *random practice* methods (Stambaugh, 2011; Carter and Grahn, 2016).

Support auditory imagery development

In addition to providing a pitch guide, the reference sound in the second measure plays a key role in facilitating the development of auditory imagery. Well-recognized in educational literature and a current subject of neuroscientific study (Zatorre and Halpern, 2005; Gates, 2021), mental imagery requires students to form a clear mental representation of a note, or musical phrase in general, before actually playing it.

The exercise structure, where the learner prepares the correct fingering in the first measure and listens to the sound before playing it in the second measure, is designed to incorporate auditory imagery in conjunction with motor imagery. This creates anticipatory imagery for the execution of musical actions. As noted by Steenstrup (2023), the presence of the audio-motor loop, which facilitates the initiation of motor responses through increased activation of the auditory cortex, underscores the critical connection between auditory and motor imagery. This association is essential for precisely refining the muscle activity necessary for accurate musical performance.

Since auditory imagery also involves timbre (Halpern et al., 2004), the decision was made to use a sound recorded by a professional musician and not a synthesized sound. This ensures that the learners are not only aware of the pitch but also the quality of the sound they are aiming to produce.

Multi-sensory approach

Playing a musical instrument inherently engages multiple senses. Generally, musicians interact with the instrument through touch, feel vibrations on the instrument and their body, listen to the sound produced, and may move in space while playing. This multi-sensory engagement is crucial for developing a comprehensive understanding and mastery of the instrument (Karakaş and Dündar, 2024).

The devised software capitalizes on this approach to enhance the learning experience. As detailed in Section 3.3.2, the software uses visual aids with fingering indications, guiding students to associate notes with finger positions, which helps reinforce tactile memory crucial for muscle memory development (Schmidt and Lee, 2020). Additionally, as described in Section 3.3.2, the software provides auditory feedback through a reference sound played by a professional musician, helping students form a mental representation of the desired tone. When playing, students feel the vibrations of the instrument and their body, contributing to a more holistic sensory experience. Such kinesthetic feedback is vital for understanding the physical sensations associated with sound production and can aid in refining their technique and embouchure (Kohut, 1985).

The exercise structure – rest, listen, play – facilitates the integration of multiple senses in the learning process. During rest, students visually and tactilely prepare fingerings, while the interface provides the information needed to understand the actions required to produce the selected note. In the listening measure, they engage their auditory senses by listening to the reference note. Finally, in the playing measure, they combine tactile, auditory, and kinesthetic feedback to produce the note. This multi-sensory engagement aligns with McCoy's model of learning in singing, which is based on the VARK framework and identifies auditory, kinesthetic, intellectual, and visual modalities (McCoy, 2004; Fleming and Mills, 1992). Given that most learners are multimodal (Fleming and Baume, 2006), this approach is likely to support diverse learning preferences effectively, enhancing memory retention and facilitating the learning process.

Imitation

The principles discussed in the previous subsections are grounded in imitation, a fundamental mechanisms of human learning (Rizzolatti and Craighero, 2004). The trumpet image with fingering indications and the reference sound act as imitation stimuli to train the learner's motor and auditory skills, progressively making the achievement of the musical goal more immediate and transparent, as discussed by Nijs et al. (2013).

Furthermore, the decision to provide recordings of a professional musician rather than using synthesized sounds is based on the found correlation between the trumpet timbre quality and sound production efficiency (Acquilino et al., 2023a). This insight is supported by numerous pedagogical sources, which indicate that efficiently produced sounds are perceived as rich and round, while less efficiently produced sounds are perceived as shrill and strained (Jacobs and Nelson, 2006; Steenstrup, 2007). Thus, the aim of the developed technology is to familiarize students from the beginning of their studies with listening to high-quality sounds, potentially stimulating them through imitation to seek better sound production efficiency.

3.3.3 Outcomes and Reflections

Maintaining the flow zone

Introducing progressively more challenging and intriguing tasks helps improve the learner's technical skills without losing interest. Csikszentmihalyi (2008) defines the flow zone as a state of mind where a person is fully concentrated and completely immersed in their current activity. This concept emphasizes the importance of assigning exercises of appropriate difficulty to promote optimal learning. Exercises that are too easy can lead to boredom and loss of motivation, while overly difficult tasks may cause tension, bad habits, and a sense of panic. Maintaining this state of attention is fundamental in musical practice, as it can be involuntarily captured or deliberately focused, influencing how learners perceive and interact with their instruments, and shaping their skills and understanding through active exploration and practice (Tullberg, 2022).

Abstraction to Practice

Music, due to its abstract nature, can be difficult for learners to grasp and translate into appropriate actions. As Nijs (2018) discusses, music is often perceived as abstract and disconnected from concrete events, challenging learners to understand and translate it into practical performance actions. Bridging the gap between abstract musical notation and tangible musical execution can help learners internalize these concepts more effectively.

For example, providing visual and auditory cues, such as displaying correct fingerings alongside a reference pitch, can strengthen the association between abstract musical symbols and the corresponding physical actions. This process helps demystify abstract concepts, enabling students to better understand and perform musical tasks. Similarly, guiding learners in the acquisition of rhythm through metronome beats and reference examples can improve timing and rhythmic accuracy by making these abstract patterns more concrete and easier to follow.

Alternated practice and reflection

The proposed exercise encourages alternated practice and learner reflection. This approach not only helps beginners in playing long tones, but also combines physical practice with observation, an efficient learning strategy. A study by Larssen et al. (2021) highlights that interleaving observation and physical practice benefits visuomotor adaptation and motor memory consolidation. They found that observation can effectively replace physical practice if supplied intermittently, providing benefits beyond mere rest periods.

The inclusion of rest and listening measures in the exercise is motivated by these findings. The rest measure, in addition to allowing learners to prepare the appropriate fingering, also serves as a reflection time on their previous performance. One of the reasons for not including immediate feedback on note performance is to encourage learners to reflect on their actions. Assuming the goal of the exercise is to learn correct fingering and produce the indicated note (without focusing on intonation at this stage), the interface provides all necessary information, including the corresponding fingering and a realistic reference pitch. This should enable students to discern whether they played the correct note, fostering what is known in motor learning literature as internal reference-of-correctness (Maas et al., 2008).

If the desired sound is not produced, corrective feedback highlighting the error could distract the learner with negative reinforcement. Instead, the current design offers an opportunity for self-reflection and self-correction. As philosopher John Dewey stated, "*we do not learn from experience... we learn from reflecting on experience*" (Dewey, 1938). This emphasis on reflection aligns with studies indicating that alternating physical practice of a new skill with short rest intervals (even for just a few seconds) leads to performance improvements during these brief rest periods, known as *micro-offline gains* (Bönstrup et al., 2020). Personal reflection enables learners to produce the suggested note and internalize the motor and auditory process, bypassing the typical ambiguity of the interpersonal feedback process between teacher and learner (Welch et al., 2005).

This does not preclude the inclusion of possible upgrades with corrective feedback and general suggestions provided by the software or input by the instructor. Additionally, the exercise could be adapted to extend the measures of rest, listening, and playing based on the student's needs, providing more time for learners to experience the musical action and reflect on it.

By combining these pedagogical principles, the developed case study interface offers a robust framework for learning trumpet fingering technique. The current software's ability to flexibly adjust different difficulty levels should help maintain learners within the *flow zone*, fostering continuous improvement and engagement. Also, it provides both learners and educators with a personalized and engaging approach that supports the learner's progression from known to unknown, in small, achievable steps.

A study involving participants has been conducted to assess the usability and social validity of this case study educational technology. The findings from this study will be presented in Chapter 4.

3.4 Suggestions for future studies

The application presented in this study is a work in progress that offers significant potential for future developments. Leveraging open-source practices, we aim to build a collaborative community of researchers, educators, and developers to combine, adapt, and expand its modular components, fostering innovation in music educational technology. The source code of the app is available in a public GitHub repository. The design of the app is modular and built around the concept of reusable components (see Appendix B). This reduces the need for advanced programming skills, facilitating the creation of sophisticated tools and prototypes, and accelerating research in fields such as pedagogical design, interactive user interfaces for music education, and real-time audio analysis.

This approach functions both as a conceptual (pedagogical) framework and as a modular software (GUI) infrastructure, allowing educators and researchers to build new applications that rest on robust pedagogical foundations while taking advantage of flexible code components. Future developments could expand this synergy to address a broader range of musical skills that are crucial for expressive performance but often lack dedicated technological support. Music education methods for wind instruments, for example, highlight the importance of mastering fundamental aspects such as articulation, attack, breathing, fingering, intonation, rhythm, tone quality, vibrato (Arban, 1982; Flesch and Mutter, 2008; Westphal, 1990; Weisberg, 2007). Yet, many of these remain underexplored in available mobile edTech solutions. By leveraging the hardware capabilities of modern mobile devices (e.g., displays, loudspeakers, microphones), future studies could create new components offering real-time visualization and feedback on these critical skills. The proposed repository offers a flexible framework that can help bridge this gap, gradually expanding into a comprehensive platform for both technical and artistic development.

Although the presented case study focuses on the trumpet and Western music notation, the application can be easily extended to other musical instruments and traditions. Future studies could explore its adaptation to diverse pedagogical contexts, embedding alternative notations – such as tablatures, graphic scores, or culturally specific systems – and incorporating feedback mechanisms suited to different instrumental techniques. By supporting a wider variety of musical practices, the repository has the potential to become a truly inclusive and adaptable educational tool.

Another promising direction for future studies is the integration of gamification elements to increase learner engagement and motivation. Gamification, which involves applying game design principles such as levels, challenges, and rewards to educational contexts, has been shown to enhance learning outcomes and enjoyment (Deterding et al., 2011). By transforming repetitive exercises into interactive and rewarding experiences, gamified elements could encourage active participation, sustained motivation, and long-term retention of musical skills. Features like progress tracking, customizable difficulty levels, and creative challenges could make practice both effective and enjoyable.

Ultimately, the application aspires to become more than a tool for technical improvement. Future studies and community contributions can evolve it into a holistic educational platform, fostering intrinsic and extrinsic motivation, supporting individualized learning, and meeting the diverse needs of musicians. By bridging gaps in existing technologies and expanding its scope, this work has the potential to redefine music education as a more accessible, engaging, and meaningful experience for learners worldwide.

3.5 Conclusions

This chapter presented an open-source application that demonstrates how pedagogical best practices can inform the design of educational technologies. Grounded in a student-centered approach, the app facilitates learning through guided repetition, structured reflection, and multi-sensory engagement, addressing critical technical and auditory skills. These pedagogical principles were exemplified through a case study on trumpet fingering, showcasing the potential of thoughtful app design to enhance music education.

The app was created using a modular, open-source code framework, designed to reduce technical barriers for educators and researchers. This repository of modular programming components enables users to build custom tools tailored to diverse musical instruments, skills, and traditions, even without advanced programming expertise. By leveraging this framework, the app integrates features such as dynamic score display, note selection, metronome synchronization, and audiovisual feedback, demonstrating how modular components can be combined into robust educational tools.

Looking forward, this modular framework offers opportunities for significant expansion. Future developments could address additional musical skills, support a wider range of instruments and traditions, and incorporate gamification strategies for heightened learner engagement. These prospects closely align with the overarching goal of delivering tailored learning experiences while laying the foundation for further empirical research and pedagogical advancements in music education technology.

Author Contributions

This study was supported by the Google Summer of Code 2024 program as project of the International Neuroinformatics Coordinating Facility.

Alberto Acquilino: Conceptualization of the study and paper, development of the initial prototype of the presented technology in Python, lead mentor for the Google Summer of Code 2024, supervising the web development of the case study technology, and original draft writing. **Mirko D'Andrea**: Conceptualization of the study, development of the technology from the initial Python prototype to a web application, repository management, mentor for the Google Summer of Code 2024 project, review, and editing of the paper. **Kambham Keerthi Reddy**: Contributor to the Google Summer of Code 2024 project, repository development, and original draft writing of Appendix B. **Jenny Jieun Park**: Supervision of the educational aspects of the study, with a particular focus on Chapter 3.3, and review and editing of the paper. **Gary Scavone**: Supervision of the overall study, providing critical review and edits.

Chapter 4

An exploratory study on perceived usefulness in music education technologies

Building upon the technological framework presented in Chapter 3, this chapter explores the usability and potential of the proposed open-source mobile application for trumpet education. The study seeks to expand on the discussion in the previous chapter, which identified limitations in existing educational tools and emphasized the need for technology that supports the development of critical technical skills in music learners. By prioritizing modularity and accessibility, the introduced framework reveals its potential to address these gaps in a practical context.

This exploratory study, conducted by Alberto Acquilino, Jenny Jieun Park, and Gary Scavone, evaluates the practical application of this framework through the experiences of adult learners. It examines their interactions with the software, gathering insights into their perceptions of ease of use, perceived usefulness, and suggestions for improvement. By involving a diverse group of participants and employing both qualitative and quantitative methods, this chapter offers an initial exploration of the tool's impact on individual practice routines and its potential for integration into formal music education.

The findings highlight the strengths of the software design while also identifying areas for refinement, offering actionable insights for future development. By bridging theoretical concepts with empirical evidence, this chapter contributes to the growing field of educational technology in music and suggests pathways for creating more effective and inclusive digital tools.

4.1 Research questions

This exploratory study involves four main research questions:

1. **RQ1: Perceived Ease of Use**: How do participants perceive the ease of use of the software application presented in Chapter 3?
2. **RQ2: Perceived Usefulness**: Which aspects of the provided educational tool are perceived as useful by learners and which are not?
3. **RQ3: Missing Features**: What do learners believe is missing in the provided technology to make it useful to them?
4. **RQ4: Desired Features**: What are learners looking for in music educational technology?

4.2 Methodology: An explorative case study

4.2.1 Participants

An exploratory qualitative study was conducted to evaluate the perception of using the developed app among adult users. Twenty participants were recruited through four different sources: amateur musicians playing at the *McGill Symphonic Band Club* (non-music students), music students taking instrumental techniques courses at McGill University, members of the *Harmonie Nouveaux Horizons de Montréal*, and through a Facebook social media post. The distribution of participants across these different sources is presented in Table 4.1. The cohort includes participants from both Europe and North America. The demographic distribution in terms of age, reported gender, years of trumpet playing experience, and self-reported level of expertise is detailed in Table 4.2. The participants are adults with an average age of 42.3 years (SD = 12.2, range = 24 - 64) and have been playing the trumpet for an average of 9.1 years (SD = 10.7, range = 3 months - 45 years). Despite the average length of time playing the trumpet, most participants consider themselves to be at a beginner or intermediate level, with only one participant identifying as semi-professional (a former performance graduate now playing as a hobbyist) and one participant identifying as a professional (currently a teacher).

The decision to focus exclusively on adult participants was driven by practical considerations related to the study timeline. While minors could be ideal candidates for the software, as it is designed with beginners in mind, obtaining approval from the McGill Ethics Board to

Table 4.1: Breakdown of Participants by Recruitment Source.

Recruitment Channel	**Number of Participants**
McGill Symphonic Band Club	2
McGill Instrumental Techniques Class	1
Harmonie Nouveaux Horizons de Montréal	5
Facebook Post Respondents	12

include minors in the study would have likely extended beyond the timeline of this PhD research and the logistics of gaining access to students in primary or secondary schools would have been challenging. Nevertheless, a future study incorporating minors would undoubtedly provide valuable insights and further enhance the understanding of the software's potential impact on younger learners.

4.2.2 Data collection

After obtaining informed consent (the consent form is available in Appendix C), the candidate conducted individual meetings of approximately twenty minutes with each participant. During

Table 4.2: Summary of Demographic Data of Participants.

Question	**Category**	**Frequency**
Age	18-30	5
	31-50	10
	Over 50	5
Gender	Male	14
	Female	6
Years of playing trumpet	1 year or less	5
	1-5 years	4
	6-10 years	7
	More than 10 years	4
Level of expertise	Beginner	8
	Intermediate	10
	Semi-professional	1
	Professional	1

these introductory workshops, participants received the download link for the software on the Android or iOS app store, ensured it was working correctly, and received standardized instructions on using the interface. Each device was registered with a password to anonymously track information on recruitment group, app usage and features accessed.

Participants were then asked to use the app over the following days and decide, based on their experience, whether to include it in their practice sessions. After approximately three weeks, eight focus groups were organized to gather feedback on the app's usability, social validity, and potential improvements. Similar methodologies regarding intervention studies of educational technologies for musicians have been reported in the literature for several decades (Capurso, 1934; Eisele, 1985; Dalby, 1992; Swift, 2003; Aksoy, 2023). The duration of the intervention period in this study was chosen in line with previous literature based on the complexity level of the proposed technology.

At the beginning of each focus group, participants completed an online questionnaire focusing on usability indicators adapted to the specific context of this study. The questions of the questionnaire followed the criteria suggested by Bevan and Macleod (1994) and Hornbæk (2006), measuring ease of use, satisfaction, learnability, engagement, usefulness, effectiveness, and efficiency through nine questions about motivation, perceived utility, and structured practice. Responses were recorded on a five-point Likert scale, allowing participants to provide neutral responses, with balanced items representing positive and negative options equally (Williamon et al., 2021).

The questionnaire also included aspects of social validity based on previous studies (Jameson et al., 2012), asking participants whether the technology changed their practice habits and whether they believed the technology could play a role in a music performance class. The list of questions with the response distribution is provided in Table 4.3.

Table 4.3: Summary of Detailed Questionnaire Responses.

Category	**Question**	**SD**	**D**	**N**	**A**	**SA**
Easiness of use	Installing the app and learning its functionalities was easy	1	0	0	4	15
	I felt competent to use the app for my own level (adaptability)	1	0	2	5	12
Structured practice	The app changed my way of practice	0	5	3	9	3
	The app helped to plan and set goals for my practice	0	5	5	7	3
Motivation	Using the app felt like an effort to me	8	7	3	2	0
	The app motivated me to practice effectively	0	3	4	11	2
	The app motivated me to practice for longer periods of time	0	5	6	6	3
Perceived usefulness	I made more progress than I normally do in two weeks of practice using the app	0	4	7	8	1
	I feel the app could be useful to be integrated in a classroom	0	1	2	10	7

SD: Strongly Disagree, D: Disagree, N: Neutral, A: Agree, SA: Strongly Agree.

Upon completion of the questionnaire (average completion time of approximately 5 minutes), participants were shown a screen with aggregated results of their responses. They were then neutrally asked to comment on various points based on their experience, opening a collaborative discussion. Questions asked during the discussion in each session also included whether they interacted with the software using headphones or the device's speaker, at what point in their practice session they used the app (e.g., at the beginning as a warm-up, in the middle as an exercise, or as a final exercise), what were the most positive and negative aspects of using the app, what improvements they would suggest, and any additional feedback.

These discussions, recorded and transcribed, provided in-depth qualitative information, offering a more comprehensive view of the interaction between learners and technology. The focus groups with participants from Italy were conducted in Italian by the candidate, as it is his native language, and later translated into English by the candidate. These reflections expand towards an empirical comparison between objective and subjective measures of the technology's usability from a macro to a micro scale over a sufficiently long period to assess learning and retention (Hornbæk, 2006). They investigate qualitative aspects of the interaction, which might be prioritized over efficiency by users (Walker et al., 1998).

The focus group transcripts are available in Appendix D.

4.3 Results

Anonymous usage data of the distributed app were analyzed and are presented in Table 4.4. Participants used the educational software variably, with an average usage duration of approximately 50 minutes (SD = 46.26 minutes, range = 4 - 169 minutes) and an average of 7.1 distinct usage days (SD = 5.1, range = 1 - 20). This indicates a diverse reception of the technology among participants, with some using it sparingly and others incorporating it into nearly daily practice. The table also provides details on the pitch range selected by each participant and the percentages of utilization of the app's features, highlighting the varied ways the technology was employed.

These collected data help in understanding how and to what extent learners interacted with the software during the intervention period. However, the frequency and duration of use can depend on numerous factors that are difficult to account for in the collected data. With the exception of one professional, all participants are amateur trumpet players who play for leisure. Consequently, their consistency in practicing the trumpet can be significantly influenced by work, study, family, and many other personal factors. To gain a clearer understanding of the quality of the interaction experience with the technology, it is crucial to consider additional aspects.

Table 4.4: Summary of App Usage Data. The table summarizes the usage data of the educational software by participants grouped by their recruitment source (Group I: McGill Symphonic Band Club, Group II: McGill Instrumental Techniques Class, Group III: Harmonie Nouveaux Horizons de Montréal, Group IV: Facebook Post Respondents). For each participant, the table includes the total duration of using the app, the number of distinct days the app was used, the mean and standard deviation (in semitones) of the low and high notes selected, and the percentage of time each feature was activated relative to the total app usage duration: hiding the fingering indication, selecting flat and sharp notes, and including dynamics in the exercise.

Group	**User**	**Total**	**Unique**	**Low Note**		**High Note**		**Hidden**	**Flats/Sharps**	**Dynamics**
		Duration	**Days**	**Mean**	**Std**	**Mean**	**Std**	%	%	%
I	User 1	4m	3	F4	5.2	E5	1.0	0	28.6	28.6
	User 2	7m	4	F4	9.2	D5	8.2	0	0	5.9
II	User 3	20m	6	B3	6.5	C5	3.1	0.0	21.1	5.3
III	User 4	2h 49m	20	D4	4.4	D5	4.7	86.2	0	0
	User 5	20m	7	C4	4.9	B♭4	5.6	8.3	29.2	12.5
	User 6	2h 33m	14	C4	3.6	D5	5.0	17.0	41.5	1.9
	User 7	24m	3	D4	3.6	C♯5	4.1	0	0	13.3
	User 8	53m	13	E♭4	4.7	D5	3.7	15.6	56.2	3.1
IV	User 9	12m	2	C4	2.7	A♭4	2.5	33.3	11.1	33.3
	User 10	8m	1	F4	0.0	F5	0.0	0	0	0
	User 11	30m	4	A♭3	3.6	E4	3.7	88.2	64.7	29.4
	User 12	1h 12m	7	B♭3	3.9	D5	6.0	0	12.7	4.8
	User 13	1h 4m	8	F♯3	1.1	E5	1.0	9.3	88.4	93.0
	User 14	1h 45m	16	A3	5.9	E5	9.3	0	7.3	5.2
	User 15	39m	5	A3	4.9	C5	5.3	0	37.9	72.4
	User 16	37m	4	G3	2.3	A4	5.5	0	80.0	50.0
	User 17	1h 13m	9	A3	2.2	F5	4.0	0	50.0	15.4
	User 18	12m	3	B3	1.9	F♯5	4.2	0	14.3	0
	User 19	1h 25m	11	B♭3	4.5	E5	3.7	5.4	80.4	21.4
	User 20	22m	2	A♭3	3.2	E♭5	5.6	0	0	0

During the focus groups, additional information was gathered from the learners. Fifteen participants were divided into eight discussion groups, each consisting of two to four members. Two participants preferred to report their experiences individually through one-to-one interviews, while three others chose to provide their feedback as open comments in the questionnaire.

The learners tested the app individually, and two participants – who were trumpet teachers at music schools – also used the app with their students. No participant reported any malfunctions with the app. The focus groups began with the administration of the questionnaire presented in Table 4.3. Table 4.5 provides a descriptive statistical analysis of the participants' responses, including the mean score, standard deviation, and one-sample t-test results compared to the neutral value.

The results summarized in Table 4.5 demonstrate a high perceived ease of use for the developed software application. Participants rated "Installing the app and learning its functionalities was easy" with a mean score of 4.60 (SD = 0.94), "I felt competent to use the app for my own level (adaptability)" with a mean score of 4.35 (SD = 1.04), and "Using the app felt like an effort to me" with a low mean score of 1.95 (SD = 1.00), all showing highly significant p-values ($p < 0.0001$). These findings indicate that participants found the software user-friendly, well-suited to their skill levels, and generally not burdensome to use, effectively addressing Research Question 1 on perceived ease of use.

Further consensus among participants was found on the question "I feel the app could be useful to be integrated in a classroom" with a mean score of 4.15 (SD = 0.81, $p < 0.0001$). These results indicate that participants believed the app had strong potential for classroom integration, providing a solid foundation for its perceived usefulness in educational settings.

There was less agreement among participants on other aspects, indicating a varied impact of the app on individual practice habits. Although the limited sample size restricts the statistical power of the quantitative analysis, the collected data serves to identify trends and set the stage for subsequent focus group discussions. These discussions yielded detailed qualitative insights, which, when integrated with the quantitative findings, offer a more comprehensive understanding of users' interaction quality, engagement, and expectations from educational technology aimed at assisting musical instrument learning.

According to the Technology Acceptance Model (Davis, 1989), the results of questions 1, 2, and 5 in Table 4.5 suggest that the success of the educational technology, as indicated by the actual use of the system, primarily depends on its perceived usefulness. The hypothesis that developing the educational technology as a mobile app, available on both Android and iOS platforms, would make it perceived as easy to use seems validated. The widespread adoption and integration of mobile technology in society have established a standard that appears to be recognized and accepted in the user experience.

On the other hand, perceived usefulness is more challenging to target objectively or standardize, as it is deeply intertwined with personal experience, preferred learning modes,

Table 4.5: Summary of statistical analysis for the questionnaire. Scores were calculated by assigning a grade of 1 to strongly disagree answers, 2 to disagree, 3 to neutral, 4 to agree, and 5 to strongly agree. The mean score, standard deviation, and t-test p-value are reported for each question. The p-values are based on a one-sample t-test against the neutral midpoint of 3.

Category and Question	**Mean Score**	**Std Deviation**	**T-Test P-Value**
Easiness of use			
1. Installing the app and learning its functionalities was easy	4.60	0.94	$<0.0001^{**}$
2. I felt competent to use the app for my own level (adaptability)	4.35	1.04	$<0.0001^{**}$
Structured practice			
3. The app changed my way of practice	3.50	1.05	0.046^{*}
4. The app helped to plan and set goals for my practice	3.40	1.05	0.10
Motivation			
5. Using the app felt like an effort to me	1.95	1.0	$<0.0001^{**}$
6. The app motivated me to practice effectively	3.60	0.88	0.007^{*}
7. The app motivated me to practice for longer periods of time	3.35	1.04	0.15
Perceived usefulness			
8. I made more progress than I normally do in two weeks of practice using the app	3.30	0.86	0.14
9. I feel the app could be useful to be integrated in a classroom	4.15	0.81	$<0.0001^{**}$

** p-value < 0.001, * p-value < 0.05 indicate strong and moderate evidence, respectively, against the null hypothesis under conventional thresholds. These p-values suggest that the results may be statistically significant but should be interpreted with caution in the context of the study.

educational strategies, and various other factors. These variables now become the focus of the following discussions: which aspects of the provided educational tool are perceived as useful by learners and which are not (RQ2)? What do learners believe is missing in the provided technology to make it useful to them (RQ3)? What are learners looking for in music educational technology (RQ4)?

The anonymized discussions from the focus groups have been transcribed and are available in Appendix D. These transcripts report the conversations on the questionnaire topics, often leading to debates on positive and negative aspects, anecdotes, and reflections. The discussions also addressed whether participants used headphones or the device's speaker while interacting with the software, at which points in their practice sessions the app was utilized, their perceptions of the most positive and negative features of the app, their suggestions for improvements, and any additional feedback.

The discussions are organized into common thematic areas identified across the different focus groups. The highlighted themes are discussed below, with quotes from participants to provide a comprehensive understanding of their perspectives and experiences.

4.3.1 Focus group findings on Research Question 1

RQ1: How do participants perceive the ease of use of the software application presented in Chapter 3?

In the previous section, the quantitative results from the administered questionnaire are presented, indicating a positive perception of the software's easiness of use. Such a perception is further confirmed by the focus group participants. To illustrate their statements, key excerpts from these discussions are highlighted, with the selected sub-themes being reported by at least two different participants. While some quotes have been slightly edited for clarity, such as removing repetitions and stuttering, the original transcripts are available in Appendix D at the bottom of this manuscript.

> *For me, the app is well made, it's really immediate. Whatever you choose to use, you use it, and there are no problems, meaning it's effective.* [FG7, P2][1]

> *I got on well with the app. It's very easy to use, truly straightforward. I'm satisfied with this aspect because, frankly, with mobile applications, with technology, it isn't always so immediate.* [FG8, P1]. *From the point of view of ease of use, it's really*

[1]Focus Group 7, Participant 2. The same abbreviations are also retained in the following quotations. In the case of one-on-one interviews, only the participant identifier (P) is provided without a number (e.g., [FG6, P]).

> *easy.* [FG8, P2]. *The interface is excellent in my opinion, in the sense that it's very clear.* [FG8, P3]

> *...the ease of use is really high, in the sense that it obviously doesn't really have any flaws, it's very easy to install and use. You open it and it works. So, let's say it doesn't have flaws. And clearly, maybe this is the strongest point because I would want even a four-year-old child to use it.* [FG6, P]

However, this does not mean that learners are completely satisfied with the overall usability of the interface. Indeed, during the discussions, several inefficiencies of the technology are reported regarding:

- The slider functionality for note range selection (5 mentions in 3 focus groups):

  > *I think the other reason why I put like "The app felt like an effort to me", I think* [it was] *because* [of] *the sliding: when you have to pick your range, sometimes it was slightly more difficult to pick the exact note that I wanted because sometimes it goes too fast or like kind of a back and forth type. But otherwise it was okay.* [FG2, P1]

  > *At first, I couldn't figure out how to change the note by dragging. I didn't find that intuitive. A part of me wanted to just type in like tap it to have the note move to that line or that space* [FG3, P1]. *... I agree with Participant 1 like, it wasn't too intuitive when selecting note. It could be made at just a little easier.* [FG3, P2]

  > *The only thing I would nitpick – just to be meticulous -– is the selection system for the range between the lowest and the highest note, because that gesture is a bit cumbersome. That's the only thing.* [FG4, P]

- The exercise finishing after every 10 repetitions of 3 measures (3 mentions in 2 focus groups). In the version of the application provided to participants for this study, each exercise consisted of 10 repetitions of 3 measures, requiring users to play 10 notes before the exercise automatically stopped. To continue playing additional notes, users needed to press the *Start* button again.

  > *Each time we start, we push on "Start". I guess we have a total of around 10 or 11 notes? And then it stops. And then we restart. Is it possible to have more? Because we have always have to click on "Start" again and maybe we*

> *have to stop when we want to stop, but it would go over. You know?* [FG2, P2]
>
> *Yeah, that was gonna be my comment. I was like, maybe you could make it like an infinite amount and then we can just stop it whenever we need to. Because otherwise we just have to like redo it again and then it's like an extra step.* [FG2, P1]

- The loss of input settings between sessions (2 mentions in 2 focus groups):

> *I rated disagree* [discussing about the question: "Using the app felt like an error to me"] *only because of the settings: I find that every time I close the app, whatever settings I select has to be done again... I believe we are being very picky on this, but that's how I felt in terms of the app feeling like an effort to me.* [FG3, P2].
>
> *Often in the note selection, the app would forget the notes that had been selected in the previous study session, so somehow this was a bit annoying indeed every time I reopened it.* [FG6, P]

- The screen turning off during app usage (2 mentions in 2 focus groups):

> *There's a technical thing I found a bit difficult to manage. I mean, well, it's not actually difficult to manage. So, when using the app, the screen turns off. Basically, it's okay, you just need to change the setting on the phone, right? But actually, it's a bit inconvenient for the user; I mean, maybe if the app could keep the screen active, it would be better.* [FG8, P3]

- The lack of landscape mode (2 mentions in 2 focus groups):

> *Is it possible to have maybe the position in horizontal? And once maybe the highest note and the lowest note are selected, the time and the input data fields can be hidden when I rotate the phone and I only have the trumpet positions and the note I'm playing... Having the possibility to rotate the screen horizontally to maybe get a bigger picture would be convenient for me.* [FG8, P4]

- The size of the trumpet fingering image (4 mentions in 4 focus groups):

> *When we have to slide the third slide when we do a D, the indication is quite small and I didn't see it the first time. Maybe it could be bigger because I missed it.* [FG2, P2]

> *I found the image of the trumpet like a little bit small, in terms of seeing which valves were down.* [FG3, P1]

We acknowledge that some feedback categories are interrelated, such as the request for larger fingering images and the activation of landscape mode. Here and in subsequent discussions, it was chosen to report the main suggestions separately to preserve the variety of detail nuances observed by the participants, adding value through diverse insights.

4.3.2 Focus group findings on Research Question 2

RQ2: Which aspects of the provided educational tool are perceived as useful by learners and which are not?

During the discussions, several aspects of the technology were appreciated by participants, mentioned across different discussion groups.

Real sound

One of the most appreciated aspects of the software, mentioned in almost all discussion groups, was the reference sound recorded by a professional musician. Participants more or less explicitly linked the presence of high-quality reference sound to many of the pedagogical principles mentioned in Section 3.3:

- Auditory imagery:

  > *The app sparked my interest for several reasons. One: some of these reasons include the fact of hearing good reference sounds. That alone does a lot: single sounds with a nice color, a nice personality...* [FG5, P2]

  > *Doing the exercise as a warm up,* [it is] *a way to train my ear with the different tone of the trumpet. So, I think it helped a lot for that...* [FG1, P1] *You listen to it at first in your head and then* [you play]. [FG1, P3]

  > *Having a good sound, of a note played by someone who is a good trumpet player, you know? I don't hear that much! So, I have to have a guide to know what the notes are supposed to sound exactly.* [FG2, P2]

> *Sometimes, I've also used it just to hear the notes alone, on their own; so not necessarily in a practice session, but just for the sake of listening to the notes played randomly by the app like that. So, maybe I didn't mention this before, it's useful for ear training, right? So to also train the ear to hear a note with eyes closed, so without reading. I had fun doing that.* [FG8, P2]

- Imitation:

> *You can imitate that sound. It's like when you get a tuning note.* [FG1, P3]

> *Usually, I try to monitor progress in terms of sound cleanliness, because very often sometimes I have a fairly clean sound, sometimes I have a really bad sound. So, I used this range and then tried to adjust to the sound I heard, the one you recorded. So, my measure of judgment and improvement was to produce a sound as close as possible to that.* [FG4, P]

> *Listening to the sound and repeating it becomes easier to reproduce... The listening is wonderful, fantastic, in the sense that you almost don't notice what you do; because you've just heard it and so you try to replicate it in the best way, absolutely. For me it is a huge added value...* [FG7, P1] *I think the listening and immediate reproduction work really well.* [FG7, P2]

- Alternated practice and reflection:

> *I don't have many chances to hear good trumpet notes, you know? Apart from my teacher, which I see once a week and he doesn't play that much when I'm with him. Hearing a good note on the trumpet is quite rare for me. So the app is doing that, you know? So, I guess that, knowing what's a good sound, maybe, that's the progress I think I made. And I'm hearing more when my sound is not good, you know? I hear that my teacher says the wind that I do with the note: "The note is windy!" But now, I hear it more like that.* [FG2, P2]

> *The fact that you can listen to a sound made well, of good quality, and so you can work precisely on the tone,* [it is] *something that at home, in the absence of the teacher who lets you hear the note properly – especially in the beginning – you don't do.* [And if you do not do it,] *at a certain point, you find yourself having completely ruined the sound simply because you focused on something else maybe. That* [this software] *keeps you straight.* [FG7, P2]

- Combined practice:

  > *I guess the app reminded me to listen more when I play, when I do my exercises.* [FG2, P2]

Structured practice

Another appreciated aspect of the app was its ability to provide a structured approach to practice sessions. For those who found it beneficial, its flexibility allowed it to complement or integrate with existing study routines, enhancing learner motivation:

> *I spent more time to practice the note with the app and to focus on the notes and everything before starting to play something else. So, yes: for me it motivated me to do something to reach the right tone every time before to play.* [FG1, P1]

> *I didn't really use it to set goals, but it did help me plan and it changed a little because I did a longer warm up than usual. I'm usually a little bit lazy with the warm up. So, it was nice to have like a structured warm up and something different than the other warm ups that I've been using for a couple of years. To add in more ear training...* [FG3, P1]

Connecting this to the pedagogical principles previously discussed, the structured practice facilitated several key areas:

- Auditory imagery:

  > *It helps me, you know, to calibrate my ear.* [FG1, P2]

  > *My teacher says to me that I have to really listen to: sometimes close my eyes and listen to the notes and take time. And this this app helped me to do this.* [FG2, P2]

  > *I think that the feature of hearing the note first is really useful.* [FG3, P1]

- Alternated practice and reflection:

  > [The app helps me with] *the coordination between what I hear, the sound I want to reach and what I actually do, you know, my breath.* [FG1, P2]

> *I don't do long tones often. So, the fact that, like, that is the goal of the app and that's all you do on it, I practiced more effectively by default because I was doing the long tones. And I think switching things up from doing only scales to doing notes that are not next to each other also is like an effective way of working on your ear, which I don't do often. So, I think it did motivate me to practice more effectively...* [FG3, P1] *This app changed my view, allowing me to think of the importance of whole notes during the warm up.* [FG3, P2]

> *In my opinion, this app allows, especially beginners, to refine their ear and play these long notes because you can adjust the metronome, so maybe you go to play really long, long notes which, in addition to the fact that there's the reproduction of a real sound shortly before, allows you to refine... At the beginning, before your app, I would say: "Yes, okay, I try to play a long C", but then I don't know how it was, whether it went well or badly. Instead, in this way I have somewhat a compass to follow.* [FG4, P]

> *When I explain* [to my trumpet students]*, I say: "Hear that C you just played. How is it? Try to start the note this way." ... Then during the lesson maybe someone asks: "But, how is it?" Then I take the app again and say: "Look here how it is! That fingering association is like this! See?" Student: "Ah, yes, yes, okay, okay." ... For the kids who say: "Let's pick up the trumpet and start playing right away!" I tell them: "No, wait. Let's warm up a bit first. Let's do this, let's play some long notes, hear your sound, hear this other one".* [FG5, P1]

> *I discovered that my pocket trumpet* [that I use to practice] *is completely out of tune.* [FG7, P2]

> *For my intonation problem, it's very very useful.* [FG8, P2]

- Introduction to combined practice:

> [Discussing about the practice goals trained with the app] *... ear training mostly, and practice the long tones. So, breath training, I would say too.* [FG1, P1] *Yeah, working on tone, on the long tones, trying to keep it nice and steady and in tune.* [FG1, P3] *And being able to hit any note.* [FG1, P2]

- Expanding skills from known to unknown in small achievable steps:

> *I think I did more progress especially in the high notes that I didn't really practice a lot before because I tried to play the music we have in the band and there are not a lot of high notes* [FG1, P1] *I think it helped me more with the high notes as well...* [FG1, P3] *Yeah, it helped me well also with the high notes that I don't reach too much and forced me to practice and, you know, play them thoroughly.* [FG1, P2]

> *I think there are ways to adapt it for different levels. I used it more to focus on tuning and stuff rather than just range or like understanding the notes. So, I think that focusing on the more nuanced aspects of practice was how I adapted it, since I am not a beginner.* [FG3, P1]

> *The application has been very useful to me both for identifying the fingering positions which I still struggle with, not with the normal notes* [of the natural scale] *let's say, but with the various sharps and flats, those slightly stranger things. I still struggle with the fingerings. Moreover, having the feedback of what the sound should be like, hitting the note, that for me was very important...* [FG8, P4] *What can help me is to hear the notes, maybe the higher ones or the lower ones, that I'm not yet doing. Listening to them, and being able to reproduce, the app is useful for that.* [FG8, P4]

> *I found very interesting the dynamics of piano, mezzo forte and forte; something that, honestly, I did not know. I mean, by playing around, you realize that if one wants to play a bit softer, it can be done, but to actually have an exercise on that, I believe could be very useful for me. Then another thing: the fact of setting up the app with the various intervals, with or without semitones, is interesting.* [FG5, P2]

> *The timing thing is very important, the metronome function that keeps time, it helps a lot! Because some of my students, who maybe lagged a bit or started too early, with this feature, I could tell them: "No! Wait! Stop!" (Imitating the metronome) And you set the tempo you want: 60, 80, 120 bpm.* [FG5, P1]

- Multi-sensory approach:

> *Since I use the app at the end of my warm up routine, it helps me make sure that the feelings are here* [indicating their lips] *and that I don't make any*

> *effort to reach the notes that I want to reach. And that my lips are buzzing in a great way, etc. So for me, it's a way to, you know, finish and wrap up the warm-up and the setting. I'm putting myself in before I start playing.* [FG1, P2]

Attention to details

The discussions with participants highlighted the learners' attentiveness to and appreciation for details. One example where the technology was positively received concerns the indication to extend the third valve tuning slide for notes that would otherwise sound sharp:

> *Another thing I've noticed: I think yours is the only app that, when you have to play the low D, signals you to extend the tuning slide of the third valve. Because other apps don't tell you that! These are small details, but they make you realize that there's a lot of study and attention to detail behind it. Definitely.* [FG4, P]

Conversely, a less favorable aspect noted by participants was the app's inability to indicate alternative fingerings for notes that can be played using different valve combinations:

> *I've another thing I thought that might be a nice suggestion which is alternate fingerings.* [FG1, P3]

> *The idea that there's the image of the trumpet indicating the keys could be improved as well, not just show the standard positions but also add alternative fingerings...* [FG5, P1] *For me, the feature of extending the tuning slide of the third valve was very useful... the fact about the keys is interesting because it not only enriches the mechanical (muscular) memory but also the visual memory* [FG5, P2] *Yes, yes, exactly, you see the actual keys to press, so I can remember. But maybe for C sharp, the one on the third space, instead of suggesting the combination of first and second valve, maybe suggest doing it with the third* [alternative positions] [FG5, P1]

Specifically, participants noted the prevalent alternative fingering for the trumpet, which involves using the third valve instead of the combination of the first and second valves.

Perceived negative aspects

Participants also identified several negative aspects of the technology, which impacted their overall experience and are outlined below:

- Absence of a long-term goal:

> *I think the app for me was like, was like good like warm up, but I didn't think it was beneficial enough for me to like set goals other than practice like the range. Because I know my range is okay, I just need to like warm up before I can get there. So for me, like there wasn't so much of like a long term goal, it was more just like a tool for me to just like warm up, I would say.* [FG2, P1]

- Limited scope of practice:

> *At this moment, my objectives are a bit more articulated than just long notes. So, in the end, I already have a plan and the app doesn't fit into this plan. It's an extra. Clearly, long notes need to be done, so yes, that's fine. But it fits very little into my plan.* [FG8, P3]

> *Yes, for me too, simply my goals are different, so. Let's say that this app with the long notes represents a means to improve intonation, sound, etcetera, etcetera. However, from the point of view of objectives, in my opinion, it's not suitable.* [FG8, P1]

> *For me, it's useful for intonation. So, from this point of view, it can help me achieve a goal, which is to tune the notes better. But then one thing is to start a single note, perhaps hearing it from the app; then another thing is to play a series of notes all in tune. So, as a first step, it's a great goal, but then it needs to be extended in my opinion.* [FG8, P2]

> *After a while of playing whole notes, the usual beats with semitones and all, at some point it also becomes a bit monotonous.* [FG5, P1]

- Repetitiveness in the exercises:

> *Maybe this was something that also discouraged me a bit: the fact that after a while the app becomes a bit repetitive, and so at a certain point, I couldn't motivate myself.* [FG6, P]

4.3.3 Focus group findings on Research Question 3

RQ3: What do learners believe is missing in the provided technology to make it useful to them?

Some participants reported using the app for a short period before returning to their previous practice routines. They indicated that the software lacked specific functionalities necessary to overcome their resistance to altering their established practices.

Resistance to changing established practices

Three participants noted that adopting a new method or technology must provide clear benefits to overcome resistance to modifying or abandoning their established approaches. Learners emphasized the necessity for the technology to integrate seamlessly with their existing practice routines or offer distinct advantages that would make the transition worthwhile. The primary reasons for resistance included a fairly strong attachment to their current methods and the absence of compelling features in the app to justify the switch:

> *I have to honestly tell you that after a while, I returned to my usual routine because somehow the test was interesting, but ultimately, I find myself better off with what I was doing before because it was more or less a groove where I was always comfortable, and so somehow I didn't have much flexibility to move around. Let's say that, if I can give such a comprehensive feedback, what I missed was that feature, that facility, in short, that particularity that would push me to make the leap and therefore to move to a different type of learning... The particularity of your app* [compared to my usual routine] *was that it actually randomized the notes for me; but in the end, I didn't perceive this as so fundamental.* [FG6, P]

> *I kind of have like a systematic approach to my practice and I felt like I've been so ingrained in my way that like maybe it didn't really change* [the way I practice]*. I can prefer like doing the long notes the way I usually do it and then the scales and then I just go into the songs.* [FG2, P1]

> *My old warm up I had goes to like sequence of exercises to do and I have like a paper that I follow. And then introducing something new, the effort was just to decide: "Okay, what do I want to warm up with today?" Because I had that routine.* [FG3, P1]

Incorporating feedback functionality

One of the primary features requested by participants was the inclusion of feedback on the sound produced. This feedback was desired in different forms:

- **As a tuner**: Participants expressed the need for a tuner to provide real-time pitch accuracy feedback. This was seen as particularly beneficial for learners to ensure they were producing the correct tone and pitch.

  > *What I would have liked is to know if my tone and pitch was right... I think it would be very useful for somebody who is learning the instrument.* [FG1, P3]

  > *Having the tuner's feedback could be useful because, based on the frequency, it tells me if I'm more or less doing well or if I'm too sharp or too flat.* [FG4, P]

  > *The tuner is very important to me because I personally am out of tune, so for me, it's very important to have someone tell me: "Look, you're too out of tune." The absence of the tuner is what pushed me to go back* [to my previous routine]*, otherwise I would have continued using the app because it was perfectly fine for me... In my opinion, if you could include a tuner, that would be fantastic.* [FG6, P]

  > [Adding] *the tuner would be useful because, for example, I now used it with my external tuner, right? So, well, having something together might be useful.* [FG8, P3]

  > *It would be useful if the app could "listen" to the player and function as a tuner, indicating whether the note is accurate or not.* [OQ][2]

- **As timbral feedback**: More general feedback on the produced timbre was also desired. This could involve graphical or frequency-based feedback to indicate how closely their produced sound matched the reference sound.

