DEVELOPMENT OF PREDICTIVE MODEL OF DEMO SINGERS' WORK EFFICIENCY USING ADAPTED PERCEPTUAL EVALUATION OF SINGING QUALITY FRAMEWORK IN THE MUSIC RECORDING INDUSTRY IN CHINA

By

HUANG XUEJIE

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

July 2024

FEM 2024 7

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

DEVELOPMENT OF PREDICTIVE MODEL OF DEMO SINGERS' WORK EFFICIENCY USING ADAPTED PERCEPTUAL EVALUATION OF SINGING QUALITY FRAMEWORK IN THE MUSIC RECORDING INDUSTRY IN CHINA

By

HUANG XUEJIE

July 2024

Chairman : Ang Mei Foong, PhD
Faculty : Human Ecology

Demo singers are vocalists who record song demonstrations in recording studios before the songs are officially sung or dubbed by contract artists for release. This category also includes singers who perform for teaching materials. The, requirement of the efficient voice selection and information gap between music education and industry, along with the integration of Artificial Intelligence (AI), have prompted an exploration into the role of “hidden” singers in the music industry. **Methods:** This study adopts a mixed method approach, involving three sessions: observation of recordings, extraction of vocal parameters, and interviews with stakeholders. The research includes ten experienced singers (N=10) with a minimum of eight years of experience, who sang the same original song at Weake Studio, China. The vocal recordings were divided into pieces (N=380), and vocal parameters were extracted using the PESnQ framework. Additionally, interviews were conducted with sixteen stakeholders from four different production teams, each providing unique

perspectives on demo singers. **Results:** The findings revealed outstanding pitch control abilities and personalized dynamic ranges among demo singers. A predictive model was proposed to improve the accuracy of artist efficiency prediction, further exploring parameter directions for future music production and AI voice humanization. Qualitative data analysis identified three levels of output from current musicians: technical output, emotional output, and cultural output. **Implications:** These discussions offer valuable insights to enhance the quality of music content production and promote sustainability in the music industry. Moreover, the study proposes recommendations for potential pedagogical research that align with the current state of the industry.

Keywords: Sing, Prediction, Efficiency, Output, Sustainability

SDG: GOAL 9: Industry, Innovation and Infrastructure

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PEMBANGUNAN MODEL RAMALAN KECEKAPAN KERJA PENYANYI DEMO MENGGUNAKAN KERANGKA PENILAIAN PERSEPSI KUALITI NYANYIAN YANG DIADAPTASI DALAM INDUSTRI RAKAMAN MUZIK DI CHINA

Oleh

HUANG XUEJIE

Julai 2024

Pengerusi : Ang Mei Foong, PhD
Fakulti : Ekologi Manusia

Penyanyi demo adalah penyanyi yang merakam demonstrasi lagu di studio rakaman sebelum lagu tersebut dinyanyikan atau disulih suara secara rasmi oleh artis kontrak untuk dikeluarkan. Keperluan untuk pemilihan suara yang efisien dan jurang maklumat antara pendidikan muzik dan industri, berserta dengan integrasi Kecerdasan Buatan (AI), telah mendorong eksplorasi terhadap peranan penyanyi "tersembunyi" dalam industri muzik. **Kaedah:** Kajian ini menggunakan pendekatan kaedah campuran, yang melibatkan tiga sesi: pengamatan rakaman, pengambilan parameter vokal, dan temubual dengan pemangku kepentingan. Penyelidikan ini melibatkan sepuluh penyanyi berpengalaman (N=10) dengan pengalaman minimum lapan tahun, yang menyanyikan lagu asli yang sama di Weake Studio, China. Rakaman vokal dibahagikan kepada kepingan (N=380), dan parameter vokal diambil menggunakan rangka kerja PESnQ. Selain itu, temubual dijalankan dengan enam belas pemangku kepentingan dari empat pasukan pengeluaran yang

berbeza, masing-masing memberikan pandangan unik mengenai penyanyi demo. **Hasil:** Penemuan menunjukkan keupayaan kawalan nada yang luar biasa dan julat dinamik yang diperibadikan di kalangan penyanyi demo. Model ramalan dicadangkan untuk meningkatkan ketepatan ramalan kecekapan artis, meneroka lebih lanjut arah parameter untuk pengeluaran muzik masa depan dan penghumanan suara AI. Analisis data kualitatif mengenal pasti tiga tahap output dari pemuzik semasa: output teknikal, output emosi, dan output budaya. **Implikasi:** Perbincangan ini menawarkan pandangan berharga untuk meningkatkan kualiti pengeluaran kandungan muzik dan mempromosikan kelestarian dalam industri muzik. Selain itu, kajian ini mencadangkan rekomendasi untuk penyelidikan pedagogi yang berpotensi yang selaras dengan keadaan semasa industri.

Kata Kunci: Nyanyi, Ramalan, Kecekapan, Hasil, Kelestarian

SDG: MATLAMAT 9: Industri, Inovasi dan Infrastruktur

ACKNOWLEDGEMENTS

Thanks to the journey that has beautifully shaped my perspective, I am joyously and wholly immersed in its embrace. It seems that I appear on this journey alone, yet a multitude of background souls propels me on. I would primarily express my gratefulness to my parents who are open mind and support my personal pursuits and dreams whenever and whatever. And thanks to my husband's patience and encouragement in leading me to the intro level of vocal data collection.

I am extremely grateful to my main supervisor Dr. Ang Mei Foong, who always guides me to critical thinking, is willing to take adventure with me and tolerates my temper, even taking risks of my failures in research. However, she shares my successes more happily than anyone else. I enjoy the process of learning quantitative research frameworks with her, publishing articles, sharing knowledge, practicing in industry, organizing international conferences, and communicating with scholars. I show my respect to the co-supervisor Dr. Julia Chieng Chin Yee, who opens my dialectical thinking between humans and AI rather than debate for one's victory or defeat. As well as Assoc. Prof. Dr. Ahmad Faudzi Musib, who always proposes observations and suggestions on audio parameters for me in detail.

I should also thank Prof. Dr. Clare Chan Suet Ching, who gives me constructive in the more logical connection of my thesis. Thank you having Prof. Dr. Chan Cheong Jan and Pro. Dr. Lee Yok Fee, guiding me to embrace different research paradigms and various schools of thought from an inclusive

perspective. Thank you, Prof. Dr. Husniyah Binti Abdul Rahim, for allowing me to enjoy the magical sense of achievement brought by data interpretation. I appreciate that Prof. Dr. Sarjit Singh Darshan Singh encouraged me to look at the function of theory which gives me inspiration for the debate toward the current industry.

What moves me is my colleagues from the recording industry in both Malaysia and China, who provide me with industry experience and make data collection and interviews practical for conducting. Thanks to my studio in China for overcoming the difficulties during COVID-19 and collecting the vocal samples.

To my research collogues, thank you for accompanying me in discussions and helping me to criticize the previous proposes. Finally, I extend heartfelt acknowledgement to myself, for the growth that has granted me the unwavering courage to confront the battlefield I may one day tread alone.

This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Ang Mei Foong, PhD
Senior Lecturer
Faculty of Human Ecology
Universiti Putra Malaysia
(Chairman)

Ahmad Faudzi bin Musib, PhD
Associate Professor
Faculty of Human Ecology
Universiti Putra Malaysia
(Member)

Julia Chieng Chin Yee, PhD
Senior Lecturer
Faculty of Human Ecology
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 7 November 2024

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Demo singers are vocalists who record song demonstrations in recording studios before the songs are officially sung or dubbed by contract artists for release. This category also includes singers who perform for teaching materials. The term “demo” is short for “demonstration”, representing the original version of a song in its early stage of development. Demos are primarily used for the quick study of contract singers, as well as for review by the production team; a qualified demo includes a well-executed arrangement and an accomplished demo voice.

The demo singer is commonly called the “hidden singer” (Long, 2018) in China. In the music recording industry, it is explained that the “side musician/vocalist” plays a significant role in the recording stream (Hull et al., 2011). While in the ethnographical research of the recording studio, “session musicians” is the description that specifically refers to the workers who play instruments and sing for the demonstration (Herbst & Tim, 2018). Additionally, in the Hollywood film industry, a “ghost singer” is a professional demo singer, who sings the song in the background for the actors in films. Marni Nixon was one of Hollywood’s most underappreciated singers. For example, Audrey Hepburn’s voice in My Fair Lady, Deborah Kerr’s in The King and I, and Natalie Wood’s in West Side Story were all dubbed by her (Barnes, 2016). As well as in India, a demo singer is regarded as the “playback singer”, who sings a pre-recorded

song for use in films (Weidman, 2014). Interestingly, for the majority of demo singers' career, though, no one knew who them was.

Unknown or unidentified musicians and hired collaborators produced huge creative compositions in "the shadow of the concert stage" (Williams, 2010a). Due to the recording and producing process "feeds on itself and becomes autocatalytic" (Paul, 2012), what we could identify and clarify came from the practice.

