Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

MASTERARBEIT

zur Erlangung des akademischen Grades M.A.

Finding My Voice: A case study on vocal production aesthetics and techniques for the self-producing female singer-songwriter in today's popular music culture

vorgelegt am 31. Juli 2025

Emily Charlotte Payant

Erstprüfer: Prof. Thomas Görne
Zweitprüferin: Clara Lucas

HOCHSCHULE FÜR ANGEWANDTE WISSENSCHAFTEN HAMBURG
Department Medientechnik
Finkenau 35
20081 Hamburg

Copyright Notice

Dedication

I would like to thank my friends and family for giving me the space to dedicate myself fully to this project over the past few months. I especially want to thank my loving partner and best friend for supporting (and tolerating) me through thick and thin — I owe you lots of cooked meals, household chores and a big holiday! My heartfelt thanks to all you precious people (you know who you are -- I've always wanted to use that phrase) for cheering me on, especially towards the finish line. I've never run a marathon, but I think I have a decent idea of what that must feel like now. If this project has taught me anything — besides a better understanding of the singer-songwriter format and mad vocal production skills, of course -- it is this: how lucky and rich I am to have you all in my life.

Abstract

This artistic research study explores how contemporary vocal production techniques can enhance emotional expression and narrative clarity in the self-produced singer-songwriter format. Focusing specifically on the female voice, it examines how technological mediation—often associated with artificiality—can instead be used to convey authenticity and vulnerability. Using a practice-as-research methodology, the project combines theoretical analysis with the production of three original songs. Each track serves as a case study to investigate the narrative and emotional impact of different vocal production aesthetics, including close-mic'ing, layering, spatial effects, and more overt processing such as pitch-correction and the vocoder. Core concepts including Hall's proxemics (1969), vocal staging (Lacasse, 2000), and the embodied/disembodied voice (Malawey, 2020) inform both the analysis and creative process. Findings show that intentional vocal processing can serve dramaturgical functions—supporting or even reshaping lyrical meaning—and that so-called imperfections often heighten emotional realism. Rather than diminishing authenticity, the use of digital tools, when guided by narrative intent, can deepen expressive nuance. Constraints such as limited gear or a home-studio environment were found to encourage innovation, resourcefulness, and more deliberate aesthetic decisions. Through iterative experimentation and critical reflection, the thesis proposes that vocal production is not a postscript to composition but a narrative act in itself. By foregrounding the voice as both a technical and emotional instrument, the research challenges traditional assumptions about the singer-songwriter genre and highlights new ways in which self-producing female artists can navigate authorship, vulnerability, and agency within today's rapidly evolving popular music landscape.

Zusammenfassung

Diese künstlerisch-wissenschaftliche Arbeit untersucht, wie zeitgenössische Gesangsproduktions-techniken den emotionalen Ausdruck und die narrative Verständlichkeit im selbstproduzierten Singer-Songwriter-Format verbessern können. Mit besonderem Fokus auf die weibliche Stimme wird untersucht, wie technische Mittel, die oft mit Künstlichkeit assoziiert werden, dennoch gezielt eingesetzt werden können, um Authentizität und Verletzlichkeit zu transportieren. Die praxisorientierte Forschungsmethodik verbindet theoretische Analyse mit der Produktion von drei selbstkomponierten Songs. Jeder Titel dient dabei als Fallstudie, um die narrative und emotionale Wirkung verschiedener Gesangsproduktionstechniken zu untersuchen — u.a. Close-Miking, Layering, und räumliche Effekte, sowie Pitch-Correction und Vocoder. Grundlegende Konzepte wie Halls "Proxemics" (1969), "Vocal Staging" (Lacasse, 2000) und die Unterscheidung zwischen „Embodied" und „Disembodied Voice" (Malawey, 2020) prägen sowohl Analyse als auch den kreativen Prozess. Die Ergebnisse zeigen, dass bewusste Stimmbearbeitung dramaturgische Funktionen erfüllen kann — etwa indem sie lyrische Aussagen unterstützt oder gar umdeutet — und dass kleine Fehler oder Artefakte oft zu größerer emotionaler Authentizität führen. Anstatt die Authentizität zu mindern, kann zudem der Einsatz digitaler Werkzeuge, sofern von narrativer Absicht geleitet, die Ausdruckstiefe verstärken.
Die Beschränkung auf ein reduziertes Heimstudio-Setup förderten Einfallsreichtum und bewusstere ästhetische Entscheidungen. Durch iterative Experimente und kritische Reflexion kommt die Arbeit zu dem Schluss, dass die Gesangsproduktion kein Nachsatz zur Komposition ist, sondern ein narrativer Akt in sich. Indem sie die Stimme sowohl als technisches als auch emotionales Instrument in den Vordergrund stellt, hinterfragt die Forschungsarbeit traditionelle Annahmen über das Singer-Songwriter-Genre und zeigt neue Wege auf, wie selbstproduzierende Künstler*innen Autor*innenschaft, Verletzlichkeit und gestalterische Handlungsmacht („agency") im Kontext heutiger Popmusik selbstbestimmt gestalten können.

Contents

Abstract ..... **i**

Contents ..... **ii**

Abbreviations ..... **v**

List of Figures and Tables ..... **vi**

0 — Introduction **1**

1 — Theoretical Framework **7**

1.1 Key Terms 7

1.1.1 Lyrical and Musical Narrative 7

1.1.2 Vocal Aesthetic and Production Aesthetic 8

1.1.3 Contemporary Pop Vocal Aesthetic 8

1.1.4 Authenticity 9

1.1.5 Authorship, Agency and Ownership 9

1.1.6 Today' Popular Music Culture 10

1.2. Conceptual Framework 11

1.2.1 Sound Interpretation in Popular Music 11

1.2.2 Analysis of the Voice in Popular Music 14

1.2.2.1 Expressive Parameters 14

1.2.2.2 Grain Of The Voice and Embodied Voice 14

1.2.2.3 Disembodied Voice 15

1.2.2.4 Register and Timbre Considerations, Performance Intensity 16

1.2.2.5 Voice as Identity, Vocal Persona and Narrative Voice 16

1.2.2.6 Lyrical Analysis 17

1.2.3 Analysis of the Technologically Mediated Voice 19

1.2.3.1 Hall's Proxemic Zones 19

1.2.3.2 Stereo Sound and Mono-Centricism 20

1.2.3.3 Vocal Staging 21

1.2.3.4 Opaque and Transparent Mediation 23

2 — Historical and Aesthetic Context 24

2.1 Defining the Singer-Songwriter 24

2.1.1 The Stereotype 25

2.1.2 Genre Formation (1960s-1970s) 26

2.1.3 Rise of the Artist-Producer (1980s) 29

2.1.4 The 'Female Singer-Songwriter' Label (1990s) 30

2.1.5 Expanding the Sonic Palette (2000s) 32

2.1.6 The Contemporary Singer-Songwriter (approx. 2018-2025) 35

2.1.7 Summary 37

2.2 Contemporary Pop Vocal Production Aesthetics and Techniques 40

2.2.1 Close-Mic'ing, Compression and EQ as Tools for Hyper-Intimacy 43

2.2.2 Double-Tracking and Layering as Tools for Emotional Complexity 46

2.2.3 Reverb and Delay as Spatial, Temporal and Emotional Staging Devices 53

2.2.4 Pitch-Modulation, Vocoder and Distortion as Transformative Tools 59

3 — Methodology 67

3.1 Artistic Research as Practice-Based Inquiry 67

3.2 Bridging Theory and Practice 68

3.3 Creative Process and Tools 69

3.3.1 Tools and Environment 69

3.3.2 Workflow and Production Strategies 69

3.3.3 Analytical Listening 70

3.4 Documentation and Reflective Journal 70

4 — Case Study: Original Song Production Process 72

4.1 Overview 72

4.2 ‘Forever Vow’ – Bronty 74
4.2.1 Summary of Lyrical Narrative 74
4.2.2 Artistic Intention and Production Concept 76
4.2.3 Technical Considerations and Reflective Process 77
4.2.4 Findings 82
4.3 ‘What Do You Expect’ – Bronty 84
4.3.1 Summary of Lyrical Narrative 84
4.3.2 Artistic Intention and Production Concept 84
4.3.3 Technical Considerations and Reflective Process 85
4.3.4 Findings 88
4.4 ‘I Don’t Mind’ – Bronty 90
4.4.1 Summary of Lyrical Narrative 90
4.4.2 Artistic Intention and Production Concept 92
4.4.3 Technical Considerations and Reflective Process 93
4.4.4 Findings 95
5 — Conclusion 97
References 102
Appendices 107
Discography 107
Comparative Overview of Singer-Songwriter Categories 109
Original Song Lyrics 110
Documentation of Production Processes 113
Note on the Use of AI Tools (‘Hilfsmittel’) 120
Statement of Originality (‘Eigenständigkeitserklärung’) 121

Abbreviations

ASMR - Autonomous Sensory Meridian Response

DAW - Digital Audio Workstation (e.g. Logic Pro X, ProTools, Ableton, Nuendo etc)

DIY - Do-It-Yourself

EQ - Equaliser / Equalisation

Hz - Hertz (KHz: Kilo-Hertz)

Lo-Fi - Low-Fidelity (i.e. of low quality/definition)