  > *On this point, though, if the input could be very interesting, however I missed a checkout, a control of the output. I mean: "Yes, I listen to a nice sound, but what about what I did?" Something that tells me if what I'm producing is correct. So, there could have been, I clearly imagined, something graphical or feedback in terms of frequencies, I don't know, or a form of timbre or something that tells me more or less: "You are very far from what the reference sound was". Because yes, the ear does a lot, but an additional check could be useful.* [FG5, P2]

[2] Open question in the questionnaire; feedback provided by three participants who opted for written comments instead of joining focus groups or one-to-one interviews.

- **As playback functionality**: Replaying the sound the learner just produced, allowing them to compare it with the reference sound and reflect on their performance.

> *There should have been a third measure to hear one's own sound on the app... If there was also a tuner included or a third measure to hear one's own sound again, then it would be really the best, because then there's the comparison from the sound of the trumpeter you chose – this great trumpeter – to the sound I make, and you say: "See the difference?" You say: "Wow, in the example provided by the app the note is like this, I played it wrong, I made it flat, sharp... And there, I think the ear would work a lot!* [FG5, P1]

> *For me, it's necessary to understand how bad what I played sounds compared to the proposed note.* [FG5, P2]

Variety in exercises

In addition to the lack of feedback on the produced sound, participants in the study also identified a limitation in the variety of the proposed exercises. The specific features they desired have been grouped and detailed below:

- Adjustable length of measures:

> *I wanted the tone to be longer.* [FG1, P3]

> *I found the app a bit too unchallenging. Maybe I would add, if possible, this functionality: decide how many measures to go forward, not just one with four beats.* [FG8, P3]

> *The pause in the middle is too long and cannot be modified.* [OQ]

- Increased randomness in the note selection:

> *It was kind of funny: it wasn't very random in the scale.* [FG1, P3]

> *Sometimes the app made me repeat the same note up to three times... it was a bit too recursive, but for the rest, it was perfectly fine.* [FG4, P]

- Inclusion of multiple notes:

> *If we have a group of notes, I think it would be fun.* [FG2, P2]

> *Is it all strictly in whole notes? Would there be an option one day like to implement half or even quarter notes? ... In terms of variety, I think it's very limited for now because it is simply whole notes that are being played. And if there would be some new additions that could replace like certain beginner methods it would be better.* [FG3, P2]
>
> *The fact that after a while the app becomes a bit repetitive, and so at a certain point, I couldn't motivate myself too much. Maybe this was something that also discouraged me a bit.* [FG6, P]
>
> *It needs something else, right? A melody... Like, for example, you make me hear the notes, I replay them with those same notes and it produces a melody.* [FG7, P2]
>
> [The software is too] *basic, let's say. In the sense that I would have preferred there to be at least intervals and chords rather than notes just thrown there... For my level, I would hope to be able to, let's say, expand it if you could manage to do it; it wouldn't be bad.* [FG8, P2]
>
> *Only one note can be played at a time; it should be expanded to work on groups of notes as well... I would expand the ability to play the same thing with duplets, triplets, and quadruplets at the very least.* [OQ]

- Expanding the dynamics range:

> *Dynamics were just mezzo forte, forte, piano. There wasn't like any like decrescendo or crescendo on it so I think to be slightly more useful for me, I think it would be beneficial, to have like a decrescendo mark.* [FG2, P1]
>
> *It would be nice to maybe change the dynamic level as well, like crescendi and diminuendi. That would be beautiful, very useful. Because, I mean, I started not long ago, and I realize that maintaining intonation in a variable-dynamic sound is much more difficult compared to terraced* [fixed] *dynamics.* [FG8, P3]
>
> [I would expand] *the dynamics, maybe from fortissimo to pianissimo. Then maybe even fortississimo and pianississimo wouldn't be too bad.* [FG8, P1]

- Expanding the note range:

> *I would improve the range, I would greatly expand the range. For example, from the pedal note to the triple high C. No, the triple high C seems a bit excessive. But anyway, I would greatly expand the range. I would also put in the key signature.* [FG8, P1]

- Focus on sharps and flats:

> *Would it be possible to have an option that would play only sharp and flat notes? Just because these are the more difficult for me and for many people, I guess, to remember the fingering... So, I would like an option just for flats and sharps.* [FG2, P2]

- Inclusion of additional effects:

> *About the effects, could there be a way to incorporate them into such a context? All the various effects that can be done with the trumpet.* [FG5, P2]

4.3.4 Focus group findings on Research Question 4

RQ4: What are learners looking for in music educational technology?

To address Research Question 4, the focus group discussions were analyzed to understand how participants interacted with the educational technology beyond its suggested use. The analysis aimed to identify the directions taken by users to adapt the exercise to their preferences and needs.

Creative adaptation and engagement

A prominent theme that emerged was the participants' creativity in adapting the exercises and engaging with the interface. This creative interaction appears to be a significant factor in maintaining learner motivation. Participants reported experimenting with various ways to extend their practice sessions and make the exercises more challenging and enjoyable:

> *I could say the app motivated me to practice a bit longer. Yes, because then I try different things on it.* [FG1, P3]

The various ways learners engaged with the technology, including both the methods they applied during the intervention and the ideas they wished they could implement, are listed below:

- Playing longer tones:

> *What I end up doing was playing in the bar of rest before the next notes make it twice as long. I thought that would work. So that worked for me.* [FG1, P3]

> *One can also play all the notes they want afterward as indicated by the app in whole notes. But if one wishes, they can also do it in unison with the app. At least that's what I tried to see if it more or less matches or not.* [FG5, P2]

- Practicing ear training:

> *What I found was interesting: I tried as well to do by closing my eyes and seeing if I could figure out what the next note was and match the note.* [FG1, P3]

> *Once I start using that, it was kind of fun. I tried to test myself with the ear training by using the app. I kind of like extended my practice a bit longer by default.* [FG2, P1]

> *It's the ear training part that I probably did more of during these few weeks. So let the app play and try to determine which note. And which I scored not too well on, I have to admit.* [FG3, P2]

- Incorporating gamification:

> *I play a little game with my trumpet teacher, you know, just for fun, but ear training also: He plays to me, let's say 5 notes random and I have to reproduce the same thing. For me, it's hard! And it's quite fun and I think it's very important and he adds notes, you know, 6 notes, 7 notes and I have to reproduce each time. And I was wondering if there would be an app that would do that, you know, just randomly plays 5, 6 or 7 notes and we have to reproduce, just like that. It's kind of an idea that I had and I would use it, you know, just to play a bit.* [FG2, P2]

- Learning staff notation:

> *Would it be possible to have the note written just under the trumpet when you play it? The name of the note, you know? Do or C or, you know, just visually to get what is the note to associate the name of the note with the sound and the fingering.* [FG2, P2]

> *I think it would be really useful, for like a teaching setting, to also have the note names somewhere. Like, I don't know, maybe under here* [pointing below the score] *you put like "this is an E". So then like students are also being learning how to read the sheet music, with the note names. Maybe this comes like at a later time in their playing and they would already know that, but to turn it ON or OFF might be interesting.* [FG3, P1]

- Tracking progress:

> *If there was maybe a way to track like the hours on the app so that like the teachers could use it as like, you know, an accountability measure.* [FG3, P1]

- Challenging established routines:

> *The technology helped me to plan to set goals in my study because, fundamentally, even just the fact of having tried it, and having tried to use a different method, was useful to try to understand also the limits of my previous one. So anyway, I find it a useful operation to try to change the method. Not necessarily, this must then actually result in a change.* [FG6, P]

- Engaging in focused and exploratory practice tasks:

> *In the beginning, I practiced while looking at the phone screen, and it was a nice motivation, also fun, to try to hear the sound without looking at the app, with a real sound, and then try to reproduce it. So maybe while I was waiting for those 3 or 4 beats of rest, I would say: "Come on, now I have to guess the note that's about to play!" And so it motivates you by teasing. It piqued my curiosity a bit, as if I were in a quiz: I have to guess the note; let's guess it 100%, especially if you then add sharps and flats, you increase the level of difficulty a bit.* [FG4, P]

Towards a Comprehensive Digital Method

Feedback collected during the interviews revealed a strong interest in enhancing the software's capabilities, aiming to transform it into a comprehensive digital tutor and method for personal use and classroom settings. Participants expressed a desire for additional functionalities that would make the software a more effective and versatile educational tool.

- **Expansion to form a digital method**: Participants emphasized the need for the software to offer more than just reproducing notes heard, offering a more in-depth method for comprehensive study.

> *It would be, let's say, useful to expand it. So, all in all, even leaving the structure as it is but I could choose to say: "Practice on the scale, I don't know, on the circle of fifths", right? Even on just the major keys: C major... Scales and chords, scales and chords, scales and chords.* [FG8, P2] *Yes, yes, I agree. Scales, arpeggios, things like that, intervals of up a third, a fourth would be very useful.* [FG8, P3]

> *To be an app completely dedicated to study, you need to add something else. That is, one expects more than just reproducing notes you've just heard, you know? So, as a warm-up it's definitely great but as a real study tool it is not, in my opinion.* [FG7, P2]

> *After playing these whole notes for a while, I'd also try to split them up, make a kind of random addition, like a study method, I don't know, like "Gatti"* [it is a trumpet method] *or whatever. I'd like to do half notes, triplets, eighth notes, quarter notes.* [FG5, P1]

- **Adapting Practice Sessions to Learner Needs**: Participants highlighted the importance of adapting exercises to their current skill level and recent practice routines. They suggested that the software could dynamically generate exercises tailored to the range of notes, rhythms, and techniques they had been working on, making practice sessions more personalized and engaging.

> *To make it a bit more fun, a module with some simple melodies or easy scales for learning based on the range practiced could be included.* [OQ]

> *It would be nice that, after you've warmed up on that range of notes, maybe you also do a more complex exercise. Maybe you also choose the beat and let's say the rhythms, if you want to go up to sixteenth notes, if you want to use thirty-second notes, triplets, quintuplets; it matches the exercise on that range you've chosen, and you do it.* [FG7, P2]

- **Developing the method as mobile app**: Participants appreciated the idea of having a comprehensive app that integrates various functionalities, including a tuner and a practice diary, in a method for structured practice.

> *It's an excellent starting point that can become even better, in my opinion, because I like the idea of having a single app where I have the tuner, can listen to sounds, and can also have a sort of method to practice with. In a single*

> *app, I have everything and basically don't have to do anything else but study. It would also be interesting from this point of view. I don't know what you think about also creating a sort of progress diary to keep track of when I'm studying, how I'm doing... It can be useful – and I already imagine it complete with tuner and everything – because it creates a sort of continuous training, meaning the student/apprentice goes to music lessons, practices freely, and can continue to do quality practice at home... A digital tutor that can be very useful.* [FG4, P]

> *Having a study companion, a help, is certainly valid. And it motivates you perhaps to do it rather than practicing by yourself. Even just the idea of having an app, seeing that the world today works on apps, is not bad.* [FG8, P2]

> *A more constructed, more structured app can absolutely help humanity.* [P2, FG8]

- **Enriching Classroom Instruction**: Participants saw potential for the app to be used as a supplementary tool in educational settings. They believed it could be a quick reference for teachers and a useful practice companion for students.

> *I find it very useful. I gave the highest rating, or something like that, in the teaching part in a class because I find that this app, together with a teacher who somehow makes it become more – how do you say? – lively, you know? It entices you, then it's good because basically it's perfect for a teacher to be a quick tool and have kids do exercises.* [FG6, P]

> *I think it's a good app to integrate in a class or, just like you said before, just like we did so at home as a practicing homework. I think it both can be very good to use.* [FG1, P1] *Yeah, I think so as homework for the new students, I think that would be very good. It might motivate them more to try it. Especially since new children all have phones and like playing with apps.* [FG1, P3]

> *I continue to keep this app because I believe my students like it. Then we have fun. In fact, I say, "Look, this is how it's done, see here on the app?" I mean, I explain it well to them, the association of fingering and rhythm. So, it's indisputable, right? What I say can be disputed because it's just one*

> *person's word; but here's the app, and nowadays, the app tells the truth. So...* [FG5, P1]
>
> [Discussing about the usefulness of the software in a classroom] *What the app does is done by the trumpet teacher in class, so it's not needed in class. If, however, we're talking about using it as a complementary tool when the teacher is absent and the kid practices at home, then absolutely yes.* [FG7, P2]

4.4 Discussion

The previous section presented a selection of quotes from participants during focus groups, organized by research questions and identified subthemes. As it is not an exhaustive analysis of all potential themes, readers are encouraged to review the interviews for additional insights and consider new categorizations and reflections based on the collected data. This discussion section aims to highlight key insights from the reported data, while acknowledging the limitations of this exploratory study, which are important to consider when interpreting the findings and their implications.

4.4.1 Discussion over Research Question 1

RQ1: How do participants perceive the ease of use of the software application presented in Chapter 3?

The quotes reported in Section 4.3.1 highlight a strong demand for educational technologies that are not only effective but also extremely easy to install and use. Participants particularly valued the mobile application available on app stores, as this can be installed and operated like any other app, enhancing perceived ease of use.

However, despite the high appreciation for the ease of use of the software provided, feedback about inefficiencies suggests that users seek even more intuitive and user-friendly interfaces. This indicates a rapidly advancing expectation for mobile technologies, driven by the seamless experiences offered by widely used apps like social media platforms. The latter applications are developed by teams of experts who optimize every aspect of user interaction to ensure simplicity, intuitiveness, and comfort, and they are available for free. This pervasive ease of use is likely raising the standard for what users expect, creating a challenge for educational technology developers to meet these high expectations.

The contrast between user expectations and the resources available for academic software development underscores the need for an open-source platform to facilitate the creation of educational technologies for musical instrument learning, especially mobile apps. Such a platform could provide researchers with modular components and services, similar to those used in the software developed in this study. These modules could include score visualization, fingering images, data input selection, synthesized or pre-recorded sounds, recording features, and any form of feedback visualization. By providing these ready-made modules, the platform library would reduce the need for advanced coding skills, allowing researchers to develop sophisticated software and prototypes more quickly and efficiently.

The importance of ease of use is further highlighted by participants' behaviors. During the interviews, many participants acknowledged that using headphones would have improved their experience by enhancing the quality of the reference sound and better following the metronome ticks. However, 95% of participants chose to use the device's speakers for convenience. This choice suggests that users often prioritize convenience and simplicity, even when they are aware that an alternative method would provide better results.

The findings suggest that while the development of mobile educational technologies leveraging embedded device sensors is a significant step forward, these tools should also prioritize advancing their intuitiveness and efficiency to meet user expectations shaped by mainstream mobile applications. An open-source platform could play a pivotal role in supporting researchers to develop high-quality, easy-to-use educational tools that align with these evolving standards.

4.4.2 Discussion over Research Question 2

RQ2: Which aspects of the provided educational tool are perceived as useful by learners and which are not?

In examining the feedback provided by participants, several key aspects of the educational technology's design and functionality emerged as critical to its effectiveness and user acceptance.

Real Sound

Participants value the reference sounds recorded by a professional musician over synthesized sounds, recognizing their role in fostering auditory imagery, aiding in imitation, and alternated practice through reflection. The inclusion of real trumpet sounds differentiates this technology from other educational tools that typically use synthesized sounds:

> *In my opinion, it's an excellent starting point, also because I believe it's the only one available that has a recorded sound – you told me it's a tone recorded by a*

> *musician – so it's not a synthesized sound. Before meeting you, I had downloaded others, where there are virtual trumpets but those sounds are, let's say, artificial.* [FG4, P]

It was also noted by a participant in the Focus Group 7 that this technology could be something that emulates what the teacher is doing and therefore, in the presence of a teacher, it is not necessary. This highlights the software's ability to support independent learning, particularly outside of formal lessons or in the absence of a teacher. This aligns with the broader goal of music education: fostering students' independence and equipping them to become lifelong learners who can continue developing their skills autonomously.

Learners appreciate the real sounds because they provide a more authentic and accurate reference, closely mirroring the actual experience of playing a trumpet. This direct and faithful mapping between the practice tool and the musical instrument enhances the learning process by offering a realistic auditory target for students to emulate.

A noted limitation is that the software provides only one sound track per note recorded by a single professional musician. Expanding the range of sounds to include different timbres could enrich the learning experience, offering a broader palette of auditory guides. For future research, this expansion could provide various sounds reflecting different techniques, dynamics and styles, giving learners a more comprehensive auditory reference.

Among brass pedagogues, there is a widely held belief that there is a relationship between timbre quality and sound production efficiency, as extensively argued in the next chapter of this thesis. Strained and shrill tones often indicate excessive muscle tension and rigidity in the musician's body, resulting in inefficient playing. In contrast, rich and round tones are associated with efficient sound production techniques. Therefore, especially in the early stages of learning, it is crucial for students to listen to and imitate efficiently produced sounds to avoid developing bad habits and muscle strains.

Consistent exposure to high-quality sounds helps establish strong neural connections that promote good playing habits. According to Steenstrup (2017), if students regularly listen to and aim to replicate poor sounds, they reinforce neural pathways associated with those poor sounds, making them habitual. Conversely, regularly hearing and striving to replicate high-quality sounds strengthens the neural connections linked to producing those sounds, making excellence the norm. This highlights that effective practice is not just about the amount of time spent but also the quality of the auditory examples internalized.

Therefore, it is essential for educational technologies designed to support musical instrument learning to incorporate these principles into their design. After all, it is a music app for learning music. By integrating high-quality, real sound references, these technologies can guide learners toward developing efficient playing techniques from the beginning of their studies, fostering better long-term practice habits and musical proficiency.

Structured practice

Section 4.3.2 highlights the benefits of the educational technology in providing a structured approach to practice sessions. Learners generally appreciate the adaptability of the exercises, which allows them to tailor their practice to their current level and progressively expand their skills. This structured practice also guides them towards alternated and combined practice through multi-sensory engagement and the development of auditory imagery.

The positive reception of adaptable exercises suggests a demand for modular practices that can be customized by both teachers and learners to meet individual needs. A one-size-fits-all approach may not be the most effective, as students have varying learning paces influenced by factors such as available practice time and age. For instance, a participant in Focus Group 4 expressed frustration with the *Arban*'s method (Arban, 1982), a widely-used comprehensive trumpet learning method, which introduces high register notes since the first exercises. Struggling to play many of the notes, the participant ultimately abandoned the method.

To overcome this issue, experienced teachers typically select exercises from various method books to suit individual learners' needs. A digital method, however, could address this by allowing exercises to be tailored to the learner's abilities and desired progression. This flexibility would benefit students, teachers, and self-taught learners by enabling quick and interactive customization of exercises. It would help reduce frustration and empower learners to understand their limitations (e.g., pitch range, metronome speed) and track their progress over time. By generating exercises appropriate to the learner's current skill level, the software can adapt to the context of music teaching, allowing teachers to set personalized input parameters for each student.

Such a structure provides learners with clear guidance on "where to go next", enhancing their ability to self-assess and regulate their learning processes. This approach aligns with the concepts discussed by McPherson et al. (2022), who emphasize the importance of self-regulated learning in music education. By offering a structured yet adaptable practice environment, the educational technology fosters a personalized and effective learning experience, enabling learners to progress at a pace that suits their individual needs and capabilities.

Attention to details

Participants' focus on details, such as appreciating the guidance to adjust the third valve slide when needed and requesting alternative fingerings, demonstrates learners' attentiveness to in-depth pedagogical content. This further validates the effectiveness of the TPACK model as a robust framework for integrating technology in education (Willermark, 2018).

The involvement of individuals with expertise in teaching the specific musical instrument targeted by the technology is crucial during the design and development phases (Michałko et al., 2022). Their input ensures that the technology is pedagogically sound and relevant, which not only enhances the learning experience but also builds credibility among educators, potentially lowering barriers to its adoption in their classes. By integrating such detailed and context-specific content, educational technology can provide a richer, more tailored learning experience that meets the needs of both students and teachers, fostering a deeper understanding of the instrument.

4.4.3 Discussion over Research Question 3

RQ3: What do learners believe is missing in the provided technology to make it useful to them?

Participants' feedback highlighted a key consideration in the adoption of new educational technologies, emphasizing that their success depends on offering clear and tangible benefits that justify the effort of integrating them into established learning routines. When asked to suggest improvements to the software, participants provided a detailed list of enhancements, as outlined in Section 4.3.3. These suggestions primarily focused on incorporating feedback on intonation, replaying the note just performed, and expanding the exercises to include multiple notes, varied rhythmic figures, and dynamic markings.

Interestingly, many of the proposed improvements reflected ideas already found in existing educational technologies, even though participants were encouraged to think of any new ideas beyond current technological limitations. This indicates that their suggestions were likely biased by familiarity with existing music pedagogical tools, rather than inspired by innovative ideas that push the boundaries of current practice.

There were, however, a few suggestions from individual participants that ventured beyond what is currently available. For example, one request proposed a more general form of timbral feedback that would assess how correctly the sound had been produced. Another suggestion involved the inclusion of "additional effects", though the formulation was somewhat vague, and the participant was unable to articulate precisely what they envisioned. While these ideas hint at a desire for more advanced features, they were limited in scope and detail.

A more innovative approach could involve integrating well-established pedagogical strategies into the software. For instance, a recent study has shown that alternating between mental, vocal, and physical practice can significantly enhance musical performance (Steenstrup et al., 2021). This type of combined practice could be easily incorporated into the proposed exercise by adding a step where learners sing the note after imagining it and before playing it on the instrument. Singing would provide immediate feedback on whether the aural image of the tone is accurate, free from the technical constraints of the instrument, and would iteratively guide learners more effectively toward their musical goals.

However, none of the participants suggested this potential improvement, nor did they experiment with it, despite their creative use of the app in other ways. This observation suggests that the conception of innovative educational exercises might benefit more from the insights of educators and pedagogical experts than solely from user feedback. While learner input is invaluable, it often reflects current experiences and expectations rather than pioneering new approaches to learning.

4.4.4 Discussion over Research Question 4

RQ4: What are learners looking for in music educational technology?

The focus group findings reported in Section 4.3.4 reveal that learners are receptive to educational technologies that offer customization, engagement, and progress-tracking. This openness, however, varies based on individual preferences and priorities.

Some participants emphasized the convenience and portability of a mobile app for educational purposes, appreciating its ease of access and the immediacy it provides:

> *And then it's also interesting that one has it at their fingertips, on their phone, and let's say it's a trivial thing, but not too much... It's a bit of an advantage for lazy men like me in this case.* [FG5, P2]

The shift from traditional, often cumbersome, paper-based methods to digital tools like mobile apps is already valued for features such as viewing sheet music, adding unlimited annotations, and navigating through multiple pages with ease. However, participants expressed a preference for technology that takes a more active role in their learning process, such as by adapting exercises to their specific needs and tracking their progress.

Different learners have different priorities and preferences. Traditional music methods used in schools and conservatories typically follow a linear path of continuous improvement, offering progressively challenging exercises that lead motivated learners to higher levels of mastery over time through dedicated practice. However, the focus group discussions revealed that many participants are satisfied with their current skill level and do not seek or have the time for intensive study aimed at significant improvement. For them, playing music is primarily a hobby, and their main goal is to enjoy playing stimulating exercises and participating in ensemble music.

This preference for ensemble playing is reflected in several quotes from the discussions:

> *I really didn't have set goals for my practice... it's just to play enough to get in shape, I guess, to stay in my band.* [FG3, P2]

> *My progress is more measured on how I can play the pieces we're playing in band rather than my technical abilities.* [FG3, P1]

> *What counts as motivation more than the app is perhaps, I don't know, when they propose you to form a music group. So, I practice a lot to avoid making a bad impression; or there's a concert coming up with your band and so on and so forth.* [FG6, P]

> *About the goals for most of my practice sessions, I'm not doing as many like rudiments or techniques.* [This is] *because I'm busy and I play to be in the band rather than play to improve as a player.* [FG3, P1]

In addition to ensemble playing, the importance of enjoyment during practice was another recurring theme:

> *For me using the app adds just a bit more time of practicing the instrument, which is good because, it is fun for me.* [FG1, P1]

> *I prefer to have a bit of fun, reward myself, rather than spend too much time studying.* [FG5, P2]

These insights point toward a demand for a comprehensive digital method that is not only easy to install and use but also flexible enough to cater to the different needs of learners. Whether a learner seeks to improve or simply maintain their current level, the technology should offer a personalized and targeted, yet engaging and enjoyable experience that aligns with their individual preferences and specific goals.

4.4.5 Limitations

One of the primary limitations is that the educational software developed in this study is specifically designed for Western classical music training, using staff notation as commonly taught in music schools and conservatories. As such, the proposed solutions may not directly translate to other musical traditions or educational contexts, which limits the broader applicability of the findings. Additionally, although the study focused exclusively on trumpet learning, the underlying principles could potentially be adapted to other musical instruments, though this remains to be explored.

Another significant limitation is the sample size and diversity. With only twenty adult participants, all from Europe and North America, the study does not represent the broader global population of trumpet learners, including minors or individuals from other cultural backgrounds. Therefore, only descriptive statistics could be generated and any statistical data are not significant. This narrow demographic scope positions the study as exploratory, providing initial insights that would benefit from further research with a more diverse and larger sample.

Additionally, the study faced potential selection bias. Participants were recruited from specific groups, such as university students and social media followers, who may be more inclined toward technology and innovation. The recruiting process reached a larger audience, but only those who responded and were willing to commit to the study participated, likely skewing the results toward individuals who are more motivated and open to new educational approaches.

Despite these limitations, the findings provide meaningful insights into the use of technology in music education, particularly for trumpet instruction. The following subsections delve into the key findings related to each research question. These insights, while rooted in a specific context, lay the groundwork for further research and potential applications in broader educational settings.

4.5 Suggestions for future studies

A recommendation for future educational music software is to consider leveraging the hardware components embedded in mobile devices, particularly when these devices are capable of delivering sufficiently accurate and robust performance. Developing such software as a mobile application can be advantageous, as it tends to increase the perceived ease of installation and use among learners, as observed in this study.

Participants in this study demonstrated a willingness to integrate new software into their practice routines, provided that the software offers clear and evident benefits. While some learners prioritize improving their technical skills, others place greater value on the software's ability to engage and entertain. For many musicians, especially those who play recreationally, educational technology that combines effective pedagogical principles with elements of enjoyment and adaptability can significantly enhance their relationship with their instrument. When educational tools cater to both extrinsic motivation (meeting the learner's needs) and intrinsic motivation (providing a pleasurable experience), they are more likely to be continuously used and accepted (Davis et al., 1992).

Achieving this level of engagement requires a strong alignment between the capabilities of the technology and the tasks learners need to accomplish (Goodhue and Thompson, 1995). This alignment can help learners enter a state of *flow*, as described by Csikszentmihalyi (2008), where they are optimally challenged, avoiding both boredom and anxiety. This promotes effective learning through personalized and stimulating exercises. The application presented in Chapter 3 sought to create such alignment by allowing users to customize metronome settings, note ranges, and other options, such as fingering indications and dynamic markings. These features address basic technical aspects while providing a tailored practice experience.

Future directions could integrate generative artificial intelligence to create adaptive learning paths in music education. As outlined in recent studies (Li et al., 2024), generative AI has shown promising applications in creating tailored content for learners, dynamically adjusting to individual needs and learning trajectories. The software developed in this study could serve as a foundational source of learning material, providing the necessary content for generative AI to develop personalized music exercises. Furthermore, as suggested by Calo and Maclellan (2024), generative AI could enable educators to create effective and engaging educational interfaces through simple textual prompts, adhering to established design principles without requiring coding skills. This code-free approach empowers educators to focus on pedagogical goals while leveraging AI to streamline the development of adaptive and interactive learning tools.

Lastly, to increase engagement, promising direction is the incorporation of gamification elements into the exercises. Gamification, which involves using game design elements in non-game contexts, has the potential to enhance students' motivation and engagement, making the learning process more enjoyable and effective (Deterding et al., 2011). By integrating playful elements such a tiers and levels within the software, educational software can transform routine practice into an engaging experience that fosters both learning and enjoyment.

Furthermore, future research could explore new interfaces and approaches that cater to diverse learner needs. For instance, expanding beyond traditional Western musical notation to include alternative notations could make the software more accessible to a broader audience. This is reflected in the feedback from a participant who expressed a preference for focusing on fingering positions rather than reading musical staff notation:

> *For me,* [having the fingering images] *was really convenient because I'm more interested in knowing the fingering positions than the musical staff, so I often look more at the valve positions than the notes marked on the staff, which means I often disregard the latter.* [FG5, P2]

Therefore, expanding beyond the musical notations on the staff could allow for more users to engage in this software. For example, expanding the software to address inclusivity and accessibility barriers through the use of digital technologies and mobile device sensors could pave the way for more adaptable and inclusive educational tools. Mobile technology could address challenges that traditional paper-based methods cannot, opening up new possibilities for a wider range of learners.

Although this study is focused on just the trumpet, the development of a new technology based on the study could be expanded to other musical instruments, including digital musical instruments. Future studies could explore ways to apply the model of this study into the settings of specific musical instruments and traditions.

In summary, future educational music software could aim to be more than just a tool for technical improvement. It can become an engaging, adaptable, and inclusive resource that meets the diverse needs of learners, fostering both intrinsic and extrinsic motivation. By considering these factors, educational technologies have the potential to create a more meaningful and enjoyable learning experience for musicians at all levels.

4.6 Conclusions

This chapter is an exploratory study aimed at assessing the usability and potential of the mobile application presented in Chapter 3. By investigating the experiences and perceptions of adult trumpet learners, the study sheds light on how such a technology can address limitations in existing educational tools and foster skill development in music education. The results highlight the app's strengths, particularly its intuitive design, use of authentic sound references, and ability to support structured and adaptive practice routines.

While the findings affirm the app's promise as a practical tool for individual and classroom use, they also reveal areas for improvement, including the need for enhanced feedback mechanisms, greater exercise variety, and the integration of features that engage learners more deeply. Participants' insights underscore the importance of aligning technological solutions with pedagogical goals, ensuring that educational tools are both effective and enjoyable.

This study contributes to the broader discourse on the role of digital technology in music education by emphasizing the importance of accessibility, customization, and inclusivity. As educational technologies continue to evolve, their capacity to foster both independent and collaborative learning will play a critical role in shaping the future of music pedagogy. By building upon the findings of this chapter, future research can further refine and expand the possibilities for digital tools in supporting musicians at all levels.

Author Contributions

Alberto Acquilino: Conceptualization of the study, principal investigator in the ethics certificate, participant recruitment, organization and facilitation of meetings and focus group interviews, data collection and analysis, transcription and translation of the focus group discussions (Appendix D), and writing of the original draft. **Jenny Jieun Park**: Supervision of the data analysis, review, and editing of the study. **Gary Scavone**: Conceptualization of the study, supervision of the overall research, and critical review and edits.

Ethics Statement

Ethical approval for the study, including consenting procedures, was granted by the Research Ethics Board Office of McGill University under REB File Number 23-06-088, following the guidelines of the Canadian Tri-Council Policy Statement.

Chapter 5

Detecting efficiency in trumpet sound production: Dataset, baseline and pedagogical implications

In Chapter 3, an open-source app designed to support trumpet fingering learning was introduced. The app was developed using a modular framework composed of reusable components and services, as detailed in Appendix B. The usability of this technology was evaluated through an exploratory study with adult participants in Chapter 4. This study suggested a perceived ease of use of the app, highlighting its strengths while also revealing areas for improvement.

A key motivation for adopting a modular programming structure was the flexibility it offers in designing diverse educational exercises and solutions. The modular components can be adapted in different ways:

1. **Modifying individual component outputs**, such as:
    - The Score Display Component can illustrate four quarter notes instead of a whole note.
    - The metronome can be programmed to gradually increase the tempo, making exercises more dynamic.
2. **Reconfiguring the assembly of components**, for example:
    - The Score Display Component can be removed to encourage users to identify notes through ear training.
    - The Fingering Display Component can be replaced by the Chromatic Tuner Component, requiring users not only to play the correct notes but also to maintain

accurate intonation.

3. **Developing entirely new components and services**, such as:
    - Introducing additional exercises to target different musical skills.
    - Implementing an alternative notation system tailored to specific learner populations.

Ultimately, the choice of configuration depends on the educational needs of both learners and instructors.

While each of these potential directions holds significant educational and research value, this chapter and the following one focus specifically on the study of machine learning models and digital signal processing algorithms aimed at providing feedback on technical skills discussed in Section 2.4.2. In particular, this chapter explores the relationship between timbre quality and efficiency in trumpet sound production, whereas Chapter 6 introduces an algorithm for robustly measuring the duration of the attack transient in wind and bowed string instruments. The algorithms presented as proofs of concept in these chapters could later be integrated in the presented app as additional components and services, significantly expanding its educational potential.

Timbre quality plays a fundamental role in brass instrument pedagogy, as it is closely linked to efficiency in sound production. Prominent brass pedagogues have reported that excessive muscle tension and inefficient playing techniques are often reflected in the timbre quality of the sound produced, which can be easily identified by an experienced ear. An automatic system capable of assessing timbre quality could provide valuable feedback to musicians, particularly during independent practice, helping them develop good playing habits.

To investigate this problem, a new extensive dataset of more than 19,000 tones played by 110 trumpet players of varying expertise was used. A subset of 1,481 tones was manually labeled by eight expert graders on a scale of 1 to 4 based on the perceived efficiency of sound production. Statistical analyses were conducted to examine inter-rater agreement and assess the consistency of expert judgments. A Random Forest Classifier was then trained using the mode of the expert ratings, with its accuracy and variability assessed relative to human grader agreement. Feature analysis identified stability of spectral peaks as a critical factor in determining timbre quality.

Finally, a pedagogical interface was designed to integrate the classifier into an interactive learning environment. This system provides real-time feedback on timbre quality, offering musicians an objective tool to monitor and refine their sound production. The proposed approach contributes both to music information retrieval research and to music education by bridging the gap between automated analysis and pedagogical application.

This chapter is based on two publications:

- A. Acquilino, N. Puranik, I. Fujinaga, and G. Scavone. Detecting efficiency in trumpet sound production: Proposed methodology and pedagogical implications. In *Proceedings of the 5th Stockholm Music Acoustic Conference*, pages 72–79, Stockholm, 2023b. KTH Royal Institute of Technology.

- A. Acquilino, N. Puranik, I. Fujinaga, and G. Scavone. A dataset and baseline for automated assessment of timbre quality in trumpet sound. In *Proceedings of the 24th International Society of Music Information Retrieval Conference*, pages 684–691, 2023a.

The content of these two articles has been integrated to remove redundancies and maintain a coherent flow, enhancing readability.

5.1 Introduction

The significance of tone quality in brass musical instruments has attracted considerable attention due to its relevance in areas such as pedagogy and musical performance. Teaching aural discrimination skills of tone quality is indeed a major component of music training (Simmons, 2005). The emphasis placed on the development of good tone quality can be attributed to its close relationship with sound production efficiency. However, the multi-variable interaction that contributes to the characterization of timbre makes defining its quality a challenging task (McAdams et al., 1995). The following subsections examine this topic from two angles: the musician's perspective, which considers tone as a direct outcome of sound production efficiency and playing technique, and the researcher's perspective, which seeks to analyze and quantify its acoustic properties to enable objective evaluation.

5.1.1 The musician's perspective of tone quality

Sound production on a trumpet involves a complex coordination and balance between the embouchure, the oral cavity, and the airflow (Campos, 2005). It is a delicate balance in which even the slightest alteration in any component contributing to the creation of a tone can result in changes to the overall timbre (Levarie and Levy, 1980). In trumpet teaching, it is widely believed that inefficiencies in playing, caused by sub-optimal coordination of the elements involved in sound production, are reflected in the quality of the sound produced.

Arnold Jacobs, one of the foremost brass pedagogues of his time, reports how the presence of excessive muscle tension, which causes rigidity in the musician's body and inefficiency in playing, is reflected in a forced and strained sound (Jacobs and Nelson, 2006). Similarly, Thompson (2003) argues that an incorrect and inefficient sound production mechanism obliges the player to force, leading to a decreased endurance and a strident tone. Analogous observations are mentioned by Campbell et al. (2021) who argue that obtaining a good sound depends on the skill of the musician as well as the quality of the instrument, and that timbre plays a significant role in the ability to project the sound.

Similar conclusions are also drawn by Steenstrup (2007), according to whom "*it is more efficient, from the point of view of both the lip and respiratory musculature, to produce a beautiful, round tone rich in harmonics than a shrill or dull tone*" and points to a lack of thorough research in this area. Steenstrup proposes a parallelism between the sound production mechanism of trumpet players and that of singers, for whom there is more research in the literature. In particular, he hypothesizes that the influence between different types of phonation (i.e., pressed, flow, breath phonation) on the acoustic spectrum – as described by Sundberg and Gauffin (1978) – may find a counterpart in the way sound is produced for brass musicians. It should be noted that this idea was proposed as a way to visualize or better understand sound production for trumpet players, without suggesting that the actual acoustic mechanisms involved are similar.

Other sources report that timbre provides most of the information to music teachers on which to base the choice of suggestions offered to their students, consequently influencing the determination of the educational path through targeted exercises (Campos, 2005; Cassone, 2009).

What is reported in the pedagogical field has also a correspondence in the context of orchestra auditions and competitions, where some of the rehearsals hide the performer from the jury's view in order to limit the introduction of bias in the selection of candidates. This suggests that an experienced ear, such as that of the jury members, is able to distinguish the level of efficiency and ease of sound production based solely on auditory information.

5.1.2 The researcher's perspective of tone quality

Helmholtz (1977) was among the first to attempt providing insight into the audio properties related to the quality of a musical tone by proposing a direct relationship to the quantity and to the relative intensity of its constituent partials. In an exploratory study using the trumpet as a case study, Madsen and Geringer (1976) identified the amplitude of the first overtone as a discriminatory feature between tones of differing sound quality. Building on this finding, a subsequent perceptual study by Geringer and Worthy (1999) analyzed the tonal quality of the trumpet by altering the content of partials in the sound.

In more recent years, the investigation of trumpet tone quality has emerged as an area of inquiry within the field of Music Information Retrieval. A pioneering study conducted by Knight et al. (2011) examined the potential of a model classifier to categorize trumpet tones into two, three, and seven classes. This research assessed 56 single- and multi-dimensional audio features, as well as their correlations with human judgments, utilizing a dataset comprised of 239 individual sounds played by four trumpeters. Despite the relatively low accuracy of the resultant model, the study provided a promising proof of concept for future research in this field.

A few years later, a research study was conducted on similar premises, targeting various musical instruments, including the trumpet (Romani et al., 2015; Bandiera et al., 2016). Having created an online platform for collecting, labeling and evaluating audio samples, the researchers proposed machine learning algorithms for the assessment of sound quality based on five attributes, namely dynamic stability, pitch stability, timbre stability, timbre richness, and attack clarity. Among these attributes, timbre richness – defined as the quality of timbre – is the one that comes closest to that considered by Knight et al. (2011) and suggested by brass teachers. Despite these efforts, the findings revealed a weak correlation between the scores generated by the trained model and the rankings assigned by human evaluators, indicating significant room for improvement in the model's performance. Additionally, the study presents notable limitations due to the lack of diversity in the dataset, which consisted of sound samples collected from only two trumpet players. In order to train the model, the two musicians – professional music performance graduates – were required to record correctly played notes, as well as intentionally alter each attribute to produce incorrect variations. This could introduce further limitations, as it is not necessarily true that the sound of a professional musician who deliberately plays badly presents the same audio characteristics as the incorrect sound of a novice trumpet player, given that the musician may not easily disassociate from years of procedural memory training. To the best of our knowledge, this represents the most recent investigation in this domain.

This chapter aims to provide a comprehensive exploration of this subject, incorporating a complete dataset of sampled sounds and expert-generated labels[1].

Section 5.2 describes the dataset collection and preprocessing. Section 5.3 presents the machine learning training process, results, and visualization based on the most important feature. Section 5.4 introduces an interactive exercise that implements the designed algorithm within the educational framework described in Section 3.2.

[1]The dataset can be accessed at: https://zenodo.org/record/8132780

5.2 Materials

The dataset employed for training the proposed model comprises auditory samples gathered by the candidate at various music institutions and master classes throughout Europe before the start of his academic program at the host institution. In total, 110 distinct trumpet performers were recorded under varying acoustic conditions. To encompass the complete spectrum of sound production efficiency levels, individuals from diverse backgrounds were recorded, including students and instructors from amateur music schools, arts universities, orchestral musicians, and international jazz and classical soloists.

The same recording system was utilized across all data acquisition sessions, specifically the IM69D130 Shield2Go evaluation board developed by Infineon Technologies, which is equipped with two Infineon IM69D130 Micro-Electro-Mechanical Systems microphones. Such a microphone exhibits an Acoustic Overload Point of 130 dB, allowing it to capture loud audio signals such as those produced by a trumpet without distortion or saturation. Moreover, the microphone offers a sufficiently flat and extensive frequency response ranging from 20 Hz to 20 kHz, thereby covering the entire audible spectrum.

The selected evaluation board was connected to a Raspberry Pi 4 Model B and a Raspberry Pi Model 3B+ for recording, as shown in Figure 5.1. To ensure portability for recordings in music classrooms and masterclass venues, the entire system was powered by a portable power bank, allowing for flexibility when traveling between different locations. The redundancy of using two Raspberry Pi units served as a safeguard against potential technical issues on-site. Additionally, a custom-built interface was implemented on the Raspberry Pi to facilitate the recording process by allowing users to input a name identifier and enabling seamless data transfer to a PC.

A sampling rate of 48 kHz and 32-bit depth were used for the acquisition of audio data. The subsequent section provides a detailed account of the recording methodology employed for audio data collection.

5.2.1 Dataset acquisition methodology

The data acquisition process involved inviting each musician into a room with a fairly low ambient noise level. A microphone was positioned approximately 50 cm in front of the trumpet bell and 10 cm from its longitudinal axis. In most instances, a set of two microphones was employed concurrently to ensure data redundancy, mitigating the risk of data loss should a device malfunction occur during the recording session.

Participants were instructed to play isolated tones of approximately one-second duration over

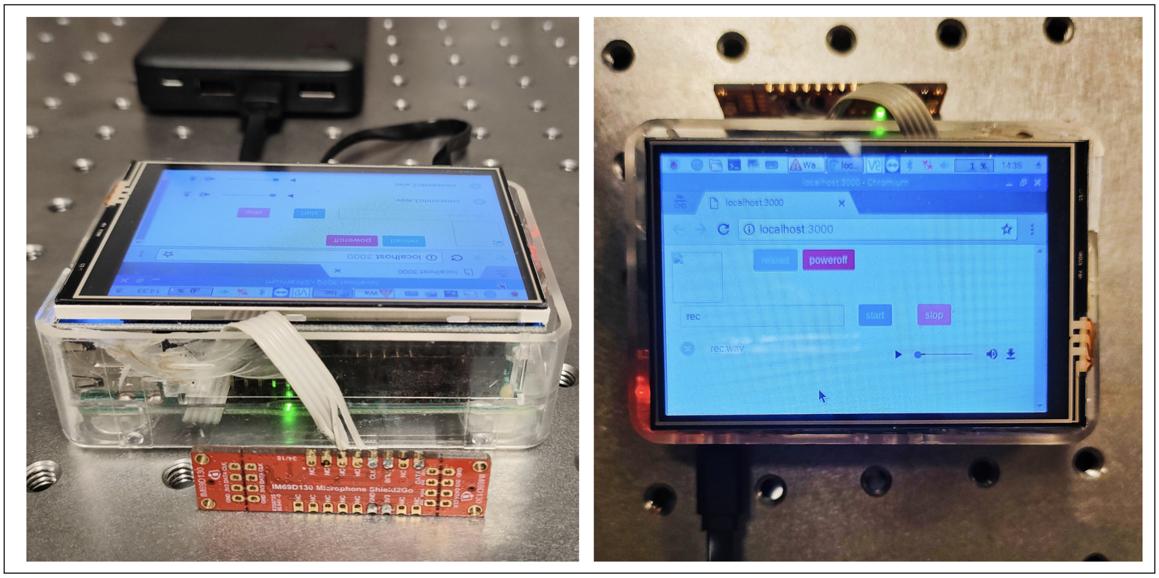


Figure 5.1: Assembled system for dataset recording. The figure on the left shows the microphone board connected to a Raspberry Pi and powered by a portable power bank, enabling flexibility for mobile recordings in different locations. The right side highlights the custom-built interface used for entering name identifiers and managing seamless data transfer to a PC.

a chromatic scale ranging from E3 to B♭5 at three distinct dynamic levels: *piano*, *mezzoforte*, and *forte*, in their preferred sequence. Musicians utilized their personal instruments and mouthpieces and were not required to adhere to a reference pitch (e.g., A4 at 440 Hz) as timbral quality concerning sound production efficiency is anticipated to be independent of a reference pitch.

The inclusion of various dynamic levels aimed to enhance the dataset's variability, as the timbre of brass instruments is significantly influenced by loudness (Luce and Clark, 1967). A digital sound level meter was positioned adjacent to the microphone, providing real-time decibel level readings during the recording. Trumpet players were given indicative reference levels of 85 dB, 105 dB, and 115 dB, corresponding to the *piano*, *mezzoforte*, and *forte* dynamic levels, respectively.

Despite the specified guidelines, the dataset exhibits several inherent variabilities:

- The sustain duration of the tones ranged from 0.7 to 4 seconds.
- The chromatic scale's range was contingent upon the performer's skill level. Generally, less proficient musicians struggled to produce tones in the high register, in which case they were instructed to play up to their highest achievable note.
- For beginner musicians, playing a chromatic scale in front of a microphone proved

challenging at times. Some participants opted to perform legato notes rather than separate tones.

- Less skilled musicians often experience difficulty in controlling the instrument's dynamic range, resulting in the recommended dynamic levels being primarily adhered to by more proficient players.