Fortunately, the pre-digitalization period music industry value chain (MIVC) (Hviid et al., 2017) provided us with a general review of the music production process. The MIVC which displayed the content production team was sitting at the beginning of the process, which implied the demo singer was situated in the primary position. Content production was a crucial phase in the digital music era, referring to all the creative work involved before the music was published. This included songwriting, composition, arrangement, demo singing, formal singing, mixing and mastering. Research about critical success factors (CSF) in the music recording industry in Indonesia found that the "high quality of content production link" was one of the CSF (H. S. Saragih, 2017a). Elena Castro- Martínez developed this concept and illustrated that content production was exactly the co-innovation process (Castro-Martínez et al., 2013), which meant the demo singer was the team member for creativity. The roles in the music creation process were compared as "gatekeepers" (Hadida, 2020). The increasing number of discussions silently underpinned that the demo singer, as a member of content production, had a duty on the pre-

released music. As the digitalized music-making process, some demo singers changed into Contemporary Commercial Music (CCM) singers (Portillo et al., 2018).

Over time, demo singers took on many different titles and identities. Their main roles focused on collaboration and evaluation. While they seemed not to be the main players in the production process, they were, in fact, an important part of it. While music was indeed a heavily mediated activity, the sense of immediacy had frequently played a crucial role in shaping people's perception of it (Fosler-Lussier, 2020). The demo singer was indicated to work as the "translator" between the music notes and the acoustics. The strong requirement of the demo singers in the pre-digitalized production related to the special condition, with the most significant factor being the lack of widespread music education. Due to the lack of knowledge in sight-reading, some vocalists relied on audio recordings to rapidly imitate and learn the melody of a new song. The cross-disciplinary performing requirement, such as movie or sports stars who were involved in vocal singing, also depended on demo singers' voices for learning purposes. In music educational material publications, demo singers played a role in demonstrating vocal performance in that they used different vocal styles to showcase songs and meet the expectations of teaching. All above, the role of demo singers in the history of the music industry is significant and lasting, prompting the question of whether their voices possess a meaningful significance that warrants deeper exploration. This inquiry served as the motivation for this study.

> “If you mess up at this stage, you will taste the failure soon. I deeply understand this anxiety, but this (the music content) is a critical task” (Takamitsu.S, 2020).

Drawing from the author’s industry experience, it was clear that musicians are highly attentive to specific feedback from the production team. Individuals like him had effectively translated industry practices into tutorials, creating valuable resources for emerging musicians. Through technology-sharing initiatives, these musicians gained insights into the significant technological demands of the industry. Additionally, demo singers were expected to enhance their ability to shape the musical identity in this context. Given the flourishing digital landscape of the industry, it was important to explore whether there were additional skills that should be imparted to singers to ensure their continuous success.

As digital music surges in popularity, the criteria for high-quality production had evolved over time. Global online music streaming subscriptions, including services like Spotify, Apple Music, and Amazon Music, grew by 32% to reach 358 million in the past year (Counterpoint Research, 2019). Online streaming platforms accounted for over half of the revenues derived from recorded music in-stream sales (IFPI, 2022). A worldwide chart of recorded music revenue indicated that digital music revenue is rising rapidly, a trend that was particularly evident from streaming music sales in the United States over the past three years. Additionally, the growth rate of subscribers to US streaming music services continued to rise (Bolduc & Kinnally, 2018). In China, the revenues of the digital music industry increased from 61.242 billion yuan in

2018 to 62.773 billion yuan in 2019, with digital albums gradually becoming a core component. With the expansion of streaming media, original music was expected to continue capturing a significant share of the music market in the coming years (Entertainment, 2023).

Currently, the Chinese recording industry set higher standards for audio quality. With the advent of 5G mobile networks, 4K resolution, and artificial intelligence (AI) technology, sound engineers were continuously exploring creative techniques like surround and panoramic sound. The pursuit of music production was to achieve a more humanized voice from digital-analogue signals. The sound engineers never paused the exploration of the vocal mastering, for the objective of more humanized voice of the digital analogue signal. However, to be attractive on hearing, the singer's voice was over-processed, even losing the original characteristics in some TV shows. More than this, the "loudness conflict" (a situation in music production that led to certain instruments or vocals being too loud to each other, resulting in an imbalanced and unsatisfactory mix) was becoming serious all over the world (Nannan, 2019). For a healthier perception organ of humans, the control of man-made loudness should be taken. While it was promoted to establish a distinct and comfortable audio-visual experience for the audience, the humanized voice parameters need to be discovered. An increasing number of demo-level compositions were being published on music platforms, highlighting that efficiency would become a crucial asset for demo singers.

A group of musicians who worked in the pre-digitalization era with huge contributions was always “hidden” without name or honour, and some practitioners could not search for any echo from the music market (Herbst & Tim, 2018). This clue provided the possibility that the demo singer had been buried by the rapid industrial development. However, back in the early days of digital music, Bates (2019) explained the appearance of work in recording work is not the actual work, especially intended the technology and art. This spirit encouraged researchers to discover the culture behind the industry and related to the work, better savouring the appearance of the music composition.

The demo singer, as one of the members of content production, was believed to be potentially achieving value in MIVC as an important link spanning the era of pre-digitalization and digitalization. It was the mystery of the demo singer that turned on this research, and further value or expectations were suggested to be found.

1.1 Statement of Problems

Music performing is the most complex task for humans. During performances, people bore different interactive effects and psychological anxieties, which could be identified in Kenny & Osborne’s (2006) research. Performance Success imparted a skill set enabling musicians to perform at their peak potential, effectively channelling energies that might otherwise be squandered in unproductive ways when they go out to sing or play. There were even specialized teams like Osborne et al. (2014) focusing on combating performance anxiety that attempted to help singers alleviate their anxiety.

However, these academic observations are predominantly centred on live performers, and in the case of demo singers, they have received comparatively limited academic scrutiny.

1.1.1 The Selection of Efficient Voice

The most challenging aspect of this research was the lack of a clear voice description of demo singers. Since the beginning of the recording industry, no one of the demo singer's working process was recorded publicly.

Of course, the reason had been noted. It is normal for the industry which was still alive in business always maintained the technical secrets. It was also understood that researching a recording studio obeyed the concept of the "private nature" of the recording studio. In this case, the main reason for the lack of "academic attention" was the difficulties to be accessed both physically and socially, which were absolutely "closed facilities" in a general sense (Watson & Ward, 2013a).

Although there was a lack of voice identification, this group shared a common goal: speed and efficiency. The recording industry required highly streamlined work processes, which meant that demo singers had to master compositions within a limited timeframe. This objective was particularly appealing to budget-constrained studios, as it could help teams reduce costs associated with vocal recording. However, in practice, it became apparent that most demo singers often needed significantly more time than music producers expected. The difficulties were still growing with musicians' idealism which often conflicted

with the aims to make profits (H. S. Saragih et al., 2019), even becoming a huge challenge for the production team to succeed in demo work. Hence, if there was a framework could provide a predictive model for identifying outstanding singers, pre-employed singers could be evaluated based on their vocal parameters to assess their potential for efficient recording. However, at present, the existing vocal frameworks mainly focus on entertainment-oriented predictions, such as rating systems in karaoke (Bohm et al., 2017; Bozkurt et al., 2017; Li et al., 2021; Nakano et al., 2006).

Different types of singers showed different vocal states when they sing, depending on the professional environment in which they work (Limarenko, 2020). In addition, the vocal performance of the demo singer in the recording studio was much different from the live one, especially for the acoustic transmission environment of the two activities.

Thus, identifying the vocal characteristics in their real work environment was the initial target.

In summary, the vocal parameters that defined the characteristics of demo singers were established for primary exploration. From a sound perspective, we should carry out a predictive model after gathering the crucial acoustic variables, which enabled us to assess the singing characteristics of demo singers from various angles.

1.1.2 The Information Gap between Music Education and Industry

Industrial sustainability appealed to keep rethinking the business models to contribute to the efficiency of the industry (Tonelli et al., 2013). Education was no different; it reflected changes in pedagogy as the industry evolved. In numerous cases, the performer's connection to the microphone, to the technology itself, took precedence over their connection to the audience, diminishing its significance (Savage, 2023).

With the opening of the music market, there were increasingly more music distribution platforms, which has given rise to many independent musicians. Since singing has always been taught by the vocal pedagogy, the aims were most primarily considered as classical singing, professional singers were becoming less adaptable to contemporary commercial music market development (Lovetri & Weekly, n.d.). A number of contradictions suggested that there was rarely a training system or model for conservatory graduates who aspired to become independent musicians, such as demo singers.

In a limited space like a recording studio, sound decayed rapidly (Strek, 2020) and the singers' general vocal skills could be less noticeable. This environment necessitates that they developed more specialized skills suited to the unique conditions of vocal production. For example, the fry vocal was commonly regarded as a non-use value in aesthetics, surprisingly, the skill of singing the fry vocal had been widely used in song recording with a transferable construction (Nix et al., 2005). In a live setting, it was often challenging for the audience to hear the details of these special techniques.

However, in recorded audio with microphone amplification, the nuances of these unique sounds were magnified to a level where they could be easily perceived by the listeners. This demonstrated that different methods of sound amplification had long facilitated the application of various vocal techniques.

The demo singer accomplished the whole vocal demo with the energy created by himself because of the absence of audiences. Watson (2013) reported that there were recognized approaches and techniques that could be employed to enhance a demo singer's vocal power. However, due to a lack of data, there were currently no extensive results available to identify the specific energy that the industry demands. Those so-called "specific energy" should be concretized in order to better understand whether education was progressing in sync with the industry.