MIDI - Musical Instrument Digital Interface

C - Centre

L - Left

R - Right

List of Figures and Tables

Figure 1: Typology Grid of Pop Song Lyrics (adapted from Randle and Evans, 2013) 18

Figure 2: Singer-Songwriter Stereotype(s), generated by ChatGPT 25

Figure 3 (left): Bob Dylan in 1961 (Source: Michael Ochs Archives/Getty Images) 26

Figure 4 (right): Joni Mitchell in 1968 (Source: Jack Robinson/Getty Images) 26

Figure 5: The famous 'Q' cover with PJ Harvey, Björk and Tori Amos.....31

Figure 6: An example of my lead "vocal chain" in "Forever Vow" 41

Table 1: Lacasse's General Categorisation of Vocal Staging Effects.....22

Table 2: Overview of Case Studies: Vocal Concepts, Aesthetics & Narrative Function 73

Table A: Comparative overview of Singer-Songwriter Categories as per Chapter 2.1.....109

0 — Introduction

Over the past two decades, the "democratisation" of music technology (Bell, 2018; Leyshon, 2009) has not only led to the decline of recording studios (Bennett, 2016; Leyshon, 2009; Watson, 2015) and the "delineation" of the producer role (Auvinen, 2020; Kealy, 1979), but it has also meant that technological tools of pop music production have become increasingly accessible to independent artists. This shift has created new opportunities for singer-songwriters to integrate the DAW into their creative process as early as the songwriting stage, allowing them to craft their overall sound and narrative voice with greater autonomy throughout the recording and mixing processes.

Notably, this development has been particularly significant for women, who have historically been underrepresented[1] in the technical domains of music production (Reddington, 2021; Wolfe, 2019). The past few years have borne witness to a new generation of female singer-songwriters, such as Billie Eilish, Phoebe Bridgers, Melina Duterte (aka Jay Som), Maggie Rogers, Clairo, Mitzki, girl in red, King Princess, Ellie Dixon to name but a few, who have evolved the 'singer-songwriter' aesthetic by embracing digital tools and pop production techniques. For some of these contemporary singer-songwriters, self-producing has become an integral part of their artist identity.

As a self-producing female singer-songwriter, this thesis emerges from a personal and artistic need to explore how contemporary vocal production techniques and aesthetics such as close-mic'ing, layering and signal processing can enhance the communicative power of a song without overshadowing the core values of intimacy, authenticity, and lyrical depth that characterise the singer-songwriter format.

[1] According to a the USC Annenberg Inclusion Initiative study, "In 2024, there were 237 producers credited across the most popular songs of the year. Of those producers, 94.1 were men and 5.9% were women. There was no change in 2024 compared to 2023 (6.4%) or 2012 (2.4%) in the percentage of women producers. Across the 10 years examined and a total of 2,209 producers, men held 96.5% of all producing jobs and women held 3.5%. This is a gender ratio of 27.3 men to every 1 woman producer" (Smith et al, 2025, p. 3).

Addressing the Literature Gap

Despite the growing body of popular music scholarship over the last decade, significant gaps remain—particularly concerning 21st-century commercial pop music and contemporary singer-songwriters who blur genre and production roles. Steinbrecher (2021) highlights this gap, noting that "21st century commercial pop music is still little represented as a research subject in [musicology], despite its cultural impact and the technological and artistic revolution it represents" (Peres, 2016, as cited in Steinbrecher). Similarly, Appel (2017) observes a scarcity of scholarly attention to contemporary singer-songwriters, especially those who self-produce and operate across hybrid creative roles (p. 11). This thesis seeks to contribute to filling this void.

While the human voice is arguably the most central and emotionally resonant element in popular music, it has often been neglected in academic discourse, especially in relation to its recorded and technologically mediated form. Victoria Malawey's *A Blaze of Light in Every Word* (2020) provides a crucial foundation for this thesis, offering an analytical framework that bridges musicology, voice studies, and technology. Her work demonstrates how meaning is shaped through pitch, timbre, prosody, and the affordances of recording and production. However, despite the voice's interdisciplinary scope, there remains a notable lack of scholarly attempts to systematically connect these disciplinary perspectives within the context of practice-based research. This thesis builds upon Malawey's work while also seeking to further bridge the gap between theory and practice by integrating analytical, technical, and reflective insights from the perspective of a self-producing singer-songwriter.

Research Aims and Hypothesis

By foregrounding the mediated female singing voice within a self-produced, singer-songwriter context, this thesis seeks to contribute to a more nuanced understanding of how vocal production aesthetics shape narrative, intimacy, and emotional expression in contemporary pop music.

The guiding central hypothesis is as follows:

Integrating contemporary pop vocal aesthetics into the traditionally sparse singer-songwriter format can support the lyrical narrative and add emotional depth to a song without compromising authenticity.

Supporting this hypothesis are several guiding thematic assumptions that inform both theoretical research and creative practice:

(1) *Songs have a narrative and emotional core that exist prior to their production;*

(2) *In singer-songwriter music, the voice carries the majority of a song's narrative and emotional meaning, making vocal production a central tool in shaping how that meaning is communicated and perceived;*

(3) *The perception of "authenticity'" in popular music production has shifted alongside technological developments, expanding to include both raw, unprocessed vocal styles and highly mediated, digitally enhanced voices as valid forms of emotional expression.*

These assumptions function as thematic lenses through which key ideas about voice, narrative, authenticity, and production are explored. Rather than aiming to prove them outright, the thesis seeks to critically examine and reflect upon these ideas through an iterative dialogue between the production of three original songs and scholarly inquiry. They will be introduced where relevant throughout the thesis.

Artistic Research Approach

This thesis is situated within the field of artistic research and takes the form of a practice-based case study. The research centres around three original songs, written and produced by the author, which serve as both artistic works and sites of reflection. Each song functions as an experiment in vocal production, an attempt to find a vocal aesthetic that supports its

narrative without surrendering to over-polished pop conventions[2]. The goal is to balance the authentic, intimate feel of the singer-songwriter format with contemporary production techniques, such as close-mic'ing, vocal layering, and heavier digital processing, whilst not applying them as formulaic tools[3], but as expressive elements selected in response to the needs of each composition.

Rather than aiming to generate wholly new findings through practice, the project applies existing theoretical insights from fields such as popular music studies, voice and sound studies, phonomusicology[4], and music production to the creative process. The creative process is thus guided by extensive theoretical considerations around voice, authenticity, and technological mediation, with the artistic work acting as a reflective practice to explore and critically evaluate these concepts.

Structure of Thesis

The structure of this thesis is as follows — Chapter 1 lays the theoretical groundwork by defining key concepts drawn from voice studies, musicology, and sound studies, which underpin the analysis throughout the project. Chapter 2 situates the study within a broader historical and cultural context, tracing the evolution of the singer-songwriter tradition, shifts in vocal production techniques, and the gendered dynamics shaping today's popular music culture. Chapter 3 outlines the artistic research methodology, detailing the practice-based inquiry approach along with tools, workflow, and documentation of the song production

---

[2] Angharad ("Anna") Thomas (2022) explores how vocal editing in 21st-century commercial pop music is shaped by a dual fetishism—both of the voice and of the technologies used to edit it. Her research highlights how producers treat the female voice as an aesthetic and commercial object, subject to intense digital manipulation through elaborate vocal chains. This process is deeply intertwined with the marketing and use of plugins, reinforcing an industry-wide ideal of the hyper-edited pop vocal.

[3] Platforms such as *Hit Songs Deconstructed* reflect the industry's increasing emphasis on vocal sound and melodic structure as key drivers of commercial success. By offering data-driven analyses of hit songs—down to vocal phrasing, hook placement, and production techniques—they provide subscribers with formulaic tools designed to emulate the sonic traits of current chart-toppers. This analytical fetishisation of the voice and melody not only highlights the commodification of vocal aesthetics, but also reveals the extent to which artistic decisions are shaped by algorithmic and market-driven logic.

[4] A term coined by Stephen Cottrell, meaning "the study of recorded music, including its contexts of production and patterns of consumption." see Cottrell, S.J. (2010). The rise and rise of phonomusicology. In: Bayley, A. (Ed.) *Recorded Music: Performance, Culture and Technology.* (pp. 15-36). Cambridge: Cambridge University Press.

process. Chapter 4 presents the three original songs as individual case studies. Each case study considers artistic intention, technical process and reflective analysis to demonstrate in practice how vocal production choices inform narrative, emotional expression, and identity construction. Chapter 5 discusses the findings across all works and concludes the research with a summary of insights and considerations for future practice.

Scope and Limitations

This thesis focuses primarily on the period from approximately 2018 to 2025, during which distinct trends in vocal aesthetics emerged in both mainstream and independent music. While this timeframe serves as the main analytical lens, examples from earlier periods are included where historically or contextually relevant, for example in Chapter 2, which traces the evolution of singer-songwriter aesthetics and production practices. These earlier references help illuminate the continuity and contrast between current practices and their artistic-technological predecessors.

While foundational areas such as physical acoustics, semiotics, and extensive narrative or music theory are acknowledged for their relevance and conceptual influence, they are not explored in detail. They may, however, appear in brief or contextualising references where necessary to support specific observations or clarify broader concepts.