During the recording sessions, the candidate, who holds a degree in trumpet performance and has professional experience as a musician and instructor, assigned a preliminary grade of the overall sound production efficiency on a scale of 1 to 100 to each player. Figure 5.2 illustrates the distribution of assigned grades divided into four ranges (i.e., 0–25, 26–50, 51–75, and 75–100), demonstrating that a substantial number of players are represented in each category.

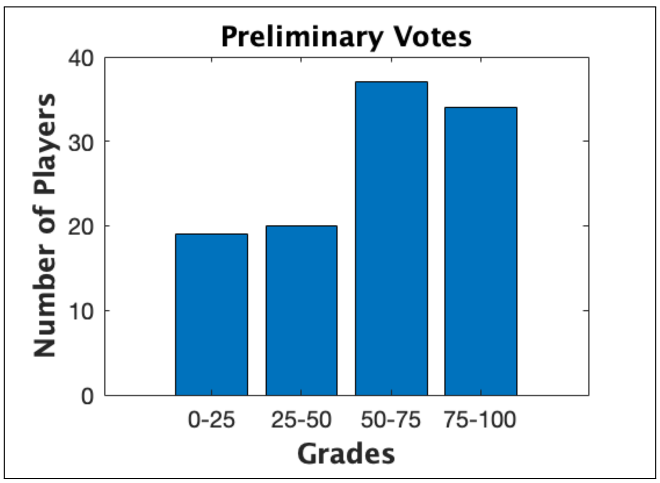


Figure 5.2: Distribution of recorded players according to the level of tone quality noted at the time of recording.

The dataset under examination was partitioned into discrete trumpet tones utilizing the *pyin vamp* plugin developed by Mauch and Dixon (2014), yielding a collection of over 19,000 tones. Although the segmentation process demonstrated a degree of inaccuracy, with certain audio segments containing noise rather than trumpet tones, it nevertheless provided a satisfactory initial categorization of the data.

The following section outlines the methodology employed to prepare the dataset for label assignment by chosen evaluators.

5.2.2 Dataset preparation

Considering the approximate accuracy of the segmentation algorithm and the extensive nature of the overall dataset, it was decided to select a representative subset of the dataset for the manual examination of audio samples. To ensure that the whole range of tone quality is sufficiently represented, the subset was constructed of seventeen trumpet players such that five individuals had received a preliminary vote between 0–25 and four individuals with a grade between the other three ranges 26–50, 51–75, and 76–100 respectively. The first category was assigned one player more as the less experienced participants only partially cover the required chromatic scale, thus compensating for the lower representation of tones within this class. The selected subset encompassed 1,712 distinct trumpet tones.

It was decided to classify each tone into four categories based on their sound production efficiency, resulting in four classification levels: 1:poor, 2:fair, 3:good, and 4:excellent. This classification into four levels was employed with the intention of simplifying the label assignment process while retaining sufficient variability, as suggested by Wesolowski (2012) and employed by Köktürk-Güzel et al. (2023) in a related research study.

The web interface shown in Figure 5.3 was subsequently developed to facilitate blind listening (i.e., without revealing the player's identity) and label assignment for each tone. The candidate listened to all 1,712 sounds in the subset under analysis through the interface and assigned a label to each tone. The "Not a note" button enabled tagging of erroneously segmented sound samples which were filtered out to yield a dataset 1,481 clean samples.

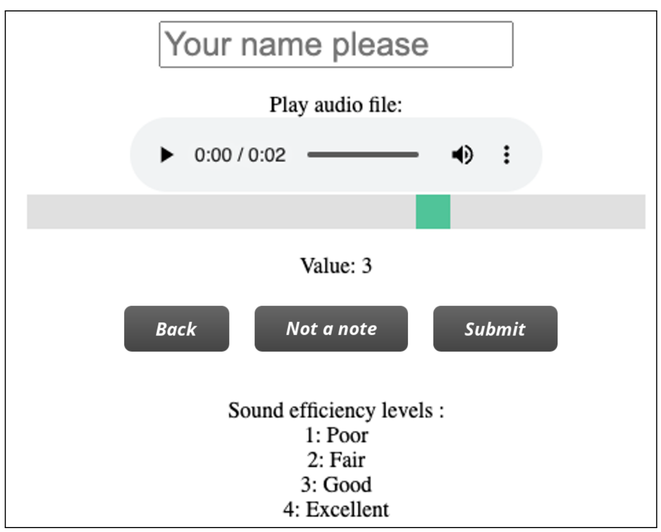


Figure 5.3: Interface for blind grading the trumpet tones.

The assignment of sound production efficiency class through anonymous listening to the audio samples in random order facilitated the allocation of a grade on a note-by-note basis, as opposed to providing an overall grade to the performance. This allowed for different grades to be assigned depending on the note if the level of sound efficiency varied along the chromatic scale. Additionally, the reliability of unbiased judgment could be assessed through a comparison with the preliminary grades assigned during the recording. The Spearman correlation coefficient between the two sets of grades was found to be 0.873 (P value$<$0.001), indicating the consistency of the candidate in assigning grades over time. This further indicates that players, in general, exhibit a consistent level of sound production efficiency along the chromatic scale.

5.2.3 Assessment labels

The cleaned dataset with 1,481 samples was subsequently presented to a panel of expert raters for evaluation via the described interface. A total of seven experts from different schools across Europe, North America and South America were chosen for the task. Among the raters, six were trumpet players, and one was a bass trombone player. All raters have professional experience as performers and/or teachers. This exploratory perceptual study was conducted online, with raters instructed to complete the task in a low-noise environment using professional headphones.

The rating sessions started with an introduction to the concept expressed by renowned brass instrument pedagogues, which asserts that rigidities in a trumpet player's body result in inefficiencies in playing, manifesting as a forced and strained sound. In contrast, a high-quality sound indicates efficiency of the embouchure and breathing muscles. Audio samples demonstrating extreme cases of this idea were presented and each rater confirmed their understanding of the concept and their ability to discern sound production efficiency in trumpet sounds based solely on audio information.

The dataset of 1,481 samples was split into two parts with 100 and 1,381 tones respectively. The raters first graded each of the 100 samples in approximately 15 minutes. After a 5 minute break, additional samples, randomly selected from the remaining 1,381 samples were presented for evaluation. The raters continued to assess the trumpet tones until they experienced fatigue or until 90 minutes had elapsed from the beginning of the experiment. Table 5.1 displays the number of audio samples rated by each grader. Grader 1 corresponds to the candidate who assigned the ratings manually by listening to all 1,481 samples in the subdataset, as described in the previous section. The set of 100 sounds was chosen such that they were equally distributed across the four classes, as determined from the labels by the candidate, and were used to ascertain the level of inter-rater reliability.

The next section describes the statistical analysis implemented on the data thus collected.

5.2.4 Data analysis

The inter-rater reliability was assessed using the subdataset containing 100 tones graded by all the experts. Table 5.2 presents the Spearman ρ correlation coefficients with the corresponding

Table 5.1: Number of individual tones evaluated by each grader.

Grader ID	**Graded Tones**
Grader 1	100 + 1381
Grader 2	100 + 401
Grader 3	100 + 206
Grader 4	100 + 312
Grader 5	100 + 383
Grader 6	100 + 366
Grader 7	100 + 564
Grader 8	100 + 491

P values for each pair of evaluators. As depicted in the table, all P values, representing the likelihood of obtaining the same results by chance, are less than 0.05.

Table 5.2: Spearman ρ correlation coefficients between each pair of graders.

Grader Pair	**Grader**						
	2	**3**	**4**	**5**	**6**	**7**	**8**
1	0.691*	0.668*	0.654*	0.645*	0.523*	0.638*	0.247***
2	-	0.701*	0.628*	0.650*	0.589*	0.650*	0.279**
3		-	0.599*	0.594*	0.496*	0.667*	0.237***
4			-	0.696*	0.650*	0.567*	0.349*
5				-	0.502*	0.637*	0.275**
6					-	0.524*	0.264**
7						-	0.353*

Legend: * $p<.001$, ** $p<.01$, *** $p<.05$.

The reported Spearman correlation coefficients range from 0.237 to 0.701. Notably, pairs including Grader 8 (the sole non-trumpet-playing expert) exhibited significantly lower correlation coefficients than all other pairs, potentially suggesting the significance of employing experts whose primary instrument aligns with the instrument under analysis for tasks of this nature. Due to the substantial differences in the ratings relative to the other raters, Grader 8 was deemed an outlier, and their results were excluded from further consideration. This adjustment increased Spearman ρ coefficients from 0.496 to 0.701, indicating fairly strong agreement among the judges (Williamon et al., 2021).

Subsequently, a confusion matrix was computed for each evaluator, comparing the ratings assigned by that specific grader to the most frequently occurring (i.e., statistical mode) value in the ratings assigned by the seven evaluators for that specific tone. Cases where the mode was uncertain on one value were eliminated, resulting in 87 overall tones. The first seven subplots of Figure 5.4 display the resulting confusion matrices for each grader and their respective accuracy values (average f1 scores).

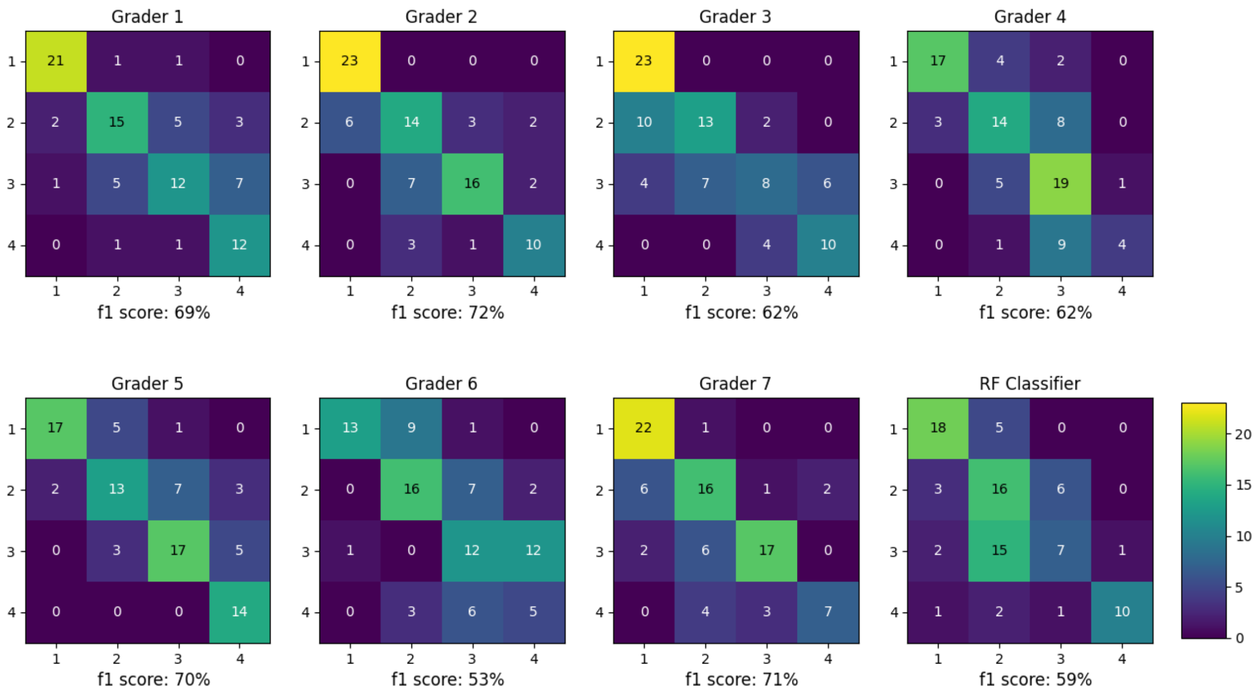


Figure 5.4: Confusion matrices with the predicted labels of each grader and of the trained RF classifier (horizontal axis) with respect to the true label as the mode of the assigned grade (vertical axis) and the corresponding f1 scores.

The confusion matrices indicate that experts generally agree on the task, with most discrepancies in assigned grades occurring between adjacent classes relative to the mode. Additionally, raters tended to agree more when evaluating poor-quality sounds, which may be because timbre-based errors in sound production efficiency are easier to identify objectively. In contrast, the assessment of higher-quality sounds could be more influenced by individual perception, making consensus less consistent.

The next section presents a model trained on the data obtained with reference to the variability of human assessment.

5.3 Methodology and Results

5.3.1 Audio preprocessing and model training

The dataset preparation process described in Section 5.2.2 yielded a clean dataset with the audio samples of 1,481 tones. As a preprocessing step, the sound samples were first normalized to have a maximum signal amplitude equal to one, assuming that loudness does not influence sound production efficiency. White noise at −60 dB was then added to the normalized audio to prevent numerical errors (division by zero) during feature extraction, without significantly altering the original signal. The audio features for each tone were then extracted using the Extractor algorithm from the Essentia library (Bogdanov et al., 2013). To reduce the computational complexity, only the statistical aggregates of the audio features (e.g., mean, variance, and mean of derivative) were utilized. Rhythm-based features were excluded since they were not deemed suitable for a timbre classification task. A total of 1,230 features were thus extracted to represent each audio sample.

As a first step, a Random Forest (RF) Classifier (Ho, 1995) was trained using the extracted audio features and labels provided by Grader 1, since Grader 1 had annotated each of 1,481 samples in the dataset. When the model was trained using the full set of audio features, a mean accuracy score of 78% was obtained in the 10-fold cross-validation. Using the model-based feature selection in scikit-learn, the top 256 features were identified from an RF-classifier model trained using a 75%-25% train-test split of the dataset. To improve performance, we experimented with different numbers of selected features and found that 256 yielded the highest accuracy, reaching 81.37% in 10-fold cross-validation.

To eliminate the bias introduced by using a single grader, it was assumed that the most frequent label given by the expert graders is the true label. Only samples with at least two votes were used and samples which had equal number of votes for two labels by the expert graders were assumed to be ambiguous and were discarded from the dataset. With this approach, 871 out of 1,381 samples were deemed unambiguous. Similarly, 87 out of the 100 samples were unambiguous. An accuracy score of 59% was obtained on the test set of 87 samples for the RF model trained using the 871 samples as training set. The confusion matrices on the test-set for the different graders and the RF classifier can be seen in the bottom right subplot of Figure 5.4. It can be observed that most of the confusion is between the adjacent classes. Since the audio samples in the adjacent classes are in fact more similar to each other than the other classes, the errors seem to be reasonable, for both the graders and the model. While an accuracy score of 59% appears low, it is within the range of accuracy scores (53%–72%) of the human expert graders and it demonstrates that the extracted audio features could be used to classify the audio samples based on timbre quality.

The trained model was tested in real time by trumpet players and on labeled datasets different from the one in this study (Bandiera et al., 2016) showing promising generalisability.

5.3.2 Feature importance

Due to a slightly subjective nature of the problem, there is considerable variability in the labels by human experts. Hence, very high classification accuracy scores cannot be achieved even with sophisticated machine learning models. However, even with a moderately accurate classifier, analysis of the most important features could help to develop an intuitive understanding of good quality timbre in trumpet sounds.

One of the main reasons to choose the Random Forest Classifier algorithm was that it gives access to the importance of each feature in the classification task. The feature importance scores for the classification are available as a model property in the scikit-learn implementation of the Random Forest algorithm (Pedregosa et al., 2011). The top 20 observed features are listed in Table 5.3.

Many of the top features are based on the mean of the derivative 'dmean' and the mean of the double derivative 'dmean2' of spectral properties, suggesting that the change in the spectrum across time is a crucial factor in the perception of the timbre quality. Notably, three of the top features – namely lowLevel.spectral_complexity.dmean, lowLevel.spectral_complexity.dmean2 and lowLevel.spectral_complexity.dvar – are related to the time varying properties of the same underlying feature of spectral complexity.

A scatter plot of the lowLevel.spectral_complexity.dmean and lowLevel.spectral_complexity.dmean2 features considering only the best and worst class samples is shown in Figure 5.5. It is apparent that just this pair of features is quite successful in discriminating between the best and worst samples. Since both features are statistical aggregates of the time-varying spectral complexity, the raw (i.e., frame-wise) spectral complexity values were analyzed to develop a visualization of sound production efficiency, as described in the following subsection.

5.3.3 Visualization based on spectral complexity

Spectral complexity is based on the number of peaks in the spectrum of a time window (Laurier et al., 2009). The Essentia implementation of this feature considers the spectral peaks only up to 5 kHz. From the spectra of the collected dataset, the presence of harmonic peaks at frequencies higher than 5 kHz was evident. It was therefore decided to implement the spectral complexity considering the entire audible frequency range. To enhance peak detection accuracy, prior knowledge of the fundamental frequency 'f0' of the tone was utilized

Table 5.3: Top 20 features ranked by importance in the Random Forest Classifier.

Audio Feature	**Score (%)**
lowLevel.spectral_complexity.dmean	1.381
lowLevel.scvalleys.mean_5	1.182
lowLevel.spectral_complexity.dmean2	1.049
lowLevel.spectral_complexity.dvar	0.897
lowLevel.sccoeffs.var_5	0.648
lowLevel.scvalleys.mean_3	0.636
lowLevel.sccoeffs.stdev_5	0.622
lowLevel.scvalleys.median_5	0.594
lowLevel.spectral_spread.dmean	0.570
sfx.tristimulus.dmean2_2	0.561
lowLevel.sccoeffs.median_4	0.531
lowLevel.sccoeffs.dmean2_3	0.496
lowLevel.scvalleys.median_3	0.492
lowLevel.barkbands.dmean_25	0.478
lowLevel.pitch_instantaneous_confidence.dmean2	0.465
lowLevel.spectral_flux.dmean	0.465
lowLevel.spectral_complexity.dvar2	0.425
lowLevel.sccoeffs.mean_4	0.424
lowLevel.scvalleys.mean_2	0.412
lowLevel.spectral_complexity.stdev	0.402

to search for spectral peaks exclusively in the vicinity of the integer multiples of the f0 frequency. For a normalized audio, peaks with signal energy less than -40 dB were discarded to reduce noise. An FFT-bin mask was generated by assigning the value of one to the FFT bin if a peak was detected in it while all other bins were assigned a value of zero, thus generating a visualization to track the peaks across the analysis time windows.

Figure 5.6 shows the visualization for two representative sounds. It is evident that for sounds rated as excellent quality, the spectral peaks consistently lie in the same FFT-bin across time, leading to flat horizontal lines in the visualization. On the other hand, for sounds rated as poor quality, the spectral peaks show unsteadiness, particularly at the higher harmonics, which leads to broken and wavy lines in the visualization. The total number of peaks could be more or less depending on the f0 frequency of the note and the loudness. However, our qualitative observations across multiple samples suggest that the perception of

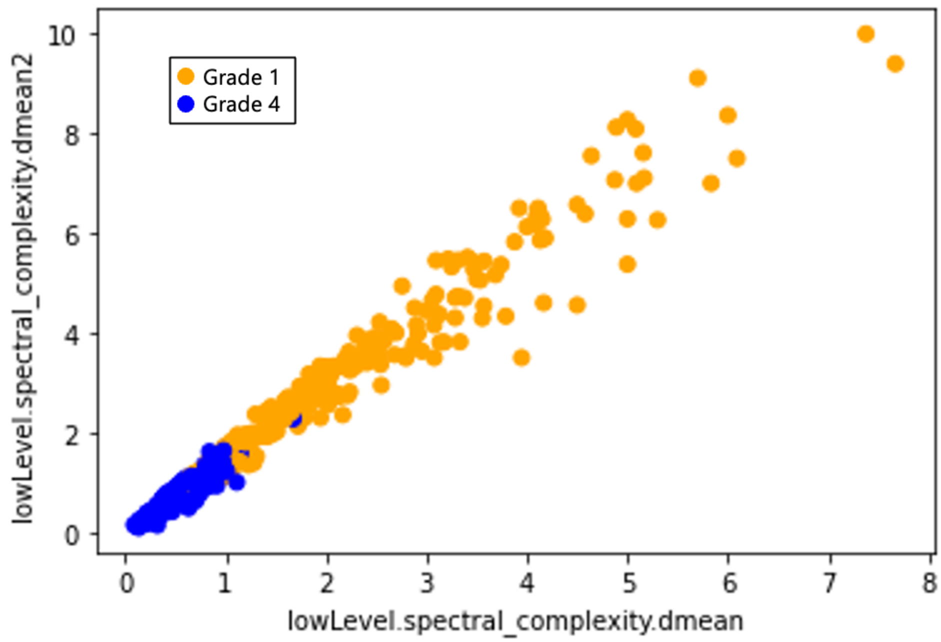


Figure 5.5: Scatter plot depicting the spectral complexity based features for best (blue) and worst (orange) class samples.

timbre quality may be more related to the steadiness of the peaks rather than their total number. While this trend was evident in manual analysis, a more systematic investigation is needed to quantitatively validate this relationship. Future studies could explore this aspect further by developing models trained on this feature or its aggregate versions. Additionally, a real-time implementation of this visualization could offer invaluable feedback on the efficiency of sound production, greatly benefiting new trumpet students who are still developing their auditory skills.

Building upon the educational exercise described in Section 3.2, this chapter introduces a machine learning component designed to provide real-time feedback on timbre quality. The model, developed and validated in the previous section, has been integrated into a Python-based interface that replicates the original structure of the educational app presented in Section 3.2. This enhancement aligns with principles of deliberate practice, a well-established framework introduced by Ericsson (1993) in music education that emphasizes structured, goal-oriented training supported by immediate feedback.

5.4 Proposed educational exercise

This section presents an interactive educational exercise designed to help trumpet learners refine their sound production efficiency through immediate, automated feedback. Building

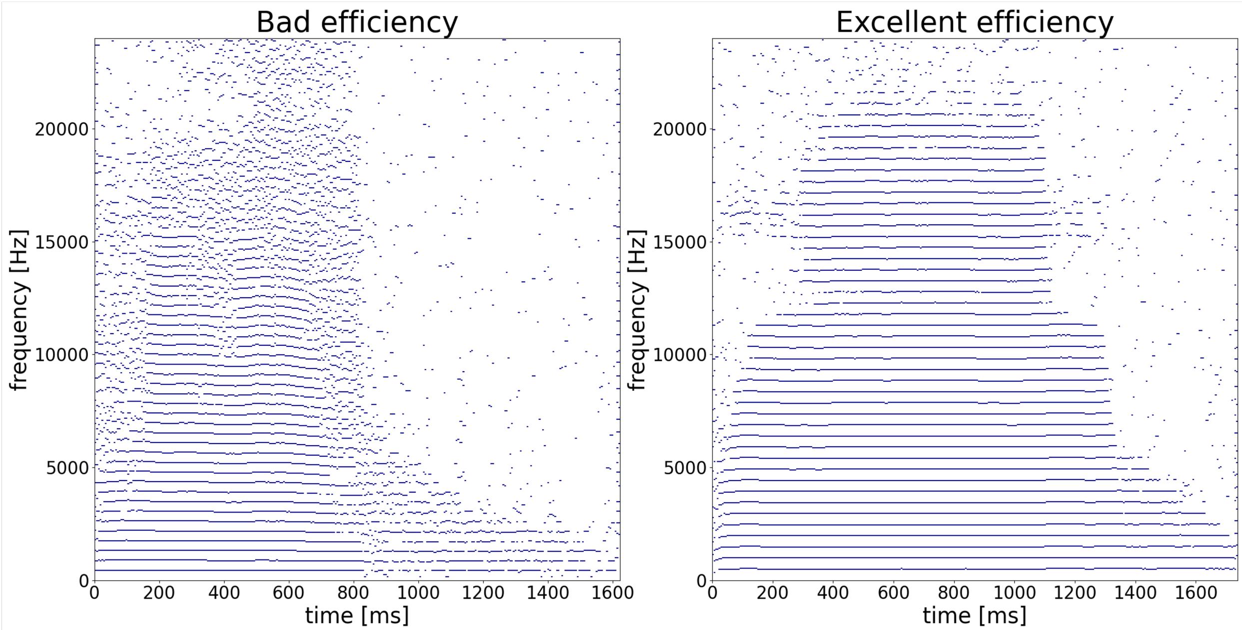


Figure 5.6: Visualization of the temporal evolution of spectral peaks for trumpet sounds rated as bad efficiency (left) and excellent efficiency (right).

upon the educational app for fingering introduced in Section 3.2, this version integrates the developed machine learning classifier as a modular component to evaluate timbre quality in real time, providing musicians with visual feedback on their sound efficiency.

The exercise follows the same structural principles as the fingering tool presented in Section 3.2, prompting the user to enter a metronome value and define a playable note range between E3 and B♭5. These values are displayed in the top row of the interface (Figure 5.7). Once set, the system generates a musical score of three four-quarter note measures in which a random pitch – within the selected range – is presented for practice. The first measure is a rest, the second measure displays the target pitch in gray while the software plays a high-quality reference recording, and the third measure shows the same pitch to be played by the musician. After completion, the system selects a new note, repeating the cycle for a given number of iterations.

To ensure proper timekeeping, the interface includes a metronomic "tic" and a synchronized numerical display of the beat count in the first column of the second row of Figure 5.7. The reference sound played by the software during the second measure was previously recorded by a professional trumpet player. It was chosen to use a recording of a fine performer, instead of simply using synthesized sounds, to induce the user to emulate a high-quality timbre by imitation. To enhance variability and prevent overfamiliarity, multiple recordings can be included for each note and randomly selected during playback. A demonstration video of the exercise is available at: `https://albertoacquilino.github.io/smac2023/`.

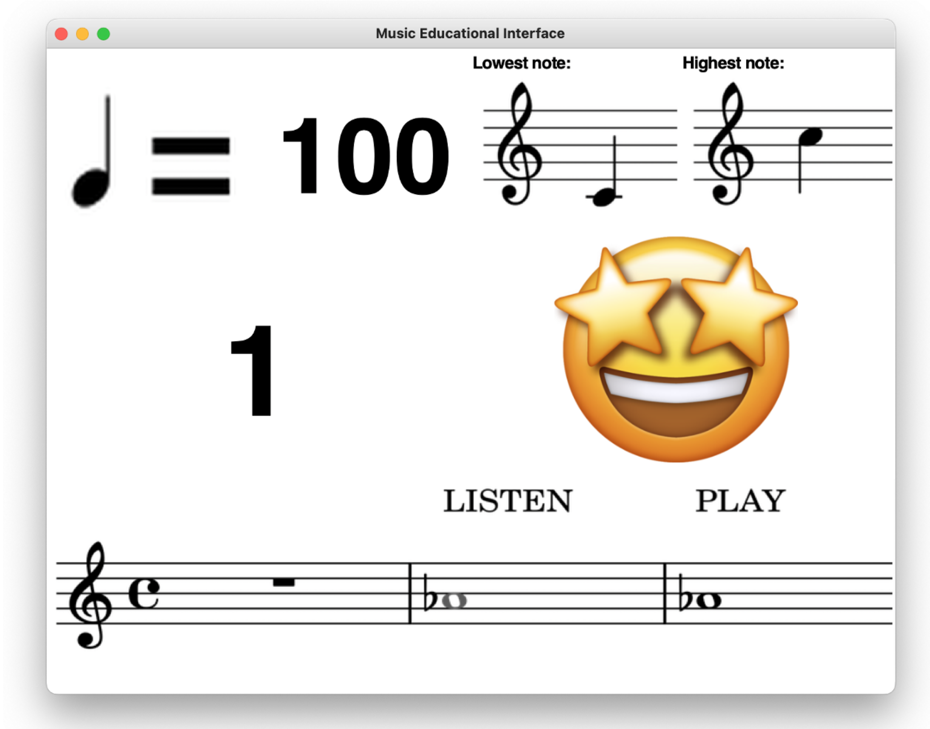


Figure 5.7: Interface of the proposed edTech system. Top row from left to right shows the input parameters: metronome value, lowest note, and highest note. The middle row shows the current beat value on the left and an emoji feedback on the quality of the timbre produced on the right. At the bottom, a three-measure score is shown cyclically updating with a random note to be played between the highest and lowest values selected.

The machine learning model, described in Section 5.3, is embedded in the software to provide near real-time feedback on the quality of tone produced by the student. The sound is analyzed by the algorithm which returns feedback corresponding to the class in which it is associated. This classification is visually represented by an emoji-based feedback system, displayed in the second row on the right side of the interface of Figure 5.7, where 🤔, 😊, 😃, and 🤩 correspond to increasing levels of timbre quality.

The exercise aligns with the principles of deliberate practice – as defined by Ericsson (1993) and Ericsson and Harwell (2019) – fostering the acquisition of high-quality timbre, which many brass pedagogues associate with an efficient sound production mechanism. The inclusion of the model within an interactive exercise with emoji feedback is intended to foster users' motivation. The possibility of defining the range of notes that can appear in the score ensures that the teacher can gradually approach trumpet students toward the high register. In fact, in the book *Brass Techniques and Pedagogy*, Weidner (2020) reports as a common problem among trumpet players the change in tone quality when playing in the high register during lip slurs. The author states that it can be counterproductive to try to play high notes with the trumpet without being able to play with good tone quality lower notes, pointing out as potential undesirable effects the development of bad habits involving excessive pressure, lip pinching, and poor air support. The exercise presented thus offers the teacher control over the students' note range development while also providing direct real-time feedback on the quality of tone played. This is intended to make the user focus on the development of this skill by stimulating deliberate practice while avoiding the aforementioned problems.

The flexibility of the developed system also allows for broader applications beyond this specific exercise. The timbre classification model could be embedded in stand-alone interfaces, such as chromatic tuners, to provide real-time visual feedback on sound production efficiency. Additionally, rather than relying on a random note selection process, instructors could define specific note sequences for students to practice, tailoring exercises to individual learning needs. These extensions further enhance the tool's adaptability, making it a versatile resource for both students and educators.

5.5 Suggestions for future studies

While the emoji-based feedback system employed in the proposed educational exercise could enhance engagement and motivation, its informativeness remains limited. The classification into four discrete categories restricts the ability of students to perceive subtle variations in their sound production efficiency. A more continuous feedback system – offering a clear, dynamic visualization of the student's timbral space – could allow learners to explore how slight adjustments in their playing technique impact their timbre. Such an approach would enable students to actively navigate and refine their sound production within a perceptual space, fostering deeper awareness of their playing mechanics.

Beyond benefiting students, this enhancement could also serve as a valuable tool for instructors, enabling them to provide more precise and targeted synchronous formative feedback. As discussed by Martin (2020), specificity in real-time feedback plays a crucial role in facilitating student growth by bridging the gap between current performance and desired learning outcomes. A refined feedback system, offering a continuous representation of timbre quality, could help teachers more effectively guide students toward optimal sound production while allowing students to independently experiment with and internalize these refinements.

To achieve such a system, future studies could explore the definition of a finite number of (psycho)acoustic dimensions that characterize sound production efficiency in trumpet performance. This could be accomplished through Principal Component Analysis (PCA) applied to the identified 256 most relevant audio features, reducing the dimensionality of the dataset while retaining critical information. Additionally, a combination of Partial Least

Squares Regression (PLSR) and nonlinear modeling via Artificial Neural Networks (ANNs) could be employed to analyze the statistical significance of various spectral, temporal, and spectro-temporal descriptors in relation to expert assessments.

5.6 Conclusions

This chapter explored the role of timbre quality in trumpet performance and pedagogy. With an aim to develop an automated tool for the assessment and visualization of trumpet tone quality, an extensive dataset of trumpet tones was collected and manually graded by expert evaluators. Analysis of inter-rater agreement demonstrated that, despite individual timbre preferences, experts generally concur in differentiating different levels of trumpet tone quality.

Random Forest Classifier models trained using extracted audio features were found to have accuracy scores comparable to the accuracy scores of human experts. Features based on spectral complexity were observed to have very high importance in the models trained for the task of trumpet timbre discrimination. A representation based on the harmonic peaks in the spectrum was developed to visualize the timbre quality. The proposed visualization suggests that the stability over time of spectral partials plays an important role in discriminating the timbre quality of trumpet sounds.

To support learning, an interactive educational tool was developed, integrating real-time feedback based on the trained model. This system provides musicians with an objective assessment of their timbre quality while allowing instructors to customize exercises to better guide students' technical development.

Beyond trumpet pedagogy, this work contributes to enriching music education by demonstrating how machine learning and visualization tools can bridge the gap between automated analysis and pedagogical application. The proposed system serves as a valuable resource for trumpet students, teachers, and researchers alike.

Author Contributions

Alberto Acquilino: Conceptualization of the study, dataset recording, organization and coordination of the labeling process with experts, development of the grading interface and educational software, writing of the original draft, and review and editing. **Ninad Puranik**: Conceptualization of the study, development of data processing and machine learning components, writing of the original draft, and review and editing. **Ichiro Fujinaga**: Supervision of the machine learning model, results processing, and review and editing. **Gary Scavone**: Supervision of the overall study, providing critical review and edits.

Ethics Statement

Ethical approval for the study, including consenting procedures, was granted by the Research Ethics Board Office of McGill University under REB File Number 430-0415, following the guidelines of the Canadian Tri-Council Policy Statement.

Chapter 6

Adaptive wavelet-based algorithm for measuring attack transients in music sounds

Articulation is one of the fundamental skills in learning to play musical instruments (Flesch and Mutter, 2008; Westphal, 1990; Weisberg, 2007). Yet, little attention has been dedicated to the development of educational technologies that could support its learning acquisition.

In this chapter, a new algorithm is presented for estimating the duration of the attack, with particular attention to wind and bowed string instruments. For these instruments, the quality of the tones is highly influenced by the clarity of the attack. Along with pitch stability, the duration of the attack is an indicator often used by teachers to assess sound quality by ear. However, the direct estimation of the attack duration from sounds is challenging due to the initial preponderance of excitation noise. To address this challenge, a more robust approach is proposed, based on the separation of the harmonic components from the excitation noise, achieved through an improved pitch-synchronized wavelet transform. In this process, a new parameter, the *noise ducking time*, is introduced to quantify the extent of the noise component during the attack.

The performance of the algorithm is evaluated using both existing sound databases and a newly created annotated dataset that includes a variety of problematic sounds. Additionally, the consistency and robustness of the duration estimates are assessed by applying the algorithm to sets of synthetic sounds with noisy attacks of known duration.

This chapter is based on the following research article:

G. Evangelista and A. Acquilino. An adaptive wavelet-based algorithm for assessing the quality of the attack transients in non-percussive instruments. In *Proceedings of the 28th International Conference on Digital Audio Effects (DAFx25)*, Ancona, Italy, 2025. A3Lab, Università Politecnica delle Marche.

6.1 Introduction

Methods for assessing the tone quality of instrumental sounds are desired in various contexts such as self-study and practice of musical instruments and searches in sound databases. Romani et al. (2015) proposed a model for the evaluation of the quality of single notes from trumpet, clarinet, and flute by analyzing five sound attributes: dynamic stability, pitch stability, timbre stability, timbre richness, and attack clarity. The characterization of the salient elements of the attack-transient could play a significant role in unassisted practice while learning to play musical instruments. The learners could check the quality of the tones they produce from the objective feedback parameters such as attack time and pitch stability.

For several musical instruments, such as wind and bowed strings, the attack phase is a very critical segment of the note that heavily influences the timbral and articulation aspects of the overall produced tone. It is the time interval in which, by exciting the right resonant modes, the noisy excitation gives way to a louder and possibly stable harmonic sound: the transition from pure chaos to ordered chaos. This chapter examines sound descriptors that are relevant to the automatic assessment of the clarity of the attack.

There is no common consensus on the definition of the attack boundaries. Luce and Clark (1965) defined this as the time from the onset of the note until the sound pressure level reaches 3 dB below the steady-state value. Their empirical approach provided a first framework for measuring attack durations in non-percussive orchestral instruments, accounting for variations in pitch, dynamics, and performer's style. However, as also noted by Peeters (2004), the tones produced by several instruments do not show clear attack, decay, sustain, and release phases. This implies that one cannot rely on the detection of a proper steady-state amplitude of the tones.

Measurements of the duration of the attack reported by Romani et al. (2015) made use of methods implemented in the Timbre Toolbox (Peeters et al., 2011). In its early releases (Kazazis et al., 2022), the duration of the attack is detected by finding the time interval from the onset of the note to the instant in which maximum amplitude, or a given percentage of it, is attained. To address the limitations of fixed threshold methods, Peeters (2004) introduced the *weakest-effort method*. This method, implemented as an option in subsequent releases of the Timbre Toolbox, uses highly smoothed versions of the signal envelopes. It computes the start and end of the attack based on adaptive thresholds estimated according to the behavior of the signal during the attack phase.

The noise present in the raw envelopes may lead to large errors in estimating the end of the attack phase. However, excessive smoothing of the envelopes results in estimated amplitudes that do not adhere tightly to the signal, which affects the detection of the duration of the attack (Nymoen et al., 2017). Experimental results indicated that both the direct envelope thresholding and the weakest-effort methods are unreliable for finding the attack durations of sounds with noisy attacks (see Section 6.4).

Hajda (1996) was one of the first researchers to propose a theoretical model that combines spectral and temporal information to better characterize the attack-transient, acknowledging the interplay between these two domains in the perception of musical sounds. This chapter introduces a methodology and algorithms for the accurate measurement of the attack duration in non-percussive instruments, based on a particular time-scale representation. The idea is to first extract two signals resulting from the separation of the harmonic components from the blowing or bowing noise, the mixing of which recovers the original signal. The amplitude envelopes of these two signals can be analyzed to detect relevant events.

Based on the envelope of the harmonic components, a new definition of attack-time is provided. Furthermore, based on the comparison of the harmonic and transient noise envelopes, the concept of *noise ducking time* is introduced as the time in which the attack noise decreases to a level below the harmonic components, which is useful to discern the quality of the attack.

Separating the noisy excitation from the resonant component of the sound also has a pedagogical benefit, since the student can be presented with acoustic feedback, which could be crucial to revealing, understanding, and correcting mistakes. Together with the pitch profile, the detected attack and noise ducking times provide sufficient time-frequency cues to determine the clarity of the tone produced.

In order to achieve an accurate separation of noise and harmonic components, a method based on the Pitch-Synchronous Wavelet Transform (PSWT) (Evangelista, 1993, 1994) is revisited, with improvements introduced through interpolation (upsampling) and period regularization based on local pitch-shifting.

This chapter is organized as follows. Section 6.2 examines possible application scenarios of the separation method to estimate the duration of the attack. Section 6.3 revisits wavelet concepts and multiplexing associated with the PSWT, highlighting improvements in the resolution of the transform based on signal upsampling and period regularization by interpolation. Section 6.4 details the algorithms for estimating the duration of the attack and of the noise ducking time, along with a consistency and robustness analysis based on synthetic sounds with known characteristics. Section 6.5 describes the acquisition of a dataset with quality annotations by an expert, used to validate the algorithm with instrumental sounds. Section 6.6 discusses the results obtained from the application of the proposed methods to the attack clarity of sounds in both the constructed datasets and other available databases. Section 6.7 presents the conclusions.

6.2 Case Study Description and Applications

This chapter focuses on determining descriptors of the attack transient in isolated monophonic tones produced by wind and bowed string instruments. The significance of this focus lies in the critical role that attack-transient characteristics play in the assessment of the quality of tones (Guettler and Askenfelt, 1997; Pàmies-Vilà et al., 2020). In pedagogical applications, the availability of an automatic objective quality evaluator is bound to help enhance the technical skills and expressive capabilities of the students. In the exploration of musical tones databases, a quality feature, such as the duration of the attack and the noise ducking time, can improve the search for the best match of the recorded tones.

Wind and bowed instruments exhibit attack transient durations that typically differ from each other (Luce and Clark, 1965). As described by Meyer and Hansen (2009), these durations depend on the physics of the resonator itself, which cannot react instantaneously to an excitation; rather, vibrations must gradually build up to reach their full amplitude. This phenomenon is related to the fact that part of the energy provided externally to the resonant system is radiated, while another part is absorbed by the instrument. The attack phase ends when an equilibrium is reached between the input energy and the total absorbed and radiated energy, allowing the oscillation to attain its steady-state condition.

Within certain limits, performers can influence the duration of the starting transient. Different types of articulation (e.g., staccato, détaché, martelé) are associated with varying rates of transient development, which musicians can utilize to make stylistic choices in their performances. Learning to control the type and duration of the attack thus becomes an important skill for instrumentalists, enabling them to select the appropriate attack style required by the performance context.

A crucial technical aspect that students should master to express a broad palette of musical ideas is to achieve an attack that is pure, i.e. uncontaminated by noise or unwanted frequency components, and accurate with respect to the desired pitch (Acquilino and Scavone, 2022). It is not uncommon for beginners to make articulation mistakes that produce sounds with excessively long attack transients, generally perceived as unpleasant. For wind instruments, these may include obstacles to the emission of airflow within the oral cavity, such as diction errors, suboptimal tongue positioning, incorrect jaw opening, or an overly constricted throat (Jacobs and Nelson, 2006). For bowed string instruments, errors can involve inadequate bow pressure on strings, irregular bow speed, incorrect bow angle, or uneven bow contact with strings, all of which can disrupt the sound production and lead to undesirable articulation (Guettler, 2010).

These considerations underline the need for a tool that provides a robust and consistent measure of the attack transient duration, much like how a chromatic tuner is essential for learning to play in tune, together with a measure of the noise extent. Such an educational system would offer teachers greater clarity and objectivity in their instructions and provide learners with objective means to verify their technique during individual practice sessions. For example, an instructor might indicate: "*For the next lesson, try to play the C4 note with an attack duration shorter than 40 ms with piano, mezzo-forte, and forte dynamics.*"

Previous studies have attempted to address this need. The results of the machine learning-based model developed by Romani et al. (2015) seem to hint that the features that are best related to the clarity of the attack of trumpet and clarinet tones were tonal descriptors deriving from the pitch, while a temporal property, the duration of the attack, scored the best for the case of flute tones. However, some mistakes, e.g. pitch instability during the attack, as blurred or breathy attacks, can be included more often than others in datasets. Thus, the distribution of various types of playing mistakes in the training dataset may well bias the final score. Moreover, the attack duration estimator proposed by Romani et al. (2015) may not be adequate for the analysis of noise-driven sounds, as demonstrated by the example in Figure 6.5 in Section 6.4 and by other examples provided at the following URL[1].

The next section introduces a new method for the estimation of attack characteristics based on transient/harmonic-components separation.

6.3 Noise + Harmonics Decomposition by means of PSWT

The excitation noise in wind or bowed instruments is wideband, while, when a steady tone is reached, most of the energy concentrates in narrow bands centered on harmonic frequencies. A simple idea to improve the attack duration estimators is to isolate the resonant signal from the noise. Intuitively, this can be realized by designing two comb filters, one peaking

[1] https://attackdurationestimator.github.io/DAFx25

on the harmonic frequencies and the other one notching these frequencies. This process inevitably introduces small gaps in the spectrum of the noisy component. However, for the intended purposes, this spectral alteration of the noise is not critical to the listening experience. Moreover, the algorithm for estimating the duration of the attack described in Section 6.4, requires detecting the time at which a full resonance develops, which happens at the end of the attack phase. To do so, after the note onset the amplitude envelopes of the separated signals are analyzed to determine when the harmonic component becomes more significant than the noisy component and when it reaches a percentage, e.g. -3 dB, of its maximum level.

While comb filters were the basic inspiration, the scheme based on wavelets revisited in this section offers many advantages. In the first place, being realized with multirate filter banks, it features a very efficient implementation of high-order comb filters. Moreover, the whole separation procedure is structured in a series expansion over a complete and orthogonal set of functions that does not introduce energy bias. As will be discussed, the bandwidth of the comb filters is controlled by the number of scales used in the wavelet transform. Next, fundamental wavelet concepts are revisited, along with an outline of a comb extension of wavelets.

The Wavelet Transform (WT) (Daubechies, 1992; Mallat, 2008) is a time-scale representation of signals which is equivalent to a time-frequency representation on a logarithmic frequency axis. It is mostly useful for separating transients at several time scales from the average behavior of signals. Properly sampling in the time-scale plane, one can arrive at a class of complete and orthogonal sets of wavelets in $L^2(\mathbb{R})$, which are suitable for the wavelet series expansion of any finite energy signal $s(t)$:

$$s(t) = \sum_{n=1}^{\infty} \sum_{m=-\infty}^{+\infty} a_{n,m} \psi_{n,m}(t), \tag{6.1}$$

where

$$a_{n,m} = < s, \psi_{n,m} > = \int_{-\infty}^{+\infty} s(t) \psi^*_{n,m}(t) dt \tag{6.2}$$

are the wavelet expansion coefficients and $<,>$ denotes the scalar product in $L^2(\mathbb{R})$.

In its canonical form, the wavelet decomposition achieves segregation of constant or nearly constant components from fluctuations from the constant behavior. This is realized by means of a generalized sum (average) and differences (innovations) encoding scheme based on Quadrature Mirror Filters (QMF) leading to band-pass wavelets with band allocation similar to that of a graphic equalizer (e.g., fractional octave bands). For the simplest case of octave band (dyadic) wavelets, one has:

$$\psi_{n,m}(t) = 2^{-n/2}\psi(2^{-n}t - m), \ n \in \mathbb{N}, \ m \in \mathbb{Z} \tag{6.3}$$

where $\psi(t) = \psi_{0,0}(t)$ is the *mother wavelet* and, due to their roles, the indices n and m are respectively called the *scale index* and the *time-shift index*.

In the construction of the wavelet sets one can show the existence of a low-pass function $\phi(t) \in L^2(\mathbb{R})$, called the *scaling function*, which in this context can be useful to express the residue of a scale-truncated wavelet expansion:

$$s(t) = s_f(t) + s_h(t) \tag{6.4}$$

where

$$s_f(t) = \sum_{n=1}^{N} \sum_{m=-\infty}^{+\infty} a_{n,m}\psi_{n,m}(t) \tag{6.5}$$

is the scale-truncated wavelet expansion and

$$s_h(t) = \sum_{k=-\infty}^{+\infty} b_{N,k}\phi_{N,k}(t) \tag{6.6}$$

is the *scaling residue*, where

$$b_{N,k} = < s, \phi_{N,k} > = \int_{-\infty}^{+\infty} s(t)\phi^*_{N,k}(t)dt \tag{6.7}$$

are the *scaling coefficients*, with

$$\phi_{N,k}(t) = 2^{-N/2}\phi(2^{-N}t - k), \ k \in \mathbb{Z}. \tag{6.8}$$

In the canonical wavelet expansion, the signal $s_h(t)$ in (6.6) represents the quasi-constant trend – nearly DC level or deep low-pass – while $s_f(t)$ in (6.5) represents the fluctuations from the quasi-constant behavior up to scale index N. However, while retaining the original form, the use of different wavelets will modify the methodological approach.