The constantly changing industries and business models were destined to cause a rethinking of the education system, and educators need to draw from the industry when formulating strategies (Tonelli et al., 2013). Tertiary education had searched for feedback on employability skills (Griffiths et al., 2018) from the close-knit industry. A Malaysian study analysed the feedback from the industry on the engineering students' industrial skills, in which the student's skills cannot meet the expectation of the industry indicated, and the absence of personal soft skills is the main reason (Azmi et al., 2019). Likewise, the absence of industrial targeted training for the demo singer would cause a negative evaluation from the music industry, leading to unhealthy development. Traditional vocal education might not align with the efficient

vocal goals demanded by the industry, and it might also overlook the evolving definition of a good voice within that context. Consequently, it was essential to understand the industry's evaluation of the current status of demo singers in order to gain deeper insights into the future competencies required for vocal talent. This understanding can facilitate the integrated development of industry and education, thereby providing valuable references for employment opportunities for students majoring in music.

1.1.3 The Survival Status and Symbiosis with AI

The rise of digital platforms as modern gatekeepers continued to maintain the elements of the traditional music industry's dominance. While having access to distribution and production tools helped musicians navigate the music scene at that time, it also created new challenges for artists trying to reach their audience. These challenges included increased competition in the market, an unpredictable environment driven by algorithms, and the need for cultural and financial resources. (Frenneaux, 2023).

Through the observation of the decreasing amount of content production musicians, the survival status of the demo singers would also be a sign of consideration (Naveed et al., 2017). The music industry was meeting the decreasing budget for music recording. Recording artists who previously relied on income from record sales were compelled to abandon their careers primarily because copyright service providers occupied the sphere of interest (H. S. Saragih et al., 2019). According to the needs of customers in the

industry, there was a reality to be accepted, that is, Artificial Intelligence (AI) demo singers were being used as a replacement for human vocals.

NEUTRINO was a voice synthesizer that utilized a neural network as its primary framework, enabling the creation of features such as timing, pitch, voice quality, and the blending of voice from music scores. The feature extraction was based on data from the actual song performed by the singer, allowing NEUTRINO to replicate not only the quality of an individual's voice but also their singing style and vocal nuances (Https://n3utrino.work/, n.d.). With the development of the substitute plugin, the human concern of its artistry spontaneously came out. The career of AI singers was flourishing, and it was imperative to study the parameters that contribute to humanized voices in this domain. From the view of independent musicians, the career of the human singer was challenged by the fresh substitute AI plugins. What were the differences between humans and AI in the singing profession, and what advantages did humans possess, inspiring us to explore the outputs of them.

The AI plugin was principally human-driven and consumed, which was pointed out by (Avdeeff, 2019). The concept of improving the treatment of AI singers was introduced. Instead of branding it as a threat to humanity, it would be more constructive to establish parameters that enable it to contribute positively to the music industry. In this scenario, the humanized vocal parameters derived from the AI singer's output were intended to emulate the structure of imitation.

1.1.4 Summary

As a unique “product” of the transition from pre-digitalization to digitalization, demo singers have both contributed to and witnessed the historical development of the music industry, much like the evolution of civilization itself. They have played an essential role in the domain of vocal art and digital music. During the pre-digitalization era, when music education was not yet systematic, demo singers served as a “bridge” that connected composition texts with contract singers in the music industry.

However, their contributions to the music industry were often not well-documented. Issues such as unclear vocal characteristics, a lack of a proper training system, and their living conditions today all hinder research on assessing demo singers. Whether or not their efforts were recognized, demo singers consistently worked behind the scenes in the music industry. While we might not have a definitive explanation of their significance right now, there was no reason to underestimate the impact they have.

1.2 Theoretical Background

The research theoretical framework of this study was based on the PESnQ framework (Gupta & Li, 2017) in Sound Ontology and the concept of Co-innovation in the music business (Samuel et al., 2019).

Sound Ontology advocated research to explain the result of sound from the original point of the ontology mechanism, which was a minority research theory that is contemporarily used in vocal art. It was argued that the sound

was just a form of existence, not the wave itself. Vocalization acted as the cause; sound was an effect. Exploring the inner logic of voicing required the development of a framework rooted in the background of ontology. Sound Ontology represented the pure study of sound without music culture which was produced mainly by Steve Goodman, Christoph Cox, and Greg Hainge, who developed it based on the metaphysical system of philosophy. The sound ontology mainly focused on the sound, the body itself (Kane, 2015).It is believed that the sound we hear is either a wave or an event, which means the sound is a kind of phenomenon (Méndez-Martínez, 2020). In the history of sound study, sound was always regarded as a hypothesis. To destabilize the powerful hypothesis, sound researchers in Europe and America promoted the development of the ontology of sound (Marlene, 2017).

The purpose of implanting sound ontology in this study was to analyse various parameters of sound and establish an aesthetic model more suitable for vocal music, which enabled more identity of sound parameters within the limited range of auditory sensitivity. Human perception of sound mainly relied on the ears, but the missing frequency bands in the sound cannot be obtained. As digital listeners, humans needed to delve into the multiple dimensions of sound and set for the parameters that the ears were not able to perceive. We intended to analyse the vocal elements of singers as comprehensively as possible, hence the analytical parameters mentioned in the modern Sound ontology was applied to the exploration of demo singers.

The theory of Co-innovation refers to the effective convergence of creative resources and elements, combined with the realization of deep cooperation by breaking through the barriers among the industry, discipline, and research (Barile et al., 2020). During the objective-setting process, collaborative members fully harnessed the vitality of innovative elements such as talent, capital, information, and technology. Considering that the business competency demands of co-innovation encompass technology, talent, and capital, this study aligned the identified value domains for singers with their corresponding abilities in vocal skills, musical talent, and social aptitude. These data served as the primary independent variables, facilitating a thorough examination of the vocal characteristics and other abilities demonstrated by demo singers.

Co-innovation theory was commonly used in business research; however, the complexity of business projects made the value generated as co-innovation more and more diversified, the unexpected values came out in the work process. (H. S. Saragih, 2017b) analysed co-innovation processes in the music industry, specifically focusing on jazz production and custom chain, the study highlighted the creation of value. Through this analysis, four dimensions of the creative process were identified, which emphasized that establishing a common goal served as a fundamental prerequisite in musical creation. Co-innovation was proven to actively promote sustainable innovation in the industry, especially when co-innovation groups maintain high capabilities (Adomako & Nguyen, 2023). Co-innovation inspired the framework shaping

this research, that was, obtaining outcomes in cross-interactions of the recording process.

By combining the concept of co-innovation, an evaluation framework for professional demo singers was established. The final direction of all variables should be the intersection, which was the final ability output of the demo singer. The framework in the perspective of the co-innovation compensates for dimensions that could not be adequately captured by acoustic parameters.

1.3 Conceptual Framework

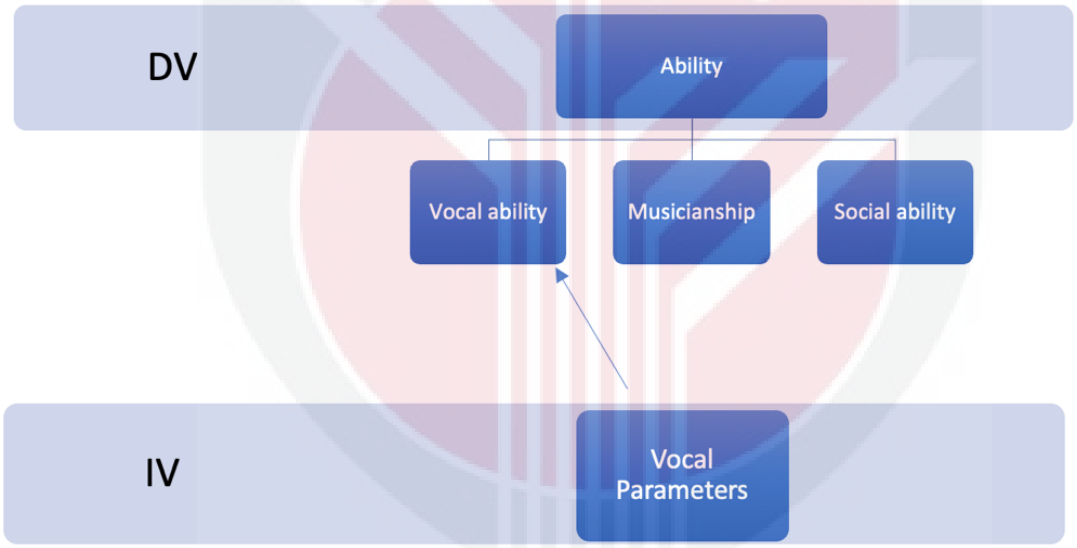

Figure 1.1 : Conceptual Framework of the Present Research

Figure 1.1 illustrates the conceptual framework of the study. The ability concludes the dimensions of the vocal, musical, and social. The study consists of three main sessions: studio recording, observation, and interview. The vocal characteristics are delineated using vocal parameters derived from the enhanced PESnQ framework, while the music industry skills are categorized

based on musicianship and social ability, drawing from the co-innovation theory in teamwork.

1.4 Research Objectives

The overarching aim of this research is to provide additional evidence to elucidate the vocal characteristics and identify the core abilities of demo singers that significantly contribute to the recording process. The specific objectives of these studies are as follows:

1. To describe the vocal characteristics of demo singers.
2. To build a predictable model for a singer's singing efficiency and qualification.
3. To assess the contribution of human singers compared with AI singers.