The scope is intentionally narrow and personal, emphasising depth over breadth. It does not aim to present an exhaustive overview of contemporary vocal techniques or generalise across all artist-producer practices. Instead, it offers a situated and practice-based perspective that foregrounds the female voice as a central narrative medium. The study excludes topics such as live performance, comprehensive gender theory, or advanced mixing and mastering considerations.

While gender and feminist discourse in music production are highly relevant, particularly in relation to authorship, agency, and representation, I have chosen not to engage with these aspects in depth, as doing so would exceed the scope of this thesis. Instead, I acknowledge my own positionality as a self-producing female singer-songwriter and refer interested readers to the work of scholars such as Leslie Gaston-Bird (2019), as well as the feminist perspectives on self-production and agency offered by Wolfe (2019) and Reddington (2021).

Wherever possible, I highlight illustrative works by women in music production and vocal performance. However, my primary aim is to use this research process to deepen my artistic and technical understanding of vocal production aesthetics, rather than to make a theoretical contribution to gender studies.

Additionally, some limitations stem from the practical reality of working within a home-studio environment. While collaborative input was selectively integrated from working with producer Rami Olsen, the project remained rooted in the author's own aesthetic, artistic voice, and workflow.

Finally, although immersive audio formats such as binaural and 3D have received growing scholarly attention in recent years, this thesis focuses solely on the standard stereo format, as experienced through headphones and loudspeakers. Stereo remains the dominant format for music production and distribution today, and this is unlikely to change in the foreseeable future (cf. Rada et al., 2025). The production and analytical framework is shaped entirely by stereo listening conditions, with no exploration of the spatial or perceptual dimensions offered by other formats.

1 — Theoretical Framework

This chapter establishes the theoretical and conceptual framework that underpins both the analytical and creative components of the thesis. It introduces key terms and interpretive lenses that inform the contextual discussion in Chapter 2 as well as the reflective case studies presented in Chapter 4.

1.1 Key Terms

1.1.1 Lyrical and Musical Narrative

Lyrical narrative centres on the lyrics of a song and the way they construct meaning, character, or emotional development over time.
While Negus pointed out in 2012 that "the popular song—one of the most pervasive narrative forms that people encounter in their daily lives—has been almost entirely ignored in the vast literature on narrative" (p.368), more recent scholarship confirms and discusses the presence of narrative structures and emotional arcs in popular song lyrics. For instance, Alberhasky and Durkee's (2024) quantitative study demonstrates that pop songs often follow identifiable narrative trajectories. Drawing on Mieke Bal's (1997) narrative theory and semiotics, elsewhere Randle (2013) proposes a typology of lyrical narrativity based on four core components: event, actor, time, and setting (p. 132). This typology is presented in more depth in the conceptual framework section 1.2.2.6.

Musical narrative, by contrast, refers to the way a sense of story, emotion, or progression is conveyed through non-verbal musical elements such as melody, harmony, rhythm, dynamics, instrumentation, and arrangement. The concept is rooted in musicology and has historically been applied to instrumental and classical music, where narrative is conveyed rather through musical form and gesture than text.

While recognising that musical narrative also exists in popular music and plays a role in shaping the overall emotion of a track, this thesis will focus primarily on lyrical narrative as the central site of narrative expression, exploring how vocal production aesthetics contribute to the communication of meaning and emotional resonance in self-produced singer-

songwriter music.

1.1.2 Vocal Aesthetic and Production Aesthetic

Vocal aesthetic refers to the combined technical, performative, and production choices that shape how a recorded voice is perceived. This includes singing style, vocal tone, use of effects (such as reverb or doubling), spatial placement in the mix, and overall sonic texture. These elements contribute to how the voice communicates emotion, narrative and intimacy — all central themes in this thesis.

The overarching term production aesthetic in music production can be placed on a scale ranging from what Moylan (1992) calls *naturalist-realistic* to *all-technological*, with a lot of popular music living in the hybrid *artistic-technological* space.
While this concept is explored throughout the thesis, it reappears in more detail in the 2.2.3 chapter on reverb, as it allows a nuanced exploration in particular with regards to the spatial staging of a song.

1.1.3 Contemporary Pop Vocal Aesthetic

In this thesis, the term *contemporary pop vocal aesthetics* refers to a set of stylistic tendencies and production techniques that, in my own observation as both listener, artist and producer, have become especially prominent in popular music over the past six to seven years (roughly 2018–2025). Rather than attempting to define a comprehensive or authoritative list, in Chapter 2 I focus on techniques that I have found to be expressive, emotionally effective, and artistically meaningful in my own listening and creative practice. This approach aligns with the auto-ethnographic nature of this research, where aesthetic observations are rooted in embodied experience and critical listening.

1.1.4 Authenticity

Within the popular music discourse, there have been long standing and ongoing debates[5] about *authenticity*, particularly in the context of recorded music. Alleyne (2020) summarises, "with each technological phase, musicians and audiences were challenged to reassess what was authentic to them" (p. 24).

Rather than equating authenticity with unedited or unprocessed performance, this thesis follows the view that authenticity is performative and context-dependent—a relational effect constructed through both delivery and production, with meaning being "ascribed" by the listener, as opposed to "inscribed" by the artist (Moore, 2002).

Even recordings that aim to sound natural or unedited rely on deliberate choices: recording environment, microphone placement, performance intensity, arrangement and editing all work together to construct a sense of authenticity (Stiegler & Campbell, 2023). In other words, even so-called "raw" recordings are carefully framed to feel unmediated. Thus, when I speak of authenticity here, I'm referring to a *perceived* authenticity. In this sense, a heavily layered or processed vocal can still feel emotionally "true" and "authentic" if the choices align with the artist's aesthetic and the song's emotional message.

Throughout this project, I explore authenticity not as a fixed standard, but as an evolving and negotiable quality shaped by artistic intention, listener expectation, and production aesthetic.

1.1.5 Authorship, Agency and Ownership

Authorship and agency in the context of creative works and digital tools, explore who is considered the originator (author) of a work and who controls the creative process (agency).

Meanwhile *(Creative) Ownership* refers to the legal rights and control that a creator or owner has over their original works. This includes the right to reproduce, distribute, display, and make derivative works from the original content, as well as the ability to license or transfer these rights. Famously, Taylor Swift brought widespread attention to the issue when she re-recorded her early albums in order to regain control over her masters — highlighting how

[5] Most often fuelled by the conservative viewpoints of Theodor Adorno and Walter Benjamin.

creative ownership can deeply affect an artist's autonomy, income, and ability to shape their legacy. Her re-recordings, released as *Taylor's Versions*, were not only a bold statement of artistic and legal agency but also achieved remarkable commercial success, topping charts and generating record-breaking streaming numbers[6].

In the context of gender discourse, the notion of "(re)claiming authorship" or "reclaiming the narrative" allows female artists to take control over how their own stories are voiced, produced, and framed, encompassing all three elements of authorship, agency and ownership. (cf. Reddington's (2021) *"Gender Ventriloquism"* and Wolfe (2019))

1.1.6 Today' Popular Music Culture

In this thesis, *today's popular music culture* refers to the dynamic and rapidly evolving landscape of mainstream and indie-adjacent music in the 2020s. It is shaped by subscription-based distribution models (such as Spotify and YouTube), short-form video social media platforms (such as TikTok and Instagram), accessible digital production tools, and shifting listener habits driven by algorithmic discovery. The result is a sonic culture marked by shortened song structures, genre hybridity, and an emphasis on mood and atmosphere (Muchitsch, 2023; Petrusich, 2023), which might be described as the "vibe era". Petrusich states that "Spotify operates from a playlist model, frequently sorting music by vibe an idea that's perhaps even more ineffable than genre, but which also seems considerably more in tune with how and why people listen to music." Within this environment, emotional immediacy and vocal distinctiveness have become key tools for cutting through the mass, with headphone listening providing the perfect playground for the emotive use of breath, texture, and proximity (or distance). Stereo imaging, *binaural*[7] mixing, and vocal production techniques are thus increasingly tailored to the solitary, mobile listener, reinforcing a sense of immersion and intimate connection.

---

[6] see https://www.rollingstone.com/music/music-news/taylor-swift-album-re-record-project-success-1235351625/

[7] *Binaural audio* (head-related stereophony) simulates natural hearing by using a Head-Related Transfer Function (HRTF), which translates how sound interacts with the outer ear to convey spatial cues. It is only effective when played back through headphones, and especially earphones (as the HRTF already accounts for the ear's anatomy) (cf. Görne, 2015, pp. 126,131).

The success of intimacy-driven content in recent years reflects a growing human desire for authentic connection as a response to our fast-paced, scroll-heavy, and often anonymous online culture. In music, the voice — as the "profoundest mark of the human" (Middleton, 1990) — has become central to meeting this need. The emergence of AI-assisted tools and machine learning in recent years have not only intensified this urge, but are blurring the boundaries of authorship and performance, complicating our ideas of vocal identity, authenticity, and creative agency even further.

Importantly, *today's popular music culture* also reflects ongoing shifts in gender dynamics within the music industry. While the systemic underrepresentation of female and non-binary music producers persist, the rise of accessible home-recording technologies, digital platforms, and DIY culture has opened new avenues for these artists to self-produce and self-release music on their own terms.

The global COVID-19 pandemic further accelerated many of these trends, forcing artists to adapt quickly to remote collaboration, virtual performance spaces, and independent production methods. This period reinforced a turn inward and helped solidify the cultural prominence of the solitary artist-producer working from the intimacy of their bedroom or home studio.