In fact, for the representation of pitched-tones it is certainly more useful to segregate the periodic or quasi-periodic components and the fluctuations from the periodic trend. Achieve this requires a modification of the wavelets, which actually results from a different computational scheme. For convenience, the discussion adopts a discrete-time framework for wavelets and signals, utilizing the scalar product in $\ell^2(\mathbb{Z})$.

If the time period P of the signal is constant, a winning idea is to arrange all periods in the columns of a matrix, as shown in Figure 6.1, and then compute a canonical wavelet transform along each of the rows (Evangelista, 1994). If the signal were perfectly periodic, then each column of the matrix would be identical, so that each row signal would be constant. Thus, the band-pass wavelets would not play a role in this case, leading to zero wavelet expansion coefficients. However, if the signal is not exactly periodic, then the wavelets will represent all deviations from the periodic behavior at several time scales.

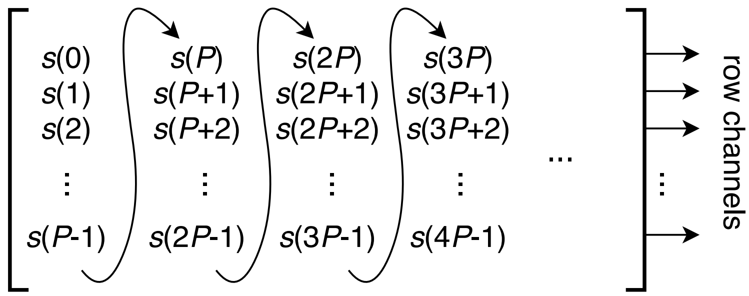


Figure 6.1: The construction of the matrix of the periods of a discrete-time pseudo-periodic signal $s(n)$ and the forming of the row channels to be represented by means of canonical wavelets.

From the flow point of view, the formation of the matrix is equivalent to demultiplexing the original signal to P channels corresponding to the rows of the matrix. Equivalently, one can define the multiplexed wavelets (Evangelista, 1994), which already incorporate the multiplexing operations and enjoy the same formal structure as in (6.1 - 6.8) but different physical interpretation. In fact, the DTFT $\hat{\Phi}(f)$ of the discrete-time multiplexed scaling function is related to the DTFT $\Phi(f)$ by P-fold shrinking:

$$\hat{\Phi}(f) = \Phi(Pf) \tag{6.9}$$

Since the scaling function of the canonical wavelets is low-pass, due to the periodicity of the DTFT the scaling function for the multiplexed wavelets is a comb covering the harmonics of the signal. Given the sampling rate f_s, the bandwidth BW_N^{tooth} of each tooth of the harmonic comb at scale level N is

$$BW_N^{tooth} = \frac{f_s}{2^N P} \tag{6.10}$$

which can become very narrow as N increases. As shown in Figure 6.2, the multiplexed wavelets are also comb-shaped with their teeth peaking on sets of sidebands of the harmonics. These sidebands become narrower and closer to the harmonics as the scale index n grows.

Scale-truncation of the multiplexed-wavelet expansion achieves the separation of the noisy excitation – the signal $s_f(t)$ in (6.5) – from the resonant part or harmonic trend – the signal

$s_h(t)$ in (6.6), which is required for the attack duration estimator and for the presentation of the acoustic feedback of the excitation for pedagogical purposes.

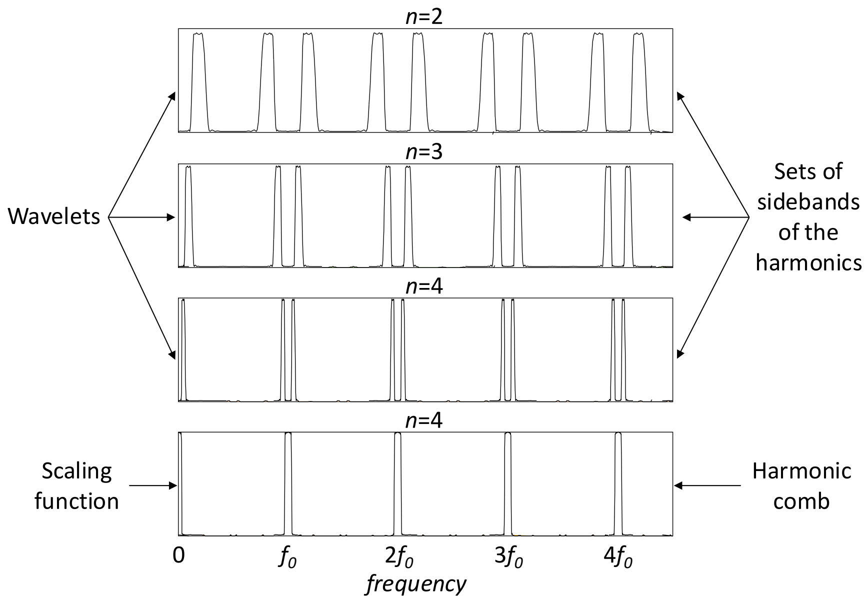


Figure 6.2: Magnitude Fourier Transforms of scaling function and comb wavelets at several scales.

Things become more complicated when the local period of the signal is not constant and one truly needs to compute a Pitch-Synchronous Wavelet Transform (PSWT) with variable pitch. Two main modes were presented by Evangelista (1993), in which the shorter periods were either zero- or constant-padded to form a matrix whose columns are the size of a pre-assigned maximum period. The extra samples are subsequently deleted in the final reconstruction.

6.3.1 Improvement of the PSWT

In the present work, a novel technique has been successfully tested, which is essentially based on time warping. Prior to multiplexed wavelet analysis, which can then be carried out with constant pitch, each period is stretched to a maximum period by using interpolation based on a polyphase antialiasing filter (Vaidyanathan, 1990). In the synthesis, the periods are decimated back to their original lengths, again with a polyphase anti-aliasing filter. Since the synthesis process separates the wavelet contribution (noisy component) from the scaling residue, period decimation has to be performed separately on these two signals.

Another improvement to the PSWT based method carried out in this study is the up-sampling of the signal prior to pitch detection and wavelet analysis. Since pitch detection is based on a sliding-window autocorrelation method, the estimated period is an integer approximation of the true period. Up-sampling has a mitigating effect on the quantization of the period estimate, which makes the separation of the noisy components of the signal from the resonant part much more accurate. In fact, in the frequency domain, the scaling function forms a comb tuned to the pitch of the tone that is supposed to trap all the harmonics. In case of mistuning, the higher harmonics could fall out of the harmonic comb and end up in the territory of the wavelets, i.e. in sidebands of the harmonics, thus contributing to the fluctuations component, which is an undesired behavior.

In general, the number of scales N at which one truncates the wavelet analysis is also limited by mistuning: at lower N the teeth of the comb are less narrow so they are more keen to cover the harmonics, but this also means that more energy from the noisy component would be covered by the scaling residue and not by the wavelets. Therefore, more accurate tuning achieves deeper analysis and better segregation of the components.

Up-sampling the signal by factor 10 adds a decimal point to the resolution of the period estimate. Polyphase filter interpolation was again used for this process. Effective segregation was achieved by pushing the number of scales to 4-5 for most sounds in the dataset and in other public databases.

It must be pointed out that, while vibrato can and must be tolerated, erratic pitch variations as in the sounds typically produced by beginners are considered to be mistakes which, besides being detected by the pitch instability indicator, can be heard in the acoustic presentation of the noisy signal sound. Since a pitch detection and tracking module is embedded in the proposed PSWT-based attack duration estimator, pitch instability measures, such as standard deviation (STD median or Tukey (Peeters et al., 2011)) are easily computed; these features are used to complement the attack duration estimate to determine the clarity of the attacks.

The pitch estimation used in conjunction with the PSWT is period synchronous, where a detection frequency range is preset. A window of length equal to 2-3 maximum periods is sliding on the signal by an amount equal to the last detected period. When no pitch is detected, the maximum frequency, corresponding to the minimum period, is output as the "pitch" of the current signal segment. Therefore, pure noise samples are arranged in short segments that are then stretched by interpolation to maximum period length P, before ending up in columns of the demultiplexing matrix of Figure 6.1. However, since the samples of adjacent noise segments greatly differ from each other and from subsequent pitched periods of the signal, they are mostly picked up by the row-channel wavelets and do not contribute to the row-channel scaling residue, which is in line with the intended separation idea.

A block diagram of the complete procedure to extract the noisy and resonant parts of the signal is shown in Figure 6.3. Sound examples and further plots and tables of noise-resonance segregation in various tones from orchestral instruments and synthetic sounds can be found

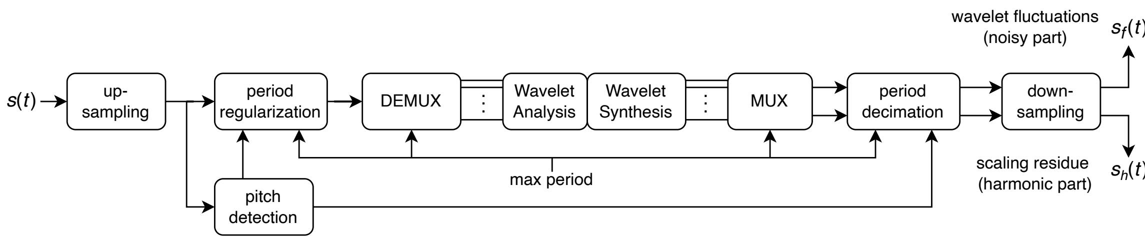


Figure 6.3: Block diagram of the PSWT-based separation of noisy and resonant parts.

at the provided URL[1].

6.3.2 Complexity

The multiplexed wavelet analysis-synthesis block has linear complexity in terms of the number of samples and, in principle, can be computed in real time using FIR QMF filters. Clearly, up-sampling and period interpolation increase the complexity factor and introduce further latency. In the foreseen applications, either as feedback for the student musician or in database quality indexing, real-time is not a strict requirement. The off-line interpreted Matlab implementation running on a basic M2 ARM CPU laptop with 8GB RAM, including signal up-sampling/down-sampling factor of 4, autocorrelation-based pitch detection, pitch regularization, and the computation of 5 multiplexed analysis/synthesis wavelet scales rooted on order 9 Daubechies' QMF filters (Daubechies, 1992), runs slightly faster than real-time, within a time factor of 0.875. The system lags behind real-time when the up-sampling/down-sampling factor is increased. For reference, when increasing this factor to 10 the computation in Matlab requires double the time required by real-time.

6.4 Attack Duration Estimation Algorithm

The accurate estimation of the duration of the attack transients in musical sounds, particularly for noise-driven harmonic instruments such as wind and bowed string instruments, presents a significant challenge due to the presence of the excitation noise, which, in some cases, overshoots above the steady-state oscillation. For these instruments, the sound can be characterized as pseudo-periodic, exhibiting a clear harmonic structure only once the note is fully developed.

This section proposes a novel algorithm that leverages the pseudo-periodic nature of harmonic instruments to reliably estimate the duration of the attack transient. The suggested

[1] https://attackdurationestimator.github.io/DAFx25

algorithm is designed for isolated monophonic pseudo-harmonic sounds, which are characterized by three distinct phases: an initial silence, a noisy attack transient, and the steady-state fully developed sound. The estimation of the attack duration can take great advantage of the separation of the noisy component from the harmonics, which simplifies the detection of the onset of the periodic behavior, i.e. the end of the attack phase. The proposed method relies on the signal separation based on the PSWT and its improvements described in Section 6.3.

The amplitude envelopes of the two signals are detected, $e_h(t)$ for the harmonic content of the sound and $e_f(t)$ for the noisy fluctuations associated with the attack transient. Various methods were tested to extract the envelopes and to interpolate them and it was found that the classical sliding-window maximum method with linear or spline interpolation gives the best results for its adherence to the signal dynamics.

The onset time t_{on} of the tone is identified as the instant when the amplitude of the input signal level lies for the first time above a threshold A_{thr}. The minimum useful threshold level depends on the SNR of the recording and is estimated, with a margin, from the recording of the silence preceding the note. Depending on the instrument and play mode, during the attack transient the amplitude associated with the fluctuations can be significantly higher than that of the harmonic signal.

At the end of the attack phase, the amplitude envelope of the harmonic signals attains higher levels. The attack offset t_{off} is detected when the level of the harmonic signal reaches a fraction α of the maximum value of the envelope of the harmonic components $e_h(t)$. In other words, given the envelope $e_h(t)$ of the harmonic part $s_h(t)$ and the input signal $s_{in}(t)$, the following definitions apply:

$$\begin{aligned} t_{on} &= \min_t \{t : |s_{in}(t)| > A_{thr}\} \\ t_{off} &= \min_t \{t : e_h(t) \geq \alpha r\} \\ t_{dW} &= t_{off} - t_{on} \end{aligned} \tag{6.11}$$

where $r = \max_t e_h(t)$ and t_{dW} is the wavelet-based estimation of the attack duration time. A block diagram showing the computation flow for the estimate of the attack duration is shown in Figure 6.4. In most experiments, α was set to $10^{-3/20} \approx 70.8\%$ which yields a 3 dB attenuation, but, in specific applications, α can be considered as a free calibration parameter.

An example of attack duration estimate that illustrates the advantage of applying the proposed PSWT-based method is shown in Figure 6.5. There, a trumpet sound signal is plotted in which the initial excitation noise peak is higher than the steady-state amplitude. The classical method, also included in the early versions of Timbre Toolbox (Kazazis et al., 2022), defines the attack time as the duration of the interval from t_{on} until the instant in

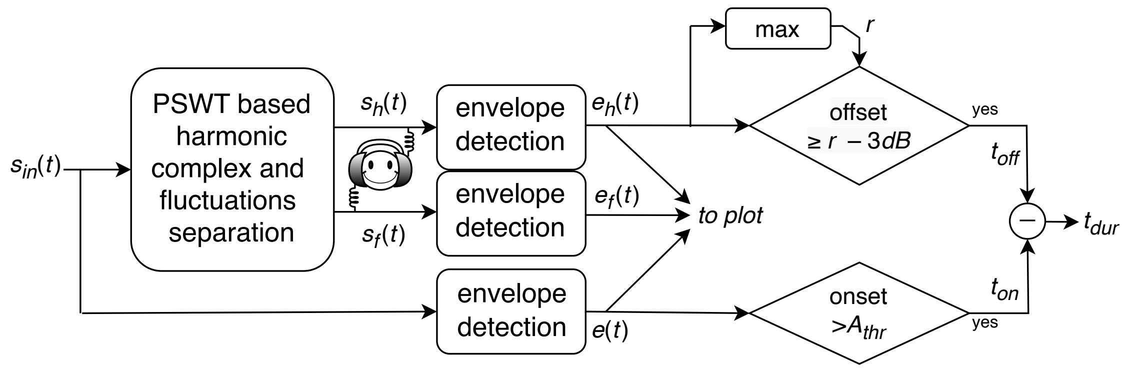


Figure 6.4: Block diagram of the attack duration estimator by means of PSWT based separation.

which the maximum amplitude is reached. This criterion clearly fails for the signal in Figure 6.5 since the maximum of the signal envelope occurs at the very beginning and is purely due to excitation noise during the attack phase.

In Figure 6.5, superimposed on the input signal are the estimates $e_h(t)$ and $e_f(t)$ of the envelopes for the harmonic and noisy components, respectively. The envelope $e_h(t)$ correctly ignores the initial noisy transient and reaches the maximum roughly when the steady-state part of the sound begins. By thresholding $e_h(t)$, the duration of the attack was correctly estimated at 180 ms, which makes more sense than the estimates for t_{off} provided by thresholding the original envelope and by means of the weakest-effort method, both of which occur when the attack is still in the noisy transitory part.

The detected t_{off} is only slightly larger than that of typical well-rated attacks (≤160 ms). However, an additional quantity derived from the separated signals can help in assessing the clarity of the attack: the *noise ducking time* t_{nd}. This is defined as the instant at which the initial attack noise starts to be overtaken by the harmonic components. A strategy for evaluating t_{nd} that works for a large class of tones is to track the most prominent local maximum of $e_f(t)$ and then find the first subsequent instant where $e_f(t)$ falls below $e_h(t)$ by a prescribed amount, set to -3 dB in the experiments.

In some sounds with badly rated attacks, such as growling wind sounds or string tones played with wrong bow pressure, the initial attack is actually short and clean, but a noisy phase is initiated immediately after it. In such cases, the value of t_{off} is not decisive, but large t_{nd} allows us to detect a prolonged noisy activity. The t_{nd} detected for the signal in Figure 6.5 is 89 ms, whereas clean attacks show much shorter noise ducking times (≈ 20 ms). Further examples of attack analysis can be found in the provided URL[1].

In order to test the consistency and robustness of the estimation algorithm for t_{nd}, statistical

[1]https://attackdurationestimator.github.io/DAFx25

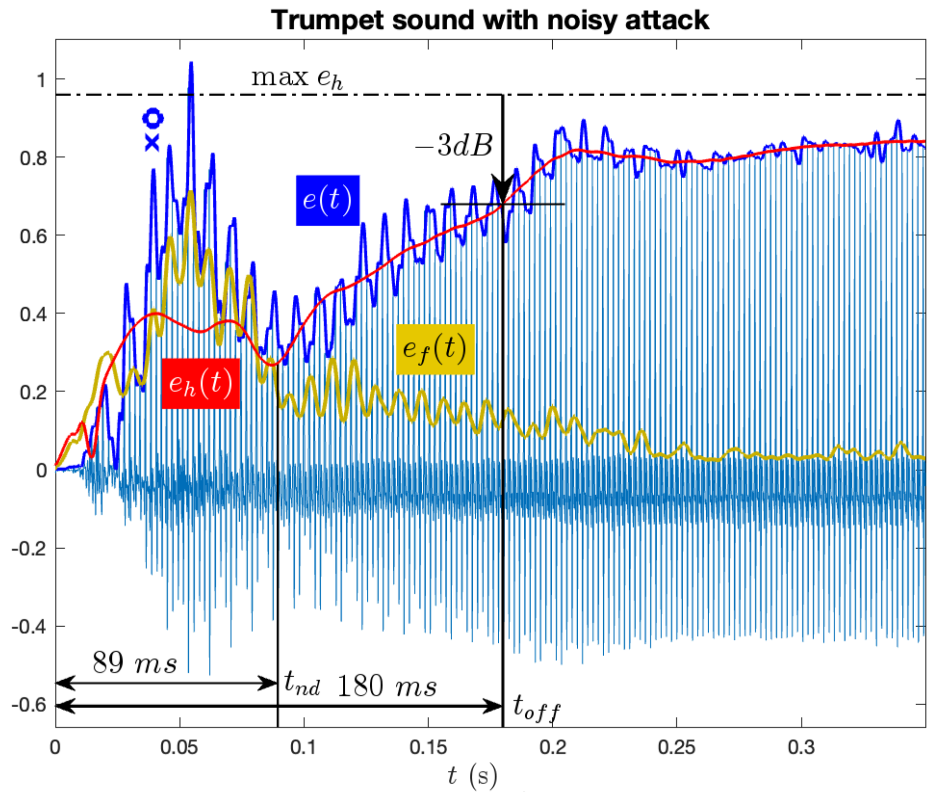


Figure 6.5: Trumpet sound signal, time-shifted so that $t_{on} = 0$. The estimate of the attack duration t_{off} obtained by thresholding the envelope $e_h(t)$ (red curve) of the harmonic part is shown by a thick black vertical line. The estimates of t_{off} obtained by means of the threshold (-3 dB below $\max_t e(t)$) and the weakest-effort methods are also shown, respectively, by means of a cross (x) and a circle (o) mark above the input signal envelope $e(t)$. The envelope $e_f(t)$ (yellow curve) of the fluctuation components is also shown, which peaks right after the onset of the signal. An estimate of the noise ducking time t_{nd} is shown, which delimits the end of the noisy part of the attack.

trials were conducted using synthetic sounds in which the noise envelope is a short, linearly or exponentially decaying pulse, which overlaps with the wave envelope. Across all synthesizers and various programmed t_{nd}, the relative standard deviation (STD) – the ratio of the STD and the mean – was less than 10%, where the mean remained within a few milliseconds from the programmed t_{nd} value.

6.4.1 Consistency and Robustness

In order to test the consistency and robustness of the PSWT-based estimate of the attack duration time and compare the proposed method with existing ones, synthetic signals were generated and analyzed. These included sinusoids, band-limited square, sawtooth, and triangular waves together with a trumpet-like sound reconstructed from its 10 strongest Fourier series coefficients, which occurred at harmonic frequencies of the fundamental. All sounds were corrupted by time-enveloped Gaussian random noise. In order to simplify the analysis, trapezoidal envelope shapes were used for both noise and signals, where the envelope of the noise largely covers the attack phase of the signal. In the tests, several Signal-to-Noise Ratios (SNR) were set, defined as $20\log_{10}$ of the amplitude ratio between signal and noise levels in the flat and overlapping part of the envelope. Additionally, the possibility to introduce vibrato was included by applying frequency-modulation to the waves.

Given that the imposed envelopes are programmed, it is easy to assess the duration of the attack phase when noise is not present. In order to estimate statistics, the estimation algorithms – PSWT-based, input envelope thresholding, and weakest-effort methods – were applied to sets of 100 test sounds of the same wave type but corrupted by different samples of statistically independent noise. For each method, a histogram was plotted and the sample mean μ and standard deviation σ of t_{off} were extracted, as illustrated in Figure 6.6. Due to the way it is defined, it is natural that the average estimates of the duration time obtained by the weakest-effort method may differ from those of the other two methods, but its high standard deviation is of concern, which shows that the method is not very robust; as such, this method will not be considered in further analyses.

Example behaviors of the standard deviation σ_{dW} as SNR grows for the wavelet-based method and $\sigma_d T$ for the input envelope thresholding are reported in Figure 6.7. It can be seen that the standard deviation of the PSWT-based method, which exponentially decreases as the SNR increases, is always much smaller than that of the classical method. Further examples of estimate statistics using synthetic sounds with and without vibrato can be found in the provided URL[1].

6.5 Dataset

To evaluate the use of the proposed algorithm for the analysis of instrumental sounds, a specialized dataset of isolated monophonic trumpet tones was recorded using high-end audio equipment in a soundproof booth to minimize ambient noise and external interference.

The testing dataset allows for precise control over recording conditions and articulation variations, providing a solid foundation for the development of algorithms computing tone-quality features. The recordings were conducted in a soundproof booth to minimize ambient noise and external interferences. A Behringer ECM8000 condenser microphone was used, connected to a Focusrite Scarlett 2i2 USB audio interface. The microphone was positioned 50

[1] https://attackdurationestimator.github.io/DAFx25

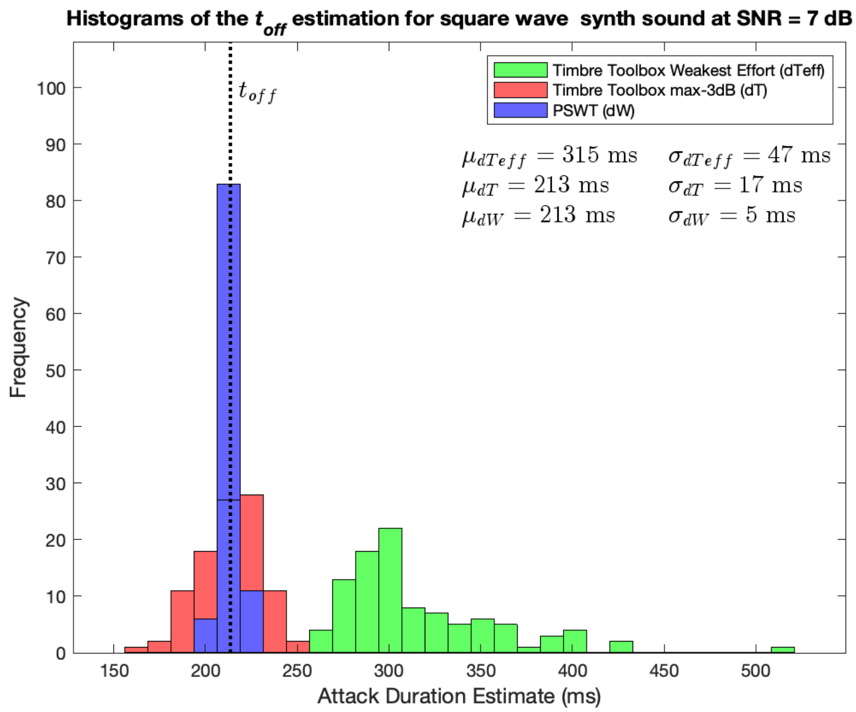


Figure 6.6: Histograms of the estimates of the duration of the attack of a noise-corrupted square wave synthetic sound, where dW denotes the estimate using the PSWT-based method, dT the estimate by means of input signal envelope thresholding, and dTeff by means of the weakest-effort method.

cm in front of the trumpet's bell, aligned at the same height and facing the instrument. This placement remained constant throughout the sessions to ensure consistency in the recordings. Audio was captured at a sampling rate of 48 kHz with a 16-bit depth.

The performer was a professional musician with a degree in music performance and a professional background in music education. The musician played isolated tones across the primary range of the trumpet, specifically targeting the notes B♭3, D4, F4, B♭4, D5, and F5. These pitches were chosen to cover a representative spectrum of the instrument's range. For each selected note, the performer was instructed to play multiple tones, exhibiting both good and poor attack clarity. The poor attack-clarity sounds were intended to simulate common articulation errors made by novice players. No specific dynamic levels were imposed. After recording, the musician provided annotations for each selected sample, focusing on four main characteristics associated with poor attack-clarity:

- *Noisy attack*: The attack contains noticeable noise, perceived as a prolonged crack-like sound at the onset.

- *Delayed stabilization of pitch*: The onset begins on a different harmonic than the

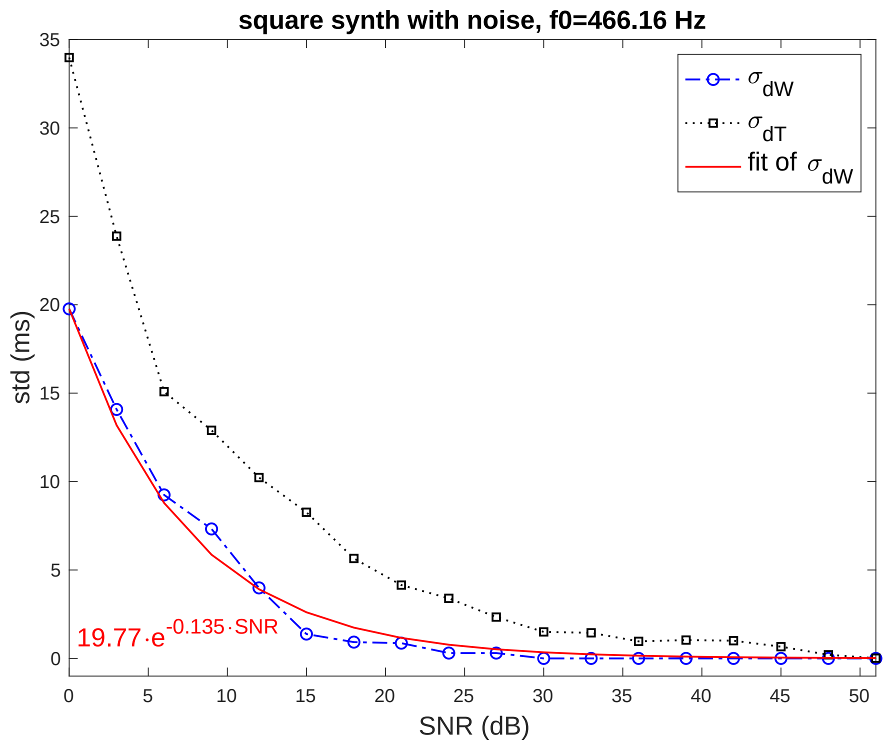


Figure 6.7: Behavior of the standard deviations σ_{dW} and σ_{dT} for the duration of the attack based on wavelets and on input envelope, respectively, using square wave synthesis and noise at increasing SNR. The curve described by a decreasing exponential fit of σ_{dW} is also plotted.

intended pitch before settling into the intended note, resulting in distinct transient sounds depending on whether the onset starts on a higher or lower harmonic. This issue is more prevalent in the high register of the trumpet because, although harmonic frequencies are equally spaced, they correspond to notes that are closer together on the musical scale in this range, requiring higher onset pitch precision.

- *Delayed stabilization of resonance*: The sound starts muffled and unsteady before reaching a more resonant timbre, creating a characteristic "ti-OH" effect at the attack discussed in the pedagogical literature (Jacobs and Nelson, 2006).

- *Attack with breath noise*: Despite tonguing, the sound does not start immediately; instead, there is an audible breathing noise as air passes through the instrument before the vibration begins uncontrollably late.

The dataset – available in the companion website[1] – comprises 149 labeled sounds, each annotated according to the identified attributes. It is important to note that individual recordings may exhibit more than one of these characteristics simultaneously. Although limited

[1] https://attackdurationestimator.github.io/DAFx25

in size, the dataset was instrumental in the development of the attack duration estimation algorithm described in Section 6.4, especially useful to attribute a physical meaning to the attack phase, which is absent from other definitions devised for generic signals.

6.6 Results and discussions

This section analyzes the performance of the proposed attack duration estimation method across different types of attack transients of the collected trumpet dataset. The results are discussed in terms of the estimated attack duration and noise ducking time, as these are found to be the salient features.

Good attacks generally show a harmonic envelope that increases quite rapidly and linearly until it reaches a flatter region, as illustrated in Figure 6.8. The onset of the note is characterized by a short peak of the envelope of fluctuations, likely due to the tongued attack, before the envelope then decays to lower levels. The estimated t_{nd} is very small as the harmonic envelope soon prevails. Depending on the slope of the attack, the estimate of t_{off} can reach a range of values that are generally smaller than in attacks of lower clarity.

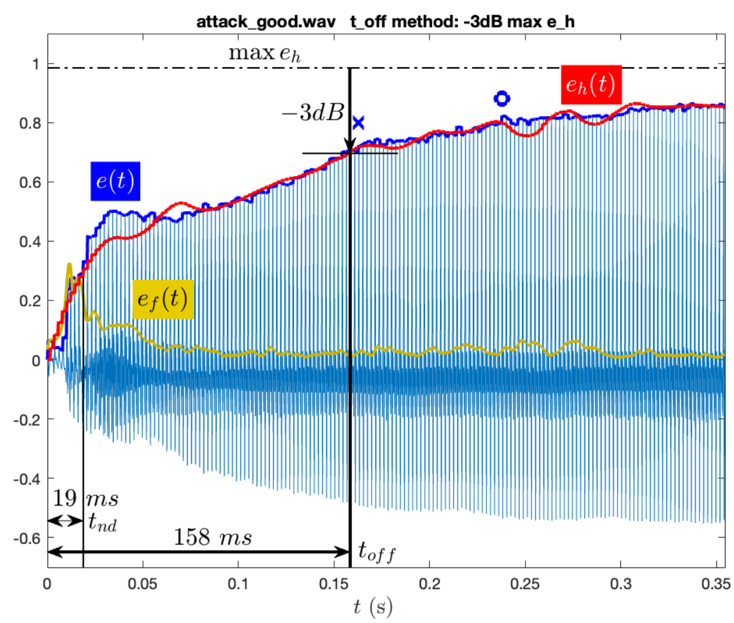
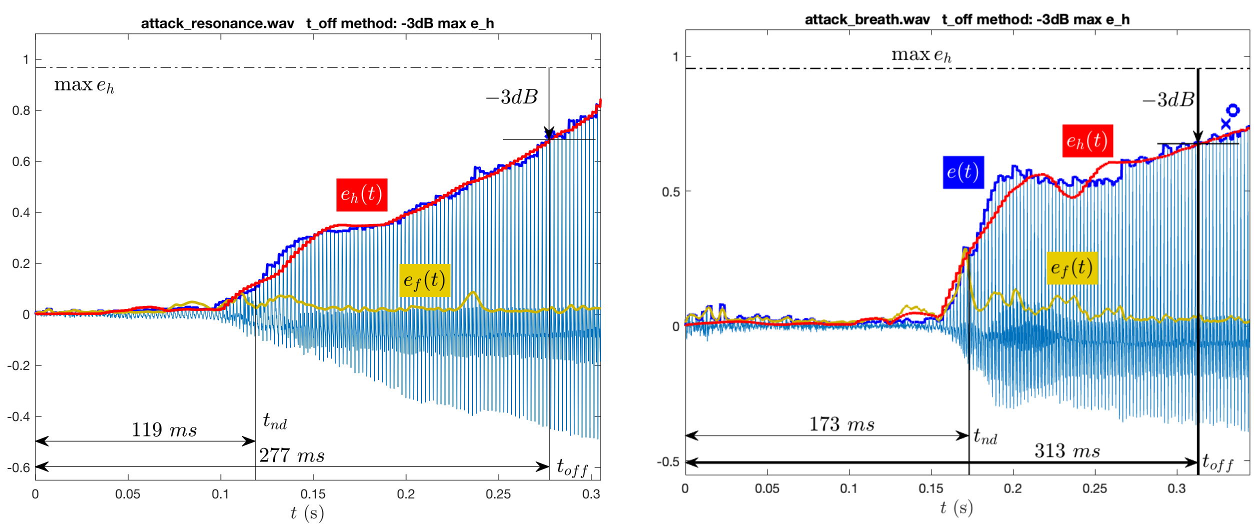


Figure 6.8: Trumpet sound signal with a good attack, time-shifted so that $t_{on} = 0$. The line colors and labels are the same as in Figure 6.5.

In *noisy attacks*, the estimated t_{off} is generally only slightly higher than in cleaner articulations. An example is shown in Figure 6.5. Here, t_{nd} emerges as the main discriminant, showing values significantly higher than in cleaner articulations. As illustrated in Figure 6.5, the separation of the noisy excitation from the harmonic components prevents false detections of the termination of the attack, which are induced by transient noise peaks in the original signal envelope. This is a net improvement over the Timbre Toolbox detection methods.

In *attacks with delayed stabilization of resonance* (Figure 6.9a), the harmonic envelope exhibits an initial lower amplitude, increasing until the sound stabilizes. The algorithm accurately reflects this transition, associating these cases with larger estimates of both t_{off} and t_{nd}. Since the transient development occurs within a much shorter time scale, this behavior is distinct from a deliberate crescendo.

In *attacks with breath noise* (Figure 6.9b), an onset detection algorithm based on a dynamic threshold estimated during silence ensures that the breath noise is not misclassified as background noise. Both t_{off} and t_{nd} are generally higher in this case.



(a) Attack with delayed stabilization of resonance

(b) Attack with breath noise

Figure 6.9: Trumpet sound signals with (a) a delayed stabilization of resonance and (b) a breathy attack. The sound signal is time-shifted so that $t_{on} = 0$. The line colors and labels are the same as in Figure 6.5.

Additional analysis material for these attack types is provided in the companion website[1].

Since the harmonic envelope may temporarily stabilize on an unintended pitch, the *attacks with delayed stabilization of pitch* do not necessarily correspond to larger t_{off} and/or t_{nd}. This type of attack error is perceptually salient and can be easily identified by observing the large standard deviation of pitch.

The developed algorithms were also tested on a wider set of recordings using the Good-sounds dataset (Bandiera et al., 2016), which includes a collection of isolated tones of wind

[1] https://attackdurationestimator.github.io/DAFx25

and bowed string instruments with a substantial number of partially annotated examples of correctly played notes and notes with attack errors. Among the annotated errors, some instances included brief descriptions of the type of attack issue, while others were generically labeled as “bad attack”.

Unfortunately, a systematic classification of attack mistakes across different instruments could not be found in the literature. However, based on the temporal and spectral evolution of the noise and harmonic components rather than the physical mechanism of sound production itself, it is suggested that the classification developed for trumpet attacks could be extended to other instruments. For example, a violin sound with a noisy attack due to incorrect bow pressure (see Figure 6.13b) is physically distinct from a noisy attack of a trumpet caused by improper embouchure articulation. However, both exhibit similar behavior in terms of the interaction between the noise and the harmonic components, such as an initial broadband onset followed by a delayed emergence of stable pitch. These shared acoustic patterns suggest that perceptual categorization may generalize across instruments, even if their physical causes differ.

It was decided to analyze the developed algorithms on sounds from two representative categories of musical instruments: the flute for woodwinds and the violin for bowed strings.

6.6.1 Analysis of flute sounds

The Good-sounds dataset provides flute sounds annotated as either *good attack* or *bad attack*, without further specification. To evaluate the applicability of the developed algorithm, three flute sound samples of the same pitch (i.e., G6) were selected as representative examples. These include:

- One sound annotated with a *good attack*.
- Two sounds annotated with a *bad attack*, specifically:
  - One resembling a *noisy attack*.
  - One resembling an *attack with breath noise*.

In the case of the *good attack* shown in Figure 6.10, the envelopes exhibit a behavior similar to that observed in trumpet sounds. The envelope of the fluctuations remains low, while the envelope of the harmonic components increases very steeply. As a result, both the estimated noise ducking time and attack duration are short.

In contrast, in the case identified as a *noisy attack* (Figure 6.11a), the envelope of the fluctuation components presents higher values, indicating an unstable sound onset. Additionally,

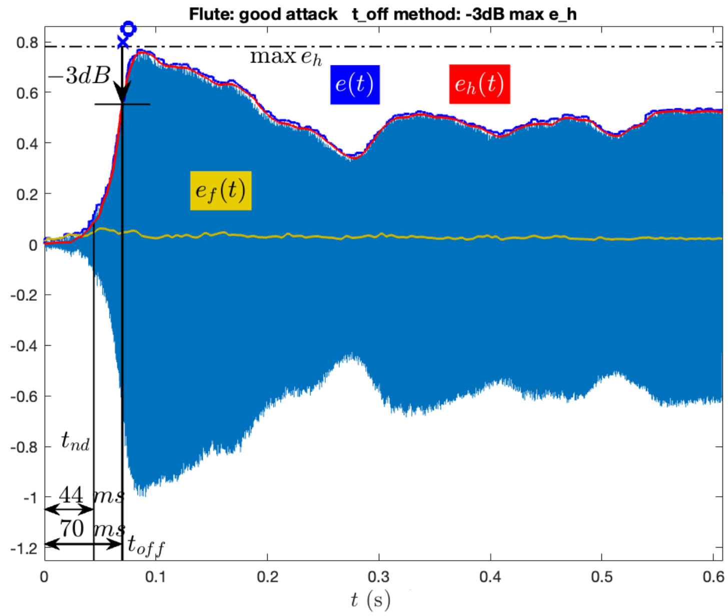


Figure 6.10: Flute sound signal with a good attack, time-shifted so that $t_{on} = 0$. The line colors and labels are the same as in Figure 6.5.

a short noise is present due to air entering the instrument before the actual sound begins, creating a hybrid between a noisy attack and a breathy attack. This is followed by a distinct growling sound characteristic in the attack, resulting in higher estimates of both the noise ducking time and the attack duration.

Similarly to the trumpet case, the attack identified with *breath noise* (Figure 6.11b) features an initial segment dominated by breath noise as air enters the instrument before the note develops. Here as well, the adaptive threshold effectively captures the attack characteristics, leading to long estimates of both the noise ducking time and the attack duration.

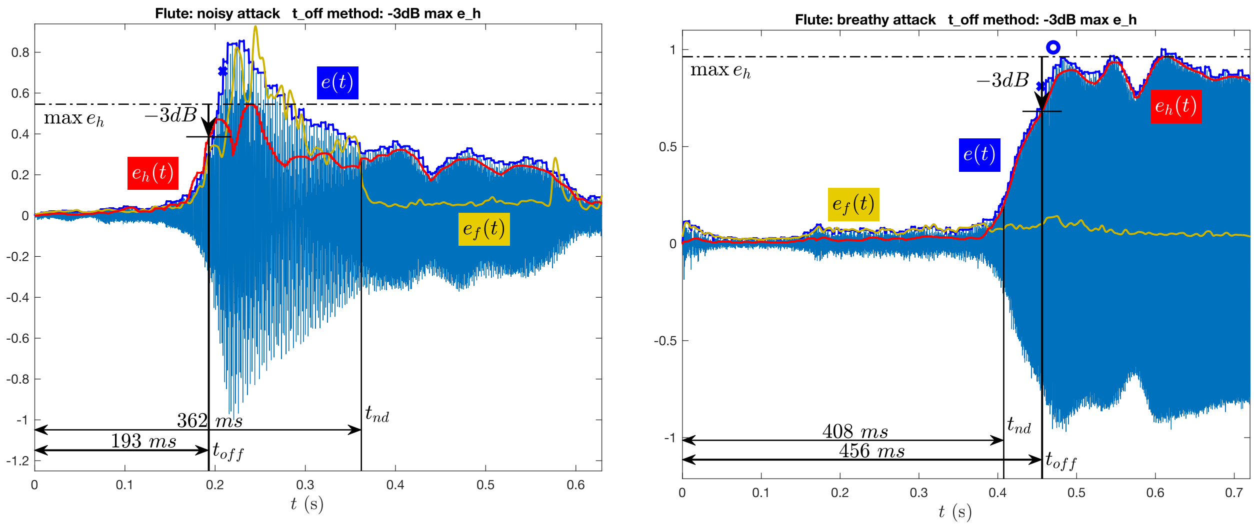
The audio tracks of the analyzed flute and violin sounds can be listened to on the companion website[1].

6.6.2 Analysis of violin sounds

The Good-sounds dataset provides violin sounds annotated as either *good attack* or *bad attack*. Among the sounds classified as *bad attack*, three distinct groups can be identified:

- A subset of sounds lacks specific annotations but exhibits a noisy transient characteristic.
- A subset is annotated as *bad pressure*, which we interpret as an indication that the

[1] https://attackdurationestimator.github.io/DAFx25



(a) Noisy attack (b) Attack with breath noise

Figure 6.11: Flute sound signals with (a) a noisy attack and (b) a breathy attack. The sound signal is time-shifted so that $t_{on} = 0$. The line colors and labels are the same as in Figure 6.5.

player applied incorrect bow pressure at the attack. These sounds feature a longer and more pronounced noise during the onset.

- A subset is annotated as *rebond*, which we interpret as the French term for "bounce", suggesting that the bow bounces on the string, producing a very short note at the onset. Since this type of error is not related to the duration of the attack transient but rather to a sequence of a short sound followed by the note, it is not considered in this analysis.

Three different violin sound samples were selected as examples to evaluate the algorithm's applicability.

For the *good attack* example shown in Figure 6.12, the envelope of fluctuations remains at low values throughout the attack, while the harmonic envelope increases smoothly until reaching a stable level.

In the examples of *bad attack* shown in Figure 6.13a and *bad pressure attack* shown in Figure 6.13b, the envelope of fluctuations exhibits significantly higher values during the attack transient, while the harmonic envelope behaves more erratically. In these cases, the noise ducking time emerges as the most salient feature distinguishing good and bad attacks, showing substantial differences in values between the two scenarios.

Despite variations in absolute attack times across different instruments, the developed algorithms consistently distinguished between properly executed attacks and faulty ones. The results confirm that the integration of both temporal and spectral information is essential for accurately analyzing transient behaviors in instrumental sounds.

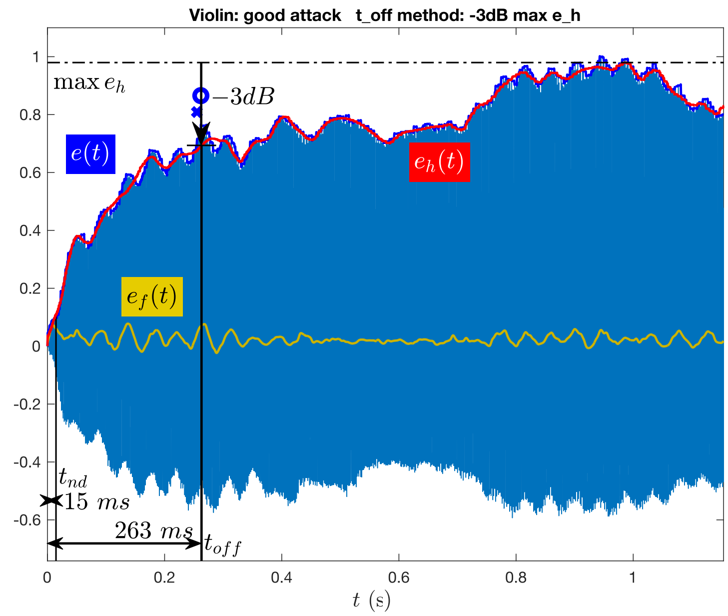
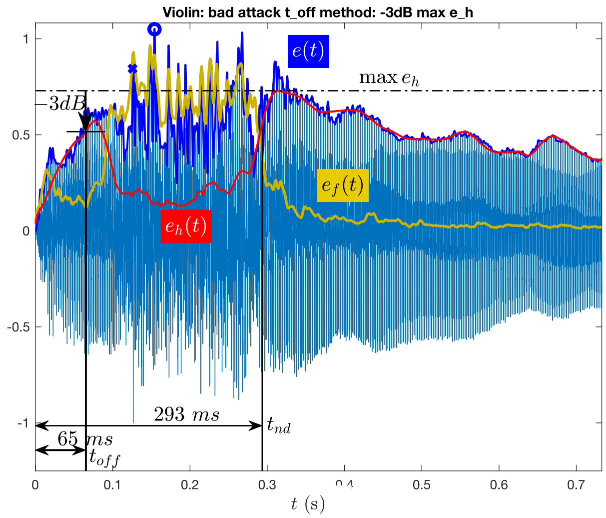


Figure 6.12: Violin sound signal with a good attack, time-shifted so that $t_{on} = 0$. The line colors and labels are the same as in Figure 6.5.