1.5 Research Questions

The research questions of this study are as follows:

1. What are the ideal parameters for a professional demo singer?
2. How to predict the singer's efficiency and qualification through the vocal parameters?
3. What is the difference between a human versus AI demo singer?

1.6 Significance of the Study

The research provided an analysis paradigm across vocal and acoustic disciplines, filling the shortcomings of sound parameters being analysed in isolation. It provided a model framework for predicting sound for some

industries, which was expected to save assessment costs or be used in the entertainment and AI industries. Three levels for evaluating the output of artistic workers were built, which was used in scientific performance evaluation practice. The output of the demo singer at three levels provided reform ideas for the music education industry and further reminded educators of the importance of setting a curriculum system that was closely associated with the music industry. As to the methodology, using acoustic analysis to predict singing efficiency and determine voice quality was a new initiative.

The vocal characteristics study contributed to enriching the standards based on evaluating the professional demo singers' singing level, which provided a glimpse of digital reference to various characteristics of the vocal and enabled teachers to give a more reliable judgment of qualified demo singers in combination with their hearing. As for the undergraduate singers, they could immediately get a detailed understanding of the music industry skills of the recording studio market. Compared to past vocal studies, this research offered more comprehensive parameters covering vocal production.

This research enhanced the design ideas in the music curriculum system, particularly in the context of training market-responsive singers. The findings could be utilized to refine and improve the training approaches and strategies employed in the education of singers, ensuring they are better prepared to meet the demands of the music industry. It created motivations for the conservatories to revolutionize the teaching of vocal music, allowing them to accurately understand the industry standards and feedback. More targeted

pedagogies especially for demo singers inspire to improve the education level of the music-applied field for conservatories, as well as the graduates' employment rate in the music industry.

Vocal predicting models utilized advanced parameters of Pitch, Rhythm Consistency, Emotional Coherence, singing duration, and so on to generalize more dimensions of singers. By considering a wider range of vocal characteristics, these models could provide a more nuanced and accurate assessment of the singer's vocal performance. Incorporating a vocal predicting model into the karaoke scoring system, enabled a more comprehensive evaluation of the singer's voice, resulting in a more detailed and precise scoring mechanism. This advancement in technology enhanced the overall karaoke experience and provides singers with valuable feedback on their vocal skills and performance.

Besides, the exploration of music industry skills contributes to the clear aim of searching for singers in the recording industry, which will further improve the efficiency of music production. Recording is the mean stream in the music industry, the growth of streaming music demands new means and work processes in recording (Hull et al., 2011). The study provides the means of the management of recording time through the co-process of demo singer and master, which attempts to fundamentally overcome the current dilemma in music production. It is suggested that the predictive model for vocal ability be implemented as the outcome to be applied in the Automatic Scoring System for Karaoke. The current karaoke scoring systems mainly extract parameters

such as pitch and duration, which provide limited information for assessing the singer's vocal performance (Lao, 2006). In contrast, a vocal predicting model can extract a broader range of vocal dimensions to evaluate the singer's vocal ability more comprehensively.

Further discussion of the developing AI plugin contributes to refreshing the function and role the demo singer is playing in the current process, which also provides market aesthetic standards for outlook. The implications of domesticating the more humane numbers but not letting the numbers replace us humanized the scholars' primary goal of the contemporary study.

1.7 Limitations of the Study

The study limitations are as follows:

Comprehensive modelling of perspectives was the most important limitation. Since the research objective is to generalize the perspectives for output model construction, the mere number of reference sentences was recorded at this level, which could not focus on a certain perspective for analysis one by one. This lied in the theoretical background that was constructed by vocal ability, musicianship, and social ability. The three-level output model could just use to explain their contribution within vocal ability, musicianship, and social ability.

Limitation of the "Parameter aesthetic model" which was now still an exploration in the analysis of vocal performance. The dimensions of the physical characteristics in vocal were currently discovered, but not complete. The judgment for the variables was generalized from different acoustic fields,

which allowed variations in the present in the data extraction process. The employed software was within the industry for vocal acoustic evaluations. This choice might result in differences in evaluation scores and standards when compared to assessments in fields such as physical acoustics or psychoacoustics.

Furthermore, the predictive model was based on assumptions derived from the professional demo singer group. The selection of vocal parameters was deliberate, and the final samples were chosen to represent the singer's peak performance capabilities. Achieving this level of performance involved multiple iterations, with the model utilizing the parameters from their most recent peak performance for prediction. The degree of accuracy for non-professional singers remained unverified.

Additionally, the realms of emotion and culture were intricately intertwined in music creation, and when determining coding terminology, the process was closely tied to the researchers' educational backgrounds, unavoidably incorporating subjective judgments. To determine whether these aspects intersect or coexist, further industry case studies were required for validation. The development of output models across these three levels should continue to adapt to the rapidly evolving industry landscape.

1.8 Definition of Terms

Demo singer: the singers who record songs demonstration in a music recording studio before the song is being sung by the contract artist for official release.

Demo: is short of the word “demonstration”, and it is the original version of a song in the newborn stage.

AI singer: It develops from visual singer, who is the virtual idol known for the visually appealing anime-style appearance, often creates using computer graphics, and perform song with their own personalities, backstories, and fanbases, through music videos, concerts, and media productions. The AI singers referred to below are human voice replacement plug-ins used in music production. Like humans, they also have a fixed identity and recognition in the industry.

Recording Studio: A specialized facility for sound recording, engineering, and production, which ranges from a small in-home studio to a large building.

Live: The singer sings through the vocal cords, without any pause or software mastering during the performance.

Online streaming platforms: refer to digital services that allow users to watch or listen to content over the internet in real-time, without the need to download

the entire file. It enables users to stream or play media files, such as movies, TV shows, music, podcasts, and more, directly from a server or cloud storage.

Music Content Production: refers to the process of creating and producing music for various purposes, such as albums, singles, soundtracks, commercials, and other audiovisual projects. It involves several stages and professionals working together including Songwriting, Recording, Editing, Mixing, and Mastering.

Vocal characteristics: refer to qualities that include but are not limited to the range, weight, tessitura, timbre, transition of vocal, image-molding, and emotional ability, which are divided into two sections, namely vocal skills, and re-creation ability.

Sound Ontology: Before the 18th century, a school of philosophy based on Descartes used reductionism and structuralism to argue for music, which is against the rise of the metaphysical theory of music in the 19th century.

PESnQ: Perceptual Evaluation of Singing Quality framework proposed by Gupta's team for comprehensive evaluation of parameter synthesis for singing assessment.

Co-innovation: refers to the effective convergence of creative resources and elements, combined with the realization of deep cooperation by breaking through the barriers among the industry, discipline, and research.

REFERENCES

Adomako, S., & Nguyen, N. P. (2023). Co-innovation behavior and sustainable innovation in competitive environments. *Sustainable Development*, *31*(3), 1735–1747. https://doi.org/10.1002/sd.2479

Adomavicius, G., Bockstedt, J. C., Gupta, A., & Kauffman, R. J. (2007). Technology roles and paths of influence in an ecosystem model of technology evolution. *Information Technology and Management*, *8*(2). https://doi.org/10.1007/s10799-007-0012-z

Aghajanzadeh, M., Saeedi, S., Jalaie, S., Esarian, K., & Barsties v. Latoszek, B. (2023). Validation of the Acoustic Voice Quality Index and the Acoustic Breathiness Index in the Persian Language. *Journal of Voice*, *25*(2), 1–7. https://doi.org/10.1016/j.jvoice.2023.07.002

Anand, S., Wingate, J. M., Smith, B., & Shrivastav, R. (2012). Acoustic parameters critical for an appropriate vibrato. *Journal of Voice*, *26*(6), 820.e19-820.e25. https://doi.org/10.1016/j.jvoice.2012.06.004

Anonymous. (1874). Science in music. *Nature*, *11*(1874), 88–89. https://doi.org/https://doi.org/10.1038/011088c0

Athanasios, L., & Alexander, L. (2015). Beat histogram features for rhythm-based musical genre classification using multiple novelty functions. *18th Int. Conference on Digital Audio Effects*, 1–8.

Atkinson, P., Coffey, A., Delamont, S., Lofland, J., & Lofland, L. (2007). Semiotics, Semantics and Ethnography. In *Handbook of Ethnography* (p. 151). Sage.

Avdeeff, M. (2019). Artificial Intelligence & Popular Music: SKYGGE, Flow Machines, and the Audio Uncanny Valley. *Arts*, *8*(4), 130. https://doi.org/10.3390/arts8040130

Azeem, M., Hussain, S., Ijaz, M., Salahuddin, N., & Salam, A. (2023). An improved version of systematic sampling design for use with linear trend data. *Heliyon*, *9*(6), e17121. https://doi.org/10.1016/j.heliyon.2023.e17121

Azizan, M. H., Sadun, A. S., & Jalaludin, A. (2022). Singing Pitch Guide for Multiple Voice Pitch Offset Detector. *Progress in Engineering Application and Technology*, *3*(1), 479–484.