1.2. Conceptual Framework

This section outlines the conceptual frameworks that inform my reflective listening and production practices throughout this study. Rather than applying these frameworks prescriptively, they serve as analytical and creative lenses—tools for reflection and artistic exploration. Their primary function here is to establish a shared vocabulary for the interpretive strategies used throughout the thesis.

1.2.1 Sound Interpretation in Popular Music

Human hearing is evolutionarily attuned to vocal cues. The ear's sensitivity peaks within the range of human speech (approx. 2KHz-5KHz), allowing for high-resolution perception of timbral detail (Görne, 2017, p. 37). This is why we can recognise voices over telephones

despite the limited bandwidth. Timbre, in particular, serves as a powerful index of emotional state, physical proximity, and identity. Zagorski-Thomas (2014) notes that breathiness or high-frequency detail, often lost with distance in real-world acoustics, can signal intimacy when foregrounded in recordings, as our cognitive associations with proximity are triggered (p.141).

While a detailed explanation of physical acoustics of sound or the fundamentals of auditory meaning-making[8] falls outside the scope of this thesis, it is important to acknowledge that vocal production aesthetics are shaped by context-dependent and listener-specific interpretation.

Within the study of popular music, Lacasse's (2000) PhD thesis provides valuable insights regarding the perception and interpretation of the singing voice within the technologically mediated realm. Lacasse argues that the technical manipulation of the voice can evoke a variety of connotations and effects that emerge in the listener's mind not randomly, but according to patterns that suggest the existence of shared interpretive codes. The reception tests conducted in his study reveal consistent patterns across participants in how vocal staging and manipulation shape their interpretations of the singing voice. While Lacasse is careful to emphasise that vocal effects do not express *fixed* meanings or emotions, he notes that they can be intentionally varied and "exploited depending on the context in which they are heard" (p.166). The results of his tests are not meant to be definitive, but rather serve as "clues, as possible tendencies within a certain context," (p. 144), i.e. western countries or countries under western influence. For example, a distorted vocal was widely perceived by listeners as very *Malevolent*, low in *Potency*, and quite *Unnatural*. It was also judged to be *Unstable*, somewhat *Profane*, and emotionally ambiguous—neither clearly *Sad* nor *Happy*. Interestingly, this setting evoked a strong sense of spatial *Distance*, yet no clear association with *Time*. These kinds of recurring listener responses suggest that certain vocal treatments carry relatively stable connotative weight.

---

[8] For a grounding in these foundational concepts, including sound as a physical, perceptual, and communicative phenomenon, readers are referred to Görne's *Sound Design* (2017), which, though focused on sound design for film, offers relevant insights applicable to music production.

Lacasse distinguishes between meanings derived from the performer's delivery and those shaped through *production*. He notes that qualities such as *Potency* and *Happiness* appear to be conveyed more by the vocal performance itself ("the sound of the source"), whereas impressions of *Naturalness* and *Distance* are more closely tied to the *technical mediation of a voice*. In these cases, "the voice itself is not able to convey this type of information on its own," and instead "needs specific settings to convey the desired effects" (p. 165). Lacasse concludes, "it is plausible to argue that such a setting conveys [these] kind of connotations when used in pop songs" (p.164).

Lacasse's findings support the idea that while unmediated voice is rich with expressive potential, it is often the combination of vocal delivery and production choices that guide the listener's emotional and spatial interpretation—offering not absolute meanings, but affective cues shaped by shared listening habits and cultural codes.

1.2.2 Analysis of the Voice in Popular Music

1.2.2.1 Expressive Parameters

While this remains a conceptual overview, a brief clarification of terminology is needed for key acoustic and expressive parameters:

- *Pitch* refers to the perceived frequency of a sound
- *Register* describes distinct vocal ranges or resonance areas, such as chest voice, head voice, or falsetto
- *Tessitura* denotes the range within which a voice lies most comfortably and is most frequently used.
- *Timbre* describes the spectral qualities of a sound that allow us to distinguish one voice from another, or one instrument from another despite their same pitch.
- *Prosody*[9] encompasses rhythm, inflection, and the dynamic contours of vocal delivery.
- *Performance intensity* refers to how loudly or softly a performance is sung during the recording process. This differs from *dynamic level*, which describes the perceived loudness of that vocal sound as it appears in the context of a recorded mix as a result of the production and mixing process (cf. Malawey, 2020 p.137).

1.2.2.2 Grain Of The Voice and Embodied Voice

Barthes' (1985) concept of the *grain of the voice* describes "the body in the singing voice" (p.276) and "the materiality of the body emerging from the throat" (p.255). This materiality can be understood as breath, texture, and timbral nuance which foreground the corporeality of vocal sound, inviting an analysis that considers the physical and affective presence of the singer behind the sound. This is also referred to as the *embodied voice*. Malawey (2020) builds on this idea to explore how the grain relates to both the physical act of singing and to its reception: listeners use their own bodily experience to interpret and emotionally respond to vocal performance. This embodied perspective opens analytical and expressive pathways that foreground the physicality of music-making and listening.

---

[9] A great example of prosody is the sentence: *"I didn't say we should kill him"* - its meaning shifts entirely depending on which of the seven words is stressed through a variation of rhythm, pitch and dynamics.

1.2.2.3 Disembodied Voice

By contrast, the *disembodied voice* can be understood as a voice whose material source is obscured, displaced, or sonically altered. While, strictly speaking, any recorded voice is already disembodied (severed from its physical origin and played back via speakers), Zargorski-Thomas (2020) argues that pop production often plays with disembodiment more deliberately. For example, in Eminem's *The Real Slim Shady* (2000), "one of his vocal takes is processed as if it is being broadcast through a tannoy system, like a public address message in a supermarket" (p.16).

Disembodiment becomes more pronounced in heavily processed vocals, such as those manipulated through vocoders or pitch correction tools (see 2.2.4). These treatments can strip the voice of its human markers, introducing a sense of distance or even the uncanny. In this way, vocal production can emphasise or erase corporeality, shaping the listener's perception of identity, intimacy, and affect.

On this topic, Middleton (1990, p. 262) comments that:

> […] vocalising is the most intimate, flexible and complex mode of articulation of the body, and also is closely connected with the breath (continuity of life; periodicity of organic processes). Significantly, technological distortion of voice-sound (through use of a vocoder, for example) is far more disturbing than similar treatment of instrumental playing (which is regarded usually as a logical extension of manual performance).

As such, disembodied voices are not just removed from the physical source, but are often shed of their humanness, complicating notions of authenticity and presence in popular music. This tension between embodied and disembodied voice forms a crucial axis for this thesis, particularly in examining how vocal production choices construct intimacy, emotional meaning, vocal identity, and authenticity.

1.2.2.4 Register and Timbre Considerations, Performance Intensity

Building on Malawey (2020), this study considers how register and timbre function as expressive and symbolic tools, particularly in shaping the emotional intensity of a performance. Higher pitches sung in the chest voice are typically delivered with greater performance intensity than those sung in the head voice. While a combination of rising pitch and intensity often signals emotional outburst or climax, singing in a higher register with softened performance intensity can create a hushed, intimate quality that draws the listener closer (Appel, 2017, p.10). These shifts between chest and head voice, or between breathy and bright tonalities, often signal emotional or narrative changes that can be further shaped or amplified by production decisions. The diffuse quality of head voice may require close mic'ing and EQ adjustments to maintain presence and clarity in a dense mix, while the richer harmonic content of chest voice naturally sounds warmer and more present. Choices around compression and reverb can further reinforce these expressive functions.
Interestingly, these sonic and affective characteristics can also connote readings of *embodied* and *disembodied voice*: chest voice is often perceived as grounded and corporal (emanating from the chest and throat, closer to the heart), while head voice (projected primarily through the head cavities) may seem less embodied, and more ethereal or stylised.

1.2.2.5 Voice as Identity, Vocal Persona and Narrative Voice

The voice is a key signifier of identity (the biographical self), conveying markers such as gender, age, ethnicity, class, and cultural belonging through timbre, register, phrasing, and other vocal qualities (Malawey, 2020). Listeners automatically ascribe identities to voices, but these associations are culturally and socially constructed rather than fixed or biological. (Malawey, 2020, pp. 19-24).

This leads to the related concepts of *vocal persona* and *narrative voice*, which shape how listeners interpret a recorded voice's expressive function. Tagg (2012) describes vocal persona as a kind of "vocal costume"—a performative layer for expressing attitudes and behavioural positions, shaped by prosody, timbre, diction, and other expressive parameters p. 343 f). Importantly, this persona is not necessarily aligned with the singer's biographical

self; it is a crafted voice that may emphasise, mask, or entirely depart from the singer's own identity.

Narrative voice, by contrast, refers to the perspective or position adopted within the song's storytelling, often framed through pronouns, lyrical address, and staging choices. While the two concepts overlap, narrative voice pertains more directly to the lyrical "I" and its relation to the story, whereas persona encompasses the overall affective and stylistic framing of that voice in relation to the listener.

These distinctions are particularly relevant in the context of singer-songwriters, whose music is often perceived as autobiographical or emotionally authentic. However, even when writing their own material, artists construct narrative voices and adopt vocal personas that may align with, exaggerate, or deliberately obscure their own identity. The recording becomes a performative site where personal identity, artistic intention, and audience expectation intersect. The impact of technological mediation on the perception of vocal identity and persona is explored further in section 2.2.4.