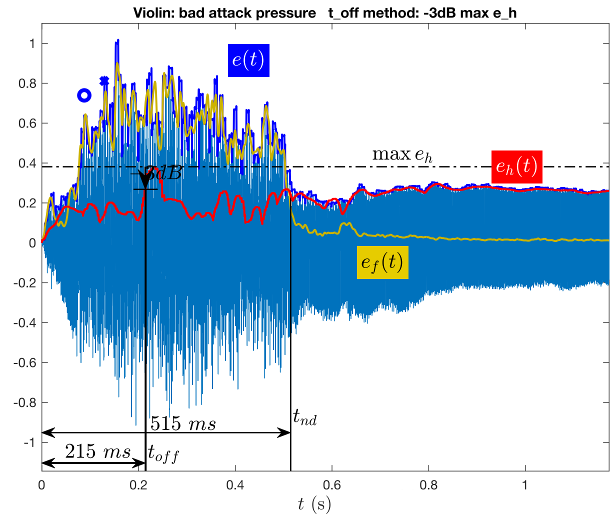
6.7 Conclusions

This chapter introduced a new method for the estimation of the duration of the attack in non-percussive orchestral instruments for which, due to the excitation bow or blow noise, the classical direct estimate and the weakest-effort method are not sufficiently robust. Such a method is based on an excitation/resonance separation by means of an improved PSWT. The proposed algorithm was checked for consistency and robustness by means of statistical trials conducted on synthetic sounds. Qualitative checks and musical interpretation could be performed in the specially created dataset and in other available databases. Usage in database indexing for tone-quality related queries and in self-assisted music practice was addressed.

Further work will extend the dataset and interpretation to a broader class of instruments with annotations by experts. Furthermore, since the attack times may vary for each harmonic of the tone, the Harmonic-Band Wavelet Transform (HBWT) will be explored, which is essentially a PSWT where multiplexing is replaced by a Discrete-Cosine Transform (DCT) (Polotti and Evangelista, 2001).



(a) Bad attack



(b) Bad pressure attack

Figure 6.13: Violin sound signals with (a) a bad attack and (b) a bad pressure attack. The sound signal is time-shifted so that $t_{on} = 0$. The line colors and labels are the same as in Figure 6.5.

Contributions

Gianpaolo Evangelista: Conceptualization of the study and method, software development, analysis, supervision of the overall study, writing of the original draft (primarily Sections 6.3 and 6.4), and review and editing. **Alberto Acquilino**: Conceptualization of the study, dataset recording, annotation and analysis, software testing and curating the companion website, writing of the original draft (primarily Sections 6.2 and 6.5), and review and editing. Other sections (6.1, 6.6, and 6.7) were written collaboratively by Gianpaolo Evangelista and Alberto Acquilino. **Gary Scavone**: Supervision of Alberto Acquilino's work, providing review and edits to the overall study.

Chapter 7

Conclusion and Future Directions

7.1 Conclusion

The field of music education technology encompasses a vast and interdisciplinary landscape, drawing from disciplines such as pedagogy and music education, acoustics and psychoacoustics, music psychology and cognitive science, sound and signal processing, artificial intelligence, and human-computer interaction, among others. These diverse fields, while interconnected, are inherently broad and multifaceted, each contributing distinct perspectives to the development of technological solutions for music learning.

Any research within this domain must inevitably grapple with defining the scope and objectives of music education. This involves reflecting on what music education encompasses and the desired pedagogical approaches that should guide it. The direction of technological development is shaped by considerations such as the intended learners, the specific musical skills and knowledge to be imparted, and the educational needs and resources available. These factors ultimately influence the design of music education technologies, determining their content, target audience, and underlying pedagogical frameworks.

These considerations have led to a variety of technological advancements in music education, including but not limited to:

- Systems that provide real-time feedback on specific musical skills.
- AI-driven algorithms that adapt the difficulty of a musical score to facilitate learning.
- Technologies that monitor and assess posture during instrumental practice.
- Tools that visualize sound features.
- Novel hardware interfaces that expand the expressive potential of traditional musical instruments, offering both creative tools and assistive technologies for learners with

diverse needs.

- AI-assisted music therapy applications, which provide new ways of engaging with music for therapeutic and rehabilitative purposes.

A critical debate in the field concerns the role of technology in shaping pedagogy. Often, technological advancements dictate trends in music education, rather than pedagogy guiding technological innovation. Ideally, the pedagogical needs of educators and learners should define the trajectory of technological development. However, in practice, educators often lack the technical expertise or resources to develop new educational technologies, leading to a situation where emerging tools may not fully align with pedagogical best practices.

Another major challenge arises from the sheer breadth of musical learning. The vast number of different musical skills, theoretical concepts, and performance techniques makes it easy to become lost in over-specialized or fragmented solutions. One possible future direction to address this issue is to build modular and effective systems for teaching individual skills, while maintaining long-term adaptability. In the future, as AI technology advances, it may become possible to orchestrate multiple learning tools, adapting them to the specific needs of individual students or even aligning them with the instructional strategies of human teachers. Such adaptive AI-driven learning environments could transform how technology supports both self-regulated learning and guided instruction.

7.2 Contributions

This dissertation pursued this vision by focusing on the development of open-source solutions. It introduced frameworks, codebases, datasets, and methodologies designed to maximize accessibility and impact in the field of music education technology. The technologies developed in this work were designed with a student-centered approach, providing educators with new tools to offer clear and precise instruction. Rather than dictating a rigid structure, the research aimed to be broadly applicable, leveraging audio analysis and algorithms compatible with web-based applications to ensure maximum accessibility.

The systems developed in this dissertation were intentionally designed to be flexible, allowing for the creation of targeted exercises that support the structured development of technical skills. At the same time, they enable exploratory practice, fostering reflection and self-regulated learning through interactive feedback mechanisms. Furthermore, the open-source modular code framework proposed in this research lowers the barrier to entry for educators and developers, making it easier to create new educational tools without requiring advanced programming expertise.

7.3 Future work

Future extensions of this work could explore expanding the framework to support a wider range of musical skills, instruments, and traditions. This could lead to the development of a digital method that adapts to the individual learner, progressively guiding them through structured exercises and exploratory practice, by incorporating interactive feedback. Such an approach would offer a more adaptive way to develop instrumental technique and broader musical competencies.

Such an adaptive method could be further enhanced by generative AI, which presents promising opportunities to dynamically create personalized learning content. The advancements in Large Language Models (LLMs) and generative AI open new possibilities for tailoring music education experiences to individual learners. AI-generated interactive exercises and adaptive difficulty levels, informed by real-time analysis of learning progress, could empower a new generation of personalized music education tools. By continuously monitoring student performance, AI-driven systems could dynamically adjust instructional content to address diverse learning needs, ensuring that each learner receives targeted support aligned with their goals and context. Rather than replacing human-led instruction, these technologies could enhance the teacher's role, offering data-driven insights that enable more informed pedagogical decisions.

Finally, real-time AI-driven monitoring of the learning process could provide adaptive feedback systems, capable of adjusting instructional content in response to students' progress and evolving proficiency levels. By integrating AI with human-led instruction, such technologies could enhance the role of the teacher rather than replace it, offering data-driven insights that allow for more informed pedagogical decisions.

7.4 Final remarks

This dissertation has laid the groundwork for more accessible, flexible, and effective technology-enhanced learning tools for music education. By leveraging open-source web technologies, it provides a scalable and inclusive approach that allows both educators and learners to engage with innovative musical training solutions. While the field continues to evolve, the ideas and frameworks developed in this work contribute to a growing ecosystem of research that seeks to bridge the gap between technology, pedagogy, and artistic expression.

By embracing interdisciplinary collaboration and fostering technological accessibility, the use of technology in music education technology holds great promise for the future. The research presented in this dissertation serves as a step toward that vision—one in which technology acts as an enabler of learning, providing new opportunities for students and educators to explore the rich and dynamic world of music.

References

A. Acquilino and G. Scavone. Current state and future directions of technologies for music instrument pedagogy. *Frontiers in Psychology*, 13, 2022.

A. Acquilino, N. Puranik, I. Fujinaga, and G. Scavone. A dataset and baseline for automated assessment of timbre quality in trumpet sound. In *Proceedings of the 24th International Society of Music Information Retrieval Conference*, pages 684–691, 2023a.

A. Acquilino, N. Puranik, I. Fujinaga, and G. Scavone. Detecting efficiency in trumpet sound production: Proposed methodology and pedagogical implications. In *Proceedings of the 5th Stockholm Music Acoustic Conference*, pages 72–79, Stockholm, 2023b. KTH Royal Institute of Technology.

A. Acquilino, M. d'Andrea, K. K. Reddy, J. J. Park, and G. Scavone. Open-source mobile apps for music education: A case study on trumpet fingering. *Submitted to International Journal of Music Education*, 2025.

Y. Aksoy. Seeing sounds: The effect of computer-based visual feedback on intonation in violin education. *International Journal of Education and Literacy Studies*, 11(2):2–12, 2023.

E. Allingham and C. Wöllner. Putting practice under the microscope: The perceived uses and limitations of slow instrumental music practice. *Psychology of Music*, 51(3):906–923, 2023.

J.-B. Arban. *Arban's Complete Conservatory Method for Trumpet (Cornet): or E-flat Alto, B-flat Tenor, Baritone, Euphonium and B-flat Bass in Treble Clef*. Carl Fischer, New York, NY, 1982.

E. C. Axford. *Music Apps for Musicians and Music Teachers*. Rowman and Littlefield, Lanham, Maryland, 2015.

S. Bagga, B. Maurer, T. Miller, L. Quinlan, L. Silvestri, D. Wells, R. Winqvist, M. Zolotas, and Y. Demiris. Instrumentor: An interactive robot for musical instrument tutoring. In

Proceedings of the 20th Annual Conference on Towards Autonomous Robotic Systems, pages 303–315, 2019.

G. Bandiera, O. Romani, H. Tokuda, W. Hariya, K. Oishi, and X. Serra. Good-sounds.org: A framework to explore goodness in instrumental sounds. In *Proceedings of the 17th International Society for Music Information Retrieval Conference*, New York, 2016.

W. I. Bauer. *Music Learning Today: Digital Pedagogy for Creating, Performing, and Responding to Music*. Oxford University Press, New York, NY, second edition, 2020.

N. Bevan and M. Macleod. Usability measurement in context. *Behaviour and Information Technology*, 13(1/2):132, 1994.

A. D. Blanco, S. Tassani, and R. Ramirez. Real-time sound and motion feedback for violin bow technique learning: A controlled, randomized trial. *Frontiers in Psychology*, 12, 2021.

D. Bogdanov, N. Wack, E. Gómez, S. Gulati, P. Herrera, O. Mayor, G. Roma, J. Salamon, J. Zapata, and X. Serra. Essentia: An open-source library for sound and music analysis. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, pages 855–858, New York, NY, USA, 2013. Association for Computing Machinery.

E. Booth. *The Music Teaching Artist's Bible: Becoming a Virtuoso Educator*. Oxford University Press, Oxford, 2009.

C. Bouras, A. Papazois, and N. Stasinos. Cross-platform mobile applications with web technologies. *International Journal of Computing and Digital Systems (Special Issue on MobiApps 2014)*, 4:153–163, 2015.

J. Brennan. Poper's game - a 21st century approach to addressing initial attacks in practice. *International Trumpet Guild Journal*, 41(1):68–71, 2016.

M. Bönstrup, I. Iturrate, M. N. Hebart, N. Censor, and L. G. Cohen. Mechanisms of offline motor learning at a microscale of seconds in large-scale crowdsourced data. *Nature Partner Journals Science of Learning*, 5(1), 2020.

T. Calo and C. Maclellan. Towards educator-driven tutor authoring: Generative AI approaches for creating intelligent tutor interfaces. In *Proceedings of the Eleventh ACM Conference on Learning @ Scale*, L@S '24, pages 305–309, Atlanta, GA, USA, 2024. Association for Computing Machinery.

M. Campbell, M. Gilbert, and A. Myers. *The Science of Brass Instruments*. Modern Acoustics and Signal Processing. Springer, 2021.

F. G. Campos. *Trumpet Technique*. Oxford University Press, 2005.

A. A. Capurso. The effect of an associative technique in teaching pitch and interval discrimination. *Journal of Applied Psychology*, 18(6):811–818, 1934.

C. E. Carter and J. A. Grahn. Optimizing music learning: Exploring how blocked and interleaved practice schedules affect advanced performance. *Frontiers in Psychology*, 7: Article 1251, 2016.

G. Cassone. *The Trumpet Book*. Zecchini, 2009.

A. Creech and H. Gaunt. The changing face of individual instrumental tuition: Value, purpose and potential. *The Oxford Handbook of Music Education*, 1:694–711, 2012.

A. Cremaschi, E. Leger, N. Smith, and K. Ilinykh. Students who quit music lessons: Recent research and recommendations for teachers. *MTNA e-Journal*, 6(15), 2015.

L. Crocco and D. Meyer. Motor learning and teaching singing: An overview. *Journal of Singing*, 77(5):693–702, 2021.

M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper Perennial, New York, NY, first harper perennial modern classics edition, 2008.

B. F. Dalby. A computer-based training program for developing harmonic intonation discrimination skill. *Journal of Research in Music Education*, 40(2):139–152, 1992.

S. Dalla Bella. Music and brain plasticity. In S. Hallam, I. Cross, and M. H. Thaut, editors, *The Oxford Handbook of Music Psychology*. Oxford University Press, 2nd edition, 2016.

R. Dannenberg, M. Sanchez, A. Joseph, R. Joseph, R. Saul, and P. Capell. Results from the Piano Tutor project. In *Proceedings of the Fourth Biennial Arts and Technology Symposium*, pages 143–150, 1993.

I. Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, USA, 1992.

F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–340, 1989.

F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology*, 22(14):1111–1132, 1992.

S. Deterding, D. Dixon, R. Khaled, and L. Nacke. From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, pages 9–15, 2011.

J. Dewey. *Experience and Education*. Kappa Delta Pi, New York, 1938.

M. J. Eisele. *Development and Validation of a Computer-Assisted Lesson in Teaching Intonation Skills to Violin and Viola Students*. Phd thesis, Indiana University, Bloomington, 1985. Unpublished doctoral dissertation.

A. Engel, M. Bangert, D. Horbank, B. S. Hijmans, K. Wilkens, P. E. Keller, and C. Keysers. Learning piano melodies in visuo-motor or audio-motor training conditions and the neural correlates of their cross-modal transfer. *NeuroImage*, 63(2):966–978, 2012.

K. A. Ericsson. The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3):363–406, 1993.

K. A. Ericsson and K. W. Harwell. Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. *Frontiers in Psychology*, 10:2396, 2019.

G. Evangelista. Pitch synchronous wavelet representations of speech and music signals. *IEEE Transactions on Signal Processing*, 41(12):3313–3330, 1993. Special issue on Wavelets and Signal Processing.

G. Evangelista. Comb and multiplexed wavelet transforms and their applications to signal processing. *IEEE Transactions on Signal Processing*, 42(2):292–303, 1994.

G. Evangelista and A. Acquilino. An adaptive wavelet-based algorithm for assessing the quality of the attack transients in non-percussive instruments. In *Proceedings of the 28th International Conference on Digital Audio Effects (DAFx25)*, Ancona, Italy, 2025. A3Lab, Università Politecnica delle Marche.

M. Fabiani and A. Friberg. Influence of pitch, loudness, and timbre on the perception of instrument dynamics. *The Journal of the Acoustical Society of America*, 130(4):193–199, 2011.

M. Fautley. The potential of audio and video for formative assessment purposes in music education in the lower secondary school in England: Issues arising from a small-scale study of trainee music teachers. *Journal of Music, Technology and Education*, 6(1):29–42, 2013.

H. Feldmann. History of the tuning fork. i: Invention of the tuning fork, its course in music and natural sciences. Pictures from the history of otorhinolaryngology, presented by instruments from the collection of the ingolstadt german medical history museum. *Laryngorhinootologie*, 76(2):116–122, 1997.

S. Ferguson. Learning musical instrument skills through interactive sonification. In *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06)*, pages 384–389, Paris, France, 2006.

N. Fleming and D. Baume. Learning styles again: Varking up the right tree! *Educational Developments, SEDA Ltd*, 7(4):4–7, 2006.

N. Fleming and C. Mills. Not another inventory, rather a catalyst for reflection. *To Improve the Academy*, 11(1):137–155, 1992.

C. Flesch and A.-S. Mutter. *The Art of Violin Playing*. Carl Fischer, New York, 2008.

M. Gall. Trainee teachers' perceptions: Factors that constrain the use of music technology in teaching placements. *Journal of Music, Technology and Education*, 6(1):5–27, 2013.

S. Gates. Developing musical imagery contributions from pedagogy and cognitive science. *Music Theory Online*, 27(2), 2021.

H. Gaunt. One-to-one tuition in a conservatoire: The perceptions of instrumental and vocal teachers. *Psychology of Music*, 36(2):215–245, 2008.

H. Gaunt. Apprenticeship and empowerment: The role of one-to-one lessons. In J. Rink, H. Gaunt, and A. Williamon, editors, *Musicians in the Making: Pathways to Creative Performance*, pages 28–56. Oxford University Press, New York, NY, 2017.

B. Gebel, C. Braun, E. Kaza, E. Altenmüller, and M. Lotze. Instrument specific brain activation in sensorimotor and auditory representation in musicians. *NeuroImage*, 74:37–44, 2013.

J. M. Geringer and M. D. Worthy. Effects of tone-quality changes on intonation and tone-quality ratings of high school and college instrumentalists. *Journal of Research in Music Education*, 47(2):135–149, 1999.

S. Giraldo, G. Waddell, I. Nou, A. Ortega, O. Mayor, A. Perez, A. Williamon, and R. Ramirez. Automatic assessment of tone quality in violin music performance. *Frontiers in Psychology*, 10:1–12, 2019.

D. L. Goodhue and R. L. Thompson. Task-technology fit and individual performance. *MIS Quarterly*, 19(2):213–236, 1995.

J. Granda Vera, J. C. Barbero Alvarez, and M. Montilla Medina. Effects of different practice conditions on acquisition, retention, and transfer of soccer skills by 9-year-old schoolchildren. *Perceptual and Motor Skills*, 106(2):447–460, 2008.

M. A. Guadagnoli and R. M. Kohl. Knowledge of results for motor learning: Relationship between error estimation and knowledge of results frequency. *Journal of Motor Behavior*, 33(2):217–224, 2001.

K. Guettler. Bows, strings, and bowing. In *The Science of String Instruments*, pages 279–299. Springer New York, New York, NY, 2010.

K. Guettler and A. Askenfelt. Acceptance limits for the duration of pre-Helmholtz transients in bowed string attacks. *The Journal of the Acoustical Society of America*, 101(5):2903–2913, 1997.

K. L. Hagglund and K. Jacobs. Physical and mental practices of music students as they relate to the occurrence of music-related injuries. *Work*, 6:11–24, 1996.

J. Hajda. A new model for segmenting the envelope of musical signals: The relative salience of steady state versus attack, revisited. In *Proceedings of the 101st Audio Engineering Society Convention (AES)*, Los Angeles, California, USA, 1996.

A. R. Halpern, R. J. Zatorre, M. Bouffard, and J. A. Johnson. Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42(9):1281–1292, 2004.

I. M. Hanken. The role and significance of masterclasses in creative learning. In J. Rink, H. Gaunt, and A. Williamon, editors, *Musicians in the Making: Pathways to Creative Performance*, Studies in Musical Performance as Creative Practice. Oxford University Press, New York, NY, 2017. Accessed 10 Jan. 2025.

S. D. Harrison and J. O'Bryan. *Teaching Singing in the 21st Century*. Springer, New York, 2014.

H. v. Helmholtz. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Trans. by A. Ellis, New York: Dover, 1977.

T. K. Ho. Random decision forests. In *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282. IEEE, 1995.

K. Hornbæk. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human - Computer Studies*, 64(2):79–102, 2006.

A. Jacobs and B. Nelson. *Also Sprach Arnold Jacobs: A Developmental Guide for Brass Wind Musicians*. Polymnia Press, 2006.

J. M. Jameson, V. Thompson, G. Manuele, D. Smith, H. Egan, and T. Moore. Using an iTouch to teach core curriculum words and definitions: Efficacy and social validity. *Journal of Special Education Technology*, 27(3):41–54, 2012.

B. Karakaş and M. Dündar. Multi-sensory learning in violin training. *International Journal of Quality in Education*, 8(2):12–30, 2024.

S. Kazazis, P. Depalle, and S. McAdams. *The Timbre Toolbox User's Manual*, 2022.

R. F. Kenny and R. McDaniel. The role teachers' expectations and value assessments of video games play in their adopting and integrating them into their classrooms. *British Journal of Educational Technology*, 42(2):197–213, 2011.

K. Kholykhalova, E. Volta, G. Waddell, A. Williamon, S. Ghisio, C. Canepa, R. Ramirez, and G. Volpe. Capturing high-quality violin performance data. In *Proceedings of the International Symposium on Performance Science*, 2017.

T. Knight, T. Upham, and I. Fujinaga. The potential for automatic assessment of trumpet tone quality. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, pages 573–578, 2011.

D. L. Kohut. *Musical Performance: Learning Theory and Pedagogy*. Prentice-Hall, Englewood Cliffs, N.J., 1985.

B. E. Köktürk-Güzel, O. Büyük, B. Bozkurt, and O. Baysal. Automatic assessment of student rhythmic pattern imitation performances. *Digital Signal Processing*, 133, 2023.

M. Lam. The physicality of music production: Investigating the roles of mindful practice and kinesthetic learning. *Music Educators Journal*, 106(3):23–28, 2020.

B. C. Larssen, D. K. Ho, S. N. Kraeutner, and N. J. Hodges. Combining observation and physical practice: Benefits of an interleaved schedule for visuomotor adaptation and motor memory consolidation. *Frontiers in Human Neuroscience*, 15, 2021.

C. Laurier, O. Meyers, J. Serrà, M. Blech, P. Herrera, and X. Serra. Indexing music by mood: Design and integration of an automatic content-based annotator. *Multimedia Tools and Applications*, 48(1):161–184, 2009.

M. Leman and L. Nijs. Cognition and technology for instrumental music learning. In *The Routledge Companion to Music, Technology, and Education*. Routledge, 2017.

M. Lennon and G. Reed. Instrumental and vocal teacher education: Competences, roles and curricula. *Music Education Research*, 14(3):285–308, 2012.

S. Leong and L. Cheng. Effects of real-time visual feedback on pre-service teachers' singing. *Journal of Computer Assisted Learning*, 30(3):285–296, 2014.

S. Levarie and E. Levy. *Tone: A Study in Musical Acoustics. 2nd ed.* Kent State University Press, Kent, Ohio, 1980.

H. Li, T. Xu, C. Zhang, E. Chen, J. Liang, X. Fan, H. Li, J. Tang, and Q. Wen. Bringing generative AI to adaptive learning in education. *arXiv preprint arXiv:2402.14601*, 2024.

Z. Li, D. Fan, H. Wang, Z. Lu, and C. Liu. Piano beginner: A glove-based finger training VR application. In *2022 International Symposium on Control Engineering and Robotics (ISCER)*, pages 262–265, 2022.

D. Luce and M. Clark. Durations of attack transients of nonpercussive orchestral instruments. *Journal of the Audio Engineering Society*, 13(3):194–199, 1965.

D. Luce and M. Clark. Physical correlates of brass-instrument tones. *The Journal of the Acoustical Society of America*, 42(6):1232–1243, 1967.

F. M. B. Lã and M. B. Fiuza. Real-time visual feedback in singing pedagogy: Current trends and future directions. *Applied Sciences*, 12(21):10781, 2022.

E. Maas, D. A. Robin, S. N. Austermann Hula, S. E. Freedman, G. Wulf, K. J. Ballard, and R. A. Schmidt. Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, 17(3):277–298, 2008.

C. Madsen and J. Geringer. Preferences for trumpet tone quality versus intonation. *Bulletin for the Council for Research in Music*, 46:13–22, 1976.

D. Malandrino, D. Pirozzi, and R. Zaccagnino. Learning the harmonic analysis: Is visualization an effective approach? *Multimedia Tools and Applications*, 78, 2019.

S. Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way.* Academic Press, Inc., USA, 3rd edition, 2008.

L. Martin. Using synchronous formative feedback to facilitate student growth. *Music Educators Journal*, 107(2):51–57, 2020.

K. Maton. Cumulative and segmented learning: Exploring the role of curriculum structures in knowledge-building. *British Journal of Sociology of Education*, 30(1):43–57, 2009.

M. Mauch and S. Dixon. Pyin: A fundamental frequency estimator using probabilistic threshold distributions. In *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 659–663, 2014.

S. McAdams, S. Winsberg, S. Donnadieu, G. Soete, and J. Krimphoff. Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58:177–912, 1995.

S. J. McCoy. *Your Voice, an Inside View: Multimedia Voice Science and Pedagogy*. Inside View Press, Princeton, N.J., 2004.

P. McLeod and G. Wyvill. A smarter way to find pitch. In *Proceedings of the International Computer Music Conference*, pages 138–141, 2005.

G. E. McPherson, J. Blackwell, and J. Hattie. Feedback in music performance teaching. *Frontiers in Psychology*, 13, 2022.

J. Meyer and U. Hansen. *Acoustics and the Performance of Music: Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instruments Makers*. Springer, New York, USA, 2009.

A. Michałko, A. Campo, L. Nijs, M. Leman, and E. Van Dyck. Toward a meaningful technology for instrumental music education: Teachers' voice. *Frontiers in Education*, 7, 2022.

L. Nijs. Dalcroze meets technology: Integrating music, movement and visuals with the music paint machine. *Music Education Research*, 20(2):163–183, 2018.

L. Nijs and M. Leman. Interactive technologies in the instrumental music classroom: A longitudinal study with the music paint machine. *Computers & Education*, 73(2):40–59, 2014.

L. Nijs, M. Lesaffre, and M. Leman. The musical instrument as a natural extension of the musician. In *Music and Its Instruments*, pages 467–484. Editions Delatour France, 2013.

K. Nymoen, A. Danielsen, and J. London. Validating attack phase descriptors obtained by the Timbre Toolbox and MIRtoolbox. In *Proceedings of the SMC Conference*, pages 214–219, 2017.

F. J. M. Ortega, S. I. Giraldo, and R. Ramirez. Bowing modeling for violin students assistance. In *Proceedings of the 1st International Workshop on Multimodal Interaction for Education*, pages 60–62, 2017.

L. S. Pardue and A. McPherson. Real-time aural and visual feedback for improving violin intonation. *Frontiers in Psychology*, 10:627, 2019.

J. R. Parker. *A Musical Biography, or, Sketches of the Lives and Writings of Eminent Musical Characters: Interspersed with an Epitome of Interesting Musical Matter*. Stone & Fovell, Boston, 1825.

M. P. Paule-Ruiz, V. Álvarez García, J. R. Pérez-Pérez, M. Álvarez Sierra, and F. Trespalacios-Menéndez. Music learning in preschool with mobile devices. *Behaviour and Information Technology*, 36(1):95–111, 2017.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

G. Peeters. A large set of audio features for sound description (similarity and classification) in the CUIDADO project. Technical report, CUIDADO 1st Project Report (IRCAM, Paris), 2004.

G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams. The Timbre Toolbox: Extracting audio descriptors from musical signals. *Journal of the Acoustical Society of America*, 130:2902–2916, 2011.

G. K. Percival. Computer-assisted musical instrument tutoring with targeted exercises. Master's thesis, University of Victoria, Victoria, Canada, 2008.

G. K. Percival, Y. Wang, and G. Tzanetakis. Effective use of multimedia for computer-assisted musical instrument tutoring. In *Proceedings of the International Workshop on Educational Multimedia and Multimedia Education*, pages 67–76, 2007.

A. Perez-Carrillo. Violin timbre navigator: Real-time visual feedback of violin bowing based on audio analysis and machine learning. In *MultiMedia Modeling*, pages 182–193. Springer International Publishing, 2019.

J. M. Perkel. No installation required: How WebAssembly is changing scientific computing. *Nature*, 627(8003):455–456, 2024.

P. Polotti and G. Evangelista. Fractal additive synthesis by means of harmonic-band wavelets. *Computer Music Journal*, 25(3):22–37, 2001.

M. Pàmies-Vilà, A. Hofmann, and V. Chatziioannou. The influence of the vocal tract on the attack transients in clarinet playing. *Journal of New Music Research*, 49(2):126–135, 2020.

R. Ramirez, G. Volpe, C. Canepa, S. Ghisio, K. Kolykhalova, S. Giraldo, O. Mayor, A. Perez, M. Mancini, E. Volta, G. Waddell, and A. Williamon. Enhancing music learning with smart technologies. In *MultiMedia Modeling*, pages 0–3. ACM International Conference Proceeding Series, 2018.

S. Raptis, A. Chalamandaris, A. Baxevanis, A. Askenfelt, E. Schoonderwaldt, K. F. Hansen, D. Fober, S. Letz, and Y. Orlarey. IMUTUS: An effective practicing environment for music tuition. In *Proceedings of the 2005 International Computer Music Conference*, pages 383–386, Barcelona, Spain, 2005.

G. Rizzolatti and L. Craighero. The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192, 2004.

M. Robine and M. Lagrange. Evaluation of the technical level of saxophone performers by considering the evolution of spectral parameters of the sound. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 79–84, Victoria, Canada, 2006.

M. Robine, G. K. Percival, and M. Lagrange. Analysis of saxophone performance for computer-assisted tutoring. In *Proceedings of the 2007 International Computer Music Conference*, pages 381–384, 2007.

O. Romani, H. Parra, D. Dabiri, H. Tokuda, W. Hariya, K. Oishi, and X. Serra. A real-time system for measuring sound goodness in instrumental sounds. In *Proceedings of the 138th Audio Engineering Society Convention (AES)*, pages 1106–1111, Warsaw, Poland, 2015.

R. A. Schmidt and T. D. Lee. *Motor Learning and Performance: From Principles to Application*. Human Kinetics, Champaign, IL, sixth edition, 2020.

E. Schoonderwaldt, A. Askenfeld, and K. Hansen. Design and implementation of automatic evaluation of recorder performance in IMUTUS. In *Proceedings of the 2005 International Computer Music Conference*, pages 97–103, International Computer Music Association, Barcelona, Spain, 2005.

C. H. Shea and G. Wulf. Enhancing motor learning through external-focus instructions and feedback. *Human Movement Science*, 18(4):553–571, 1999.

Y. Shi. The use of mobile internet platforms and applications in vocal training: Synergy of technological and pedagogical solutions. *Interactive Learning Environments*, 31(6): 3780–3791, 2023.

A. L. Simmons. The relationship between prospective teachers' tone quality evaluations and their knowledge of wind instrument pedagogy. *Applications of Research in Music Education*, 23(2):42–51, 2005.

L. M. Squire. Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4(3):232–243, 1992.

L. A. Stambaugh. When repetition isn't the best practice strategy: Effects of blocked and random practice schedules. *Journal of Research in Music Education*, 58(4):368–383, 2011.

K. Steenstrup. *Teaching Brass*. Det Jyske Musikkonservatorium, 2007.

K. Steenstrup. *Blow Your Mind*. Aarhus University Press, 2017.

K. Steenstrup. *Deep Practice, Peak Performance: The Science of Musical Learning*. The Royal Academy of Music, Aarhus, 2023.

K. Steenstrup, N. T. Haumann, B. Kleber, C. Camarasa, P. Vuust, and B. Petersen. Imagine, sing, play: Combined mental, vocal and physical practice improves musical performance. *Frontiers in Psychology*, 12, 2021.

J. Sundberg and J. Gauffin. Waveform and spectrum of the glottal voice source. In *Frontiers of Speech Communication Research*, pages 301–320, London: Academic Press, 1978.

D. G. Swift. Improving harmonic intonation skills of high school band students using Coda Music technology's *Intonation Trainer*. Master's thesis, University of Louisville, Kentucky, 2003. Unpublished Master's Thesis.

G. Tambouratzis, S. Bakamidis, I. Dologlou, G. Carayannis, and M. Dendrinos. The IMUTUS interactive music tuition system. *The Journal of the Acoustical Society of America*, 111(5): 2348–2348, 2002.

J. Thompson. *The Buzzing Book Complete Method; Trumpet or Other Brass Instruments*. Editions BIM, 2003.

M. Tullberg. Affordances of musical instruments: Conceptual consideration. *Frontiers in Psychology*, 13, 2022.

P. P. Vaidyanathan. Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial. *Proceedings of the IEEE*, 78(1):56–93, 1990.

G. Waddell and A. Williamon. Technology use and attitudes in music learning. *Frontiers in ICT*, 6, 2019.

M. A. Walker, J. Fromer, G. Di Fabbrizio, C. Mestel, and D. Hindle. What can I say? evaluating a spoken language interface to email. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 582–589, 1998.

P. R. Webster. Key research in music technology and music teaching and learning. *Journal of Music, Technology and Education*, 4(2–3):115–130, 2012.

B. Weidner. *Brass Techniques and Pedagogy.* Palni Press, 2020.

A. Weisberg. *The Art of Wind Playing.* Meredith Music Publications, Galesville, MD, and Milwaukee, WI, 2007. Exclusively distributed by Hal Leonard.

G. Welch. Variability of practice and knowledge of results as factors in learning to sing in tune. *Bulletin of the Council for Research in Music Education*, 85:238–247, 1985.

G. Welch, D. M. Howard, E. Himonides, and J. Brereton. Real-time feedback in the singing studio: An innovatory action-research project using new voice technology. *Music Education Research*, 7(2):225–249, 2005.

B. C. Wesolowski. Understanding and developing rubrics for music performance assessment. *Music Educators Journal*, 98(3):36–42, 2012.

F. W. Westphal. *Guide to Teaching Woodwinds.* Brown Publishers, Dubuque, IA, 5th edition, 1990.

S. Willermark. Technological pedagogical and content knowledge: A review of empirical studies published from 2011 to 2016. *Journal of Educational Computing Research*, 56(3): 315–343, 2018.

A. Williamon, J. Ginsborg, R. Perkins, and G. Waddell. *Performing Music Research: Methods in Music Education, Psychology, and Performance Science.* Oxford University Press, Oxford, UK, first edition, 2021.

J. Yin, Y. Wang, and D. Hsu. Digital violin tutor: An integrated system for beginning violin learners. In *MULTIMEDIA '05: Proceedings of the 13th Annual ACM International Conference on Multimedia*, pages 976–985, 2005.

R. J. Zatorre and A. R. Halpern. Mental concerts: Musical imagery and auditory cortex. *Neuron*, 47:9–12, 2005.

Appendix A

Software links

Table A.1: List of considered software with corresponding URLs accessed 11 December 2021.

Software	Reference Link
Anytune Pro+	https://anytune.us/
Amazing Slow Downer	https://www.ronimusic.com/
EarMaster	https://www.earmaster.com/
Estill Voiceprint Plus	https://store.estillvoice.com/products/voiceprint
forScore	https://forscore.co/
GNU Solfege	https://www.gnu.org/software/solfege/solfege.html
Guitar Pro	https://www.guitar-pro.com/c/15-guitar-pro-ios-android
GuitarToolkit	http://agilepartners.com/apps/guitartoolkit/
GuitarTuna	https://yousician.com/guitartuna
KORG cortosia	https://www.korg.com/us/products/software/cortosia/
Knock Box Metronome	https://www.pfeiferdrumco.com/knock-box-metronome.html
Modacity	https://www.modacity.co/
liveBPM - Beat Detector	https://play.google.com/store/apps/details?id=com.DanielBach.liveBPM
Piascore	http://piascore.com/
Pickup Music	https://www.pickupmusic.com/
Play with a Pro	https://www.playwithapro.com/
QuantiForce	https://www.bonsai-systems.com/music-tech/
Rec'n'Share	https://usa.yamaha.com/products/musical_instruments/drums/el_drums/apps/rec_n_share/index.html
Rhythm Teacher	https://play.google.com/store/apps/details?id=net.gamya.rhythm
Rhythm Trainer	https://play.google.com/store/apps/details?id=ru.demax.rhythmerr
Riyaz	https://riyazapp.com/
RTFactory Rudiments	https://rhythmtoolsfactory.de/appsios/rudimentsios/
SkyNote	http://telmi.upf.edu/
SmartMusic	https://www.smartmusic.com/
Tempo	http://www.frozenape.com/tempo-metronome.html
Soundbrenner	https://www.soundbrenner.com/
TonalEnergy	https://www.tonalenergy.com/
tonestro	https://www.tonestro.com/
TrueFire	https://truefire.com/
Visual Note	https://www.visual-note.com/
Yousician	https://yousician.com/
Youtube	https://www.youtube.com/

Appendix B

Software Structure

This research work presents an application designed to provide state-of-the-art solutions for music education, leveraging modern web technologies and reusable components. The app is released under the Affero GPL license[1] and is built using the *Ionic*[2] framework and *Angular*[3]. Those technologies allow to create a hybrid mobile app that can run directly in a browser or be deployed as mobile apps for both iOS and Android devices. The use of widely adopted web technologies ensures compatibility across platforms, broadens accessibility, and allows more developers to contribute to and customize the codebase effectively.

The source code is organized into a set of modular Angular components and services that can be combined to easily create various educational tools. These reusable components address music educational elements such as score display, note selection, and audio processing (e.g., playing or recording sound). By combining these building blocks, developers can create more sophisticated features, such as a chromatic tuner or customized audio feedback systems. Each component automatically adapts its interface to changes in input or screen size, ensuring a responsive layout across platforms and devices. The entire application relies on a service that establishes timing for triggering events. The service uses RxJS[4] to emit events at intervals defined by the user.

For detailed information on the components and how to use the repository, we invite the reader to refer to the online documentation[5].

[1] Available at: `https://www.gnu.org/licenses/agpl-3.0.html` (accessed December 14, 2024).

[2] *Ionic* Framework. Available at: `https://ionicframework.com/` (accessed December 14, 2024).

[3] *Angular* Framework. Available at: `https://angular.io/` (accessed December 14, 2024).

[4] *RxJS*, a library for reactive programming in JavaScript. Available at: `https://rxjs.dev/` (accessed December 14, 2024).

[5] `https://github.com/albertoacquilino/music-education-interface-ionic`

B.1 Reusable App Components

B.1.1 Tempo Selector

This component allows users to select the tempo in beats per minute (bpm) in a GUI optimized for mobile devices.

B.1.2 Note Selector Component

This component offers an intuitive interface for selecting individual musical notes for a given exercise. The current note is displayed as an image, and tapping it opens a scrollable modal gallery of note images, enabling specific note selection. This interface supports various clefs (e.g., treble or bass) and can be configured to display subsets of notes (e.g., specific scales like chromatic or C major).

B.1.3 Score Display Component

The *Score Display* component renders dynamic and responsive musical scores using VexFlow[6]. It accepts a score object containing detailed Western music notation elements such as measures, clef, key signature, time signature, dynamics, and notes. It can be fed with predefined scores or generate them on the fly via random or deterministic algorithms, or even AI-based approaches. Parameters like the lowest and highest notes or specific scales are also configurable, enabling a wide range of educational scenarios.

B.1.4 Fingering Display Component

The *Fingering Display* component embeds static or vector images within an application's interface. Taking trumpet as an example, a GUI created with SVG displays a detailed illustration of a trumpet along with interactive buttons. These buttons simulate sliding mechanisms and turn red when activated, visually indicating the corresponding trumpet fingering for the note to be played. To enable interactivity, visual indicators or feedback can be synchronized with metronome data or performance metrics. This feature allows developers to provide real-time visual feedback on technical aspects of music learning.

[6] *VexFlow* library for rendering music notation in the browser. Available at: https://github.com/vexflow/vexflow (accessed December 14, 2024).

B.1.5 Exercise Guide Component

This component visually guides users through different phases of the musical exercise: rest, listen, and play. It uses a traffic light representation with red, yellow, and green lights representing each phase, helping users follow a structured flow of activities. By providing clear visual cues, this feature enhances focus and rhythm during practice sessions.

B.1.6 Audio Output

The application uses this service to manage the timed playback of sounds. These can be pre-recorded or synthesized sounds managed through the MIDI protocol and allow real-time modulation of musical parameters such as duration and dynamics.

B.1.7 Audio Input

The *Audio In* service acquires real-time audio from the device microphone. While its simplest use case is straightforward audio recording, the component is designed to serve as a foundation for more advanced sound analysis tools.

B.1.8 Chromatic Tuner Component

The *Chromatic Tuner* component uses the *Audio In* service for real-time pitch detection. It implements the pitch-tracking algorithm proposed by McLeod and Wyvill (2005), identifies the detected pitch, and compares it to a reference frequency (default equal temperament with A4 = 440 Hz). An interface displays note names, cent deviations, and visually indicates whether the current pitch is in tune, sharp, or flat via an animated SVG pointer. Emoji-based feedback further enhances user engagement, showing "in tune", "slightly off", or "out of tune" states. The tuner can function as a standalone feature or be embedded within more comprehensive exercises to provide immediate and precise intonation feedback.

User Performance Tracking

This service leverages the *Firebase Realtime Database*[7] to log user performance data (e.g., number of errors) alongside multiple usage statistical metrics. An optional user registration system enables personalized tracking while preserving participant anonymity. The collected data can inform pedagogical research or guide the development of targeted application

[7] Google Firebase, *Firebase Realtime Database*. Available at: https://firebase.google.com/ (accessed December 14, 2024).

features, bridging the gap between industry-grade tracking tools and academic research goals.

Appendix C

Consent form: Evaluating efficacy and usability of music education technologies

Participant #:_______

McGill

INFORMED CONSENT FORM
COMPUTATIONAL ACOUSTIC MODELING LABORATORY
SCHULICH SCHOOL OF MUSIC, McGILL UNIVERSITY

REB 23-06-088
Evaluating efficacy and usability of Music Education Technologies

WHY ARE WE DOING THIS RESEARCH? The aim of this study is to understand how well a developed mobile application assists trumpet practice in learning fingering, and sound production techniques. We are particularly interested in how this software proposes interactive exercises on specific technical skills and helps in practice activities. The main goal is to make practicing more targeted, individualized, and effective. Ultimately, we want to see how these technologies potentially enhance motivation, student-teacher interaction, and the overall learning experience for trumpet players.

STUDY PROCEDURES: By participating in this study, we will provide you with a mobile app specifically designed to support trumpet learning. This app, named "MEI-Trumpet", is available for free download on the App Store for iOS devices and the Google Play Store for Android devices. The app is intended to teach users specific technical aspects of playing the trumpet in order to enhance musical skills through deliberate practice.

At the start of the study, you will partake in a 30-minute introductory workshop. Here, the mobile app will be introduced, distributed, and its functionalities explained in detail.

Following the workshop, you will be provided with the opportunity to use the app during your individual practice sessions over a span of 2-3 weeks. It's important to note that while we encourage you to explore and use the app, there is no obligation for you to do so. Whether you find the app helpful or not, or choose not to use it at all, provides valuable insights for our research. Your genuine interaction with, or lack thereof, the app helps us understand its usability and areas for improvement.

At the end of this intervention period, you will be asked to participate in an in person or remote focus group with other participants. In this setting, you'll have the chance to discuss and share your experience with the software. This focus group can take place either in person or remotely:

- **In-Person**: If the focus group is in person, the session will be video-recorded using a camera. If you're uncomfortable being on camera, you have the option to request not to be in the frame of the video, ensuring your visual privacy.
- **Remotely**: If the focus group is conducted remotely, it will be hosted on Microsoft Teams. You have the full discretion to decide whether you'd like to keep your camera on or off during this session.

Regardless of the format, the focus group will last approximately 60 minutes. The video recording, whether in-person or remote, is exclusively for research analysis to ensure all feedback is captured comprehensively. Rest assured; the video recording is solely for research purposes: it won't be shared publicly. Only the principal investigator of this research - the PhD candidate Alberto Acquilino - will have access to the footage, which will be securely stored on a password-protected hard drive, kept in a key-locked room at the Computational Acoustic Modeling Laboratory of McGill University.

At the end of the study, you will be asked to fill an online questionnaire to assess the usability and social validity of the provided software.

Regarding the app's data policies: it doesn't ask for any permissions like location or microphone access and it won't gather personal details such as phone numbers or contacts. It will, however, collect anonymized data on user engagement, like how long the app is utilized. This data aids us in objectively understanding how users interact with the app. To understand more about this, the app's Privacy Policy is available at the following link: https://github.com/albertoacquilino/mei-privacy-policy/blob/cf95cf2737f909b5ac441c99fb89e463e38a3c01/privacy.md

VOLUNTARY PARTICIPATION: Participation in this research study is entirely voluntary. You reserve the right not to answer any question or engage in any procedure, and you may withdraw from the study at any point and for any reason. Should you choose to withdraw during or immediately after the study, all data related to you up to that moment will be destroyed, unless you express a different preference during the withdrawal process. After the study's findings have been published, any data that has already been shared cannot be erased. Moreover, once data have been anonymized, they also cannot be deleted. I can, however, exclude your dataset from subsequent analysis and from being featured in future publications. The video recordings from the focus group will be retained for a duration of two months. During this span, conversations from these recordings will be transcribed, with participants' identities being replaced by anonymized tags such as "Participant 1," "Participant 2," and so on. After this two-month period, the video recordings will be permanently deleted. With the video recordings gone and data anonymized, I will no longer be able to recognize or retrieve your specific data.

POTENTIAL RISKS: There are no known or foreseeable harms or discomforts associated with your participation in this research study. In the event that you experience distress, resources for stress management will be provided.

POTENTIAL BENEFITS: The potential benefits of this study include the opportunity to make use of new tools to enhance your music learning experience and potentially improve your musical skills. These findings will be disseminated through academic publications, benefitting both researchers, and practitioners in the field of music education.