Azkarate, A. A. (2017). *Emotions as Aesthetic Properties of Absolute Music* [Universidad del Pais Vasco]. https://addi.ehu.es/bitstream/handle/10810/24189/TESIS_AGIRREBALZATEGI_AZCARATE_ANA.pdf?sequence=1

Azmi, A. N., Kamin, Y., Md Nasir, A. N., & Noordin, M. K. (2019). The engineering undergraduates industrial training programme in Malaysia: Issues and resolutions. *International Journal of Engineering and Advanced Technology*, *8*(5), 405–419. https://doi.org/10.35940/ijeat.E1058.0585C19

Banse, R., & Scherer, K. R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

Barile, S., Grimaldi, M., Loia, F., & Sirianni, C. A. (2020). Technology, value co-creation and innovation in service ecosystems: Toward sustainable co-innovation. *Sustainability (Switzerland)*, *12*(7). https://doi.org/10.3390/su12072759

Barnes, M. (2016). Marni Nixon, famed 'Ghost Singer' in Classic Musicals, Dies at 86. *The Hollywood Reporter*. https://www.hollywoodreporter.com/news/general-news/marni-nixon-dead-ghost-singer-914202/

Barsties v. Latoszek, B., Lehnert, B., & Janotte, B. (2020). Validation of the Acoustic Voice Quality Index Version 03.01 and Acoustic Breathiness Index in German. *Journal of Voice*, *34*(1), 157.e17-157.e25. https://doi.org/10.1016/j.jvoice.2018.07.026

Barsties V Latoszek, B., Ulozaitė-Stanienė, N., Petrauskas, T., Uloza, V., & Maryn, Y. (2019). Diagnostic Accuracy of Dysphonia Classification of DSI and AVQI. *The Laryngoscope*, *129*(3), 692–698. https://doi.org/10.1002/lary.27350

Bates, E. (2019). Technological encounters in the interculturality of Istanbul's recording studios. *Oido Pensante*, *7*(1).

Battcock, A., & Schutz, M. (2021). Individualized interpretation: Exploring structural and interpretive effects on evaluations of emotional content in Bach's Well Tempered Clavier. *Journal of New Music Research*, *50*(5), 447–468. https://doi.org/10.1080/09298215.2021.1979050

Bergquist, A. K. (2019). Business and Sustainability. In *The Routledge Companion to the Makers of Global Business*. Springer. https://doi.org/10.4324/9781315277813-40

Bishop, L. (2018). Collaborative musical creativity: How ensembles coordinate spontaneity. *Frontiers in Psychology*, *9*(July), 1–17. https://doi.org/10.3389/fpsyg.2018.01285

Black, K. (2010). *Business Statistics: For Contemporary Decision Making* (6th ed.). Wiley.

Blake, P. (2020). *Music Industry Skills* (Vol. 2). Pearson Education. https://doi.org/https://qualifications.pearson.com/content/dam/pdf/btec-level-2-skills/music-industry-skills/2017/specification-and-sample-assessments/specification-music-industry-skills.pdf

Bohm, J., Eyben, F., Schmitt, M., Kosch, H., & Schuller, B. (2017). Seeking the Super Star: Automatic assessment of perceived singing quality. *Proceedings of the International Joint Conference on Neural Networks*, *May*, 1560–1569. https://doi.org/10.1109/IJCNN.2017.7966037

Bolduc, H., & Kinnally, W. (2018). Examining the Impact of Social Identification with Music on Music Streaming Behavior. *Journal of Radio and Audio Media*, *25*(1), 42–61. https://doi.org/10.1080/19376529.2017.1362893

Bonassi, G., Lagravinese, G., Bove, M., Bisio, A., Botta, A., Putzolu, M., Cosentino, C., Mezzarobba, S., Pelosin, E., & Avanzino, L. (2023). How Music Moves Us: Music-induced Emotion Influences Motor Learning. *Neuroscience*, *526*(July), 246–255. https://doi.org/10.1016/j.neuroscience.2023.06.023

Bozkurt, B., Baysal, O., & Yüret, D. (2017). A Dataset and Baseline System for Singing Voice Assessment. *The 13th International Symposium on Computer Music Multidisciplinary Research*, *c*, 1–9. http://cmmr2017.inesctec.pt/wp-content/uploads/2017/09/43_CMMR_2017_paper_31.pdf

Buren, V., Müllensiefen, D., Roeske, T., & Degé, F. (2021). What makes a child musical? conceptions of musical ability in childhood. *Early Child Development and Care*, *191*(12), 1985–2000. https://doi.org/10.1080/03004430.2020.1866566

Butkovi, A., Vukojevi, N., & Carevi, S. (2022). Music performance anxiety and perfectionism in Croatian musicians. *Psychology of Music*, *50*(1), 100–110. https://doi.org/10.1177/0305735620978692

Campayo-Muñoz, E., Cabedo-Mas, A., & Hargreaves, D. (2020). Intrapersonal skills and music performance in elementary piano students in Spanish conservatories: Three case studies. *International Journal of Music Education*, *38*(1), 93–112. https://doi.org/10.1177/0255761419873782

Campbell, W., Toulson, R., & Paterson, J. (2010). The effect of dynamic range compression on the psychoacoustic quality and loudness of commercial music. *39th International Congress on Noise Control Engineering 2010, INTER-NOISE 2010*, *4*, 2580–2588.

Cao, C., Li, M., Wu, X., Suo, H., Liu, J., & Yan, Y. (2009). Automatic singing performance evaluation for untrained singers. *IEICE Transactions on Information and Systems*, *E92-D*(NO.8), 1596–1600. https://doi.org/10.1587/transinf.E92.D.1596

Castro-Martínez, E., Recasens, A., & Jiménez-Sáez, F. (2013). Innovation systems in motion: An early music case. *Management Decision*, *51*(6), 1276–1292. https://doi.org/10.1108/MD-11-2011-0433

Chalmers, D., Matthews, R., & Hyslop, A. (2021). Blockchain as an external enabler of new venture ideas: Digital entrepreneurs and the disintermediation of the global music industry. *Journal of Business Research*, *125*(September 2019), 577–591. https://doi.org/10.1016/j.jbusres.2019.09.002

Chitralekha Gupta, & Rao, P. (2011). Objective Assessment of Ornamentation in Indian Classical Singing. *Speech, Sound and Music Processing: Embracing Research in India*, *LNCS 7172*(1), 1–25. https://doi.org/https://doi.org/10.1007/978-3-642-31980-8_1

Choong, H. (2023). The learning experiences and musical proficiencies of formal and informal popular musicians in Malaysia. *International Journal of Music Education*, *41*(3), 358 –370. https://doi.org/10.1177/02557614221107917

Counterpoint Research. (2019). *Combined Global Market Share of Huawei, OPPO, Vivo, Xiaomi and Realme Reaches Highest-Ever Level in Q2 2019*. https://www.counterpointresearch.com/combined-global-market-share-huawei-oppo-vivo-xiaomi-realme-reaches-highest-ever-level-q2-2019/

Davidova, J., Zavadska, G., & Rauduvaite, A. (2019). Levels and Indicators for Assessment of Singing Voice of 6–8Year-old Children. *Pedagogika*, *133*(1), 149–160.

de Bruin, L. R. (2019). Improvising musicians' self-regulation and distributed creativities: A phenomenological investigation. *Thinking Skills and Creativity*, *32*(April 2017), 30–41. https://doi.org/10.1016/j.tsc.2019.03.004

Demmer, T. R., Kühnapfel, C., Fingerhut, J., & Pelowski, M. (2023). Does an emotional connection to art really require a human artist? Emotion and intentionality responses to AI- versus human-created art and impact on aesthetic experience. *Computers in Human Behavior*, *148*(November), 107875. https://doi.org/10.1016/j.chb.2023.107875

Diani, M., & Sacchetti, S. (2023). Embedded performers: The relational foundations of record production. *Social Networks*, *74*(April), 206–215. https://doi.org/10.1016/j.socnet.2023.04.003

Diaz, J. A., Rangarathnam, A. B., Medina, A., Khare, V. D., & Fisher, H. R. (2023). Preliminary Perceptual Findings of Singing After Flow Phonation Voice Exercises for Singers in Training. *Journal of Voice*, *37*(5), 1–8. https://doi.org/10.1016/j.jvoice.2023.08.014

Dorschel, A. (2018). Music as play: a dialogue. In *Virtual Works - Actual Things* (pp. 115–134). Leuven University Press. https://doi.org/10.11116/9789461662521.ch05

Doxat-Pratt, S. (2021). Musical communities in the society of captives: Exploring the impact of music making on the social world of prison. *Musicae Scientiae*, *25*(3), 290–302. https://doi.org/10.1177/1029864921101761 6

Duvvuru, S., & Erickson, M. (2012). *The Effect of Timbre, Pitch, and Vibrato on Vocal Pitch Matching Accuracy*. (Doctoral dissertation). https://doi.org/10.1016/j.jvoice.2015.05.011

Eidsheim, N. S. (2011). Sensing Voice-Materiality and the Lived Body in Singing and Listening.pdf. *Science&Society*, *6*(2), 133–155.

Eidsheim, N. S. (2018). The Race of Sound. In *The Race of Sound*. Duke University Press. https://doi.org/10.2307/j.ctv11hpntq

Entertainment, T. M. (2023). *Tencent Music Entertainment Group 2023 First Quarter Earnings Webinar*. https://www.ncbi.nlm.nih.gov/books/NBK558907/

Fosler-Lussier, D. (2020). Sound Recording and the Mediation of Music Book Title: Music on the Move. In *Music on the Move* (p. 95). University of Michigan Press Chapter.