1.2.2.6 Lyrical Analysis

Drawing on Mieke Bal's (1997) narrative theory and Barthes' (1974) semiotics, Randle and Evans (2013) proposes a typology for analysing pop song lyrics based on four core components of the *fabula*[10]: event, actor, time, and setting (p. 132). The more clearly these elements are present, the more defined the song's lyrical narrative becomes. This becomes particularly relevant in the context of this study, as these narrative components—especially actor, time and setting—can be actively shaped, reinforced, or even subverted through vocal production choices, effectively contributing to the construction of narrative meaning beyond the words themselves.

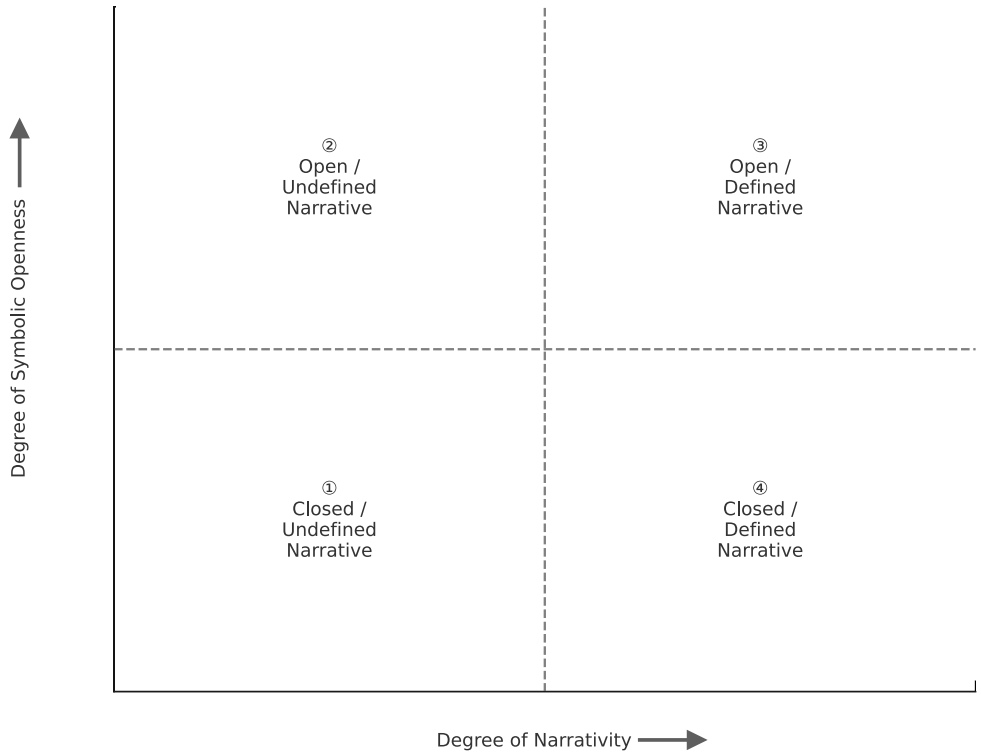
To visualise the interaction between narrative structure and interpretive openness in pop lyrics, Randle and Evans introduce a two-dimensional grid. The x-axis charts the degree of *narrativity*, ranging from *undefined narrative*—where events are static, fragmented, or absent—to *defined narrative*, characterised by identifiable characters, sequential events,

[10] *Fabula* is a narratological term drawn from Mieke Bal's (1997) work, referring to the "raw material" of a story (Randle and Evans, p.132).

and clear temporal or spatial grounding. The y-axis maps the degree of *symbolic openness*, from *closed* (literal, singular readings) to *open* (metaphorical or ambiguous lyrics that allow for multiple interpretations), drawing on Barthes' narrative codes (cf. p. 133). This matrix results in four quadrants:

(1) **Closed / Undefined Narrative** – literal lyrics with no clear story. LMFAO's *Party Rock Anthem* (2011), for instance, presents a static scenario of partying with minimal character development or progression.
(2) **Open / Undefined Narrative** – symbolic, abstract lyrics that evoke mood, imagery, or metaphor without a concrete narrative arc. Adele's *Set Fire to the Rain* (2011) is emotionally intense but abstract, leaving much open to interpretation.
(3) **Open / Defined Narrative** – structured storytelling that also employs symbolic language or metaphor. Keane's *Somewhere Only We Know* (2004) fits here, blending narrative with emotional and poetic ambiguity.
(4) **Closed / Defined Narrative** – a clear, linear story in a straightforward, literal fashion. Miley Cyrus's *Party in the U.S.A.* (2009) chronicles a moment in the singer's life with a clear setting, character, and progression, leaving little room for multiple readings.

Figure 1: Typology Grid of Pop Song Lyrics (adapted from Randle and Evans, 2013)



To illustrate the defined narrative quadrant in greater detail, Randle and Evans discuss Taylor Swift's *Love Story* (2008) as a textbook example of high narrativity. The song presents a clearly identifiable actor—a young girl in love—situated in a specific setting (*"on a balcony in summer air"*) and timeframe (*"We were both young…"*). The lyrics narrate a sequence of interrelated events, from initial longing (*"I sneak out to the garden to see you"*) to a climactic proposal *("He knelt to the ground and pulled out a ring"*). The presence of additional characters—most notably the overprotective father (*"my daddy said, 'Stay away from Juliet'"*)—further deepens the song's fabula. The spatial and temporal cues not only locate the narrative but also enhance the listener's sense of emotional immersion and progression, resulting in a tightly constructed lyrical story.

1.2.3 Analysis of the Technologically Mediated Voice

The recorded voice has always been shaped in some form by technology. From the invention of the microphone, even attempts to capture a "natural" performance involved mediation. In this context, the voice in recorded music is always shaped and staged[11] through a range of technical choices that influence perception and emotion. Today, listeners are not only accustomed to, but often expect deliberate sonic shaping of the voice as part of a genre's emotional and aesthetic identity.

The following selected frameworks from sound and phonographic studies offer critical tools for understanding how production shapes vocal affect, and have informed how I approached these decisions throughout the analysis and production of my own original songs.

1.2.3.1 Hall's Proxemic Zones

Edward T. Hall's theory of proxemics (1969) outlines four zones of interpersonal distance: intimate (0–0.5 m), personal (0.5–1.2 m), social (1.2–3.6 m), and public (3.6+ m). Originally developed to describe spatial behaviour in face-to-face communication, these zones have since been applied to the mediated voice in recorded music.

[11] Consider how most "MTV Unplugged" performances were never actually "unplugged".

Dibben (2013) adopts Hall's framework to argue that *"acoustic information about interpersonal distance influences the listener's perception of emotional content and intimacy"* (p. 116). She suggests that spatial cues such as vocal clarity, reverberation, and the proximity effect carry affective significance by simulating physical closeness between performer and listener.

Lacasse (2000) similarly draws on Hall's proxemic theory in his discussion of phonographic staging, where production tools like reverb, EQ filtering, and microphone placement construct a sense of spatial positioning. These techniques not only determine how close the voice feels, but also guide the emotional and narrative interpretation of a performance.

1.2.3.2 Stereo Sound and Mono-Centricism

Stereo sound refers to the use of two distinct audio channels — left and right — to create the illusion of spatial distribution in recorded or amplified audio. Unlike mono recordings, which deliver the same signal through all speakers or headphones, stereo enables listeners to perceive directionality, depth, and dimensionality. When experienced via stereo loudspeakers, this creates a *phantom image*: a virtual stage positioned in front of the listener where each sound source can be placed with intentionality.

Dockwray and Moore's (2010) *sound-box* model offers a framework for analysing this phenomenon in music, mapping sound placement across three key axes: stereo width (left to right), frequency height (low to high), and spatial depth (near to far). This three-dimensional space enables nuanced sculpting of emotion, affect, and perceived presence.

Tagg (2012) highlights the analytical importance of such spatial design, noting that "aural staging is still often overlooked as a vital parameter of expression to consider in the analysis of the vast majority of music produced since the mid 1960s" (p.30). Despite the rise of immersive formats, stereo remains the standard both in production and consumption. All of this thesis's analytical listening and case study reflections likewise refer to music experienced through stereo loudspeakers or headphones.

Young (2019) expands the conversation into headphone listening, which shifts spatial perception: Through headphones or earphones, sound is often perceived to originate inside the listener's head — a psychoacoustic phenomenon known as in-head localisation (IHL). As

Görne (2015) explains, headphone playback bypasses spatial cues from the outer ear and the room, causing the auditory image to collapse into the space between the ears (p. 131). This "internalisation" of sound can heighten the perceived intimacy of a vocal performance. In this context, the use of stereo reverb becomes particularly important: when elements are hard-panned to one side of the stereo field, they may disappear entirely if only one earbud is used. Stereo reverb helps preserve spatial cues and prevents such losses, ensuring a fuller listening experience.

In popular recordings, the normative placement of the lead vocal follows what Philip Tagg (2012) describes as "mono-centric mix" — where the lead vocal is "panned centre and mixed louder than the other sources" (Lacasse, 2000). This central positioning is not just an aesthetic convention but a marker of the voice as the primary conveyor of meaning. While the mono-centric mix remains the dominant convention in pop music, panning the lead vocal to an unexpected position — such as hard left — can draw heightened attention precisely through its contrast and destabilising effect (p. 139).