CONFIDENTIALITY: Your participation will be confidential. Video recordings will be stored password protected on two hard drives in a key locked room at the Computational Acoustic Modeling Laboratory at McGill University for a period of two months during which the contained information will be transcribed in an anonymized way. After that, video recordings will be deleted. The end-of-the-study online questionnaire is anonymous. Consent forms will be safely stored password protected in a key locked room at the Computational Acoustic Modeling Laboratory at McGill University for a period of 7 years. Anonymized and anonymous data might be shared externally with other researchers for academic purposes.

DISSEMINATION OF RESULTS: The outcomes of this research will be shared through academic publications. Though you will not be identified, your anonymized data and feedback may be used in these publications to increase the knowledge and understanding of how technology can support music pedagogy.

CONSENT FOR FUTURE RESEARCH: Making research data available to others allows qualified researchers to reproduce scientific findings and stimulates exploration of existing data for further research. To ensure confidentiality, any shared data will be stripped of any information that could potentially identify the participant.

Consent Statement:

☐ Yes

☐ No

You consent for your de-identified data to be used for future, unspecified research.

The Research Ethics Board II of McGill University has reviewed this study for compliance with ethical standards. If you have any ethical concerns or complaints about your participation in this study, and want to speak with someone not on the research team, please contact the Associate Director, Research Ethics at 514-398-6831 or lynda.mcneil@mcgill.ca citing REB file number 23-06-088.

PARTICIPANT'S STATEMENT:

"I have read the preceding details and agree to participate. I understand that by consenting, I do not waive any legal rights."

______________________ ______________________ ______

Signature Printed Name Date

I would like to receive a summary sheet of the experimental findings ☐

E-mail Address: ____

Alberto Acquilino, Ph.D. student, Schulich School of Music, McGill University alberto.acquilino@mail.mcgill.ca

Prof. Gary Scavone, Schulich School of Music, McGill University gary.scavone@mcgill.ca

Appendix D

Focus groups transcriptions

Transcription Focus Group 1

Legend:

- **AA**: Principal investigator.
- **P1**: Participant 1.
- **P2**: Participant 2.
- **P3**: Participant 3.

Transcription:

AA: Let's see the responses. Okay, so. Easiness of use. There is unanimity about how easy it is and at this point, maybe we aggregate the responses about both of the answers, like if it was easy to install the app and learn the functionalities. I felt competent to use the app for my own level. Do you have comments on it?

P1: No, I think it was very clear to install the app, actually that is not something difficult. And the app itself is very easy to understand how it works. I mean. There is no problem, no difficulty on that. So for me that was a good.

AA: Thanks. And for example – that's another question that I'm interested in – so we had one on one meetings to install the app and to learn the functionalities. What do you think it's the most effective way to explain the functionality like in a more scalable way, for example, I want to distribute it to a music school of 600 students. And they cannot have meetings one on one. And maybe we want the information to be embedded into the app. What do you think?

P3: Yes. Hmm. Yeah, because you'd already explained to me before we even got the app how it works. I already knew everything and then I installed it on my own and started playing before we had our meeting one on one. So I'm thinking.

P1: And you sent a YouTube video that was very clear also.

P3: Yeah, Yes. Yes, that's true. YouTube video, yes, is very good, yes.

AA: So a video at the introduction as soon as you installed the app that could be helpful. And, but, yeah.

P1: And I think even without anything, without the meeting, without the video, just installing the app and just going through it and just experimenting with the app. I think it's, it's easy.

P3: Lay with it. Yeah.

P1: Even with someone that never used an application or whatever, technology or what. I think it's kind of easy to understand.

AA: Thanks a lot. I will wait 1 min if you have other things to add.

P2: Oh, it's really easy and straightforward.

P3: Yeah.

AA: Okay. We can go back on that. So let's focus on deliberate practice. This is more interesting because we have reported different experiences. Let's focus on the first question: The app changed in my way of practice. We have one neutral, one agree, and one strongly agree. And maybe, yeah, we can go one by one or as you prefer, if you want to share your experience with that.

P1: Yeah, so for me, I began trumpet and music so one year ago now. So I was a beginner. So for me, yes, it changed my work practice. So now I'm spending a bit more time by using the app to focus more on long notes and try to practice more long notes and different notes before to start playing something else. So yeah, kind of changed my work practice for that.

AA: Thanks.

P3: I had hoped it would. I wanted the tone to be longer. So I tried bringing it down and down to make it a longer tone. It did help when I tried the high register because I want to get more into the high register and I put from G on the staff to G above the staff. And it was kinda funny, it didn't work very randomly. I would get maybe the G and then everything. GFGF E up on top and then. One A and then GFEF FFFF FFFF FFFF FFF FFF, FFF, and it seemed to just be the high notes. It was kind of funny: it wasn't very random in the scale. But, when I did it on a lower range, then it was very random. So, but I would have liked to have had it longer and I guess because I already play what I would have liked is to know if my tone and pitch was right, but that would have been a whole different thing and very hard to do.

AA: Hmm. What do you mean by longer? Like, to go below the 40 bpm?

P3: Yeah, I guess it was 40 the lowest? Hold on [*checking the app*]. Yes. I guess because the 4 bars of rest, so what I end up doing was just playing. There's a bar of rest and then you play and then you listen and then you play. So what I end up doing was playing in the bar of rest before the next notes to make it twice as long.

AA: Nice. Nice.

P3: Yeah, so I thought that would work. So that worked for me.

AA: Thanks.

P3: Okay.

P2: Yeah, I integrated it in my practice in my warm-up. So, it's an additional thing that I do when I'm practicing at the beginning, so it didn't radically change what I do, but it's another activity. And I don't know if I can say it helped to plan and set goals for my practice. So, it's part of my warm-up routine, and you know I finished with those long tones. And it helps me, you know, to calibrate my ear, but that's pretty much it.

P3: Yeah.

AA: Okay, so we pass to the following question. Yeah, if the app helped to plan and set goals for my practice. Please share your thoughts.

P1: Yeah, so for me, also it helped to plan and set a goal for my practice because, as I said, I am a beginner of trumpet. So of course doing that as a warm-up and, how can I say, like I mean a way to train my ear with the different tone of the trumpet. So, I think it helped a lot for that.

P3: What I found was interesting: I tried as well to do it by closing my eyes and seeing if I could figure out what the next note was and match the note.

AA: Okay, so, just to have a resume. What were your goals with the app that were most helpful for you? One is the ear training...

P3: Yeah. Yeah.

P1: Yeah, ear training and yeah, ear training mostly. And just yeah like practice the long tones. So, breath training, I would say too. And yes.

P3: Yeah, working on tone, working on tone, the long tones, trying to keep it nice and steady and in tune.

P1: Yeah. Hmm.

P2: And being able to hit any note. So, you know, the coordination between what I hear, the sound I want to reach and what I actually do, you know, my breath, etc.

AA: Thanks. Let's go on motivation. Let's see the results. So, of course there is overlap between these questions. So, I try to minimize the participants' fatigue, but still I had to compromise. So, let's focus on the first. Using the app felt like an effort to me. If you have comments on that.

P3: I didn't feel it wasn't effort at all.

AA: So you normally when you go to practice, you have your phone with you, phone tablet or anything.

P2: Yes, I have my phone because I have my tuner and the metronome and my phone so it's just another app on the phone.

P1: The same for me.

P3: No, yeah, I have a mechanical metronome and a mechanical tuner. Yeah.

AA: So did you have Participant 3 like to remember: "Oh, I need to bring my phone for the practice?"

P3: No, no, no, I was, when I used your app, it was focused to use your app and to practice using your app. I didn't do it as: "Okay, I'm gonna go practice. I'll go get your app!". It was more like: "Okay, I'm gonna go use your app to analyze your app". And I'd have my phone. So, I did it kinda not how you intended, but yeah.

AA: No no, I mean no problem.

P3: And plus the timing because I was away for a week in Dominican Republic and the whole holidays. There was very little trumpet playing going on. So that was another problem: the period.

AA: Okay, other comments or we switch to the second question? So, motivation: the app motivated me to practice effectively. We have a strong agree and agree and a neutral. I like that. Let's dig more into that.

P1: So yeah, for me, I'd say I spent more time to practice the note with the app and to focus on the notes and everything before to start to play something else. So yes for me it motivated me to do something to reach the right tone every time before to play.

AA: Thanks a lot.

P2: It didn't change much for me since, you know, it's one of the things that I do during my warm-up routine.

P3: Yeah, me too. I was kinda neutral, yeah.

AA: Okay, okay.

P3: And I guess I could say the app motivated me to practice a bit longer. Yes. Cause then I try different things on it.

AA: Okay, so let's switch so then to the third question.

P1: So yeah, for me using the app adds just a bit more time of practicing the instrument, which is good because it is fun for me. So I play, I add more time every time of practicing, before to do something else. I would say so.

AA: Yeah, yeah, yeah. So that routine was not in your routine before.

P1: No, so I played like a certain amount of time and I do something else, and then I add practicing with the app before my routine. I would say this allowed me to practice more.

AA: Yeah, yeah, yeah. No, it's interesting. You know, someone is like: "Okay, I fully integrate the app in my routine". So it's like not changing much the practicing time. While someone is like: "Oh, this is an additional thing", and they are different perspectives, both valuable. And that's very interesting to have you to have this discussion.

P3: Yeah, I guess I know I should be doing long tones and I have no time, so I skip them. Because I am in 5 different bands and there's too much music to learn for each week, so I just practice the music and so this forced me to do the long tones before, so that was good.

AA: Okay, okay. And last session. So perceived usefulness: I made more progress than I normally do in 2 weeks of practice using the app. One neutral, one agree and one strongly agree.

P1: For me, I think I did more progress especially in the high notes that I didn't really practice a lot before because I tried to play the music we have in the band and there are not a lot of high notes. So for that I think I did more progress than I normally do.

P3: Yes, I think it helped me more with the high notes as well. Yes. Yeah.

P2: Yeah, it helped me well also with the high notes that I don't reach too much and forced me to practice and, you know, play them thoroughly. And also, since I do it at the end of my warm-up routine, it helps me make sure that the feelings are here and that I don't make any effort to reach the notes that I want to reach. And that my lips are buzzing in a great way, etc. So for me, it's a way to, you know, finish and wrap up, you know, the warm-up and the setting, I'm putting myself in before I start playing.

P3: That's the end of your warm-up, not the end of your playing. Okay, that makes sense.

P2: Yeah.

AA: Okay. Yeah. And so, yeah, let's focus on the last one: I feel the app could be useful to be integrated in a classroom.

P3: May I ask you how you meant this question? Do you mean it as in a school like where the students are just starting to learn their instrument and you give it to them to practice at home or you're actually going to use it in the classroom and the students all play together the notes after they hear it is that how you mean?

AA: I mean more like in music education, so in the classroom in the sense that teacher could use it. So that question should be more clearly formulated, thanks.

P3: So like teacher one on one or like a whole school like you know when I started playing?

AA: Also in whole school, but more like, if the app could be integrated in the classroom teaching. That was the sense of the question.

P3: Okay. I guess when I saw, like when I started in high school, that was the first time we had a classroom because you had the whole band. So because it's very trumpet oriented, then I don't know how it would work. I guess they can all, on their instruments, listen to the note and just ignore the fingering and try to match the note. That might be useful. Yes, yeah, because I find that, when people start, some of them are tone deaf, and they don't know they are tone deaf, and they weren't put on a clarinet. So, it's hard for them, especially French horn: it's very hard to know what note you're playing on a French horn. So, I could see that being very useful, because I started French horn after I played trumpet and I was ready to quit after a month and a half because I could not tell between my C, my E, and my G, and there was so little lip movement that I couldn't tell what note I was playing if I'm by myself. But, if with this app, then yes, I could have heard and it's like: "Oh no, I'm in the wrong register". So I would think that would be useful, yes, for like a French horn and maybe yes in a classroom for when people are trying to play for the first time to children then yes, because a lot of them don't have an ear or haven't a trained ear yet.

P1: In the classroom, it depends on the number of students because if they are a lot, it could be a bit hard to have all of them playing the same note at the same time if they are not experienced.

P3: But that's how it starts. That's how it is in grade 9. It's exactly: you all have a rest and then everybody plays the note for 4 beats and then everybody rests for it and everybody plays a note again. That's how you start. That's how it works in the classroom, and it works fine. Hmm.

P1: I think it's a good app to integrate in a class or, just like you said before, just like we did so at home as a practicing homework. I think it both can be very good to use.

P3: Yeah, I think so as homework for the new students, I think that would be very good. It might motivate them more to try it. Especially since new children all have phones and like playing with apps.

P1: Yeah.

AA: Okay, so I would open for more general feedback. So I would reconnect to what Participant 3 said, given your experience, if you think it can be helpful also for other musical instruments to be applied.

P3: Yeah, definitely for trombone it would be another one that would be very good because for the movement of the position, it'd be very hard, so if they know how to match the note then they know: "Oh I got to move my slide!" So I think it'd be very useful for like French horn and trombone. For myself, I would have loved to have had feedback if my tone is right, but I could just put my tuner next to it so it's okay.

P1: I think it's something you can do. Maybe it's more difficult because you need to integrate the microphone.

P3: And the quality of the microphone, does it get all the frequencies... because you have the same when you're playing a note or music on your laptop. It's awful because it doesn't have the low frequencies and the speakers are not good. But I think it would be very useful for somebody who is learning the instrument. Yeah, definitely.

AA: Okay, okay. And in terms of functionalities, did you use the whole functionalities? What is the worst one that was not helpful?

P2: I didn't try sharps and dynamics.

P3: No, me neither, cause I usually do that anyway. Yeah. Yeah, I do try.

P2: Yeah, I didn't use dynamics. But I used the other with the sharps and things. But no, I didn't do the dynamics because first I'm in an apartment, so I can't really play loud.

P3: You can put a mute. I used to play in the bathroom whenever I was in apartments because I figured nobody sits in the bathroom that long that they'll be that irritated. So, because the floor is above and below all the same thing, they'll all be in the bathroom. That was always my strategy in apartments. Acoustically, it's not very good, but...

AA: But that's also because you found that this was not very helpful, the dynamics, or that it can be structured better?

P3: I'm just looking at the app again. So you have all the speeds, yes. No, I'm trying to figure out where did I get. How do you get the dynamics now? I don't see it anymore.

AA: It's in the options. You go to the options you can hide.

P3: Okay, hold on. Yeah, there it is. Okay, now, dynamic, and hide trumpet.

AA: Okay, okay.

P3: Yeah, that's right. I think the hide trumpet is a nice little feature for them to know if they remember if they've got it down path and I thought, yeah, that's kind of a nice feature, actually, for new students. Yeah.

AA: And do you foresee like other features? So, you were mentioning "I wish that to have longer tones", so in this way like how would you do like to integrate that?

P3: Yeah, I guess. Instruct it to say fortissimo and then go back to pianissimo or yeah. And I would have liked to have the feedback, because sometimes when I'm playing, my ear is not the greatest. When I was a child, I was completely tone deaf. But I'm much better now. My brother had very good pitch, but I didn't. So, like when you play the high D, and like I don't have to think: "Do I have to pull up my slide or not?" Like sometimes, you know "Am I sharp?", "Am I flat?", so it'd be nice to know like on these other ranges how much I need to pull it out. Like the low D, I have it set, but if we've had the feedback, then I would know "Okay, I'm always sharp on that note". Or "I'm always flat on that note". Or "Is it the mouth? Is it my embouchure, or is it the instrument that needs to be changed?" you know, with the first valve slide.

AA: So as a feedback more like a tuner.

P2: Yeah.

P3: Yeah, yeah. Saying you're sharp, you're flat, yes, you're nice and you're the same pitch as what you heard. That's a lot, that's a whole other project for you. That's just for me that would have helped but yeah.

AA: Yeah, no, of course, of course. Everything is possible.

P1: Yeah, I don't know if it was the same for you, but for me, I think it stopped after like around 10 different notes. Then it just stopped. Is that normal?

AA: It was programmed like this. It was programmed like this.

P1: No, okay, okay, good to know.

AA: So, would you remove this stop and just go on and go on?

P1: Yeah, until you want to stop.

P2: Yeah, or something that the user can set.

P3: Yeah. Oh yeah, okay. Yeah, because it's nice I think as you use, you can set because then you go: "Okay, I've done my so many notes, right, so many minutes, so then I was like on to the next thing". If you want to do 20 or 30 a day or yeah that's a good idea.

AA: So based on that I don't know if you would combine feedback with this. What do you think? I mean, a gamification, like how many notes can you play right without making errors? Could that be a motivation?

P3: That's true, yeah.

P1: Yeah.

AA: Do you think like this gamification could be helpful for you?

P3: I don't know. Hmm.

P1: Yeah, so you kind of, if you can at the same time, change the app to kind of a game that can be nice for beginners, no? It's true! I have to reach the note at that beat and start to make it always perfect. It can be nice, especially for beginners, you know, I think.

P3: Yeah, in that way 10 would be kind of not bad and you see if you can get 10 out of 10. Yeah, if you do it as a game.

P1: It's a kind of a reward that yes, I made the 10 perfect.

P3: Yeah. No, I think it is very good and the tone that we hear is very nice, like of the instrument. I know you said you had it recorded by a professional, so the quality was very good, so it was nice too. Yeah.

AA: One of the objectives of course is to learn appropriate fingering, but again, you are not that beginner. But I feel like one of the intended purposes was to help users create a mental imagery of the sound before playing it.

P3: No, and that's very good. That's very good. Yes. Yeah, that's very important. So you can imitate that sound, yes. It's like when you get a tuning note. Right. Same thing: you listen to it first in your head and then...

AA: And based on that, that's a feedback I received, would you invert: so it's like rest, you play, and then you hear. Do you think it could be helpful or not?

P3: Oh, the opposite. Oh!

AA: This is a functionality which you would have liked? Or...

P3: I'm thinking. It might be an interesting option that you could select either way. But I'm just thinking, okay, you play the note and then you hear it after, then you: "Okay, was it sharp or flat?" But then you have to know the note ahead of time before you play it, so you'll get it in your head to learn it. So it is just copying. It might be interesting, yeah, I don't know. What do you think Participant 2?

P2: I don't know how I would integrate it in my routine, so I don't think I would use it.

P3: Yeah, I don't think it's useful. Yeah.

P1: And if you put the other option with the feedback, I don't know, this one would be more helpful.

P3: Can you tell us, Alberto, why did those people that suggested it—what angle were they looking at? Why did they suggest it?

AA: I think it's to foster like: "Okay, I play, was I playing the right note? Let's check". Like more as a check.

P3: As opposed to playing the note first and imitating the note. Yeah. Okay, like that they might have been in the wrong register and then they hear the note and they're like, oops, I was wrong? Is that the kind of thing? Or, I played the right register, but I'm just slightly flat or sharp? Like, what were they thinking? They're completely off in the wrong register or. . . ? Yeah, it might be kind of interesting. I'm just thinking the French horn —- if you do it after, they would think: "Oh, I got it wrong again and I got it wrong again". I don't know if that's too discouraging. As opposed to, you get the note, you imitate it like, okay, that's how I get the G, that's how it is, okay, that much pressure. And then, yeah. No, I can go either way. Did you get any more comments from people about features, Alberto, that we can discuss?

AA: So, this is the first focus group. The comments come from a presentation I did at university with my colleagues. And one of them is a trumpet player and he was proposing this idea. But you know, that's why I'm doing this because a lot of people talk about how to develop educational technologies, but then there are not a lot of studies with participants that include open discussions and reflections.

P3: Yes.

AA: That's why I decided to change this format. I realized -— I don't know if you remember your answers —- that's an error of mine. I was thinking: "Oh, I would have liked that you could collect some demographics". So, I don't know if you remember, but I was thinking, oh no, it could be helpful to know demographics in the sense, like for how long you have been playing the trumpet, or how much of a beginner you are, and then associate that with the answers. But I should have done that earlier.

P3: Well, that'll be a good improvement for your next study.

AA: Yeah, exactly. For the next focus group, I would send that.

P3: I've another thing I thought that might be a nice suggestion, which is alternate fingerings.

AA: Yep.

P3: That you have, like for the A the first and second valves, but you didn't show the third valve. Because then sometimes you're doing, you know, A flat to A. Did anybody ever bring that up? That might be beneficial, or did you think that might confuse beginners?

AA: What's your opinion?

P3: I find it very useful. I think at the beginning, yes, give them that and then maybe you have an option for alternate fingering or something. I don't know. So that they can learn it because I think the third valve is very useful for you when you have to do things very fast. If you're doing one and 2 and then 3, 2 or 3, one and 2, 2 and 3, then it sounds muddled, but if you just have to flip the 2.

AA: Yeah, yeah, yeah.

P3: It's just a thought and in the high register there's different things as well too because I remember I was playing a note and everyone says: "Oh, that's completely off, use this alternate fingering".

AA: Yeah, yeah.

P3: And so, in the D or E-flat sharp or something like that, I had to use an alternate fingering and that did sound much better. It's a suggestion.

AA: Yeah, yeah, yeah. No, my thoughts on alternate fingerings—there are some that are really used, like 1+2 or the 3.

P3: Yeah. I wouldn't do ones that are obscure and you don't need very often in the music.

AA: Yeah, sometimes it's more like a refinement. Like, I don't know, the G over the staff, you can play it with nothing, with 1+3, with just the 3, or with 1+2. But again, you just play normally with nothing pressed and you would go, but that could be an impression.

P3: Yes.

AA: So maybe I can, like, again, set like a random thing to show just the 3 or 1+2 for those notes.

P3: Yeah, like when it gets higher up, for a beginner I wouldn't want to confuse them each time that's like: "Oh, I have the wrong fingering". Cause they wouldn't remember, you know what I mean? I would do like the third valve for the A and the E, maybe start with that as an alternative option. And then it would be interesting for them to hear because it doesn't sound exactly the same as 1+2. So then they maybe hear it and it's like, oh, it's not perfect, but if I play it fast, it's fine.

AA: Thanks a lot.

P3: Alright. You're very welcome and good luck. With your PhD. Yeah.

Transcription Focus Group 2

Legend:

- **AA**: Principal Investigator
- **P1**: Participant 1
- **P2**: Participant 2

Transcription:

AA: Yeah. Okay, here we have the answers: Easiness of use. I would like to go through the answers and discuss about it together. That's more informative, so I will have both quantitative data and qualitative data, getting your opinion. So, easiness of use: installing the app and learning its functionalities was easy. And we have agree. Do you have comments on that?

P1: For me, I think I didn't put strongly agree because I think when I tried to put like the lowest register note like sometimes I would put like F sharp but then when I press play like it wouldn't be F sharp. I think there's like a like a little glitch when I try to use the low register. But otherwise, yeah, it was completely fine for me other than that.

P2: Yeah. I think I had the same problem also. Yeah, but it's not bad, you know, I guess it's only the way I scroll, but it was not that big of a problem.

AA: Let's go on the second one: I felt competent to use the app for my own level.

P1: Yeah, I think I put strong agree for this one. I think it was okay for me, like it was easy for me to use.

P2: Oh, I guess that I didn't see strongly agree, but you know, well, I kind of agree. Yeah. I guess it's really easy to use it for me. Yeah, I could strongly agree also. I'm sorry.

AA: Ok, yeah, I should have explained it better.

P2: There's a strongly agree what was on the very far right, yes, and I don't remember seeing it. Sorry.

AA: Oh, okay, no problem, no problem.

P2: I could I could say, I strongly agree for this one.

AA: Okay, okay. Let's go on deliberate practice. So the app changed my way of practice. We have two neutral.

P2: Hmm. Yeah. For this one, it's difficult for me because I have a private course, a trumpet course every week. And I know that my practice changed a bit, but I don't know what is the influence of the app versus the influence of my course, you know. Because, so I know that now I try to do longer notes, but I'm not sure if it is the app or, you know, my teacher saying that it's very important to do long notes very often? So, you know, it's difficult to differentiate the effect of my course or the use of the app.

AA: Of course.

P1: And for me, I guess I picked neutral just because I think I kind of have like a systematic approach to my practice and I felt like I've been so ingrained in my way that like maybe it didn't really change. I can prefer like doing the long notes the way I usually do it and then the scales and then I just go into the songs.

AA: Okay, okay. Perfect. Let's go with the following one: the app helped to plan and set goals for my practice. And I like it because we have an agree and we have a disagree. So let's go feedback.

P1: I think I put disagree for this one. I felt like, I mean, I think the app for me was like, was like good, like warm up, but I didn't think it was beneficial enough for me to like set goals other than practice like the range. Because I know my range is okay, I just need to like warm up before I can get there. So for me, like there wasn't so much of like a long-term goal, it was more just like a tool for me to just like warm up, I would say.

P2: Yeah, well, what it helped me with is the long notes and listening to the notes, you know? So it helped me, you know, just to take time to listen. That's what I wanted to do with the app. And my teacher says to me that I have to really listen to: sometimes close my eyes and listen to the notes and take time. And this app helped me to do this.

AA: Perfect. Just a curiosity: were you using the loudspeaker of the device or were you using headphones?

P2: No headphones. No. Yeah. Yeah.

AA: No headphones: so the loudspeaker open?

P2: Yeah.

P1: Yep, same thing, just the loudspeaker.

AA: Perfect. Let's go on motivation. I like it again. So, using the app felt like an effort to me. We have an agree and a disagree.

P1: Well, I think for me, maybe my attention span is short, but basically I think like when I use an app it's like: "Oh, I have to find the app", it's like: "Oh, I have to find the app, I have to open it, I have to..." it's just, it's very little, like it probably takes like a few seconds. I wouldn't say like it's that much of an effort, but I turned to like my motivation. Yeah, I took a little bit more.

P2: Well for me it was just a little a short kind of break into my practice, you know? So it wasn't an effort because yeah, it's like a short break. And I have to, I can, relax and listen. And I didn't use it for 30 minutes, you know, it's quite short. So, I felt it like it was like a break, a good one!

AA: Yeah, yeah, yeah, I mean, given the experience with participants, it's very nice to see also how they use the app. Because some participants, they add the app to their routine to their work. Others, they like to integrate. So, it doesn't change much. It's just something that I can use to do something that I was doing before already. So, depending also on the approach. The answers are different, no?

P2: Yeah.

P1: Okay. I think the other reason why I put like the app felt like an effort to me, I think also kind of like the sliding, like when you have to pick your range, sometimes it was like slightly more difficult to like pick the exact note that I want because like sometimes it like goes too fast or like kind of a back-and-forth type, sliding scale. But otherwise it was okay.

AA: So probably, you know, I should modify this function because the app starts that only selects in the natural scale and all with no flats and sharps.

P1: Yeah.

AA: And, adding flats and sharps is an option that you have to choose. So for example, when you go all the way down and you put F sharp and at the end you will see G because that sharp is not a choice, but that can confuse.

P1: Yeah, or maybe I was referring to like—oh, go ahead. Sorry.

AA: You know, and these can confuse that and probably I should change that. That's the way.

P1: Okay. I think I was referring more to like the sensitivity of moving the notes up and down and selecting because sometimes I felt like it goes too fast. That's like a personal preference.

AA: Yeah, yeah, yeah. It's a functionality. Of course, I can try to select like a sort of sensitivity parameter. That's good, that's good. Let's go to the following: the app motivated me to practice effectively. And again, we have a disagree and an agree. Let's discuss! Good.

P2: Well, I think I'll say the same thing again and again, but, you know, listening was so important for my teacher and I guess the app reminded me to listen more when I play, when I do my exercises. So...

P1: Hmm. Yeah, for me, yeah, I think the app is a good tool, but I don't think it like affected my motivation as much. I felt like, I don't know, maybe I'm just so busy that it's like: "Oh, I have to practice!" So maybe it just comes from that, but I don't feel like the app really pushed me to practice. I mean, it was a good tool to do ear training, stuff like that, but yeah, I don't know what else to say.

AA: Yeah, of course. I think that also depends on the level of the musician, I would say. It depends on, you know, also the tuner, no? For example, sometimes it's useful to use it, sometimes it is not. Also too much, it's not good, no?

P1: Yeah.

AA: Because they also want to be independent and it will be nice to check all the data. And thanks a lot, thank you very much.

P2: Okay.

AA: Let's go with the last one of motivation. So: the app motivated me to practice for longer periods of time. We have a neutral and an agree.

P1: Well, I think I put agree for this one because I think like once I start using it, it was kind of fun. Like I tried to test myself with the ear training. And like also by using the app, I kind of like extended my practice a bit longer by default, so that's why I put agree.

P2: For me it's hard to say because I have so much practice to do these days: I have many pieces to practice because of the Harmonie New Horizons and I'm also in the lab band at McGill, so I think I have 12 or maybe 13 pieces to practice and I have the scales to do. So, I don't know if really the app made me practice for longer or if it's all the pieces that I have to do these days. Because I practice around, you know, sometimes 1 hour a day, maybe sometimes 45 minutes, but I'm not sure if the app is the cause of the, you know, the practice I do. It's hard for me to say because I have the Harmonie every week and the lab band 2 days a week, so I had to practice because of that. So I don't know. And I have a private course, so I'm not sure if it's that the app did motivate me to practice for longer because it's already a long period for me to practice.

AA: That's very good. Let's go with the perceived usefulness. So: I made more progress than I normally do in 2 weeks of practice by using the app.

P1: Hmm. I think I put disagree, but maybe it's more of a neutral answer. Basically, I couldn't really tell whether I made more progress or not within the 2 weeks. I felt like it was about the same, like I could tell. So I just disagree with the statement, but maybe my answer is more of a neutral. Like I could, I would say it's about the same.

AA: Okay.

P2: Yeah, I agree. It's hard to know because sometimes there's so many things that can make me progress. It's hard to say, but I think, listening! I'm back again with that, but I don't have many chances to hear good trumpet notes, you know. Apart from my teacher, which I see once a week and he doesn't play that much when I'm with him. Hearing a good note on the trumpet is quite rare for me. So the app is doing that, you know? So, I guess that, knowing what's a good sound, maybe, that's the progress I think I made. And I'm hearing more when my sound is not good, you know, I hear that my teacher says the wind that I do, you know, with the note: "The note is windy!" But now, I hear it more like that. Maybe it's because of the app, I am not sure, but maybe it is, you know.

AA: That's very good. No, I mean, again, it is difficult to know because it's qualitative and you don't know if you would have done the improvement anyway without the app or maybe not. So this is just something that is perceived. And that's super nice.

P2: Yeah. Hmm.

AA: Let's go with the last one. So: I think the app could be useful to be integrated in a classroom.

P1: Yeah, so for me, I think I put strongly agree because I feel especially for students who are like beginners and especially with like the technology these days and then incorporating the use of apps like the phone, I think it could be like a good motivator for them to learn notes. And I think also not only notes, but also the fingering. I think that could be really helpful for like a beginner level who's starting off maybe in like middle school or even really high school. So I think it'll be good, worthwhile in the classroom.

P2: Hmm. And I think, again, I didn't see to write the option strongly agree, but I could strongly agree also because, I would say, when I started I was looking for something like that to learn the fingering. That was very hard for me and, if I had this, I would have been very happy. And I think students would like that a lot. It's so easy to use for me, so it's something that could be very helpful, I guess.

P1: Yeah, and I know like a lot of students, like if they have like a sheet of music in front of them and they don't know like a particular fingering, it's easy for them to like just put their phone up like on the side and just try to like see what their note is and what fingering it gives them. I know a lot of my friends like in the past, they would always have to like search up on Google and like wait for it to load and find like the right chart to find the fingering. So, I think it's an easier way for them to learn.

P2: Yeah. Yeah.

AA: Thanks. I would open to more general feedback, I mean. Like overall experience and maybe some questions.

P1: I think for me, I think the app is very like well made. I think if you like want to like expand the app, like I don't know if that will be the point of the project, but like if you add like rhythm stuff to it, like I'm still like long notes and change the duration of the bars or something, if you play around with it, I think it could like be better for learning rhythm as well. But it depends on what the end goal of the app will be.

P2: And also, maybe it's not useful, but would it be possible to have the note written just under the trumpet when you play it? You know, the name of the note, you know, Do or C or, you know, just visually to get what is the note to associate the name of the note with the sound and the fingering. Do you understand?

AA: That's very nice.

P1: Okay, yeah, okay, I see, yeah. Yeah.

P2: And also the other thing I saw just lately that when we have to slide the third slide when we do a D, the indication is quite small, and I didn't see it the first time. Maybe it could be bigger because I missed it.

AA: Yeah, yeah, yeah. Yeah, it's D and C sharp. And so probably redesign the trumpet so that it's more evident.

P2: Yeah, it's too small, you know.

P1: Yeah, but I think what participant 2 said with the notes, especially for learners, like having like solfege, and also like the concert note and maybe, I think would help them to learn the notes faster.

P2: And also, I have the notes in French: Do, Re, Mi, Fa, Sol, La, Si, Do. And I would like to learn the English notation. So, if we could turn on the English or the other notation, it could be nice to learn for someone who wants to learn English notation. To have this functionality also.

AA: Yeah. Very good. These never came up and it's super nice. And it's easy not to implement also.

P2: Okay. Yeah. And also, it's just a little thing like that: I play a little game with my trumpet teacher, you know, just for fun, but ear training also: He plays to me, let's say 5 notes random, and I have to reproduce the same thing. For me, it's hard! And it's quite fun and I think it's very important, and he adds notes, you know, 6 notes, 7 notes and I have to reproduce each time. And I was wondering if there would be an app that would do that, you know, just randomly plays 5, 6, or 7 notes and we have to reproduce, just like that. It's kind of an idea that I had, and I would use it, you know, just to play a bit.

AA: Did you use the app like, for ear training? Like not looking at the display and just listening to the note and trying to hit that note.

P2: No, I didn't, but it's a good idea. Yeah, well it's the same as I would say if we could group a few notes.

AA: Yeah.

P2: It would be best because it's easier to know what is the note if we have another one just close, you know? If we have a group of notes, I think it would be fun. It would be easier. So, for me, it's an app that helps a lot with ear training and fingering also, and having a good sound of a note, played by someone who is a good trumpet player, you know? I don't hear that much, so I have to have a guide to know what it is supposed to sound like. What the notes are supposed to sound exactly?

AA: Very good. Yeah, that was my question, like also: What is the best functionality? And I guess you mentioned fingering and also you mentioned a lot about listening to the notes.

P2: And also I was wondering, I started to learn the fingering for the main notes without alterations and after a few years I started only to practice with alterations: flat and sharp notes. And these are the notes I have more difficulty remembering their fingering. So, I was wondering, would it be possible to have an option that would play only sharp and flat notes? Just because these are the more difficult for me and for many people, I guess, to remember the fingering. So only flats and sharps. Because they don't come quite much, I guess; when I select the range, I know it's random, but sometimes I don't see many flats and sharps. So I would like an option just for flats and sharps.

AA: And just as an idea, would you do mixed flats and sharps? Or maybe only sharps, or only flats?

P2: Yeah, mix, mix. Well, that's a good question. I would definitely do mix, but maybe separate also? Could it be possible? Yeah, it depends on how you decide to learn. Because I remember I learned only the flats first and then the sharps.

I don't know why but I was very systematic like that. So, I didn't mix at first. But now I'm in the phase in which I'm mixing both.

AA: Very good. And the worst functionality? Like something that you didn't use or that was not good. So you mentioned the GUI, like swiping, was not the best? And it could have been improved.

P2: Is there a way that, you know, each time we start, we push on start? I guess we have a total of around 10 or 11 notes? And then it stops. And then we restart.

AA: Yes.

P2: Is it possible to have more? Because we always have to click on start again and maybe we have to stop when we want to stop, but it would go over. You know?

AA: Yeah.

P1: Yeah, that was gonna be my comment. I was like, maybe you could make it like an infinite amount and then we can just stop it whenever we need to. Because otherwise we just have to like redo it again and then it's like an extra step.

P2: Yeah.

AA: That's good. That also came up from other groups. So that was a bad idea of mine. And, yeah, my other question is about the dynamic functionality; if you ever used it, because it was not helpful at all?

P2: I didn't use it much, I guess. No.

P1: Yeah, I don't think I used the dynamic. Oops.

AA: Yeah. But do you think it's because it's not effective or because it was already good, what was proposed?

P1: Hmm.

P2: Hmm. To me, it's like something that I can understand what it is. So, I don't need to practice this. Unless it is in the score. No, I don't know how to say that but on an individual note, it's not to me that helpful.

P1: Yeah.

P2: I need to have a score with many notes and dynamics. There it's more helpful. Well, you know, there's a big difference to me between on one note and in the score. On one note, I don't think it's that useful for me.

P1: Hmm. Yeah, I think, correct me if I am wrong, but the dynamics were just like the letters like mezzo forte, forte, piano. There wasn't like any decrescendo or crescendo on it so I think to be slightly more useful for me, I think it would be beneficial to have like a decrescendo mark so you can practice —- since it's long tones anyway -— just practice quiet to loud, loud to quiet.

P2: Yeah, maybe if you would integrate an option where you could generate 4 or 5 notes. And then I would put something like that. Under only one, I don't think it's very useful for me.

AA: Yeah. It needed to be put into a musical context. Or to add some time variations like Participant 1 was saying so that you can see the range and you can develop the range.

P2: Yeah.

AA: So, that was probably not effective, one of the worst functionalities. And about hiding the trumpet? What do you think about the option to remove the image of the trumpet?

P2: Yeah, I think this is useful if you want to learn the fingering. But for me, if I want to learn, I would just not look at the trumpet. That's it. So it depends on the people. If you want to learn the fingering and you think you will cheat, so hide the trumpet. But to me, it's not that important; I won't look.

P1: Yeah, yeah, I agree. I think it's something for like people who are learning the notes, but for like, someone like myself who can play for a while, just the notes are okay.

P2: But to me, a more clear image of the third valve slide is important because I'm not used to doing that and I forget. I'm re-learning to do that. So that is why it is important to see it. So maybe if you don't have the trumpet, I would not think of doing it. The sliding for the notes is a big problem for me.

AA: Oh yeah. Another question was about possible improvements. You gave already a lot of inputs, so of course I mean I won't bother, but if you have anything to add... I mean, I have the recordings so I don't need to list because I can go adding all these improvements that you already said. If there is anything else you would like to add. Or...

P2: Hmm. No, I took a bit of notes, but I've been there. So, yeah. Nope.

P1: Yeah, I think, what we said already was covering everything.

P2: Yeah.

AA: Thanks a lot.

Transcription Focus Group 3

Legend:

- **AA**: Principal Investigator
- **P1**: Participant 1
- **P2**: Participant 2

Transcription:

AA: Okay, yeah, so I will share the screen. So the idea is that we watch together the accumulated answers. And I was thinking in this way they can guide the discussion if you can comment on the different questions. At the end, there will be like open feedback. So, I would start with the easiness of use. And so the first question is: installing the app and learning its functionalities was easy. So, we have an agree and a strongly agree. You want to add your experience?

P1: I think like as an app in terms of downloading like that's normal it was easy. I at first couldn't figure out how to change the note by dragging. I didn't find that intuitive. Part of me wanted to just type in like tap it to have the note move to that line or that space. But like I figured it out so I still said agree because it didn't take long but I didn't feel that intuitive to me, in that sense.

P2: It was the same for me. Also, if I noticed that if we choose a note that had a sharp or a flat to it, after choosing a B flat, it may show up in the main menu as a note without the sharp. So, it doesn't throw it off by a lot, but still, like it would be nice to show the exact note that you picked for either the lower or upper range. But I agree with Participant 1 like, it wasn't too intuitive when selecting note. It could be made just a little easier.

AA: Yeah. Thanks. And by tapping, do you mean like some arrows, one up and one down?

P2: It could, yeah. That'd be good.

P1: I think that would be good or like what I tried to do initially was just like: "Oh, I want it to be a G".

AA: Ah okay. Tap the G, tap the note. Oh.

P1: But yeah. But I think that would maybe be difficult because the lines are so close together. So arrows would also be helpful, I think.

AA: Yeah, yeah, yeah. That's good. Yeah, that's most of my problem in GUI. But that's very good feedback! So next one: I felt competent to use the app for my own level. We have agree and a strongly agree.

P1: I think there are ways to adapt it for different levels. I used it like more to focus on tuning and stuff rather than just range or like understanding the notes. So, I think that focusing on the more nuanced aspects of practice was how I adapted it to not being a beginner.

P2: Yes, it was more on the pitch. It was very accurate, played by the sampled recording. But Alberto, for this, I don't know if it relates to this question, but is it all strictly in whole notes? Would there be an option one day like to implement half or even quarter notes?

AA: Yeah, that could be an option. Before, it was like an image. Now it's generated by an API. That's interesting. So my idea was to develop an app that is as simple as possible, that is following criteria that are in literature. So if you look in the academic literature, there are a lot of papers that explain how to do educational technologies, and I was trying to do something as simple as possible following these criteria. So that hopefully like especially this one, that is the parameter of usability easiness of use, most of the people are towards agree and strongly agree. I never had strongly disagree. So this is not a barrier. So the discussion that I am having with you is very helpful in the terms that there is not a lot of academic literature with that, and so the feedback I'm having with participants is to trace new guidelines for future research. So that's the gap I am trying to fill. That's also a possible improvement. So at the end I will ask for other questions. Thanks. So, the app changed my way of practice. So it's that we have a neutral and an agree. Can you comment?

P2: Maybe I didn't view this app as a beginner, but it is really just a note that is played and probably I'm not to say I'm a very good player or even good at all, but I think this app would be just great for the warm-up but not for the entire practice for myself. But for a beginner, I would say this app is very beneficial. It gives them the ability to hear the exact sound that has to be played and obviously starting with whole notes. But I think it would be great if it offers feedback on the sound that the student makes. Because right now, it would be up to the student himself or herself to decide whether his or her pitches are accurate or not. So maybe like if there could be a recording feature by the app and play back to it.

P1: Yeah, or even like a built-in tuning feature. That's an interesting idea.

AA: But I mean, in which sense the app doesn't change at all your way of practice? Was it like you integrated it in your routine, it took the place of something else that you were doing? Or you were not using it? Or. . .

P2: To begin with, to start off the practice, I think it's great. But beyond this, intermediate players or even beginners would probably explore something else than playing fixated whole notes. But, I don't know if I worded it correctly, I apologize for that.

AA: Oh, no please.

P2: In terms of variety, I think it's very limited for now because it is simply whole notes that are being played. And if there would be some new additions that could replace certain beginner methods, it would be better. But right now it's just one note that's being chosen and if you don't even change the range, it's actually that one single note on the G on the second line. I think a beginner and intermediate players will be playing something else during their practice as well. But this app, I would say, it's beneficial at the beginning of the session maybe, just to warm up, just to play in tune and that would be good.

P1: I think I had a similar experience. I think I answered neutral or maybe agree, but I guess to answer both questions in one [*about changing and planning and setting goals*], I didn't really use it to set goals, but it did help me plan and it changed a little because I did a longer warm-up than usual. I'm usually a little bit lazy with the warm-up. So, it was nice to have like a structured warm-up and something different than the other warm-ups that I've been using for a couple of years. To add in more ear training I guess because even though you hear it you have to feel it whereas most of my warm-ups are scales so you're not having to jump between notes as much. So, I think in that sense, it changed my way of practice, but it was replacing my other warm-ups. So in that sense, it didn't really change, like I used it as my warm-up and then I moved on to my pieces. And then same for the goals: I think my goals for most of my practice sessions -— because I'm busy and I play to be in the band rather than play to improve as a player -— I'm not doing as many like rudiments or techniques. It's like I warm up and I practice my pieces. So, that's why I didn't really use it to set goals but I do think it changed a little bit how I warm up and things like that.

AA: Yeah, I would pass to the second one: the app helped to plan and set goals for my practice.

P2: I really didn't have like set goals for my practice. Again, it's just to play enough to get in shape, I guess, to stay in my band.

AA: Yeah.

P2: So, I think I gave the disagree with this one, but not to say the app is not good, but it's just in terms of how it relates to my goals for my practice. I found the app did not help out that much.

AA: Yeah, yeah, so you take very different: someone is like using the app to add stuff, someone is like integrating. So, if you're integrating it's like it's not changing: you are doing long tones anyway.

P2: Yes.

AA: That's a way that doesn't mean it's not helpful and of course maybe it means that it's not helpful and of course this is a sort of something we are very interested in. Because we can also disprove the previous literature and that's still valuable research. It's interesting to see how people engage with the app. Let's go on. Motivation: I like it. Agree and disagree. Using the app felt like an effort to me.

P2: I rated disagree only because of the settings: I find that every time I close the app, whatever settings I select has to be done again. So, I don't know if you noticed, but it reverts back again to the G to G, 80 beats per minute.

AA: Yeah.

P2: But if I were to change something... So maybe you could have a feature that remembers the settings. So, when the app gets reopened the next time, it will have the same settings, yeah.

AA: Okay, yeah, yeah, to pre-save it. Save like pre-sets.

P2: Great. Right. Because, as we both mentioned at the beginning of this session, selecting notes was not easy. If you wanted a B or a B flat, it could take you a few times like you could go down to an A or up to a B. But then again, like I believe we are being very picky on this. But that's how I felt in terms of the app feeling like an effort to me.

AA: That's good feedback. Thanks!

P1: Yeah, I think I had a similar experience: I think the effort for me was more in the decision-making because I was using it to replace my warm-ups. My old warm-up I had goes through a sequence of exercises to do and I have like a paper that I follow. And then introducing something new, the effort was just to decide: "Okay, what do I want to warm up with today?" Because I had that routine. So, I think if I went longer and I used the app for longer, it would ultimately become less effort because it would be part of the routine rather than something new.

And then yeah, I think I agree with the note functionality. But other than that, it is just the effort to integrate the app itself. And then other than the effort to integrate it and the notes, there wasn't much difficulty or effort.

AA: Very good.

P1: For the second point, I think I said agree. I feel like I should remember my answers better, but... [*laughing*] I sort of said this before, but I don't do long tones often. So, the fact that like that is the goal of the app and that's all you do on it -— because I was testing it out for the purposes of the study -— I practiced more effectively by default because I was doing the long tones. And I think switching things up from doing only scales to doing notes that are not next to each other also is like an effective way of working on your ear, which I don't do often. So, I think it did motivate me to practice more effectively.