Frenneaux, R. (2023). The rise of independent artists and the paradox of democratisation in the digital age: Challenges faced by music artists in the new music industry. *DIY, Alternative Cultures & Society*, *1*(2), 125–137. https://doi.org/10.1177/27538702231174200

Goldmacher, P. (2016). *The Songwriter's Guide to Recording Professional Demos*.

Griffiths, D. A., Inman, M., Rojas, H., & Williams, K. (2018). Transitioning student identity and sense of place: future possibilities for assessment and development of student employability skills. *Studies in Higher Education*, *43*(5). https://doi.org/10.1080/03075079.2018.1439719

Gupta, C., Haizhou, L., & Ye, W. (2018). A technical framework for automatic perceptual. *APSIPA Transactions on Signal and Information Processing*, *7*(2018), 1–11. https://doi.org/10.1017/ATSIP.2018.10

Gupta, C., & Li, H. (2017). Perceptual Evaluation of Singing Quality. *APSIPA Annual Summit and Conference 2017*, *December*. https://doi.org/10.1109/APSIPA.2017.8282110

Gupta, C., Li, H., & Goto, M. (2022). Deep Learning Approaches in Topics of Singing Information Processing. *IEEE/ACM Transactions on Audio*

Speech and Language Processing, *30*, 2422–2451. https://doi.org/10.1109/TASLP.2022.3190732

Gupta, C., Li, J., & Li, H. (2021). Towards Reference-Independent Rhythm Assessment of Solo Singing. *2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2021 - Proceedings*, *April*, 912–919.

Hadida, A. L. (2020). Managerial cognition and the value chain in the digital music industry. *The Grants Register 2021*, *44*(0), 927–928. https://doi.org/10.1057/978-1-349-95988-4_915

Hak, C. C. J. M., & Wenmaekers, R. H. C. (2010). The impact of sound control room acoustics on the perceived acoustics of a diffuse field recording. *WSEAS Transactions on Signal Processing*, *6*(4).

Hallam, S. (2010). 21st century conceptions of musical ability. *Psychology of Music*, *38*(3), 308–330. https://doi.org/10.1177/0305735609351922

Hamilton, J. (2021). Across the Great Divide: Popular Music Studies and the Public. *Twentieth-Century Music*, *18*(1), 29–44. https://doi.org/10.1017/S1478572220000250

Han, Q., & Zhang, R. (2017). Acoustic Analyses of the Singing Vibrato in Traditional Peking Opera. *Journal of Voice*, *31*(4), 511.e1-511.e9. https://doi.org/https://doi.org/10.1016/j.jvoice.2016.11.016

Hargreaves, D. J., & MacDonald, R. A. R. (2012). Designing Improvisation: Intercultural Collaboration and Musical Imagination. *Procedia - Social and Behavioral Sciences*, *45*(2012), 14–20. https://doi.org/10.1016/j.sbspro.2012.06.538

Herbst, J., & Tim, A. (2018). The Work Realities of Professional Studio Musicians in the German Popular Music Recording Industry: Careers, Practices and Economic Situations. *IASPM Journal*, *8*(2), 18–37. https://doi.org/10.5429/2079-3871(2018)v8i2.3en

Hindrichs, G. (2019). Towards a General Theory of Musical Works and Musical Listening. In *Essays in Music Ontology*. Leuven University Press. https://doi.org/10.2307/j.ctv4rfrd0.6

Hong, S. M., Oh, S. J., & Lee, E. J. (2020). Musical Storytelling. *Journal of Physiology & Pathology in Korean Medicine*, *34*(2020), 299–308. https://doi.org/10.1080/07494469608629689

Hosokawa, K., Barsties, B., Iwahashi, T., Iwahashi, M., Kato, C., Iwaki, S., Sasai, H., Miyauchi, A., Matsushiro, N., Inohara, H., Ogawa, M., & Maryn, Y. (2017b). Validation of the Acoustic Voice Quality Index in the Japanese Language. *Journal of Voice*, *31*(2), 260.e1-260.e9. https://doi.org/10.1016/j.jvoice.2016.05.010

Howard, D. M., Daffern, H., & Brereton, J. (2012). Quantitative voice quality analyses of a soprano singing early music in three different performance styles. *Biomedical Signal Processing and Control*, *7*(1), 58–64. https://doi.org/10.1016/j.bspc.2011.05.013

Howes, P., Callaghan, J., Davis, P., Kenny, D., & Thorpe, W. (2004). The relationship between measured vibrato characteristics and perception in Western operatic singing. *Journal of Voice*, *18*(2), 216–230. https://doi.org/10.1016/j.jvoice.2003.09.003

Https://n3utrino.work/. (n.d.). *NEUTRINO*.

Huang, X., & Ang, M. F. (2022). The Pitch Control Ability: Discovery on the Reality of Demo Singers compared with Music Industry Criterion. *4th International Music & Performance Art Conference*, 18–23.

Hull, G., Hutchison, T., & Strasser, P. (2011). *The music business recording industry*. Routledge, Page223, ISSBN978-0415875615.

Hutchins, S., Larrouy-Maestri, P., & Peretz, I. (2014). Singing ability is rooted in vocal-motor control of pitch. *Attention, Perception, and Psychophysics*, *76*(8), 2522–2530. https://doi.org/10.3758/s13414-014-0732-1

Hutchins, S., & Moreno, S. (2013). The Linked Dual Representation model of vocal perception and production. *Frontiers in Psychology*, *4*(November), 1–12. https://doi.org/10.3389/fpsyg.2013.00825

Hviid, M., Jacques, S., & Sanchez, S. I. (2017). Digitalisation and intermediaries in the music industry [University of Glasgow]. In *CREATe Working Paper* (Vol. 07, Issue March). https://doi.org/10.2966/scrip.150218.242

IFPI. (2022). *Engaging with Music 2022*.

Jayakumar, T., & Benoy, J. J. (2022). Acoustic Voice Quality Index (AVQI) in the Measurement of Voice Quality: A Systematic Review and Meta-Analysis. *Journal of Voice*, *36*(2). https://doi.org/10.1016/j.jvoice. 2022.03.018

Jeremić, B., & Stanković, E. (2020). The Methodical model of teaching songs by ear and its effects on the development of students' vocal abilities. *Croatian Journal of Education*, *22*(specialedition), 151–166. https://doi.org/10.15516/cje.v22i0.3838

Jong, de, J. B. (2006). Collective Talent: A Study of Improvisational Group Performance in Music. In *Dutch Ministry of Justice*. Amsterdam University Press. https://doi.org/10.5117/9789056294427

Kadiri, S. R., Javanmardi, F., & Alku, P. (2023). Investigation of self-supervised pre-trained models for classification of voice quality from speech and

neck surface accelerometer signals. *Computer Speech and Language*, *83*(July), 101550. https://doi.org/10.1016/j.csl.2023.101550

Kane, B. (2015). Sound studies without auditory culture: a critique of the ontological turn. *Sound Studies*, *1*(1), 2–21. https://doi.org/10.1080/20551940.2015.1079063

Katok, D. (2016). *The Versatile Singer: A Guide to Vibrato & Straight Tone*. City University of New York.

Kayes, G., & Welch, G. F. (2017). Can Genre Be "Heard" in Scale as Well as Song Tasks? An Exploratory Study of Female Singing in Western Lyric and Musical Theater Styles. *Journal of Voice*, *31*(3), 388.e1-388.e12. https://doi.org/10.1016/j.jvoice.2016.09.015

Kenny, D. T., & Osborne, M. S. (2006). Music performance anxiety: New insights from young musicians. *Advances in Cognitive Psychology*, *2*(2), 103–112. https://doi.org/10.2478/v10053-008-0049-5

Klarlund, M., Brattico, E., Pearce, M., Wu, Y., Vuust, P., Overgaard, M., & Du, Y. (2023). Worlds apart? Testing the cultural distance hypothesis in music perception of Chinese and Western listeners. *Cognition*, *235*(February), 1–12. https://doi.org/10.1016/j.cognition.2023.105405

Kristeva, R., Chakarov, V., Schulte-Mönting, J., & Spreer, J. (2003). Activation of cortical areas in music execution and imagining: A high-resolution EEG study. *NeuroImage*, *20*(3), 1872–1883. https://doi.org/10.1016/S1053-8119(03)00422-1

Kumar, S., & Kandoi, M. (2018). Co-Creation A New Method of Building Brand Loyalty & Innovation. *International Journal of Management (IJM)*, *9*(6), 13–23. https://doi.org/10.13140/RG.2.2.34584.65281

Kwak, D. H. (2010). *Auditory loudness perception and the consumption of popular music between 1980 and 2010* [University of Pretoria]. http://hdl.handle.net/2263/51284

Lao, Z. (2006). *Method and Device for Determining Karaoke Scores*. https://doi.org/https://patentimages.storage.googleapis.com/2d/26/42/cdeb35e7b7cf0c/CN108008930B.pdf

Larrouy-Maestri, P., Magis, D., & Morsomme, D. (2014). The Evaluation of Vocal Pitch Accuracy. *Music Perception*, *32*(1), 1–10. https://doi.org/10.1525/mp.2014.32.1.1

Larrouy-Maestri, P., & Pfordresher, P. Q. (2018). Pitch perception in music: Do scoops matter? *Journal of Experimental Psychology: Human Perception and Performance*, *44*(10), 1523–1541. https://doi.org/10.1037/xhp0000550

Latartara, J., & Gardiner, M. (2007). Analysis, Performance, and Images of Musical Sound: Surfaces, Cyclical Relationships, and the Musical Work. *Current Musicology*, *84*(84), 53–78. https://doi.org/https://journals.library.columbia.edu/index.php/currentmusicology/article/view/5098

Lee, S. M., Olson, D. L., & Trimi, S. (2012). Co-innovation: Convergenomics, collaboration, and co-creation for organizational values. *Management Decision*, *50*(5), 817–831. https://doi.org/10.1108/00251741211227528

Lerch, A., Arthur, C., Pati, A., & Gururani, S. (2020). An Interdisciplinary Review of Music Performance Analysis. *Transactions of the International Society for Music Information Retrieval*, *3*(1), 221–245. https://doi.org/10.5334/tismir.53

Li, J., Gupta, C., & Li, H. (2021). Training Explainable Singing Quality Assessment Network with Augmented Data. *2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2021 - Proceedings*, *December*, 904–911.