1.2.3.3 Vocal Staging

Lacasse's (2000) concept of *phonographic staging* explains how production situates the voice within a virtual space, actively shaping the song's narrative by manipulating the listener's sense of space, time, timbre or loudness. He distinguishes between three categories:

> (*1) spatialisation of the voice (e.g. reverb, stereo location);*
> *(2) electrical alteration of vocal Timbre (e.g distortion, telephone effect, vocoder);*
> *(3) temporal manipulation of the voice (e.g. echo, reiterations)*

Table 1: Lacasse's General Categorisation of Vocal Staging Effects

Categories		Corresponding Settings
Spatialisation	**Environment**	Flat voice; Voice with different types of reverberation; Short echoes; etc.
	Stereo Location	Centre; Left; Right; Bilateral; Movement; etc.
	Distance	Close; Far; Background; etc.
Timbre Modification		Flanging; Distortion; Telephone Effect; etc.
Temporal Manipulation		Echo; Reiteration; etc.

Drawing on Moylan (1992), Lacasse elaborates that *spatialisation* in recording can be understood through three key perceptual dimensions:

> (1) *The perceived location of how the sound source is positioned along the horizontal left-to-right axis in a stereo mix;*
> (2) *The perceived distance of the sound source from the listener; and*
> (3) *The environmental characteristics in which each sound source appears to be situated* (2000, p. 176).

Lacasse emphasises that his proposed classification of vocal staging into spatial, timbral, and temporal dimensions is not meant to be exclusive, as these aspects are "usually mutually intertwined." For example, reverberation—a primarily spatial effect that places the sound source in a given environment—also alters the timbre by enhancing certain frequencies and extends the duration of the sound, thus engaging all three categories simultaneously. Similarly, a vocal treatment might involve a flanging (timbre), be panned left (space), and feature a long echo (time), making it difficult to isolate effects cleanly. Some settings even evolve over time, further complicating analysis. Ultimately, Lacasse presents this framework not as a definitive system, but "simply to help us orientate the examination process" (p. 169).

In combination with Hall's theory of proxemics, these concepts become powerful tools in the analysis and production of popular songs.

1.2.3.4 Opaque and Transparent Mediation

Brøvig-Hanssen and Danielsen (2016) distinguish between *transparen*t and *opaque* mediation to describe how listeners perceive the role of technology in recorded sound. When mediation is *transparent*, the technological layer goes unnoticed, and the listener experiences the sound as immediate and emotionally direct. *Opaque mediation*, by contrast, draws attention to the fact that the sound has been shaped or constructed—whether through obvious use of tools like pitch correction, sampling, or other highly stylised effects. As Brøvig-Hanssen and Danielsen note, this often occurs when listeners are confronted with unfamiliar technologies; initially, their attention is drawn to the audible traces of the technology instead of the musical content itself. Over time, however, these features may become "naturalised", no longer perceived as intrusive but as part of a new aesthetic norm. For example, few listeners today would consider 1940s crooning[12] as profoundly intimate, though it was once perceived that way (p. 7). A more recent example of this naturalisation process is explored in Chapter 2.2.4 in relation to Auto-Tune: once heard as a hallmark of artificiality, the effect is now often used in ways that do not compromise authenticity, even within the singer-songwriter tradition.

[12] see chapter 2.2.1

2 — Historical and Aesthetic Context

Having established the analytical lens through which this thesis approaches vocal production and analysis, this next chapter situates my practice within a broader historical and aesthetic context of singer-songwriter and pop vocal traditions. Starting with a brief depiction of the singer-songwriter stereotype, section 2.1 then traces the evolution of the singer-songwriter from the 1960s to the present day, highlighting key stylistic shifts, technological developments, and gendered dynamics that inform current vocal aesthetics.
The following section 2.2 explores contemporary production techniques that shape the emotional and narrative power of the recorded voice, providing a contextual grounding for the reflective analysis of my own production practice.

2.1 Defining the Singer-Songwriter

> *"The emergence of the singer-songwriter was not just a moment in the early 1970s, but the start of a new formation that continues to this day. "* - David Shumway (2016)

While the term *singer-songwriter* is often used interchangeably to describe a type of artist, a musical genre, or a stylistic approach, these categories are not synonymous. Each has evolved along its own path, often overlapping but not always aligning. Malawey (2020) recognises this ambiguity, noting that "defining specific features of the 'genre' of 'singer-songwriter' is wrought with difficulty. The wide breadth of styles and lyrical themes within the umbrella category 'singer-songwriter' makes defining the genre challenging" (p.167n).

To trace how technological, aesthetic, and cultural changes have shaped the singer-songwriter tradition, particularly in terms of vocal production and perceived authenticity, I have chosen to map its development through a broad chronological timeline. Beginning with a brief overview of the historical emergence of the singer-songwriter as a recognisable formation in popular music, this structure allows for a clearer understanding[13] of how evolving recording practices and vocal aesthetics have shaped the notions of intimacy and authorship over time.

---

[13] To further clarify this multifaceted concept, a comparative table summarising key distinctions of the singer-songwriter as genre, identity, and style is included in the appendix (see *Table A*).

2.1.1 The Stereotype

In its most condensed form, the term *singer-songwriter* typically evokes the image of a solo artist who writes and performs their own material and accompanies themselves on an acoustic guitar or piano, with "the acoustic guitar being the most emblematic" (Appel, 2017, p. 7). This image has often been reduced to the stereotypical figure of the (bearded), white, Anglo-American male, rooted in folk, pop, or country traditions (cf. pp. 11–13). However, this reductive image reinforces racialised and gendered assumptions, highlighting the importance of broadening our analytical framework to include artists outside this dominant stereotype.

Figure 2: Singer-Songwriter Stereotype(s), generated by ChatGPT

Description: I asked 12 different people to use the prompt ChatGPT as follows: "Generate a picture of THE singer-songwriter stereotype". These are their results: 2/12 pictures were of women, all were white. This is a representative collage of the results.

2.1.2 Genre Formation (1960s–1970s)

Overview

According to Shumway (2016), the singer-songwriter emerged as a distinct genre around the early 1970s. While the first songs that would later come to define the genre appeared around 1968, the label 'singer-songwriter' was not widely used until the early 1970s, when the figure became a recognised presence in popular culture[14]. The retrospective categorisation helped frame a distinct tradition that emerged as a counterpoint to both commercial pop and politically charged folk movements of the time, marking a shift from collective voice to individual introspection (Shumway, 2016). Central to this transformation was the genre's perceived authenticity and confessional mode, which was heavily influenced by confessional poetry from the likes of Robert Lowell, Sylvia Plath and others of the late 1950s and 60s (p. 16).

In the US and Canada, artists such as Bob Dylan, Joni Mitchell, Leonard Cohen, James Taylor and Carole King set the tone for what would become a new musical genre defined by its values: personal storytelling, lyrical intimacy, (perceived) authenticity and understatement[15].

Figure 3 (left): Bob Dylan in 1961 (Source: Michael Ochs Archives/Getty Images)
Figure 4 (right): Joni Mitchell in 1968 (Source: Jack Robinson/Getty Images)

[14] While the roots of the singer-songwriter can be traced back to historical figures such as the medieval troubadour, the formation most relevant to this study begins with its emergence as a phenomenon of Anglo-American popular music culture in the late 1960s.

[15] Alexandra Apolloni (2017) describes these early figures as "folk songbirds" with "naturalistic images: they dressed unpretentiously, and their voices often sounded untrained" (p. 214). While many scholars, such as Apolloni, tend to conflate the folk and singer-songwriter genres, Shumway, by contrast, draws a clear distinction between the two—emphasising the shift from a collective voice to one of individual introspection.

As Williams & Williams (2017) put it, these artists "laid bare" their lives through their music, often using a raw, unpolished singing style (p. 1). Appel (2017) further describes the genre's aesthetics as typically involving acoustic instrumentation, a direct and seemingly "unmediated" vocal delivery, and lyrical content that is often "poetic," "personal," and "confessional" (p. 7). These qualities extended into both their stage presence and studio production, with singer-songwriters often favouring smaller live performance venues that complemented the intimate character of their music.

However, the assumption that the singer-songwriter is inherently a *solo* performer is misleading. Challenging this common notion, Shuker (2008) points out, most well-known figures associated with the label regularly perform and record with backing bands (p. 53). Similarly, Wise (2012, as cited in Williams and Williams, 2014) defines singer-songwriters as artists who compose and perform their own material—typically solo and accompanied by guitar or piano—but also notes that *recordings often involve additional musicians, especially to flesh out studio versions* (p. 2, emphasis mine). For example, Carole King, who is widely considered one of the genre's most successful icons, used full band arrangements in the studio.[16]

Meanwhile in the UK, the figure of the singer-songwriter developed in parallel with the US and Canada, following slightly different aesthetic and cultural trajectories. Artists such as Elton John and David Bowie gained major popularity, each with their own distinctive style. Their careers demonstrate that the genre did not always imply autobiographical writing, nor was it necessarily confined to a quiet, introspective mode.[17] In Elton John's case, it was a widely known fact that the lyrics were written by Bernie Taupin, while John composed the music. Both Bowie and John performed with elaborate stage personas, challenging the notion that singer-songwriter authenticity required minimalism or personal introspection. What remained essential was the impression of authorship and direct address—a coherent persona or narrator that invited the listener into an emotionally rich world.