AA: Yeah, yeah, yeah.

P2: Yes. Also, when I play – because I'm just really like a total recreational player – I just pick up and play a few notes. And I play rarely like whole notes. So, it was just last term at my band that they recall and like emphasize the importance of it, so yeah, this app, again, changed my view allowing me to think of the importance of whole notes during the warm-up.

Jumping, yeah, to the next one for the practice for the longer periods of time. I disagree with that only because of my personal choice, which is just that I have limited time so it didn't allow me to actually extend/improve my practice time. But, it has nothing to do with the app, but I just answered disagree on the spot. It's not to disrespect the app in any way.

AA: Of course, of course.

P2: Okay. Yeah, yeah [*laughing*].

P1: I think it motivated me to warm up for longer. So maybe that answer's a bit skewed, but I think my overall practice time was the same, for the same reasons of Participant 2. Like it's, you know, within a certain parameter, but I definitely warmed up for longer.

AA: Thanks. Let's go to the last one, perceived usefulness: I made more progress than I normally do in 2 weeks of practice using the app. We have a disagree and an agree.

P1: I think I said disagree and if I didn't, that's what I meant. I think that, just because of the amount that I practice, which is not much, I made the same amount of progress, like especially because I was focusing on things that it's not necessarily what the app was focusing on. My progress is more measured on how I can play the pieces we're playing in band rather than my technical abilities. So, I don't necessarily think it was long enough for the structure of the app to actually impact my playing.

P2: Same here. I felt I didn't make more progress. On the good side, it didn't make my playing any worse. So again, maybe we could tie this to the second point. Like this app is great for beginners, newbies to the instrument.

AA: Yeah, yeah, yeah, and then you need to find a way to not get bored.

P2: Correct. And this is where perhaps the variations of different notes could come in. So, I have a question: when we select a range, let's say from a G to G for one octave, how is the next note chosen? Is it something that you could tell us about during this discussion? Or is it totally random?

AA: Yeah, it's totally random, yes. And if you use the dynamics, also the dynamics, they are totally random between the three: between piano, forte, and mezzo forte. This is how it has been selected. And so that was like to me to the least choices as possible to see how people deal with that and what people are looking for. You know? So it's a compromise. And that is why this question is tricky, because it is subjective and then it depends on so many factors.

So, let's go with the last one and then we'll be open for more general feedback. So: I feel the app could be useful to be integrated in a classroom. So, in like music education for the teacher. What do you think? We have agreement here.

P1: Yeah, I think it could. I think that it would be easier to integrate with some feature changes that I think will fit better in the open-ended question, but I think that it could be both a teaching tool and maybe like a motivator for beginners to do their warm-ups. So, like: "Okay, I'm warming up with the app" and then we move on. And I think that the feature of hearing the note first is really useful for both beginners and then for like ear training. So yeah and like if there was maybe a way to track like the hours on the app so that like the teachers could use it as like, you know, an accountability measure. Like I think there's lots of things that could be done but overall, I think with minor tweaks, it would be very useful.

P2: Yep, it's the ear training part that I probably did more of during these few weeks. So let the app play and try to determine which note. And which I scored not too well on, I have to admit, but it would be a good thing, yes, for the ear training as well. This could be implemented.

AA: Yeah, so one of my questions is like what is like the best feature and the worst features of the app? So some of them you already mentioned there if you want to expand more...

P1: I think the best is the built-in tempo, like the built-in metronome. Because that and the fact that it's meant to have an example of the note before and combined. That's what makes the app unique and like it contributes to the space.

P2: So Alberto, for the dynamics, is it also totally random if that's enabled with the notes?

AA: If that's enabled, yes. It's all random and what I try, I just multiply the amplitude of the signal. So, if it's forte, it's unmodified. If it's mezzo forte, the signal that they listen to is more attenuated, and even more attenuated for piano.

P2: Okay, I think it's great too for the beginning student to know the difference between the dynamics. But is it considered dangerous if it lands on the high note and the student is required to play really soft on it? Is it dangerous? Is it not good for a student to play on a high register on a very soft volume? This is what I heard. Is there any danger to that?

AA: Oh, I mean, if you want my answer, this app is not trying to substitute the teacher. It is thought as a tool for the teacher, like a metronome and tuner. How the app was thought is more, of course, to learn the fingering technique, but also to create this mental imagery of the note. So that you "listen" to the note in your head before you have to play it. And that's very important to play brass instruments, and that's how I was thinking that the app could help in that way. And that's why I try to use good recordings of a trumpet player as a reference instead of using MIDI.

So of course, it can be dangerous to play high notes if the student doesn't have enough force in the embouchure to achieve those. It's like going to the gym and lifting 70 kilos, while you can lift safely 20. So, that's why you need a teacher. That's my whole answer.

P2: Yeah, because for sure the student will try to mimic the sound completely, even at the volume that's being played. So it happens that if the dynamic was enabled and then the range was set very high, the student would definitely try his or her best to mimic that note.

AA: Yeah. But pedagogically, it's a good feature. Because imitating a good note is a good and safer way of learning...

And another question: were you using headphones or just the loudspeaker of your device?

P2: I was using the speaker on my device.

P1: Yeah, me too.

AA: And then, potential improvements?

P1: I think it would be really useful, for like a teaching setting, to also have the note names somewhere. Like, I don't know, maybe under here [*pointing below* *the score*] you put like “this is an E”. So then like students are also learning how to read the sheet music, with the note names. Maybe this comes like at a later time in their playing and they would already know that, but to turn it ON or OFF might be interesting.

I think too I found the image of the trumpet like a little bit small. In terms of seeing which valves were down. So maybe also having – like I know in a lot of learner books – they do like 1 2, 1 3; maybe putting the numbers somewhere. I also thought being able to change the time signature. Like does it really matter when you're doing whole notes? Not really. It just makes them shorter or longer, but again for like teaching, a teacher can use it as a tool to go over like explaining why it's different and things like that.

And I also thought – just because how I was using it was for my warm-ups where I usually do like long-term scales or things like that – being able to have the option of doing a scale instead of doing only random notes within a range would be interesting as well.

I think those are like the main, and then we already talked about the choosing the notes. So, like little improvements, not necessarily like major features that I think are make or break.

AA: Thank you so much. Of course, a lot of stuff comes out, also from other discussions, so you will see where usually users are struggling.

P1: Hmm.

P2: I have nothing to add. Participant 1 summed it up really, really well. Yeah.

AA: Yeah, yeah, yeah, no, that's super nice everything. If you have more open feedback, if you want to add something, otherwise thank you very much!

P1: No, it's great. I appreciate the opportunity and good luck with developing it. I know the grind of qualitative research myself, so. It looks good. I think it's a really good start.

P2: Keep up the good work Alberto with the app.

Transcription Focus Group 4

Legend:

- **AA**: Principal Investigator
- **P**: Participant

The participant preferred to have an informal individual discussion, without other people.

This focus group discussion was originally conducted in Italian and translated into English by the candidate.

Transcription:

P: Unfortunately, I had to study a lot here in Italy for the teacher certification exam useful for my current job and I didn't have much time since the deadlines are always very tight for public sector positions.

AA: Yes.

P: I took some notes: I didn't find any flaws in your app, nothing like: "Oh God, this thing doesn't work!". The only thing I would nitpick – just to be meticulous – is the selection system for the range between the lowest and the highest note, because that gesture is a bit cumbersome. That's the only thing. And then I noticed – but I don't know if it's programmed that way – that sometimes the app tended to make me repeat the same note maybe 3, 4 times in a row, even though I had already heard it played.

AA: Yes, of course.

P: And this is, let's say, the only aspect I've pointed out as negative, which isn't really so negative upon reflection, just the minor things I've noticed. However, I was thinking about some suggestions, because when I used it I thought to myself: "I am a beginner, I mean I am just starting out, and how would I want the perfect app to be?" I did some research, and there are some apps that integrate a tuner. For someone like me, for instance, who is just starting out and also learning on my own, having the tuner's feedback could be useful because, based on the frequency, it tells me if I'm more or less doing well or if I'm too sharp or too flat.

And then let's say that your app is pretty much ready and complete, or do you plan on developing it further? Because, in my opinion, it would also be interesting to expand it a bit like those language apps available on the PlayStore or App Store, like Babbel, etc. Maybe divide the learning path a bit like, let's say, Cambridge or IELTS does for English levels: based on your level as a trumpet player, you can do certain exercises. Or, what do I know – I also bought a trumpet method – maybe include exercises with irregular groups like triplets, in such a way that maybe the apprentice or the student learns to tongue correctly and reproduce triplets, do staccato, etc. I sincerely hope that you can expand this app. Maybe also in a way where you offer a subscription, I don't know.

AA: Yes. Yes.

P: In my opinion, it's an excellent starting point, also because I believe it's the only one available that has a recorded sound – you told me it's a tone recorded by a musician – so it's not a synthesized sound. Before meeting you, I had downloaded others, where there are virtual trumpets, but those sounds are, let's say, artificial.

AA: Yes. Yes. Kind of lifeless.

P: It's an excellent starting point that can become even better, in my opinion, because I like the idea of having a single app where I have the tuner, can listen to sounds, and can also have a sort of method to practice with. In a single app, I have everything and basically don't have to do anything else but study. It would also be interesting from this point of view. I don't know what you think about also creating a sort of progress diary to keep track of when I'm studying, how I'm doing. And I think that way it could really make a splash in the educational app field. Anyway, congratulations.

AA: Oh no, great, great. Thank you.

P: And another thing I've noticed: I think yours is the only app that, when you have to play the low D, signals you to extend the tuning slide of the third valve. Because other apps don't tell you that.

AA: No, no, of course. Of course.

P: These are small details, but they make you realize that there's a lot of study and attention to detail behind it. Definitely.

AA: Absolutely, absolutely. No, precisely, the idea is to continue it. How to say, this is the most basic app one could think of, that I could think of, something like this. From what you're telling me, indeed many things I also find in experiences of other participants and will surely be results that I then go to analyze in my publications. The idea, yes, is to continue the app, so as I tell you now that the study for you is finished, on one hand, how to say: Feel free to delete it, obviously.

P: No, I won't delete it [*laughing*].

AA: However, if you don't delete it, keep in mind that the project can go forward, integrating improvements. But this means that I might ask you for the use of the microphone to enable the tuner.

P: Ah yes.

AA: Can I still ask you some quick questions about the results?

P: Yes yes, of course.

AA: So, I'm now looking at the questionnaire responses. So, aside from the first ones that are demographic data: "Installing the app to learn its functionalities was easy. I felt competent in adapting the app to my level". I saw strong agreement. If you have any comments, aside from those you've already given which I will definitely write down...

P: Well, I have to say that I had taken some notes, but mine were more suggestions, because the only complaint was related to that gesture and then I repeat that sometimes the app made me repeat the same note up to three times. Now, I don't know, I'm not involved in this so I don't know how it works, how you can make an app from this point of view, but sometimes it was a bit too recursive, but for the rest, it was perfectly fine.

AA: Ok, yes, yes, yes. No, I mean it's a random algorithm, truly random. But you can do a penalization, so if a note has already been selected, the probability that it will be reselected is lower; doing something like that. Or that it has to go through all of them until the cycle is completed.

So, deliberate practice is, how to say, a study in which you already have a clear idea beforehand of what you want to do, right? The aspects you want to improve on. Here it says: "The app has changed the way I practice with the trumpet". And a question my professor suggested I ask is: in what ways have your study, your practice, your goals, or the way you select goals changed?

P: Let's say that, in my opinion, this app helps you to be somewhat consistent, right? Because you know, you have those 10 minutes and you dedicate them every day to practicing... And my teacher, who is a clarinetist, told me: "Look, you need to practice, especially in this initial phase, producing very long and clean sounds". So, in my opinion, this app allows, especially beginners, to refine their ear and play these long notes because, if I remember correctly, you can adjust the metronome, so maybe you go to play really long, long, long notes which, in addition to the fact that there's the reproduction of a real sound shortly before, allows you to refine. So, in this way, it has changed my practice. Because at the beginning, before your app, I would say: "Yes, okay, I try to play a long C", but then I don't know how it was, whether it went well or badly. Instead, this way I have somewhat a compass to follow.

AA: Yes, yes, yes. So, then: Motivation. It says: "Using the app felt like an effort." You put: strongly disagree. So, it was easy, I mean, you didn't need to find motivation to open the app and...

P: No, no.

AA: Then: "The app motivated me to practice effectively": agree. And "The app motivated me to practice for longer periods of time": agree. So, what my professor asked me to suggest is: in what ways has the app played a role in supporting your motivation? I mean, if you can comment.

P: Well, in my opinion, this app manages to motivate you because, in my opinion, you have been able to blend studying with "gaming" because it was very interesting. In the beginning, I practiced while looking at the phone screen, and it was a nice motivation, also fun, to try to hear the sound without looking at the app, with a real sound, and then try to reproduce it. So maybe while I was waiting for those 3 or 4 beats of rest, I would say: "Come on, now I have to guess the note that's about to play!" And so it motivates you by teasing. It piqued my curiosity a bit, as if I were in a quiz: I have to guess the note; let's guess it 100%, especially if you then add sharps and flats, you increase the level of difficulty a bit.

AA: Yes, yes, it's more granular. Fantastic! Then, last questions: perceived usefulness says: "Using the app, I made more progress than I normally do in two weeks of practice". Here there was a neutral response, also because it's not easy to define.

P: Ah, I gave you a neutral response simply because in the last two weeks I focused on studying for the competition for my own work reasons with stuff I had never studied before like psycho-pedagogy, etc., because the mode of the competition changed here. No, actually, I think I even regressed, but because I didn't use it, I didn't practice. I abandoned everything. I had to do something else, that's all.

AA: Yes, yes. And normally, how do you observe or monitor progress? If you think the app can be useful for this...

P: Usually, I try to monitor progress in terms of sound cleanliness, because very often sometimes I have a fairly clean sound, sometimes I have a really bad sound. So, I used this range and then tried to adjust to the sound I heard, the one you recorded. So, my measure of judgment and improvement was to produce a sound as close as possible to that of the musician you collaborate with.

AA: Yes. And one last thing: in your opinion, do you think the app could be useful if integrated into a trumpet class, that is, for use by the teacher. In what way? What do you think?

P: Well, first of all, it can be useful – and I already imagine it complete with tuner and everything – because it creates a sort of continuous training, meaning the student/apprentice goes to music lessons, practices freely, and can continue to do quality practice at home. Now, I've had this fixation with the trumpet for years and some time ago I saw a video on YouTube of a quite famous trumpeter, I think he was Spanish or Chilean, not the type who had done a masterclass, I think his name is Arturo and something, he wears glasses...

AA: Arturo Sandoval.

P: Exactly, exactly! And he said something very interesting: it's not the amount of time you spend practicing the trumpet, but the quality. I mean, I could spend 4 hours trying to do buzzing and that's it, but if I do it for 10 minutes, but do it well...

AA: Certainly, certainly.

P: So, in my opinion, it's useful because it allows the student continuous training, otherwise you lack that reference figure who is the experienced trumpeter, the music teacher. For me, it should be recommended to students for practicing. It's also a way to encourage them. For example, many years ago, I attended and took a music course here in Sicily in the band. But the practice was: "Okay, I played, went home, and that was it". I might put down the instrument and if I picked it up again, there was no one to tell me: "No, wait, you're doing it wrong". This way, you have, let's say, a digital tutor that can be very useful.

AA: Certainly, yes, yes. And did you usually use it with headphones or using the phone's speaker?

P: No, no, I used the speaker. But because of a problem with tinnitus of mine.

AA: Ah okay.

P: I've used headphones too much in the past, so now I need to limit their use a bit. No, no, I used it comfortably with the speaker. It has quite a powerful sound so I didn't need to use headphones to hear it better.

AA: Ok, and did you use all the features? If you have anything to add beyond the thousand things you've said, for which I thank you...

P: No, I used them all: the range one, the bpm, and then the sharps and flats. That, especially for a memory issue, to try to memorize as many positions as possible.

AA: Did you by any chance use the dynamic feature? The one to add piano, mezzo forte, and forte?

P: No, no, I didn't use that. But I didn't use it simply because I said: "Let's start with the basic simple things..."

AA: Ok, yes, yes, but, how to say, almost no one used it and so probably, indeed, it's not so important.

P: In my opinion, it is important, but in my case, as a beginner student, it's not that I don't consider it important, but I preferred to focus on the other trumpet functionalities.

AA: Yes, yes, of course. And one last thing is when and how you used it. I mean, if you used it at the beginning of your practice as a warm-up, if you preferred to use it at the end or in the middle.

P: So, my practice was basically the app. I tried, as I told you, I bought this method, the Arban's.

AA: That huge book!

P: Well, huge, right?

AA: Beautiful. It's beautiful.

P: I started with this. But I already saw that in the first exercises especially, I had some difficulties, especially because, I'm speaking to you also as a trumpet teacher, up to the high E5 – so the last space on the pentagram – okay, fine, I can get there easily. From the high E5 upwards nothing, the F5 doesn't come out, the G5 doesn't come out, let's not even talk about the A5.

AA: Also because the Arban's, after 3 pages, already goes up to A5.

P: Exactly, and I said: "But I barely make an E5!" So I said: "Let's set aside the method. Let's focus, maybe I'll keep to a bit lower octave range, do it well, and then maybe with practice, I'll manage to produce these sounds". Because up to E5 okay but then from E5 upwards I can't go higher. The Arban's at the sixth exercise shoots me…

AA: It shoots you up to the high A5!

P: Exactly! What am I supposed to do?! So, I gave up. The method is just sitting there. I'll get back to it but I need to see, maybe I need to improve a bit…

AA: Yes, as in Arban's is not one of those books you start at page 1 and go through to the end. It's a book that is divided into sections, so you can do a first exercise from that set, the first exercise from another set... but yes, it's tough. Plus, it's an old method. That's why I say: "My goodness, can we use some digital technologies in a bit smarter way?" Because then the quality with which you come to play the E5 is also important. Because if the F5 doesn't come out, probably the E5 okay, it does come out, but it's not the best, well-sustained E5.

P: Yeah, I thought that, I was saying it's not one of the best E5s out there, mine...

AA: I also compare it a lot to weightlifting, you know? As in, everyone can lift 5 kg. And it's not given much thought. But when you have to lift 50, you can notice the differences between a professional, a bodybuilder who can do the exercise with the correct technique, and one who cannot. Also because it's not so much about lifting 50 kg, it's about how! How you lift it so it's sustainable. Because if you lift 50 kg, but to do it you break your bones, you break your back, and you get 3 hernias, it makes no sense. It makes sense to reach that note in a way that's correctly built.

P: Yes. Yes. And I would like to do this: to also take the higher notes, without maybe turning red like a bell pepper. In our band, there's a former marshal of the fanfare with a real bersagliere sound who takes those high notes, inflates his cheeks impressively like Armstrong. And then I immediately saw on YouTube a lot of musicians who very relaxedly hit the high notes without any effort at all. So that would be the goal for me. It would be nice.

AA: Thank you very much.

P: You're welcome.

AA: I will write to the various participants once the results are out, and the article is published. Of course, the article will be completely anonymized. What I'm going to do with this recording now is: transcribe everything we've said – then I also have to translate it into English – and as I transcribe, I anonymize the names of the participants. Then I analyze the data and what I share externally with other researchers are the anonymized transcriptions of the results. Then I will use nVivo or probably some software that allows for qualitative analysis and so on. Thank you very much!

P: Not at all, thank you and I wish you good luck with your work!

Transcription Focus Group 5

Legend:

- **AA**: Principal Investigator
- **P1**: Participant 1
- **P2**: Participant 2

This focus group discussion was originally conducted in Italian and translated into English by the candidate.

Transcription:

AA: Great. So, first of all, I would ask if you have any general comments. Then we'll go on to discuss the questionnaire and so forth.

P2: Excellent, if I may share some thoughts, I'll start by saying that unfortunately, I didn't have as much time as I wanted to try out the app. However, the app sparked my interest for several reasons. One: some of these reasons include the fact of hearing good reference sounds. That alone does a lot: single sounds with a nice color, a nice personality. And then also the fact of repeating them twice, that was interesting too. On this point, though, if the input could be very interesting, however I missed a checkout, a control of the output. I mean: "Yes, I listen to a nice sound, but what about what I did? Was it awful?" Definitely! So, there could have been, I clearly imagined, something graphical or feedback in terms of frequencies, I don't know, or a form of timbre or something that tells me more or less: "you are very far from what the reference sound was". You know...

P1: Like a sort of tuner? Is that what you mean?

P2: Exactly, just that.

P1: You mean something that tells me if what I'm producing is correct.

P2: Exactly.

P1: Because yes, the ear does a lot, but an additional check could be useful.

P2: Yes, yes. But in fact, a long time ago I used to do this: I would put a small tuner – the one for the guitar – where the notes would come out on the display; I placed it there for reference, more or less because then there's the issue of registers, where with the same fingering of the pistons you can play more notes and sometimes you can't – at least I can't – always tell: "What was this? Was it a D or perhaps the sound of the next harmonic register?" So, I personally missed that feedback. Then, if one wants to, one can record themselves, analyze that sound, or do a frequency analysis. But I don't do it, there's no time! However, it could be a good idea for self-calibration.

Then another thing: I found very interesting the dynamics of piano, mezzo forte and forte; something that, honestly, I did not know. I mean, by playing around, you realize that if one wants to play a bit softer, it can be done, but to actually have an exercise on that, I believe could be very useful for me. Then another thing: Ah, the fact of setting up the app with the various intervals, with or without semitones, various things, is interesting.

P1: I agree with you and Participant 2 that there should have been a third measure to hear one's own sound on the app.

P2: Ah.

P1: But then, after playing these whole notes for a while, I'd also try to split them up, make a kind of random addition, like a study method, I don't know, like Gatti [*it is a trumpet method*] or whatever. I'd like to do half notes, triplets, eighth notes, a quarter note... because yes, you play the whole note, you hear it in the tempo you want... Moreover, the timing thing is very important, the metronome function that keeps time, it helps a lot! Because some of my students, who maybe lagged a bit or started too early, with this feature, I could tell them: "No! Wait! Stop!" [*Imitating the metronome*] And you set the tempo you want: 60, 80, 120 bpm... But then after a while of playing whole notes, the usual beats with semitones and all, at some point it also becomes a bit monotonous.

AA: Sure.

P1: Always playing whole notes... let's split them up a bit!

P2: Ah, but there already is a Pro version, it just needs to be purchased! [*making a joke while laughing*]

P1: I didn't know, sorry [*laughing*].

AA: No, no, it's a joke.

P2: No, no, it's a joke. Like: "Now you want the Pro version? Pay 5 €!" [*laughing*]

AA: I would like to continue this project. I'm in my final year of doctoral study. Actually, it depends a lot on whether I find a university that likes my project; if I can get a role as a researcher or a university professor to be able to carry it forward. In fact, now with this session the study ends so you can decide: either uninstall the app or if you keep it installed, it is very likely that I will add features. So the initial consent form I sent you will not be valid anymore. I mean, it's likely that I'll ask for the use of the microphone, because I want to implement the tuner. So, you're all free to do what you want. Obviously, it will remain free.

P2: What else can be said: nothing. The graphics are fine, for what it needs to do.

P1: Yes, also the idea that there's the image of the trumpet indicating the keys. That could be improved as well, not just show the standard positions but also add alternative fingerings.

AA: Yes, yes, alternative positions.

P1: To add a little something, we're talking about improving this aspect.

P2: And then for me, the feature of extending the tuning slide of the third valve was very useful.

P1: Ah, yes, nice, nice.

P2: Because for my level, since I usually play just to pass some time now and then, I prefer to have a bit of fun, reward myself, rather than spend too much time studying. So that part is interesting: instead of going to get the slides I've printed, if the app tells me, it's a bit more convenient.

AA: Of course, of course, it's right there.

P2: It's a bit of an advantage for lazy men like me in this case. [*laughing*] But yes, it's interesting, a nice project.

P1: I continue to keep this app because I believe my students like it. Then we have fun. In fact, I say, "Look, this is how it's done, see here on the app?" I mean, I explain it well to them, the association of fingering and rhythm. So, it's indisputable, right? What I say can be disputed because it's just one person's word; but here's the app, and nowadays, the app tells the truth. So...

P2: And then it's also interesting that one has it at their fingertips, on their phone, and let's say it's a trivial thing, but not too much. Then it doesn't seem to take up much space, so one, if they want, can also leave it installed even if they don't use it very often, because then when they need it, they have it.

P2: Then another thing I was also thinking: regarding the effects, clearly this is a question, it's my curiosity, I throw it out also as a provocation: about the effects, could there be a way to incorporate them into such a context?

AA: What do you mean by effects?

P2: All the various effects that can be done with the trumpet.

AA: Like bending, mute, articulation... That's definitely something I want to work on.

P2: Also because, a bit like with the guitar, when you add some nuances thanks to the effects, it gives an interesting color, right?

AA: Of course, of course.

P2: Maybe that's what even catches the attention of the general public. It's a bit like makeup for women [*this is a joke and laughs*].

AA: Sure. Sure. I definitely want to do it.

P2: Nothing else comes to my mind.

AA: How have you typically used it? At the beginning as a warm-up, or at the end of your practice? Or...

P2: I've used it a few times. But as a warm-up, I think, it's interesting. And, as far as I'm concerned, also as practice for certain things that maybe I don't remember anymore, or don't know.

P1: At the beginning, definitely at the beginning, as a warm-up. When I explain, I say: "Do it like this! Do this like that. Hear your sound? No?" Maybe they make a sound. "No. Now let's listen again from the beginning. Hear that C you just played. How is it? Try to start the note this way." So, absolutely at the beginning as a warm-up. Then during the lesson maybe someone asks: "But, how is it?" Then I take the app again and say: "Look here how it is! That fingering association is like this! See?"

"Ah, yes, yes, okay, okay." And I: "Okay, okay". And we proceed like that. But absolutely at the beginning as a warm-up. Now, I know those notes by heart, so when we need to start, we play long notes and do warm-ups. But for the kids who say: "Let's pick up the trumpet and start playing right away!" I tell them: "No, wait. Let's warm up a bit first. Let's do this, let's play some long notes, hear your sound, hear this other one". Rightly, as Participant 2 said, if there was also a tuner included or a third measure to hear one's own sound again, then it would be really the best, because then there's the comparison from the sound of the trumpeter you chose – this great trumpeter – to the sound I make, and you say: "See the difference?" You say: "My goodness, in the example provided by the app the note is like this, I played it wrong, I made it flat, sharp...

AA: Yes, yes.

P1: And there, I think the ear would work a lot!

AA: So, propose the note that has been played after, okay.

P2: This topic of psychoacoustics [*talking about listening to ourselves*] is really a hot topic, because many groups use it, even for professional musicians who need to listen to themselves while playing. It's not taken for granted, and even for them, feedback is interesting. For me, it's necessary to understand how bad what I played sounds compared to the proposed note. But for them, it's serious. So yes, it's interesting. But those reference sounds that were sampled, could they be available? To possibly do a frequency analysis and see how bad mine are compared to those?

AA: They are sounds that I recorded. I've recorded a lot of sounds, to be honest. I have this database of about 20 Gb of different people, from beginners to professionals.

P2: I mean those from the app. Or rather, I ask specifically: I'm interested in maybe having some high-level sounds to take as a reference, then I compare them with some of my sounds and see what happens, to observe them a bit.

P1: Alberto told me that it was a famous trumpeter from whom you took the sounds for this app.

AA: This subject is also something I'm interested in; in my latest research, I presented an artificial intelligence algorithm that can distinguish how efficiently a trumpet sound was produced with performances similar to those of a professional, an expert human, and so these are all things that I would anyway like to implement in the app, like a tuner. These are the researches I am doing. The idea is: this app is the starting point, the most basic thing that came to my mind.

So, I indeed have some questions, but more based on the answers that were provided in the questionnaire. I can share the screen again. Now, regarding ease of use, just to discuss, if you have anything to say... It seems that, indeed, the app is perceived as easy to use. "I felt competent in adapting the app for my level". There's a completely agree and an agree. Feel free to add any comments...

Then in deliberate practice: The app has changed the way I practice with the trumpet.

P2: Yes, a bit, because of the dynamics like forte, piano, mezzo forte. Then what else? I was saying the addition of the slide on the third valve is interesting, because for various reasons I always paid a little less attention to it, clearly a mistake. Then another thing, the fact about the keys is interesting because it not only enriches the mechanical [*muscular*] memory but also the visual memory, so it was interesting.

P1: Yes, yes, exactly, you see the actual keys to press, so I can remember. But maybe for C sharp, the one on the third space, instead of suggesting the combination of first and second valve, maybe suggest doing it with the third [*referring to an alternative fingering position*], but there it tells you to do it with first and second…

AA: So, add alternative positions.

P1: Yes, I think so. Also the alternative fingering positions, because there are maybe some passages that you can't do easily with the first and second valve, but with the third, you can easily.

AA: Sure, sure, depending on whether there's a trill, for example, the alternative fingering positions are very useful.

P1: Yes, instead of always doing first and second, first and second, doing it with the third might be more convenient.

P2: Then, can I say something? For me, it was really convenient because I'm more interested in knowing the fingering positions than the musical staff, so I often look more at the valve positions than the notes marked on the staff, which means I often disregard the latter.

AA: Yes.

P2: Because, for example, I'll give the opposite case: I have this musical score where both the notes and positions on the keys are marked, but maybe then you can't follow it well in speed, so if one uses the app, it turns out to be handy. I'm clearly talking about a trivial matter, but for what I need, it was very, very useful.

AA: Sure. Did you use it with headphones or just the phone's speaker?

P2: This [*i.e., using headphones*] is something I wanted to try, but didn't manage to.

AA: Ok.

P1: I used the phone's speaker.

P2: Besides, I had an idea, but I don't know if I'll be able to do it. Where I usually go in Parma, there's a room with 50, 60 speakers arranged 360 degrees, so listening there might offer a more immersive experience. But I want to do it out of curiosity.

AA: Yes, yes.

P1: But with that thing, you're really in the middle of it.

P2: Yes, yes, yes. Clearly, with this type of sound, there's no question of directionality, as there might be with other sounds, like environmental or natural ones, but it could still be interesting, to see what's perceived.

AA: Participant 1, did you use it with or without headphones, with the phone's speaker?

P1: Without headphones, without headphones. I listened to the sound, I used it by playing the sound for my students at school, and then had them repeat it. But then, I had to go through the steps. I always had to repeat: listen, play, then listen again and play again, then listen until they were able to connect well.

AA: So, this is how you think it could be useful in a trumpet class.

P2: Ah no, no. On this, I'd make a clarification: I tried it and listened to it with the phone, but what I thought is that it might have been even better with headphones, only I haven't tried that yet.

P1: No, I haven't tried with headphones either. Just with the phone.

P2: I think the result could be better.

P1: Absolutely yes. Well, with the phone, the audio isn't the best.

AA: Yes, but, as you say, it's interesting to see how people interact, how they manage to find a way to integrate technology into their practice. Which then, of course, doesn't have to be something to use every day. So, at this point, I would say if you have anything to add based on what we've discussed, please...

P2: I'll add another thing. One can also play all the notes they want afterward as indicated by the app in whole notes. But if one wishes, they can also do it in unison with the app. At least that's what I tried to see if it more or less matches or not.

AA: Yes.

P1: Yes.

P2: Doing it afterward is one thing, and it's useful, but doing it together as well. Clearly, not the first or the second time, but if that note comes up again and maybe you can do it together.

P1: But in my opinion, Participant 2, using headphones you can do it in unison with the app, because listening from the phone's speaker you can't do it because you cover the note, but if you have headphones in your ears and that note is played there, you go in unison with the note, I think it would be very interesting.

P2: Also, I'm using bone conduction headphones, so my ears are free. These headphones I have are awful, not really to be considered, but who knows what the result would be to do something like that.

AA: Yes, yes.

P1: I think it should be interesting.

AA: So, the more negative aspects of the app are the fact that, as we said, it's repetitive, always requiring whole notes, and then one gets bored, and if there are others...

P1: Eh, it's hard to remember what we said before right now.

AA: No, of course. That's why I recorded, so everything is saved.

P1: I was just kidding. No, otherwise, in my opinion, the app is positive, it's all good. You have to work on it. That's your task, Alberto.

AA: Sure, sure. I'm now collecting the results. Everything you say I'm transcribing, I have to then translate it into English and then analyze it. So, I say: "There are n people who said this feature is nice; m people who say this other feature is not good." And then I'll report it in a scientific article, and if you're interested, I'll send it to you by email, it will be the result of the study.

P2: Do you already know which journal it will be published in?

AA: Not yet, no.

P1: Keep us informed. We are all with you, Alberto.

AA: Great. Exactly, when this publication comes out, I'll obviously share it with you, and in the acknowledgments, I'll thank the participants, without naming them to ensure privacy.

P2: The important point is that you are one of the authors, that's it [*laughing*].

AA: Thank you very much!

P2: Thank you for involving us and we look forward to updates.

AA: Sure, sure. Great, great, and thanks again.

P1: See you. Bye Participant 2, it was a pleasure to meet you.

P2: Bye Participant 1, it was a pleasure!

AA: Thank you very much!

P2: Bye Alberto, until next time.

AA: Great, thanks again. Ciao!

Transcription Focus Group 6

Legend:

- **AA**: Principal Investigator
- **P**: Participant

The participant preferred to have an informal individual discussion, without other people.

This focus group discussion was originally conducted in Italian and translated into English by the candidate.

Transcription:

AA: So, first of all, I would ask for feedback on the experience.

P: Well, regardless, I find it very useful to try and explore new ways in the field of learning. For me, the most useful thing was to change from my usual routine. That said, I have to honestly tell you that after a while, I returned to my usual routine because somehow the test was interesting, but ultimately, I find myself better off with what I was doing before because it was more or less a groove where I was always comfortable, and so somehow I didn't have much flexibility to move around. Let's say that, if I can give such a comprehensive feedback, what I missed was that feature, that facility, in short, that particularity that would push me to make the leap and therefore to move to a different type of learning. And so, in practice, being somewhat close to what I already did using another app but... What did I do exactly? Previously, I used any tuner, and in the tuner, I also put a metronome underneath; and so, since somehow I already knew what the long notes could be, there were some long ones to practice with. Let's say the particularity of your app was that it actually randomized the notes for me; but in the end, I didn't perceive this as so fundamental. While the tuner instead is very important to me because I personally am out of tune, so for me, it's very important to have someone tell me: "Look, you're too out of tune".

AA: Ok.

P: There's the tuner with about ten cents off, I think something like that, and when I'm out of tune it gives me a red light, and this thing was important to me. So basically, after a while, the absence of the tuner is what pushed me to go back, otherwise I would have continued using the app because it was perfectly fine for me.

AA: Yes, yes, yes, okay.

P: This is really said with the utmost frankness, actually.

AA: Of course, of course, that's what I'm looking for in the end. Because in the end, what I want to do is collect data not only to possibly improve this software, but also then to provide guidelines to anyone in the academic field – but really anyone in the community – who would want to do such things. So, I would just quickly go to see the responses.

P: Of course.

AA: Then, sharing the desktop screen… Okay: ease of use. If you have any comments based on these statements…

P: So, the ease of use is really high, in the sense that it obviously doesn't really have any flaws, it's very easy to install and use. You open it and it works. So, let's say it doesn't have flaws. And clearly, maybe the strongest point because I would want even a four-year-old child to use it. I mean, if they start making music at four, this app can be used at that age because you just click... So, this is only the highest rating.

AA: Sure, sure, good, good. "Pratica deliberata", which is then an Italian translation of deliberate practice.

P: What is it?

AA: It's simply, let's say: "Ok, when I play, I already set beforehand what goals I want to achieve in my practice session". Like, simply setting goals to say: "Ok, this is what I want to achieve at the end of my session, or by the end of this week, and so in this case, I structure my practice to achieve that goal". And so, it's related to the environment. So, the related questions are indeed: "The app has changed the way I practice with the trumpet". Yes or no? Here, in this case, no; you've basically answered a bit earlier…

P: But excuse me for a moment, sorry if I interrupt you, but are these the average responses or mine? Because I didn't remember them like this.

AA: In this view, they are yours. Then, of course, I will put the overall average.

P: Well, indeed. So no no it's clear, it's clear. Ok, yes, no, I answered like that. Fundamentally, it hasn't changed because, in fact, after a while of using it, I returned to the previous method, so honestly, it can't be said that it has changed it. And it helped me to plan to set goals in my study because, fundamentally, even just the fact of having tried it, and having tried to use a different method, was useful to try to understand also the limits of my previous one. So anyway, I find it a useful operation to try to change the method. Not necessarily, this must then actually result in a change.

AA: Absolutely, also because then evaluating change remains very subjective, especially over a very short period of time, let's say. And anyway, this is just my personal judgment, this here is just one of the warm-ups; then there's a lot to work on.

AA: Motivation: using the app seemed like an effort…

P: So, no it didn't seem like an effort. I gave a neutral judgment because basically, the app is very easy to use, and here I mixed the ease of use of the app as software with also the educational use. Otherwise, it would seem in contradiction with the first response, right?

AA: Exactly, exactly.

P: I perceived the question as referring to the use of the app in improving my daily training activity, and so it didn't seem like an effort to me; but then, in the end, since there wasn't that extra motivation, I stayed halfway, and in the end, between being halfway, you can go one way or the other, and I fell back into the previous practice routine.

AA: Sure, sure.

P: It motivated me, still neutral here because yes, it motivated me then somehow in trying, then not trying, then at a certain point in returning... Also take into account another thing: in my case in particular – this might somehow be useful to you – what counts as motivation more than the app is perhaps the fact that, I don't know, when they propose you to form a music group. So I practice a lot to avoid making a bad impression; or there's a concert coming up with your band and so on and so forth. So, since particularly in this period

I've had changes in terms of groups, I was motivated but more because I was afraid of making a bad impression on the other members of the group.

AA: I understand, I understand.

P: The app motivated me to practice for longer periods: objectively no. I mean, maybe this was something that also discouraged me a bit, the fact that after a while the app becomes a bit repetitive, and so at a certain point, I couldn't motivate myself too much.

AA: Ok, ok.

P: Forgive me if I'm very blunt but you told me to be as honest as possible.

AA: Of course, that's what I'm looking for. I mean, in the end, it's not like you have to say: "How beautiful we are, how amazing we are!" You have to try to see what can be useful for…

P: Yeah [*laughing*].

AA: Shall I also comment on these last two statements?

P: Yes.

P: So, I didn't make more progress. However, somehow, I was close to making more in the sense that, as I told you, it took little to make me make the leap [*to incorporate it into my routine*]. And instead, I find it very useful. I gave the highest rating or something like that, in the teaching part in a class because I find that this app, together with a teacher who somehow makes it become more – how do you say? – lively, you know?

AA: Yes, yes, yes.

P: It entices you, then it's good because basically it's perfect for a teacher to be a quick tool and have kids do exercises, when she/he herself/himself [*the teacher*], then maybe takes it away, puts it back, does another thing. This is a bit my point of view.

AA: Yes, yes. Are you studying with a teacher or more on your own at this time?

P: Actually, I had one, but since my daughter also studies trumpet and she is progressing very quickly, the choice was to dedicate the teacher to her so I gave up the teacher for her and somehow…

AA: Noble.

P: So, I hear lessons, but I don't participate because the teacher rightly can't teach two students for the price of one, it wouldn't even be practical, right?

AA: Yes, yes, yes.

P: And in the end, no.

AA: Another question: did you usually use the app with headphones or with the phone's speaker?

P: No, without headphones. I have one of those phones that have speakers, which now seems like a stereo for how loud it is, so I didn't need headphones.

AA: Ok. So, we've more or less covered the positive and negative aspects...

P: The positive aspect, let's say, is the ease of use, the learning curve is practically zero to one, that is, you go immediately, you learn it after five seconds. Regarding the negative aspect, then I perceived something strange. I'll tell you, but I don't think it's possible: but is it possible that, especially when you go with low tempos, there is a very slight inaccuracy in timing? Because it seemed to me that it did: one, two, three, four, one, two, three, something like that [*imitating an irregular rhythm*]. I mean, maybe on my phone, which doesn't have any particular hardware, it seemed that the four quarters were in time, but then the subsequent four quarters were not exactly. And this thing, I don't know why, I noticed it sometimes yes and sometimes no. I don't know. If no one else has told you, it's definitely my psychoacoustic perception, so...

AA: No, it might be that that model of cellphone... so what I see is that just the first beat, when you click on the "Start" button, right at the beginning the first time, the first two beats are a bit off, but this is my perception. For the rest, I have no idea, but as they say, I document everything.

P: Well. It could be the way it [*i.e., the software*] manages the processes, right? Because it has to simultaneously execute the sound and check that you're not touching buttons or something like that. Couldn't it be that maybe it has a "spasm" at a particular point that systematically makes that "tick", it seems to me maybe late, not early. But it's really a minimal thing.

AA: No, no, but it's annoying if it happens.

P: More than anything else, you know who it bothers? Someone like me who may not be so sure in tuning and neither in timing [*rhythm*]. So if I start to fear that the app itself doesn't give me the right time, then I enter a negative loop.

AA: Paranoia...

P: Yeah, because anyway... I mean, maybe if Riccardo Muti listens to it, he tells you right away if it's half a bpm ahead or behind, because he understands immediately. But I don't, so...

AA: No, of course, of course. And the features, have you used them all normally? The features were like removing the image of the keys, adding the flat...

P: I tried all of those.

AA: What about dynamics? Piano, mezzo forte, and forte?

P: No, not the dynamics because I really can't do them; in the sense that, being at a very low level, right now I'm straining to produce a sound and to build a bit of endurance. Because then if you go to a rehearsal room with a group and they make you play the offbeat with the trumpet for two hours, you've worn out your lips. I'm in a group where I play ska; so you do "PA PA PA PA PA" continuously. So in the end, you wear out your lips. For me, dynamics are just not an issue at this moment. For me, it was important to make it alive to the end of the second hour, so no, I didn't really try that feature. That was the only thing, even though I didn't have much feedback anyway. I mean, you hear it, right? But then you do it as you want, so I'm not even able to tell if I did it that well.

AA: And so, as for possible improvements?

P: In my opinion, a tuner would be very important. Now, this is something that you as a trumpet teacher need to decide if it's the right direction, because I imagine there can be two approaches: I should hear the note, think about it in my head and replicate it. And judge for myself if I played it in tune. Or not hear it, try to play it and see if I hit it on the tuner. I guess there might be two different approaches and maybe one is not even as good as the other. The tuner was convenient for me because it somehow gives me the gratification of seeing that I've hit it right from the first go. So that stimulates me, you understand? Since I get quite bored studying music, the challenge with myself, apart from those silly little programs that I told you in our first meeting I had made with Matlab, was to see immediately if I can stay right in the center of intonation with the tuner.

AA: Yes, yes, of course.

P: So in my opinion, if you could include a tuner, that would be fantastic. But I guess that's a bit more complicated…

AA: It's already there.

P: Oh, you've already done it?

AA: It's in another branch, but it's already there ready to go.

P: Ah yes, you had told me, yes it's true you had told me it was already there, yes, yes.

AA: I can implement both: one that replays the note you just played, so you can listen to it again; and one is the tuner and then I even wanted to do some exercises where I basically ask the user to put on headphones to listen to a note and play in tune with that note, or with that chord. That's another exercise I'd like to do. I was thinking, if any other feedback comes to mind in general, if you think of something in terms of improvements…

P: No, it was just that, it was just that. The thing about the tuner clearly has the problem that, if it [*i.e., the software*] also acts as a metronome underneath, unfortunately, it messes up. Indeed, with separate metronome and tuner, obviously, the tuner every time it hears a tick from the metronome you see it messes up and then depends etc. I think it's physiological because while it records, it also hears what it itself is sending; it's not a dual-channel audio card. Then nothing else came to mind.

AA: How and when did you use the app? I mean, did you normally do your warm-up? Did you use it at the beginning, during the middle, or at the end of the session?

P: Well, here already... I mean, in my opinion, there's only one phase. Because I actually do so little... so I do fifteen minutes a day.

AA: Ok, ok.

P: One day I started to use it, the next day I started to use it again, a few more days, and then after a while, I just said: "Well, okay". It came naturally to return to the old way, and then I didn't think about it anymore, that was a bit of the experience.

AA: No, but it's interesting for that reason, right? To understand what the context is...

P: It just lacked a feature that would make it the "killer app" for that purpose and somehow force me to uninstall the others I had. But actually, if you already have other things in mind, this will surely happen sooner or later. Ah, often in the note selection, the app would forget the notes that had been selected in the previous study session, so somehow this was a bit annoying indeed every time I reopened it.

AA: Yes.

P: Thank you very much for the experience and good luck.

AA: Thank you, thank you very much for everything!

P: No, thank you, it was very interesting, and anyway, I tell you, I'm not uninstalling it; on the contrary, I'm interested in seeing if there are updates so maybe I'll happily try it a second time. It was still interesting.

AA: Great, thank you. Thanks again for your time and everything.

P: Don't mention it, bye bye.

AA: Ciao.

Transcription Focus Group 7

Legend:

- **AA**: Principal Investigator
- **P1**: Participant 1
- **P2**: Participant 2

This focus group discussion was originally conducted in Italian and translated into English by the candidate.

Transcription:

AA: Welcome Participant 1. Let's begin.

P1: So, a suggestion I have for you: I find the app effective for warming up before starting to play. I also see results; like instead of doing the usual scales – which in the end are just scales after all...

AA: Yes, scales must be done, of course. I mean, it's a complementary thing.

P1: Yes. The suggestion I have: if it's possible to insert a box where you choose: "I want to hear only that note", without selecting a range of notes.

AA: Just one note?

P1: Yes, just one note. Let's say, if you insert the range between the same note, that note comes out in the end.

AA: Yes, yes, yes.

P1: It's additional, but let's say visually, you say: "I want to hear this note, can you play just this one for me?" without having to select the range. Well, it's a refinement that can also be done without, you know. If you want to hear that note, you insert the range on the same level, and you hear the note. That's it.