Limarenko, V. (2020). Methodical competence of the future leader of the vocal ensemble: categorical analysis of the problem. *Academic Notes Series Pedagogical Science*, *1*(189), 189–192. https://doi.org/10.36550/2415-7988-2020-1-189-189-192

Liu, Y., Liu, G., Wei, D., Li, Q., Yuan, G., Wu, S., Wang, G., & Zhao, X. (2018). Effects of musical tempo on musicians' and non-musicians' emotional experience when listening to music. *Frontiers in Psychology*, *9*(NOV), 1–11. https://doi.org/10.3389/fpsyg.2018.02118

Long, C. (2018). *The mystery of the Demo singer behind the pop music. 隱身流行曲背後 Demo 歌手的神秘面紗*. https://doi.org/https://www.metropop.com.hk/

Lotze, M., Scheler, G., Godde, B., Erb, M., Grodd, W., & Birbaumer, N. (2000). Comparison of fMRI-activation maps during music execution and imagination in professional and non-professional string players. *NeuroImage*, *11*(5), 2000. https://doi.org/10.1016/s1053-8119(00)91001-2

Lotze, M., Scheler, G., Tan, H. R. M., Braun, C., & Birbaumer, N. (2003). The musician's brain: Functional imaging of amateurs and professionals during performance and imagery. *NeuroImage*, *20*(3), 1817–1829. https://doi.org/10.1016/j.neuroimage.2003.07.018

Lovetri, J. L., & Weekly, E. M. (2003). Contemporary Commercial Music (CCM) Survey: Who's Teaching What in Nonclassical Music. *Journal of Voice*, *17*(2), 207–215. https://doi.org/10.1016/S0892-1997(03)00004-3

Lozada, N., Arias-Pérez, J., & Perdomo-Charry, G. (2019). Big data analytics capability and co-innovation: An empirical study. *Heliyon*, *5*(10), 1–7. https://doi.org/10.1016/j.heliyon.2019.e02541

Manfredi, C., Barbagallo, D., Baracca, G., Orlandi, S., Bandini, A., & Dejonckere, P. H. (2015). Automatic Assessment of Acoustic Parameters of the Singing Voice: Application to Professional Western Operatic and Jazz Singers. *Journal of Voice*, *29*(4), 517.e1-517.e9. https://doi.org/10.1016/j.jvoice.2014.09.014

Mauro, D. A., & Valle, A. (2019). *Sound Ontologies Methods and Approaches for the Description of Sound*. 283–302.

McCrea, C. R., & Morris, R. J. (2005). Comparisons of voice onset time for trained male singers and male nonsingers during speaking and singing. *Journal of Voice*, *19*(3), 420–430. https://doi.org/10.1016/j.jvoice.2004.08.002

Méndez-Martínez, J. L. (2020). If sounds were dispositions: A framework proposal for an undeveloped theory. *Jorge Luis Méndez-Martínez*, *27*(4), 446–479. https://doi.org/10.31577/ORGF.2020.27403

Millington, I. (2021). What Is AI? In *An Introduction to Ethics in Robotics and AI*. Springer. https://doi.org/10.1201/9781003124047-2

Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. A. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences*, *20*(1), 56–60. https://doi.org/10.1016/j.lindif.2009.11.003

M.Vergara, V., Norgaard, M., Miller, R., E.Beaty, R., Dhakal, K., & Dhamala, M. & V. D. C. (2021). Functional network connectivity during jazz improvisation. *Scientific Reports*, *11*(2021), 1–12. https://www.nature.com/articles/s41598-021-98332-x

Nag, S., Basu, M., Sanyal, S., Banerjee, A., & Ghosh, D. (2022). On the application of deep learning and multifractal techniques to classify emotions and instruments using Indian Classical Music. *Physica*, *597*(2022), 1–20. https://doi.org/10.1016/j.physa.2022.127261

Nakano, T., National, M. G., & Hiraga, Y. (2006). An Automatic Singing Skill Evaluation Method for Unknown Melodies Using Pitch Interval Accuracy and Vibrato Features. *INTERSPEECH 2006*, 1706–1709. https://doi.org/10.1080/17450128.2012.729871

Nannan, L. (2019) Acoustic Context and Vocabulary in Music Recording. *Technology and Art*, 7, 87–89.

Narang, J., Miron, M., Srinivasamurthy, A., & Serra, X. (2022). Analysis of Musical Dynamics in Vocal Performances. *The 18th International Symposium on Computer Music Multidisciplinary Research*, *3*, 33–39.

National Technology Supervision of China. (1995). *Acoustic Language Clarity Testing Method* (Patent GB/T 15508-1995).

Naveed, K., Watanabe, C., & Neittaanmäki, P. (2017). Co-evolution between streaming and live music leads a way to the sustainable growth of music industry – Lessons from the US experiences. *Technology in Society*, *50*, 1–19. https://doi.org/10.1016/j.techsoc.2017.03.005

Neubauer, J., Edgerton, M., & Herzel, H. (2004). Nonlinear phenomena in contemporary vocal music. *Journal of Voice*, *18*(1), 1–12. https://doi.org/10.1016/S0892-1997(03)00073-0

Nix, J., Emerich, K., & Titze, I. R. (2005). Application of Vocal Fry to the Training of Singers. *Voice Research and Technology*, *September/October*, 53–59.

Oliveira, A. L. G. de. (2019). From sound Ontology to auditory Ontology-criticism of imaginary worship in modern western music.pdf. *Revista Música Hodie*, *19*(1977), 1–11. https://doi.org/10.5216/mh.v19.51678

Pallant, J. (2011). SPSS Survival Manual. In *National Library of Australia*. Routledge.

Paquette, S., Peretz, I., & Belin, P. (2013). The 'Musical Emotional Bursts': A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, *4*(AUG), 1–7. https://doi.org/10.3389/fpsyg.2013.00509

Paul, D. (2012). A musicianship for the recording studio. In *Sound musicianship: Understanding the crafts of music* (pp. 1–6). Cambridge Scholars.

Peter Q, P., & Brown, S. (2022). Poor-pitch Singing in the absence of 'Tone Deafness'. *Journal of Voice*, *10*(1), 2005. https://doi.org/10.1016/j.cogpsych.2013.12.005

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of 'Tone Deafness'. *Music Perception*, *25*(2), 95–115. https://doi.org/10.1525/MP.2007.25.2.95

Podesva, R. J., & Callier, P. (2015). Voice quality and identity. *Annual Review of Applied Linguistics*, *35*(2015), 173–194. https://doi.org/10.1017/S0267190514000270

Portillo, M. P., Rojas, S., Guzman, M., & Quezada, C. (2018). Comparison of Effects Produced by Physiological Versus Traditional Vocal Warm-up in Contemporary Commercial Music Singers. *Journal of Voice*, *32*(2), 200–208. https://doi.org/10.1016/j.jvoice.2017.03.022

Quinto, L. R., Thompson, W. F., Kroos, C., & Palmer, C. (2014a). Singing emotionally: A study of pre-production, production, and post-production

facial expressions. *Frontiers in Psychology*, *5*(APR), 1–15. https://doi.org/10.3389/fpsyg.2014.00262

Rao, H. (2013). Recording of pop music vocals. *AV Technology*, *20*(2013), 45–48. https://doi.org/CNKI:SUN:YXJS.0.2013-02-020.

Rink, J. (2002). Analysis and (or?) performance. In J. Rink (Ed.), *Musical Performance: A Guide to Understanding* (pp. 35–58). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511811739.004

Ruggiero, M. E. (2015). University and Industry: An Experience in Teaching, Research and Artistic Expression. *Procedia - Social and Behavioral Sciences*, *174*(2015), 4038–4045. https://doi.org/10.1016/j.sbspro.2015.01.1152

Samuel, H., Mangihut, T., & Sunitiyoso, Y. (2019). Co-innovation processes in the music business. *Heliyon*, *November 2018*, e01540. https://doi.org/10.1016/j.heliyon.2019.e01540

Saragih, H., & Pelita, U. (2018). Co-Innovation in the World Music Business: a Conceptual Framework. *The 3rd International Conference on Management, Economics and Business (ICMEB)*, 1–15.