---

[16] In fact, there are only a select few songs on her first two albums Writer and Tapestry, where the arrangement is more sparse — e.g. *Child of Mine* and *You've Got a Friend*.

[17] This is not to say that there weren't any UK singer-songwriters that fit the introspective, confessional label: See Cat Stevens, or Nick Drake (whose music received wide acclaim posthumously).

Shumway acknowledges this development, noting that while the genre evolved beyond strict confessionalism, it retained a strong association with "authentic individual expression" (p. 19). This idea of "'true' auteurship," as Shuker (2008, p.53) frames it—placing "true" in quotation marks—gets to the heart of the matter: it is not necessarily factual autobiography or compositional autonomy that defines the genre, but rather the *perception* of sincere, individual authorship.

By the late 1960s and early 1970s, recording technology began to shift from 4-track to 8- and 16-track tape machines, allowing greater control over vocal placement and stereo imaging, which had become the dominant format by this time. Studios were increasingly experimenting with multitrack layering, analog reverbs, and stereo panning.[18]

Early typical production aesthetics of the singer-songwriter genre and techniques reflected its 'unpolished', 'raw' and naturalistic[19] ideology through minimal processing, sparse instrumentation, and careful spatial placement, to preserve a sense of authenticity. A single lead vocal was typically foregrounded in the centre, placed close to the listener in the mix to enhance the feeling of direct address and emotional intimacy. Effects like reverb or compression were applied subtly, if at all. The resulting vocal sound was intimate, dry, and closely tied to the performer's physicality in the recording space. Lacasse (2000) refers to rooms with little to no natural reverb as "flat" environments. He states "the type of subtle, 'natural' reverberation used in such songs, or its complete absence, helps the listener focus on [the] voice itself and the uttered words rather than on sound effects that could be considered disturbing" (p. 180 f).

Listening Examples

Joni Mitchell's album *Blue* (1971) is often cited as a landmark in confessional songwriting. The ninth track of the album, *A Case of You*, uses open tunings and minimal acoustic instrumentation to create a sense of unfiltered emotional exposure. The recording, produced on an 8-track tape machine and a "rudimentary HAECO console with just eight

---

[18] During this time, some artists, such as the Beatles, are known to have experimented heavily with the use of panning as a storytelling device. (See *Sgt. Pepper's Lonely Hearts Club Band (*1967))

[19] see 1.1.2

busses and an EQ section with only three fixed midrange bands"[20], is a prime example of how production choices allow Mitchell's voice to maintain a raw and untreated character while still being foregrounded in the mix. Her production choice to simply pan the accompanying instruments to either side of the stereo image allows her delicate vocal to take centre stage without having to compete with the rest of the instrumentation. By designing the stereo image this way, Mitchell's singular voice appears 'naked', offering a delicate sense of clarity and precision that renders audible every breath and vocal nuance, without the need for overt use of EQ, compression and reverb.
Carole King's *So Far Away* (1971) showcases similarly understated production values, with warm piano accompaniment and a dry, upfront vocal that conveys emotional intimacy without resorting to heavy processing.

Leonard Cohen's *One of Us Cannot Be Wrong* (1967) begins with a naturalistic production style that foregrounds his voice, in a dry, intimate mix. However, in the outro (from around 3:20 onwards) Cohen begins to whistle and vocalise the main melody in a sequence of overdubs. His voice (he appears to be shouting the melody "lalala") is treated with more distance and panned slightly to the right of the stereo field. It then starts moving around the stereo image, as if captured by walking around the room, arriving on the far left at 3:58, before returning once again to the right just a second later. This movement across the stereo image contributes to a feeling of emotional fragmentation or disorientation further highlighting how even within a minimalistic framework, production choices can be used expressively.

2.1.3 Rise of the Artist-Producer (1980s)

In the 1980s, advances in recording technology expanded the creative possibilities for artists working independently or with greater control over the studio environment. The wider availability of multitrack recording, digital synthesisers, drum machines, and outboard effects enabled artists to craft increasingly complex sonic worlds—often acting as composers, performers, and producers in one (Burgess, 2013; Kealy, 1979).

[20] According to Mitchell's official website: https://jonimitchell.com/library/view.cfm?id=4622

Bruce Springsteen and Prince for example, embodied this new type of agency. Prince, in particular, wrote, arranged, recorded, and produced nearly all of his music, using the studio as an extension of his artistic identity. This shift marked a move away from the sparse and "unfiltered" singer-songwriter sound of the previous era toward a more constructed and manipulated layered aesthetic. Tagg (2012) notes how the "range of timbral variation […] radically expanded […] with the spread of electro-acoustic […] devices […] for treating audio input signals" (p. 309).

Importantly, the 1980s also witnessed pioneering contributions from women who used technology to shape a distinct authorial voice. Kate Bush, who produced much of her work herself, offers a compelling example. Her 1985 album *Hounds of Love* demonstrates meticulous control over arrangement, vocal layering, and spatial design.

Rather than undermining authenticity, this shift reflected a changing understanding of artistic agency that included conceptual production and sonic architecture. The singer-songwriter model began to expand, accommodating voices that were audibly shaped by technology.

Examples: Kate Bush *Running Up That Hill (1985);* Prince *Kiss (1986)*

2.1.4 The 'Female Singer-Songwriter' Label (1990s)

In the 1990s, the singer-songwriter label began to be more frequently associated with female artists. On the one hand, artists such as Suzanne Vega and Tracy Chapman continued to work in a relatively traditional format, foregrounding voice and lyrics, with supporting accompaniment. Their music retained the intimate and stripped-back aesthetic of the 1970s, while incorporating updated production techniques (at times leaning away from the *natural-realistic*[21] aesthetic and incorporating synthesisers) and engaging with broader lyrical themes, including political and social commentary. Shuker (2008) remarks that "this female predominance led some observers to equate the [singer-songwriter] 'form' with women performers" (p.53).

[21] see 1.1.2 and 2.2.3

At the same time, a different group of diverse women emerged—Björk, Alanis Morissette, PJ Harvey, Tori Amos, Ani DiFranco, and Fiona Apple—who, despite drawing from vastly different genres (including electronic music, jazz, folk, rock, baroque pop and punk) were frequently grouped under the singer-songwriter label. As Boak (2016) argues, this classification had more to do with gender and perceptions of authorship than shared musical traits. What particularly distinguished this group was their explicit engagement with sexuality, gender, and the female body. Drawing on the confessional tradition, they infused political messages with a newly embodied and often "angry" energy. The infamous 1994 *Q* magazine cover featuring PJ Harvey, Björk, and Tori Amos under the headline *"Hips. Lips. Tits. Power."* exemplified the way their music was framed—and often sensationalised—as expressions of female physicality and power.

Figure 5: The famous 'Q' cover with PJ Harvey, Björk and Tori Amos.

This group of women exercised a high degree of creative control, often collaborating with independent labels or managing their own releases. Increasingly they assuming roles in the production process; a shift facilitated by the widespread adoption of digital recording technologies throughout the 1990s. Tools such as MIDI, synthesisers, and affordable multi-track cassette recorders and digital samplers allowed artists greater freedom to record and

produce outside of commercial studio settings. This enabled more women to claim the production space and experiment with sound on their own terms, shaping how their voices were heard. DiFranco, for instance, released music on her own label, *Righteous Babe Records*, combining DIY methods with a distinctive vocal sound (employing vocal techniques such as vocal fry and growling), examples of which can be heard on her album *Not A Pretty Girl* (1995).

Rather than fitting neatly into a genre, the 1990s singer-songwriter category thus became a flexible framework for women to explore authorship, intimacy, power, and physicality. Whether embracing raw, DIY aesthetics or pushing the boundaries of technological experimentation, these artists reaffirmed the singing voice as a site of both emotional truth and embodied resistance.

Alanis Morissette's *You Oughta Know* (1995), in particular, was "an anthem for female rage" that achieved mainstream success (Boak, 2016). The album's success hinged on "rawness" that was, in fact, carefully crafted: production choices amplified vocal imperfections — cracks, screams, breath sounds — to stage emotional intensity.

2.1.5 Expanding the Sonic Palette (2000s)

The early 2000s marked a pivotal moment for the singer-songwriter tradition, as a new generation of artists radically expanded its sonic palette. This shift was driven largely by the rise of affordable home-recording technology and online platforms for digital distribution, such as Napster, iTunes and MySpace which enabled a new wave of independent musicians to bypass the traditional gatekeepers: major labels; A&R executives; and radio programmers. Online video distribution platforms such as YouTube further fuelled a new DIY culture resulting in an explosion of diverse approaches to songwriting, vocal production, and genre hybrids:

On one end of the sonic palette, introspective male artists such as Sufjan Stevens, José Gonzales, Iron & Wine, Damien Rice, William Fitzsimmons, Bon Iver[22], and Bright Eyes[23] embraced hushed, emotionally fragile vocal deliveries and lo-fi, handmade textures —

[22] band with one lead vocalist

[23] band with one lead vocalist

aesthetics that foregrounded intimacy and vulnerability, aligning with the indie-folk ethos. Ian Biddle (2007) observes that these artists exhibit "a kind of openness to vulnerability, a commitment to social and sexual intimacy, and a tendency to want to avoid the overt spectacularization of masculinity" (p.125).