AA: Sure, sure, yes, yes, yes. But then I also need to work on that interface; I mean, it's a bit cumbersome.

P1: Ok, thank you.

Participant 2 enters the room.

AA: Welcome Participant 2.

P2: Hi, can you hear me?

AA: Loud and clear, how are you?

P2: Good morning, I mean good afternoon.

P1: Good afternoon, hello.

AA: Great, so we were discussing a bit about the overall experience. Now I'm sending a link here in the chat to complete an online questionnaire to collect some demographic questions at the beginning, and then there are questions that collect objective data about your experience with the app.

[*Compiling the questionnaire*]

AA: So, first of all, I'd like to ask if you have any general comments. To discuss right away. Otherwise, then I can share my screen, look at the answers, and try to comment point by point, to gather a bit more data on what your experience was like.

P2: Please, proceed with the sharing…

AA: Perfect. So: Easiness of use. Ok, let's say it was more or less easy for you to understand how to install it, understand the rules, and if you felt competent to adapt the exercises that you wanted, for your technical level.

P1: Clearly, I answered not completely agree [*about adaptability*] for the simple reason that I consider the app as something one independently downloads and might have some difficulties with.

AA: Sure, sure. As to say, we also didn't have this individual section.

P1: Let's say, at the beginning of the study, when I downloaded it right away – then it may also be that I wasn't paying much attention – I didn't find the app exactly straightforward, you know?

AA: Sure, sure. Certainly, if I want to do it on a large scale, I would probably need to include an introductory video, maybe that's the simplest thing.

P1: Ah, a tutorial, you mean.

P2: Yes, something quick, 30 seconds.

AA: Yes, yes, yes.

P1: 30 seconds seems too short to me, but okay.

P2: No, not more, otherwise people get bored.

AA: Great. So now: "Pratica deliberata". Ah, I wasn't sure how to translate this because in English this concept is "deliberate practice". So it's, how to say, when one studies, setting the goals of the study. Having them very clear beforehand like saying: "OK, now I'm studying and so in this session I want to achieve this goal". Right? "I want to learn to improve this aspect". To see if the app can help in this, to have more clarity. So: has the app changed the way I practice the trumpet? I would be curious to have some more information. In what way has it changed you?

P1: In the sense that listening to the sound and repeating it becomes easier. It's likely that if you play a scale and miss the exact note; instead, listening to it makes it easier for me to reproduce.

AA: Thank you.

P2: Yes, yes, for me too absolutely the listening is wonderful, fantastic, in the sense that you almost don't notice what you do; because you've just heard it and so you try to replicate it in the best way, so absolutely on this. I mean for me it's a huge added value.

AA: Excellent. And as for possibly supporting planning and setting goals for my study? This is also very subjective I would say. If you have any comments or otherwise. . .

P1: Yes, let's say yes, it's very subjective, eh.

AA: And in what way did you use the app?

P1: No, in the sense that it's stimulating, let's put it that way. If you manage to reproduce it, it boosts your morale. Because then you notice the difference when you go to perform pieces afterward. I consider it as a warm-up in the end.

AA: Ok, ok.

P1: Clearly, it needs to be experimented with several times.

AA: Sure, sure. So, you used the app at the beginning of your study as a warm-up, or at the end, or. . . ?

P1: Yes, yes. At the beginning.

AA: Participant 2?

P2: Yes, it's definitely a great warm-up. Actually, I didn't use it that much, I didn't have all that much time, I have to tell you the truth, also because my kids [*i.e., my children*] would bother me when I came back and tried it. Surely, for it to be an app completely dedicated to study, you need to add something else. That is, one expects more than just reproducing notes you've just heard, you know? So, as a warm-up it's definitely great but as a real study tool, no, in my opinion.

AA: In terms of completeness?

P2: Yes, it needs something else, right? A melody, something to replicate with the notes that... Like, for example, you make me hear the notes, I replay them with those same notes and it produces a melody.

P1: And well, clearly then there are also books that reproduce melodies, so it adds to things that already exist. Instead, let's say, this remains an end in itself, at least that's how I see it, eh.

AA: Sure, sure, but as to say, it only addresses a small aspect. For me, this is very interesting as a starting point to understand. . .

P1: No, then it's clear. If you expand it and can insert melodies, it can also make the app more appealing, that's for sure.

AA: Sure, sure. And in terms of motivation, did it seem like an effort, so to say, to remember to use the app each time and set it up? To what extent?

P1: No, no, the only difficulty is that you have to turn off the silent mode. Otherwise, it doesn't work.

AA: Ok.

P2: No, not for me. That is, I mean, I would have used it much more often if I had had the time. I found it very useful. I discovered that my pocket trumpet is completely out of tune.

AA: Great.

P1: What? I didn't catch that, can you repeat?

P2: I have a pocket trumpet. It's of little value, eh?

P1: Ah yes, those ones. Yes, I have one too. Well, yes, yes.

P2: Just no way, eh.

P1: Is that the only one you have, Participant 2?

P2: No, no. I have several others, actually; but that was the one I had most at hand when I came home with the kids around. So I took that one and said: "Oh no, let's change the trumpet".

P1: I also have the flugelhorn. Only that in the end, I always end up on the trumpet.

P2: Ah, I also use the flugelhorn because I like it a lot more.

AA: Good, sure. So, this other point: "The app motivated me to practice for longer periods of time". This is also very subjective, depending on how you integrate the app into your study, whether you consider it something additional, or whether you consider it something that you integrate into your routine or not, in my opinion.

P1: Yeah. I too had little time; I was sick, so during that period I didn't use it because of the cold. Generally, when I'm sick, I don't use the trumpet, because otherwise, I strain my vocal cords, I already have a bit of a sore throat. And so it went like that.

AA: Sure, sure, and thanks again for the effort to integrate it. The last aspect: Perceived usefulness. So then: “Using the app I made more progress than I normally do in two weeks of practice”.

P1: We were generous here. [*joking*]

AA: Very generous.

P2: No, but it’s true. Look, I think the listening and immediate reproduction work really well.

P1: Yes, I’ve noticed it too, I already told him.

P2: The idea is good.

AA: Ok, ok. Well, it’s really useful for me. Since there are so many ideas on how to develop an app. The possibilities are many, so I decided to say: “Let’s do something really simple and see what’s liked and what’s not liked”. Precisely in terms of perceived usefulness, things like that. For now, the goal is to improve this, but then especially to provide guidelines to the entire community, absolutely, on what is perceived as useful, if we then want to do something that has an impact. And the last point, indeed: “I believe the app could be useful if integrated into a trumpet class”. Here we have a disagreement and a neutral. What do you think?

P1: I already told you when I asked for an explanation, in the sense that at an educational level I don’t know to what extent it can be useful. In the sense that the app is good in itself, but let’s say whether it can be useful I don’t know. In a trumpet class, there must be more, let’s say.

AA: Well, sure, yes, yes, yes.

P1: It can be complementary to study, but to a trumpet class it’s just additional, you know. In the sense that inserted in an educational context I don’t know; you should ask, I don’t know, a teacher, someone who is and it’s their profession, let’s say. I, as a student, don’t know how to answer, that’s why I put neutral.

AA: Sure, sure. Participant 2?

P2: No, for me, no, because what the app does is done by the trumpet teacher in class, so it’s not needed in class, if that’s what you meant. If, however, we’re talking about using it as a complementary tool when the teacher is absent and the kid practices at home, then absolutely yes. Always for the reason mentioned above, there's the fact that you can listen to a sound made well, of good quality, and so you can work precisely on the tone, let's say; something that at home, in the absence of the teacher who lets you hear the note properly – especially in the beginning – you don't do. That is, at a certain point you find yourself having completely ruined the sound simply because you focused on something else maybe. That keeps you straight, let's say. But at home. In music school, the teacher takes care of it, in my opinion.

AA: Sure, sure. And do you think it could be a useful tool for the teacher to give indications like homework? Or not?

P1: Well, music teachers are all unique, now I wouldn't know. Each goes their own way. Music teachers are artists, if you impose something on them, they say it's not valid, out of professional deformation. [*joking*]

P2: If you enrich it a bit, I think it could fit, it could be useful. But then you risk replication, right? Because there are many other applications, like trivially those that make you perform the exercise and then correct you; they point out which notes you got wrong, which you played early... I mean, I don't know how was it called, there was a super cool program...

AA: SmartMusic?

P2: SmartMusic!

AA: SmartMusic, Tonestro, etc. Yes, yes, yes, they do all that.

P2: Those are awesome, right? So, I don't think you want to redo the same thing, so either you invent something different... but those, in my opinion, are useful because the teacher assigns homework to be performed on any SmartMusic; the student gets immediate feedback, even without the teacher. And theoretically, therefore, they could improve. Plus, in SmartMusic, there's also the stimulus of the whole orchestra underneath in the accompaniment and other things. So, it becomes a bit fun too. And you – to avoid replicating – should try to invent something that still has this value. Even without the accompaniment, but where maybe you have a correction, or simply you play the exercise together with the virtual teacher who plays the same notes. Well, that could already be used at home; and so the trumpet teacher could use it to give you assignments, I guess.

AA: Nice, nice, thank you.

P1: But I don't know SmartMusic. How does it work? Just for information.

AA: It's a software. Usually, it's used with a PC, sometimes with an external microphone, but now I think you can also use it with the PC's microphone because the quality of built-in microphones has increased a lot. There are various exercises. Actually, I've used it little, but basically, it gives you feedback. You can set different tempos, like: "I want to do this exercise at 80 bpm". "Now no, I want to do it at...". And there are backing tracks and you play along with those. If there are backing tracks, ideally, you should listen to them through headphones. This way, the microphone picks up only your sound and gives you feedback if you are playing the right notes and if you are playing them in time, or if you are ahead or behind. While you play, it gives you feedback. This is also the basis for other technologies, like Tonestro, which does more or less the same thing but on an app…

P1: And are the pieces varied? Can you choose the pieces?

AA: There are various ones. I think there's also the opportunity for the teacher to compose it and you can import it, I would say. It requires a bit more effort, let's say on the part of a teacher. And it works in this way.

P1: I see. No, I wasn't aware of its existence. Well, but this is all paid for, clearly?

AA: I think there's a basic free plan and then if you want to unlock additional functionalities you pay. Yes, yes, yes. There are prices for classes, so if you buy it individually it costs 10, if you're a teacher and want to share the license with your students it costs 15 or something like that.

P1: 15 euros a month?

AA: No, no, I mean 10 or 15, but I don't know, it was just to give an idea of the proportion.

P1: Yes, yes. The range, yes, yes.

AA: On this note, I have a question: did you use it with headphones or without headphones, with the phone or iPad speaker?

P1: With the phone speaker, for me.

AA: Ah, ah.

P2: And me too.

P1: And so you think it's better with headphones?

AA: It depends, it depends on the exercise, right?

P1: Well, in the end, if the result is the same that you notice, that the note you made is the same one you heard...

AA: Exactly, it depends. Then there are other exercises for which the use of headphones can be crucial.

As for the app, did you use all the functionalities? Is there any functionality that you really didn't like or any particular aspect you didn't like? The functionalities are, namely...

P1: Dynamics.

AA: Dynamics, removing the trumpet images, adding sharps and flats.

P1: Ah, the image is useful if you want to experiment to see if you can identify the positions without seeing the trumpet. But well, you acquire this over time. So, no, I didn't use it because it's facilitated. After all, you already see the position. However, clearly, since visual memory then is what it is, it's preferable that you know the position regardless.

AA: Sure, sure.

P1: Ah, because otherwise it's not like while you're playing you say: "Let me see the app how to make this note." There's this risk, but well, then with practice it's overcome, over time. I never imagined that I would be able to play the trumpet; to reach an intermediate result. I'm not saying I'm a professional, but at least I manage to play some pieces, you know?

AA: Sure, sure, with the band or with other ensembles. And the positive aspect, therefore, was listening?

P1: Yes, listening to the sound and reproducing it, yes.

AA: Ok? Participant 2, what do you think? Pros and cons? Well, we've already discussed quite a bit about that. So, do you have anything to add about the most positive aspect and the most negative aspect?

P2: No, no. I too used them all, that is, they are quite straightforward; for me, the app is well made, it's really immediate. Whatever you choose to use, you use it, and there are no problems, meaning it's effective. Clearly, I removed the image of the trumpet right away. But then, I tried to get them [*indicating the small children*] to use it, but they refused. Now I'll try again. [*joking*]

AA: Exactly, exactly. One last thing: possible improvements? If you have any ideas on how you would improve the app.

P1: The only thing – that I mentioned before – if you could introduce the choice of the single note.

AA: Perfect.

P1: Also, including dynamics like forte would be nice. Asking: "Play a high loud C", and there you have it, it plays a high C. Without having to select the range. Very simple, you choose the same note. If you could put a box specifying: I want to hear only this note, without even putting the measure of exercise, and the measure of rest.

AA: Ok, yes, yes, directly.

P1: "I only want to hear C". A bit like when you go to music courses where they use a piano to tune all the instruments. You play a single note, and the class must reproduce that note. You could introduce this system. Without exercises, reproduce the single note. With dynamics as well if you want to include dynamics.

AA: Yes, yes, yes, interesting. Participant 2, any improvements? If you have ideas...

P2: For me, seeing it as a warm-up, then it would be nice that, after you've warmed up on that range of notes, maybe you also do a more complex exercise. Maybe you also choose the beat and let's say the rhythms, if you want to go up to sixteenth notes, if you want to use thirty-second notes, triplets, quintuplets; it matches the exercise on that range you've chosen, and you do it.

AA: Yes, yes.

P2: Maybe it's not difficult. With all the artificial intelligence going around, if you upload everything to a server, it might be doable.

P1: No, no, I don't understand, sorry. Help me understand, Alberto, ok, it was surely explained well, but I didn't catch it – maybe I got distracted – what Participant 2 meant.

AA: So, once you've done the long notes according to that range, if you can introduce, I imagine, different rhythmic figures; so not only whole notes but also introduce quarter notes, eighth notes, triplets... and then do exercises based on those, which can be suggested or simply have them directly created by a generative Artificial Intelligence. Now there are AIs that really do anything: you can also specify the chords, and they create songs in a certain style following that chord pattern. Thank you very much. If you have further feedback, please do not hesitate.

P1: Thank you for this opportunity you've given us, I was happy to participate and contribute to your experiment.

AA: Thank you very much.

Transcription Focus Group 8

Legend:

- **AA**: Principal Investigator
- **P1**: Participant 1
- **P2**: Participant 2
- **P3**: Participant 3
- **P4**: Participant 4

This focus group discussion was originally conducted in Italian and translated into English by the candidate.

Transcription:

AA: So, first of all, I wanted to ask you for a general opinion on what it was like using the app; if you encountered any technical problems or anything like that in your overall experience.

P1: I got on well with the app. It's very easy to use, truly straightforward. I'm satisfied with this aspect because, frankly, with mobile applications, with technology, it isn't always so immediate. So, I personally used it sparingly, I mean for myself, because, being advanced in my studies, I didn't feel the need, precisely.
[*Participant 1 is a graduated trumpet player.*]
However, the little that I did use it, I used it with my students because I believe it's more useful for students who are a bit more behind, those who need to read; especially for those not familiar with the trumpet fingerings. And then, it's comprehensive because I noticed it includes dynamics too. It seems to me also an original app because, as you explained, it was recorded by a trumpeter from the Santa Cecilia orchestra. That's not something everyone can boast, in short. It strikes me as an original app, indeed.

AA: Thank you.

P1: What I mean is, it's not the usual trumpet application you find elsewhere. Congratulations.

P2: You can tell from the sound, which is definitely more of a classical sound. I'm passionate about jazz, so let's say. From this perspective, I appreciated it less. Can I speak directly?

AA: Of course, of course.

P2: I am instead a kind of unusual trumpeter, in the sense that I started dabbling with the trumpet as an adult, a bit self-taught, a bit studied with a teacher here in my area. So, I know how to play some things, others I can't manage. I have serious issues with intonation and I must say the app was very helpful precisely for this. Problems with fingerings? I can read something simple from the score, in essence. In treble clef I do the direct transposition, in short, I manage to do a bit. Because of my passion, I tried to learn jazz pieces, to play with friends, etc. But for my intonation problem, it's very very useful. However, for my level it is "basic", let's say. In the sense that I would have preferred there to be at least intervals, I don't know, chords, intervals rather than notes just thrown there.

AA: Of course, of course.

P2: So, extremely useful for intonation and I believe, as Participant 1 said, for beginners it's a godsend; because they also look at the fingerings, which is not trivial, however, I dare say, you could make a second step by adding, I don't know, intervals of a scale, then in this discussion group there is a teacher, there's Participant 1 who is a teacher, so, surely he could tell you much more sensible things than mine. From the point of view of ease of use, it's really easy. Beautiful, really the notes are well made; because then I have my tuner so I also checked. Anyway, it's good from this point of view. For my level, I would hope to be able to, let's say, expand it if you could manage to do it; it wouldn't be bad.

AA: Of course, of course, how to say, with this I wanted to try making a basic app. In my doctoral research, I'm exploring various technologies, and this was really to say: "Let's try to evaluate an interface, if it can be good, and then collect feedback to understand how to implement new features". But, how to say, if instead the interface is bad, it needs to be rethought. And, from the research point of view, discussions with you are very useful, in the sense of even the criticisms, in the sense of saying: "If I have about twenty participants and they all say the same thing, it can be really useful as guidelines, for people around the world who want to develop educational technologies, to start from here, to have this research. And from here, be able to more or less have guidelines on how to develop theirs."

P3: Can I speak? Ok, so I agree both with Participant 1 and Participant 2. The interface is excellent in my opinion, in the sense that it's very clear. The idea of putting the fingerings didn't personally serve me but anyway, it can be very useful indeed for a beginner. Now, it must be said, I've been playing for a year: not much. But I already have enough familiarity both with the fingering and also say with the issues related to intonation. So this wasn't a problem. And in fact, I found the app a bit too unchallenging, in the sense that it's possible to set the notes you want to do etc. But, simple example, if I set the metronome to the minimum, it gives me 4 beats at 40 bpm, I normally do like 5, 6, I do three times, four times that. So, I found it a bit too unchallenging. Here also for what...

AA: Sorry, I didn't understand what you meant. In the sense, keeping the microphone to the minimum, like 40 bpm?

P3: So, I set the metronome low to 40 bpm, it gives me four, it gives me four beats: TAC, TAC, TAC, TAC. I normally do much more.

AA: Ok. As in sustaining the note?

P3: I mean, for me long notes go about 20 seconds each, that's it.

AA: Ok, ok, yes, yes, yes, that there. Four beats at 40 bpm are 6 seconds.

P3: Yeah, it's a bit little. So maybe I would add, if possible, this functionality: decide how many measures to go forward, not just one with four beats.

AA: Yes, yes, yes.

P3: Then there's a technical thing I found a bit difficult to manage. I mean, well, it's not actually difficult to manage. So, when using the app, the screen turns off. Basically, it's okay, you just need to change the setting on the phone, right? But actually, it's a bit inconvenient for the user; I mean, maybe if the app could keep the screen active, it would be better. I don't know if I've made myself clear.

AA: Ah, ok, ok. So, you mean, when one uses the app, the screen turns off after a certain number of seconds. Ah, thank you.

P3: With this I mean, one can change the setting on their phone, but you understand that I need to change it at the beginning, change at the end. It's a bit inconvenient, let's say.

AA: Yes, yes, yes, no, certainly, certainly. Annoying, it's a hassle.

P2: However, this has little to do with the trumpet. This is an Android or Apple app interaction, right?

AA: Yes, yes, yes. But, how to say, the feedback goes on the whole technology. Yes, yes.

P2: No, no, absolutely. No, no. Participant 3 has thought of something very, very smart, absolutely.

AA: And of course, of course. No, no, I have to make an update.

P3: Yes, as for the sounds, they are really beautiful. Yes, clear, it's the classical setting, but they are really useful. I mean, I realize that for someone who is just starting to play now, they are extremely useful. I say for me a bit less, because anyway I'm used to it; I listen to many hours of trumpet. So, this part was not useful to me. For example, I also found the fingering chart useful, but I didn't use it. It's useful though, because actually having it there with the fingerings, for a student who is just starting, it's excellent, in my opinion, it's excellent.

AA: Yes.

P3: As for the dynamics – you had already told me this in our first call – actually, the app only manages the volume of the note, it wasn't really recorded with dynamics, right? Did I understand correctly?

AA: What I did basically for the dynamics in sound is: when it appears *forte*, it's the original sound; when it's *mezzo forte*, I multiply the signal by something like 0.6, or something like that; and when it's *piano*, I multiply it by 0.3, or something like that. Well, I had done a few tests like that. So, it changes a bit; logically, in reality, I would have had to record a trumpeter playing *piano*, playing *mezzo forte*, playing *forte*; in the sense that the timbre itself also changes with the dynamics. But I had only that sound and, in the end, I did it that way.

P3: No, no, but no. But that's perfectly fine, that's perfectly fine. I wanted to tell you, the dynamics are terraced, let's say.

AA: Yes.

P3: It would be nice to maybe change the dynamic level as well.

AA: Ah, like, a crescendo?

P3: Or a diminuendo, yes.

AA: Ok, ok, so be dynamic in the dynamics?

P3: That would be beautiful, very useful.

AA: Yes, yes, yes.

P3: Because, I mean, I started not long ago, and I realize that maintaining intonation in a variable-dynamic sound is much more difficult compared to terraced dynamics.

AA: Yes, yes, yes. Thank you, thank you.

P4: So, I think I'm really the beginner of the group...

P2: So you are our goal, our primary goal. [*jokingly*]

P4: Exactly. Because I started at 46 years old to approach the trumpet having practically not practiced music since I played the recorder in middle school, so I really didn't have an idea, in short, of what it entails. I have a person, a teacher, who follows me. The application has been very useful to me both for identifying the fingering positions which I still struggle with, not with the normal notes [*i.e., the notes of the natural scale*] let's say, but with the various sharps and flats, those slightly stranger things. I still struggle with the fingerings. Moreover, having the feedback of what the sound should be like, hitting the note, that for me was very important. The only thing, I repeat: I have had glasses for a few years, so I'm not even used to those, therefore having the possibility to rotate the screen horizontally to maybe get a bigger picture would be convenient for me. Because, I repeat: seeing the finger positioning on the trumpet helps me in certain cases. Then maybe I do a session with the fingerings and one without looking at the fingerings in order to try to reproduce the sound. That there, I always used the application before warming up, then after that, I did my 10-minute warm-up and then I played my pieces that are played in the band and so I do the band pieces. But it's useful as an application. Also, the various steps that can be enabled and disabled inside. It was useful to me in short.

AA: Perfect. Thank you, yes.

P2: Is it alright, if I may just add something?

AA: Of course.

P2: I'll tell you something I see many musician friends do – I go to hundreds of concerts in my life – so recently I've noticed that everyone around my age has a nice tablet and they download apps there, the Aebersold [*it is a musical method*] and everything. And then during the evenings – well, you play in a band, so carrying a tablet around is not really practical – but at concerts, everyone now puts their tablet on the music stand. So...

P4: Yes, I've seen, I've seen it.

P2: Maybe it's the only solution for the glasses, probably.

AA: Now I'll share the screen where there are various aggregated responses. If we can quickly comment point by point. So, let's see... Share... Here we are! So, easiness of use. I'd like to open a brief discussion among the various points to gather feedback. So then, on ease of use, yes, there were more or less similar opinions, but some discordant. Then "installing the app and learning its features was easy". We have two completely agree, one agree, and one completely disagree. I'm interested, especially for those who didn't find it so easy, in understanding how the app can be improved.

P2: I believe I was the one who agreed on ease, so I think that's what I did.

AA: Ok, ok.

P3: No, I also think I put completely agree because it was very easy to use.

P4: Yes.

AA: Well, then, I don't know what happened with these forms.

P4: No, about the ease, that it's easy to install and use, I don't think there are problems.

AA: Ok.

AA: And regarding adapting the app to your own level? We have completely disagree, neutral, agree, and completely agree. So here too there's a bit of discordance in opinions. What did you find, how to say, challenging? It's cool here to open the discussion, feel free to express your opinion. Indeed, this is a situation where criticisms are welcome!

P2: Maybe me – which is what I said at the beginning when I started speaking earlier – let's say it's not suited to my level. I say "level" in quotes, because it's not like I'm that skilled, but let's say I'm a bit ahead of this and so maybe that's why I said I don't agree on this point.

AA: Ok, ok. And to adapt an exercise to how you wanted it? Was it harder?

P2: I think I also said it before, like doing scales and chords.

AA: Yes, yes, yes. Okay.

P2: I think that's the fundamental thing about music. Always because, since I love playing jazz, if you're in a jam session and clearly if you have the scales and chords memorized, you can join in, otherwise it's better to sit out and not play, let's say.

AA: Yes, yes, yes.

P2: At least that's what happens to me. So, it would be, let's say, useful to extend it. So, all in all, even leaving the structure as it is but I could choose to say: "Practice on the scale on the circle of fifths", right? Even on just the major keys: C major... Scales and chords, scales and chords, scales and chords. From my point of view.

AA: Cool, cool.

P3: Yes, yes, I agree. Scales, arpeggios, things like that, intervals of up a third, a fourth would be very useful. Although then, in reality, putting them on all scales etcetera etcetera would become a huge job anyway. Even like considering sevenths, ninths, it would really become a big job.

P2: But Alberto can do it. [*jokingly*]

P3: No, but I don't doubt that, however, I mean, it becomes a really long thing, also because then you need to find someone who can play them, that is...

P2: Eh, but Alberto knows how to play.

AA: We'll see, we'll see.

P2: I saw on your profile that you've also done something with bands, right? Originally with the band from your town. Then you did other things as well, right?

AA: Yes, yes, yes. I started in a band. Then, of course, I went to the Conservatory and then did everything playing around.

P2: And so you could, you could absolutely record the scales, right?

AA: Yes, but it can also be done via MIDI or recreate with the sounds I have, try to recreate something.

P2: Ah, excellent, even better, even better.

AA: Ok. So if we've finished this, let's move on to "practica deliberata", well it's an Italian translation a bit, I wasn't sure how to translate "deliberate practice". Deliberate practice is a bit the concept, when one starts a study session, to have already quite clear the objectives. Like saying: "Okay, in this study session I want to achieve this goal, I want to try to improve, for example, staccato". And this can be short-term, like today's lesson, or more long-term, like "In a month, or in a week, I want to be able to do this". And so how could the app have influenced this? So: "The app has changed my way of practicing the trumpet". Actually, it hasn't changed much. [*looking at the results of the questionnaire*]

P1: No, not that.

P3: No, not that, you know why? Because, as the others said, maybe it's used for a little warm-up, but then it also depends a bit on our level, on what we want to do, right? It becomes a bit difficult to plan the entire session with the app. I personally used it for a few minutes just as a warm-up and that's it.

AA: Of course, of course.

P2: I agree with Participant 3. Yes, absolutely.

AA: Perfect.

P1: Yes, me too.

AA: Ok, ok. With this question, I really wanted to understand how people eventually integrate the app into their own routine. So, a question, when did you use the app? Especially as a warm-up at the beginning, in the middle, or at the end of your practice?

P3: So, I play a few notes just to warm up, then I use the app for 5 or 10 minutes and then I start with my normal exercises.

AA: Perfect, yes, yes.

P1: Mostly at the beginning, to warm up the kids [*i.e., my students*]. Also there, around 5 or 10 minutes more or less. Now, I wasn't watching the clock...

AA: Yes, sure, I mean, actually, 10 minutes are almost too much. At least for me, I would get bored. Also because, then, playing long notes is one of the parts of the warm-up.

P1: Yes, yes.

P4: Yes, me too, as a warm-up. Before the warm-up, that is, before my warm-up sessions. But not more than a quarter of an hour of warm-up in total between the application and what I usually do, to then move on to the normal pieces.

AA: Participant 2, did you use it more at the beginning?

P2: Ah, sorry, was I the last? Yes. Given that, guiltily, I practice little. Yes, clearly at the beginning; sometimes I've also used it just to hear the notes alone, on their own; so not necessarily in a practice session, but just for the sake of listening to the notes played randomly by the app like that. So, maybe I didn't mention this before, it's useful for ear training, right? So to also train the ear to hear a note with eyes closed, so without reading. I had fun doing that. I've even guessed the right note a few times. So, from this point of view, it's also interesting to use it like that.

AA: M-hm. And just out of curiosity, did you use it with headphones or the device's speaker (e.g., phone, tablet)?

P2: Well, I used headphones.

P1: No, no, I used the normal speakers of the phone.

P4: I also used the normal phone, without headphones.

P3: I also used the speaker.

AA: Perfect. Another point, then: "The app helped to plan and set goals for my study". Here it's interesting because there is one completely disagree, one disagree, one neutral, and one agree.

P3: So, I believe I put completely disagree. Because at this moment my objectives are a bit more articulated than just long notes. So, in the end, I already have a plan and the app doesn't fit into this plan. It's an extra. Clearly, long notes need to be done, so yes, that's fine. But it fits very little into my plan.

AA: Sure, sure.

P1: Yes, for me too, simply my goals are different, so. Let's say that this app with the long notes represents a means to improve intonation, sound, etcetera, etcetera. However, from the point of view of objectives, in my opinion, it's not suitable.

AA: Yes, sure, sure.

P4: No, what can help me is to hear the notes, maybe the higher ones or the lower ones, that I'm not yet doing. Listening to them, and being able to reproduce, the app is useful for that. And still, let's say, my range [*i.e., my extension in terms of pitch*] isn't that great, in short.

AA: Sure, sure.

P2: As for me – I'm probably repeating myself – it's useful for intonation. So, from this point of view, it can help me achieve a goal, which is to tune the notes better. But then one thing is to start a single note, perhaps hearing it from the app; then another thing is to play a series of notes all in tune. So as a first step, it's a great goal, but then it needs to be implemented extended in my opinion.

AA: Ok, ok. Motivation: "Using the app felt like an effort".

P1: That's not the case.

AA: If you have any comments. . .

P3: No, that's not an effort. We just know that long notes need to be done, so maybe that's the effort. [*laughing*]

P2: Right.

AA: Then: "The app motivated me to practice effectively".

P1: Not so much. That's not the case.

AA: Ok.

P1: Because I already have a practice structure of my own. So perhaps for those who still don't have a basic study structure, that yes.

AA: Mhm.

P4: Yes indeed, I repeat, it's useful to see the fingering positions for those who still, like me, struggle to find them all and to have the feedback of what the sound should be like to hit the right note, in short.

AA: Mmm hmm. "The app motivated me to practice for longer periods".

P1: No, not that.

P4: More or less the time for studying is the same.

P2: Yes, for me, yes, because having a study companion, a help, is certainly valid. And it motivates you perhaps to do it rather than practicing by yourself. Even just the idea of having an app, seeing that the world today works on apps, is not bad.

AA: Ok, ok. So even, how to say, the alternation: "Note-rest. Note-rest". Well, here it was more "Rest-Listen-note".

P2: No, that's fine. It's absolutely useful.

AA: Ok, so: "Perceived utility". So, "Using the app I made more progress than I normally do in two weeks of practice". This, of course, is very subjective. We have, neutral and disagree.

P1: No, not that, I put disagree. Because I had already made progress before, so...

AA: Mmmh.

P4: I think I put neutral because it's more or less in line with the warm-up I usually do. I start with long notes, then I do shorter notes and all the various patterns that my teacher gives me, since I'm being followed by someone. So, the progress is similar.

AA: Is the way you measure progress based, how to say, on exercises? I mean, do you realize that maybe before some exercise you couldn't do it and then you can? Or is it more something of an internal feeling? Or there could be a thousand variables...

P3: Are you asking in general, you mean?

AA: In general, yes.

P3: Well, I generally record myself and listen back to realize. Because sometimes when you play with headphones, you might not realize if you're actually in tune or not; but if you record yourself and listen back, maybe you realize if you got some notes wrong, etc. So more or less I measure my progress like this: intonation and the beauty of the sound and being able to do the exercise I set out to do.

AA: Mhm.

P3: Normally in two weeks – except for quite difficult things, like for example the more advanced lessons of Clarke [*it is a method for trumpet*] that maybe you don't do in two weeks – but the other somewhat simpler things, two weeks for me, that is for what I'm studying, are great progress. That's it. I don't know how others do it, indeed I might also be curious to hear what others do.

P4: I manage to do little in two weeks because the time I can dedicate to the trumpet, let's say, is about those 4–5 hours a week. I have an hour a night from Monday to Friday, and when it goes well, I also manage Saturday, but more or less more than an hour a day, it is difficult.

AA: And how do you measure your personal progress with the trumpet?

P4: The progress? Based on the ease with which I can play certain pieces we have in the band, for example. My problem is that of the positions: associating the note with the position. If before I couldn't perform a sequence and then later I can do it more easily, there I see progress. For me, even the Italian Anthem, to play it all as second trumpet, I still struggle; but, little by little, I've seen that last year I could do a few notes only, now I can play almost the whole piece and keep up also with those who have been playing for many years. For me, this is progress. But in the long term, practicing a little at a time, I evaluate my progress over several months, to say. In two weeks, I see little.

AA: Yes, yes, yes. So, this is also very useful for me then to make these questionnaires and understand how to conduct future studies. So, the last point: "I believe that the app could be useful if integrated into a trumpet class". So, the teacher gives some directions on how to use it. For example, I don't know, decides the input data, (i.e., metronome, lowest note, highest note), gives some guidance.

P1: Yes, I agree. Because I essentially used it that way, that's it.

AA: Yes, no, of course, of course. I see just one disagreement, I'd be really interested in the feedback.

P2: No, I agreed, so.

P3: I agreed too.

P4: I agree, but I don't know why there's one disagreement, then if we all agree there's one who doesn't remember what they put.

AA: I want to understand what the glitch was, maybe someone mixed up strongly agree with strongly disagree.

P2: Could it be me who did the questionnaire in three seconds because I joined a bit late? It could be that I made a mistake, but absolutely I agree.

AA: Yes, yes, yes. I'm thinking, if you don't mind, I might resend it. Again, give very honest answers. It won't take long to create a new questionnaire. So I need to duplicate... Let's call it 5 bis. Collect Responses... I'll resend it in the chat. Well, the first part is just repetitive.

P2: Are you making us do homework on a Sunday? [*joking*]

AA: Exactly.

P2: Ah, there we go. So Alberto, the problem is mine because, viewing on the mobile phone, I missed some of the response options, so I think it was me who skewed the results a bit earlier.

AA: Ok, ok. No, not at all. For me, the most important thing indeed is to receive feedback, and then especially verbal feedback. But then surely I think in any publication I will insert the various data and so on. Now what I'm going to do is I need to transcribe what we've said, and then translate it into English, to then make an analysis among all the various results and analyze the results from that perspective. So now feel free, logically, to uninstall the app, in the sense that the study ends here for you. In the future, I want to make changes based on the feedback I have received. This will probably require the use of the microphone, so asking for additional permissions, which for this study were not necessary. So you can very well uninstall it; if you keep it installed there will probably be, I hope during the summer, new updates adding features. Logically, what is perceived as useful I will not change but will maintain. I like doing research on this, I find it fascinating, and I would like then to continue the academic path on this. And then of course, it doesn't depend only on me. I have to find a university that gives me positions as a researcher or professor.

Just one more question on the fly: we've already touched quite a bit on the positive aspects and the negative aspects, if you have anything to add...

P1: I wouldn't have anything to add. Maybe, the crescendo, what Participant 3 was saying a bit. Crescendos and diminuendos could be useful, but for the rest it's fine.

AA: Mhm. And as for improvements? Do you perhaps have any ideas on how to improve the app, if it were yours and you had a computer engineer by your side?

P1: If it were up to me, maybe I would improve the range, I would greatly expand the range. For example, from the pedal note to the triple high C. No, the triple high C seems a bit excessive. But anyway, I would greatly expand the range and then also the dynamics, maybe from fortissimo to pianissimo. Then maybe even fortississimo and pianississimo wouldn't be too bad. Then ah, I would also put in the key signature. That's what I would do, those things there.

AA: Yes, in my question I also take into account adding functionalities. Besides expanding those already present.

P1: Yes, yes, yes. App functionalities, that I wouldn't add, that's it. [*thinking*]

P3: The tuner if you could with access to the microphone. Yes, with the tuner it would be useful because for example I now used it with my external tuner, right?

AA: Yes.

P3: So well, having something together might be useful.

AA: Yes, yes, yes.

P2: Well, the improvements that would be useful for me, I've already mentioned before. What Participant 1 said, he doesn't want to be a teacher anymore, so at this point. [*jokingly*]

I believe that, for that little experience I had with someone who helped me, maybe those extreme things, I believe it's useful indeed to really have a teacher by your side. But many of the things said by Participant 1, I agree with. Absolutely this app can be improved because I believe it can also have a great commercial success. Because I'm sure there are many people who love the trumpet, but the difficulty is extreme. I give a striking example, I don't know if you follow Fiorello's show *Viva RaiDue* [*It is an Italian television show whose Fiorello is the conductor*]. There's Fiorello who often in the morning pulls out the trumpet – which Enrico Rava also gave him, by the way – but he plays worse than me. So, if you gave Fiorello an app to practice, probably he – who is certainly busier than all of us – would probably use it. I said it as an example, but indeed a more constructed, more structured app can absolutely help humanity. Because I believe that the trumpet is a beautiful instrument, but alas – as maybe Participant 4 can tell us – difficult to approach. I started about ten years ago, maybe even a bit more. It's really difficult. If one starts as a young person I believe it's a whole other story. Absolutely.

AA: Participant 4?

P4: Yes, as I was telling you, from what I've seen it was useful to me. Again, if it's possible to have maybe the position in horizontal. And once maybe the highest note and the lowest note are selected, the time, the input data fields can be hidden when I rotate the phone and I only have the trumpet positions and the note I'm playing.

AA: Yes, yes, yes.

P4: It was very useful to have feedback on the type of sound, centered, as it should be. For other things: the range that is already on this app, for me as a beginner, is already more than enough. Then after, if I can hit higher notes, it means I will be much more advanced.

AA: Yes, yes, yes.

P3: Ah, another piece of feedback, I don't know if it's implementable. But it would be nice if, for example, you could upload small video scores and the app plays them with a realistic sound. Because, for example, sometimes a person may have problems with solfeggio or understanding how to find the pitch of notes, maybe in a rather difficult octave interval. But in this case, it would really be about inserting a file into the app, or making it readable, or maybe even retrieving it from an external file and playing it. Because currently there are people, sometimes I do it too, maybe using MuseScore, we reproduce it, we listen to it. But if it were integrated into an app it would be a great advantage because MuseScore is basically on the computer, so it's not very practical.

AA: Yes, on MuseScore you can export the file in XML format and import it into the app. Then obviously it must be first created on the computer by the teacher and then imported. But that certainly yes, it can be done. Ok, I have to work on these feedbacks. Regarding the tuner, having something that provides feedback – a red or green light, something like that – to indicate if the played note is right, do you think it makes sense? Or, I mean, is it enough to hear the reference note first?

P3: No, it makes sense, it makes sense. Because I now use this Korg tuner which basically has lights [*picking up the Korg tuner and showing it on camera*]. Can you see it? There's the red light to indicate if it's flat or sharp and the green light if it's in tune, with also the indicator of the lever.

AA: Yes, yes, it can be seen.

P3: This is very useful because at a glance, even if you don't look at the display, you still see the LED.

AA: Yes, yes, yes. So, doing something similar.

P4: And that's why I use a free application that tells you if you're on the right note or not, called Soundcorset. This one [*showing it*]. But obviously, you should always have the application in sight, whereas if it were already implemented in your app, it could be useful.

AA: Yes, yes, yes. I have the tuner ready; I just have to press the button and upload it. Well, a functionality that tells you the fundamental frequency. Then, well, just need to build a graphical interface and do it. Thank you very much.

P3: Ah, one thing: the tuner, okay to make it chromatic, but maybe make it tailored for the trumpet because with this one I use it basically gives me the real notes. So every time I have to calculate which note I'm playing etc. [*laughing*]. Because it's not for the trumpet and it can't be set..

AA: Ok, ok. So if you play G it tells you F? Because it's a real F of the piano.

P3: Yes, yes.

AA: Ah, ok, ok. And at this point, I had more this set of questions to try to cover different aspects of your experience, but if you have additional feedback speak now; or if you think you've said more or less everything, that's great.

P2: I think I've said it all.

P1: Me too.

P3: Yeah, I think so too.

P4: Me too.

AA: Then, have a great Sunday, thank you very much again.

P1: Have a good Sunday to you too.

P3: Thank you.

AA: Thanks again.

P1: Thank you.

P3: Bye. Have a good Sunday, ciao.

Answers online

Legend:

- **AA**: Principal Investigator
- **P1**: Participant 1
- **P2**: Participant 2
- **P3**: Participant 3
- **P4**: Participant 4

Transcription:

AA: Have you typically used the app with headphones, or have you used the phone/tablet's speaker?

P1: I used the phone/tablet's speaker.

P2: I used the phone/tablet's speaker.

P3: I used the phone/tablet's speaker.

AA: When and how did you use the app? For example, as a warm-up at the beginning, at the end, or in the middle of your study session. For example, for taking long notes, producing the first sounds of the day, improving your high register...

P1: Warm-up.

P2: Warm-up.

P3: Warm-up with long tones.

AA: What have been the positive aspects of the app that you found most useful? (if any)

P1: Sound playback, tempo/speed control.

P2: The ability to choose to play at targeted low, high, or medium pitches and, at the same time, to practice very challenging leaps [*i.e., register changes*], which is fundamental for this type of instrument.

P3: The possibility to listen to the note before playing it and guidance on reading the notes.

AA: What have been the negative aspects, or the least useful features, of the app? (if any)

P1: None.

P2: That only one note can be played at a time; it should be expanded to work on groups of notes as well.

P3: The pause in the middle is too long and cannot be modified. I found the dynamics function less useful for a beginner, which I believe is the target audience more interested in using the app. It is already a step further.

AA: What improvements would you make to the app?

P1: Rhythms, for articulation practice. Trumpet fingering instead of the trumpet image (as an option).

P2: I already answered this earlier: I would expand the ability to play the same thing with duplets, triplets, and quadruplets at the very least.

P3: Enable the option to position the phone horizontally or vertically; Disable the screensaver during the use of the app; Make the trumpet valves more visible (a graphic limited to just the valves would be sufficient, in my opinion).

AA: If you have any additional feedback about your experience, please add it below:

P1: It's all great!

P2: Positive experience, but still needs further development. As it is, it can be useful for beginners and musicians at the next level, but only for certain aspects.

P3: Suggestions and ideas: It would be useful if the app could "listen" to the player and function as a tuner, indicating whether the note is accurate or not. To make it a bit more fun, a module with some simple melodies or easy scales for learning based on the range practiced could be included.

Appendix E

Consent form: Detecting efficiency in trumpet sound production

Participant #:_______

McGill

INFORMED CONSENT FORM
COMPUTATIONAL ACOUSTIC MODELING LABORATORY
SCHULICH SCHOOL OF MUSIC, McGILL UNIVERSITY

REB 430-0415
Detecting efficiency in trumpet sound production

WHY ARE WE DOING THIS RESEARCH? Our aim is to establish scientific knowledge about the perception of sound efficiency in trumpet playing. We seek to understand the relationship between muscular tension and sound efficiency on the produced timbre. Additionally, we aim to develop machine learning and deep learning models that can provide automatic feedback on the degree of efficiency of trumpet sounds, similar to an expert teacher.

PRIVACY. We know that you value your privacy. All data will also be held securely in a password-protected computer system (for electronic data) or a locked office (for paper-based data). Only the principal investigator of the study and participating students will have access to identifiable information. You will not be identified as an individual in any scientific report of this research, and your name will not be linked to your responses in this study unless we have explicitly asked you to provide written consent to be named and quoted. You may discontinue your participation in this study at any time either during the interview process or in the future.

WHAT WILL HAPPEN DURING THE EXPERIMENT?
The experiment will take approximately 1 hour. Once you have completed the experiment you will be paid $***XX*** for your participation. We are interested in how people in general respond to sound played by different players of different technical levels. There are no specific benefits or risks associated with your participation in this study. The stimuli will not be loud enough to cause you discomfort or to adversely affect your hearing. You will be free to discontinue your participation at any time without penalty. **If you are currently taking prescription drugs, over-the-counter drugs (such as antihistamines, cold or flu remedies, sleeping aids), or recreational drugs (such as marijuana, etc.) that would affect wakefulness and attentional focus, please don't participate in the experiment without informing the experimenter that you are taking something.** If you do not feel comfortable discussing this with the experimenter, you are of course free to withdraw from the experiment at any time without prejudice. If you withdraw, you can ask to have the data collected be destroyed. However, once data has been aggregated or published, it can't be withdrawn. It can only be removed from use in further analyses. Identifiable materials will be kept for 7 years following publication. Once the data has been de-identified, it can't be withdrawn. You have the right to refuse to do anything that you find disturbing or uncomfortable in this study. You will be free to discontinue your participation at any time during the experiment or in the future. In this case any data already recorded will be deleted.
Feel free to ask any questions you may have of the experimenter.

This research is funded by the Centre for Interdisciplinary Research in Music Media and Technology.

The Research Ethics Board II of McGill University has reviewed this study for compliance with ethical standards. If you have any ethical concerns or complaints about your participation in this study, and want to speak with someone not on the research team, please contact the Associate Director, Research Ethics at 514-398-6831 or lynda.mcneil@mcgill.ca citing REB file number 430-0415.

PARTICIPANT'S STATEMENT:
"I have read the preceding details and agree to participate. I understand that by consenting, I do not waive any legal rights."

_____________________ _____________________ ______
Signature Printed Name Date

I would like to receive a summary sheet of the experimental findings ☐

E-mail Address: ____

Alberto Acquilino, Ph.D. student, Schulich School of Music, McGill University alberto.acquilino@mail.mcgill.ca
Prof. Gary Scavone, Schulich School of Music, McGill University gary.scavone@mcgill.ca