Saragih, H. S. (2017a). Critical success factors of recording studio in Indonesia. *Asia Pacific Journal of Marketing and Logistics*, *29*(3), 686–699. https://doi.org/10.1108/APJML-07-2016-0123

Saragih, H. S., Simatupang, T. M., & Sunitiyoso, Y. (2019). Co-innovation processes in the music business. *Heliyon*, *5*(4), 1–24. https://doi.org/10.1016/j.heliyon.2019.e01540

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, *44*(Aug), 1–22. https://doi.org/10.1017/S0140525X20000333

Savage, S. (2023a). Studio Study: Lipsmacks, Mouth Noises, and Heavy Breathing. In *Repurposing Music in the Digital Age* (pp. 48–60). University of Michigan Press. https://www.jstor.org/stable/j.ctvc5pf3g.9

Scherer, K. R., Sundberg, J., Fantini, B., Trznadel, S., & Eyben, F. (2017). The expression of emotion in the singing voice: Acoustic patterns in vocal performance. *The Journal of the Acoustical Society of America*, *142*(4), 1805–1815. https://doi.org/10.1121/1.5002886

Schiavio, A., & Benedek, M. (2020). Dimensions of Musical Creativity. *Frontiers in Neuroscience*, *14*(November), 1–23. https://doi.org/10.3389/fnins.2020.578932

Seashore, C. E. (1912). The Measure of a Singer. *Science*, *35*(No.893), 201–212. https://www.jstor.org/stable/1638656

Silva, G. (1922). The Beginnings of the Art of " Bel Canto ": Remarks on the Critical History of Singing. *The Musical Quarterly*, *8*(1), 53–68. https://www.jstor.org/stable/737912

Slevc, L. R., Davey, N. S., Buschkuehl, M., & Jaeggi, S. M. (2016). Tuning the mind: Exploring the connections between musical ability and executive functions. *Cognition*, *152*, 199–211. https://doi.org/10.1016/j.cognition.2016.03.017

Spreadborough, K. (2010). Emotional Tones and Emotional Texts: A New Approach to Analyzing the Voice in Popular Vocal Song. *Music Theory Online*, *29*(2010), 1–14. https://doi.org/10.30535/mto.28.2.7

Stern, R., Nix, J., & Antonio, S. (2020). Listener Ratings of Male Singer Technical Ability. *Journal of Voice*, *35*(3), 501.e1-501.e9. https://doi.org/10.1016/j.jvoice.2019.11.004

Stratton, S. J. (2021). Population Research: Convenience Sampling Strategies. *Prehospital and Disaster Medicine*, *36*(4), 373–374. https://doi.org/10.1017/S1049023X21000649

Strek, T. (2020). Modeling of sound power transmission within duct systems. *Vibrations in Physical Systems*, *31*(3), 1–6. https://doi.org/10.21008/j.0860-6897.2020.3.18

Sundberg, J. (1974b). Articulatory interpretation of the “singing formant”. *The Journal of the Acoustical Society of America*, *55*(4), 838–844. https://doi.org/10.1121/1.1914609

Sundberg, J. (1994). Acoustic and psychoacoustic aspects of vocal vibrato. *Dept. for Speech, Music and Hearing, Quarterly Progress and Status Report*, *35*(1994), 45–68.

Sundberg, J., Gu, L., Huang, Q., & Huang, P. (2012). Acoustical study of classical peking opera singing. *Journal of Voice*, *26*(2), 137–143. https://doi.org/10.1016/j.jvoice.2011.01.001

Tahiroğlu, K. (2021). Ever-shifting roles in building, composing and performing with digital musical instruments. *Journal of New Music Research*, *50*(2), 155–164. https://doi.org/10.1080/09298215.2021.1900275

Takamitsu.S. (2020). To be professional composers, arrangers, and lyricists: Ways to succeed. In *Rittor Music Press*. https://doi.org/10.1145/379866.379872

Thorley, M. (2018). The role of failure in developing creativity in professional music recording and production. *Thinking Skills and Creativity*, *30*(June 2017), 160–170. https://doi.org/10.1016/j.tsc.2018.05.002

Tian, L., Alaei, R., & Rule, N. O. (2022). Appearance Reveals Music Preferences. *Personality and Social Psychology Bulletin*, *48*(12), 1635–1650. https://doi.org/10.1177/01461672211048291

Tinker, C. (2002). A singer-songwriter's view of the French record industry: The case of Léo Ferré. *Popular Music*, *21*(2), 147–157. https://doi.org/10.1017/S026114300200209X

Tonelli, F., Evans, S., & Taticchi, P. (2013). Industrial sustainability: Challenges, perspectives, actions. *International Journal of Business Innovation and Research*, *7*(2), 143–163. https://doi.org/10.1504/IJBIR.2013.052576

Truax, B. (2012). Sound, Listening and Place: The aesthetic dilemma. *Organised Sound*, *17*(3), 193–201. https://doi.org/10.1017/S1355771811000380

Tzanaki, P. (2022). The Positive Feedback Loop of Empathy and Interpersonal Synchronisation: Discussing a Theoretical Model and its Implications for Musical and Social Development. *Music and Science*, *5*(5), 1–16. https://doi.org/10.1177/20592043221142715

Vaquero, L., Rodríguez-Fornells, A., & Reiterer, S. M. (2017). The left, the better: White-matter brain integrity predicts foreign language imitation ability. *Cerebral Cortex*, *27*(8), 3906–3917. https://doi.org/10.1093/cercor/bhw199

Vidwans, A., Gururani, S., Wu, C. W., Subramanian, V., Swaminathan, R. V., & Lerch, A. (2017). Objective descriptors for the assessment of student music performances. *Proceedings of the AES International Conference*, *June*, 116–123.

Waddell, G., & Williamon, A. (2019). Technology use and attitudes in music learning. *Frontiers in ICT*, *6*(MAY). https://doi.org/10.3389/fict.2019.00011

Wang, S. (1985). Singer's high formant associated with different larynx position in styles of singing. *Journal of the Acoustical Society of Japan (E)*, *7*(6), 303–314. https://doi.org/10.1250/ast.7.303

Wanlu, Z., & Shiqi, G. (2020). A Brief Overview on the Expressive Characteristics of Vocals in Music Recording and Production. 浅谈音乐录音创作中人声的表现性特征. *Research on Music Art.艺术大观*, *11*(3), 11–12. https://doi.org/CNKI:SUN:YSDG.0.2020-03-008.

Wapnick, J., & Ekholm, E. (1997). Expert consensus in solo voice performance evaluation. *Journal of Voice*, *11*(4), 429–436. https://doi.org/10.1016/S0892-1997(97)80039-2

Watson, A., & Ward, J. (2013a). Creating the right 'vibe' Emotional labour and musical performance in the recording studio. *Environment and Planning A*, *45*(12), 2904–2918. https://doi.org/10.1068/a45619

Weidman, A. (2014). Neoliberal Logics of Voice: Playback Singing and Public Femaleness in South India. *Culture, Theory and Critique*, *62*(3), 175–193.

Wen, O. X., & Krumhansl, C. L. (2019). Perception of Pitch and Time in the Structurally Isomorphic Standard Rhythmic and Diatonic Scale Patterns. *Music and Science*, *2*(2019), 1–24. https://doi.org/10.1177/2059204319873308

Wesolowski, B. C., Wind, S. A., & Engelhard, G. (2016). Examining rater precision in music performance assessment: An analysis of rating scale structure using the Multifaceted Rasch Partial Credit Model. *Music Perception*, *33*(5), 662–678. https://doi.org/10.1525/MP.2016.33.5.662

Williams, A. (2010a). Navigating Proximities: The Creative Identity of the Hired Musician. *Journal of the Music and Entertainment Industry Educators Association*, *10*(1), 59–76. https://doi.org/10.25101/10.3

Wooding, R., & Nix, J. (2016). Perception of Non-Vibrato Sung Tones: A Pilot Study. *Journal of Voice*, *30*(6). https://doi.org/10.1016/j.jvoice.2015.10.005

Wu, Y., & Bryan-Kinns, N. (2019). Musicking with an interactive musical system: The effects of task motivation and user interface mode on non-musicians' creative engagement. *International Journal of Human Computer Studies*, *122*(July 2018), 61–77. https://doi.org/10.1016/j.ijhcs.2018.07.009

Wurtz, P., Mueri, R. M., & Wiesendanger, M. (2009). Sight-reading of violinists: Eye movements anticipate the musical flow. *Experimental Brain Research*, *194*(3), 445–450. https://doi.org/10.1007/s00221-009-1719-3

Xianbo, C. (2020). The discussion on Vocal Expression from the Perspective of the Sound Engineer. 录音师视角谈音乐表演者的声音表现问题. *Research on Transmission Competence.传播力研究*, *100*(4), 132–133.

Xueyu, W. (2013). Inquiry on Singer Psychology Analysis in the recording practice. *Journal of Gansu Lianhe University*, *26*(4), 185–197. https://doi.org/10.13804/j.cnki.2095-6991.2013.06.021

Yang, M. (2005). The recreation in music performance. *Chinese Music*, *3*, 165–166.

Yu Chan, M. P., & Do, Y. (2021). Vowel Modification (Aggiustamento) in Soprano Voices. *Music and Science*, *4*(5), 1–15. https://doi.org/10.1177/20592043211055168