By the mid 2000s, video streaming platforms such as YouTube had given rise to a more playful and upbeat aesthetic, particularly notable among female singer-songwriters like Ingrid Michaelson or Julia Nunes, who gained popularity with ukulele-driven songs, layered harmonies, and handmade video content. These artists embraced the DIY ethos, creating acoustic, melodic, and pop-leaning recordings that resonated due to their authenticity and individuality and quirkiness. Regina Spektor, Feist, Angel Olson, Kate Nash and Agnes Obel brought even more idiosyncratic singing styles and eclectic musical influences into the mix.

Others — notably Alicia Keys, Norah Jones, Missy Higgins, Sara Bareilles, and Adele — brought a more polished, virtuosic vocal presence into the genre, blending singer-songwriter intimacy with elements of soul and jazz, rock and pop— often delivering technically polished, mainstream-radio-friendly recordings.

Large-scale commercial breakthroughs from artists like Taylor Swift and Ed Sheeran continued in the early-mid-2010s, demonstrating that the singer-songwriter could thrive within mainstream pop. Swift in particular embodied the evolution from country-rooted storytelling to global pop stardom while maintaining a consistent narrative voice. During this decade, the indie-folk lineage continued through artists like Laura Marling, Lucy Rose, Lisa Hannigan, Ben Howard and Passenger.

On the other end of the sonic spectrum, the 2000s had also witnessed pioneers like Imogen Heap fully integrating music technology into their creative identities, blurring the lines between producer, songwriter, and performer through innovative use of looping, sampling, and custom-built interfaces.

Along this stylistic spectrum, artists such as James Blake (e.g. *Limit To Your Love* (2011) / *The Wilhelm Scream* (2011)) and Lykke Li (e.g. *Little Bit* (2008) / *I Follow Rivers* (2011)) further expanded the sonic palette by adopting more hybrid approaches to technology, integrating ambient layers and electronic beats into their work.

This period also saw the rise of the *singer-songwriter aesthetic* as a distinct *stylistic marker*, best demonstrated through cover songs: especially in acoustic covers of well-known pop songs. These stripped-down versions often aimed to highlight emotional depth or a sense of "authenticity" through soft vocals, minimal instrumentation, and foregrounded lyrics.

Examples:

Radiohead[24] *Creep - Acoustic Version* (1992) — Original: Radiohead, 1992

Eva Cassidy (2000) and Iron & Wine (2016) *Time After Time,* Original: Cindy Lauper, 1983

Obadiah Parker *Hey Ya - solo version* (2008) — Original: Outkast, 2003

Run River North *Mr Brightside* (2017) — Original: The Killers, 2004

Callum Scott's *Dancing On My Own* (2018) — Original: Robyn, 2010

William Fitzsimmons *I Kissed A Girl (2010)* — Original: Katy Perry, 2008

Interestingly, because the original versions of theses songs were so well known, the covers did not challenge authorship but instead reframed the emotional core of the songs. This illustrates how the singer-songwriter aesthetic can function as an interpretive lens, applied as a distinct *style*.

In summary, the 2000s witnessed an explosion of vocal and production aesthetics within the singer-songwriter genre. The democratisation of music-making tools, combined with internet distribution, enabled a generation of artists to radically stretch the genre's boundaries. Figures like Sara Bareilles, Adele, Ed Sheeran, and Taylor Swift brought the singer-songwriter image firmly into the mainstream.

Listening examples to illustrate the diversity of the sonic palette (some examples were already provided in-text):

Imogen Heap *Hide and Seek* (2005)

Ingrid Michaelson *Be OK* (2008)

---

[24] Technically not a cover, just a stripped down version by the same band. Still a good example of singer-songwriter as *style*!

Adele Chasing Pavements (2008)

Sara Bareilles Gravity (2007) — features soft compression and reverb that enhance the vocal's emotional pull.

Sara Bareilles *Lovsong* (2007) — compressed vocals

Sara Bareilles *King of Everything* (2010) — stereo panned alternating intro backings

Feist *How Come You Never Go There* (2011) — layered 'whoa' harmonies, example of rise of 'acoustic' vocal layering in the 2010s

2.1.6 The Contemporary Singer-Songwriter (approx. 2018-2025)

From the mid-late 2010s onward, the singer-songwriter format underwent a renewed transformation, shaped by a new generation of self-producing artists who fuse introspective storytelling with the aesthetics and workflows of digital production. This shift is not merely technological, but also aesthetic and ideological, reflecting evolving constructions of authenticity, artistry, and vocal presence within contemporary popular music.

Artists such as Billie Eilish, Phoebe Bridgers, King Princess, and Gracie Abrams foreground a hyper-intimate vocal aesthetic, marked by close-mic'ed, whispery delivery and ambient, lo-fi textures. These sonic choices convey not only fragility and introspection but also subtle forms of resistance, as themes of gender, queerness, and mental health permeate their lyrical narratives. Central to this aesthetic is the embrace of imperfection; breath sounds, vocal fry, and unpolished textures are treated not as flaws, but as intentional expressive devices. Increasingly, the DAW functions as a compositional instrument in its own right, with tools such as comping, layering, and automation deeply embedded in the songwriting process[25], with vocal settings often shifting every few seconds within one track, reflecting a micro-level aesthetic that treats the voice as a dynamic, sculptable element.

Furthermore, social media platforms such as Instagram and TikTok reshaped how singer-songwriters connect with audiences. Artists like dodie, corook, Lizzy McAlpine, and Victoria Canal cultivated large followings through intimate, self-recorded snippets that often felt

[25] Charlie Puth, for instance, exemplifies a pop-facing artist for whom these processes are deeply interwoven from the outset.

more emotionally "authentic" than studio-polished releases. Within these digital spaces, the lo-fi aesthetics and unfiltered presentation function as *renewed* markers of sincerity and relatability (cf. Stiegler & Campbell, 2023).

The COVID-19 pandemic further accelerated these trends as musicians were confined to domestic recording environments, leading to a surge in home-produced music, live-streamed performances, and remote collaborations. This period also encouraged a blurring of genre boundaries, as a "post-genre" ethos took hold in particular among Gen Z artists (Muchitsch, 2023). This shift builds on and diversifies the developments outlined in section 2.1.5 and is also reflected in the career trajectories of singer-songwriters with mainstream success, such as Ed Sheeran, whose work increasingly spans pop, folk, R&B, hip-hop, and afrobeat influences. Like Sheeran, many artists prioritise mood, atmosphere, and vocal identity as central to their aesthetic, rather than pursuing a strictly polished or homogenous sound.

While established singer-songwriters like Ed Sheeran and Taylor Swift exemplify post-genre fluidity (between and within individual songs) and personal expression within mainstream success, many contemporary singer-songwriters—particularly those from marginalised groups—challenge traditional conventions of authorship, agency and gender in music production. Reflecting a broader cultural shift toward hybridity, vulnerability, and self-definition, these artists emphasise emotional intimacy and digital autonomy as central to their work.

Below I list a few listening examples, some of which will be analysed in more depth in the following chapter.

Taylor Swift *cardigan* (2020)
Taylor Swift *Vigilante Shit* (2022)
Ed Sheeran *Shape of You* (2017)
Burna Boy (feat Ed Sheeran) *For My Hand* (2022)
Ed Sheeran *Sapphire* (2025)
Billie Eilish *i love you* (2019)

Phoebe Bridgers *Scott Street* (2018)

Girl in red *bad idea!* (2019) — self-produced

Gracie Abrams *Rockland* (2021)

Dodie *She* (2019)

King Princess' *If You Think This is Love* (2019) — co-produced

King Princess *Cursed* (2022) — co-produced

Lizzie McAlpine *Ceilings (2022)*

Mitski Bug *Like an Angel* (2023)

Ellie Dixon *Green Grass* (2022) — self produced and heavily sound designed!

2.1.7 Summary

This chapter has shown that while the term 'singer-songwriter' is often used interchangeably to describe an artist type, a musical genre, or a stylistic approach, these dimensions are distinct and have evolved along their own trajectories. By tracing the historical emergence and ongoing transformation of the singer-songwriter tradition, it becomes clear that the label functions less as a fixed category and more as a flexible set of values centered on personal expression and authenticity.

As a *genre*, the label refers to a tradition that crystallised in the 1960s and 70s, rooted in confessional lyricism, acoustic instrumentation, and a focus on emotional truth. As a *style*, it typically denotes a sonic approach that foregrounds the voice — often minimally produced — and relies on sparse acoustic arrangements that support rather than overshadow the lyrical content. In this stylistic sense, the label is not necessarily dependent on self-authorship; it can also encompass acoustic reinterpretations and covers that reflect the sonic markers of sincerity. As an *artist identity* the term is applied far more flexibly for any musician who simply writes and performs their own material, regardless of genre affiliation or production approach. This includes artists with a wide range of vocal aesthetics (trained or untrained, natural or digitally processed), who may blend styles or employ modern production tools, yet still uphold the self-authored ideal that defines the tradition.

Given the fluidity of the term, it becomes useful to think of the singer-songwriter less as a fixed category and more as an ethos: a set of values and practices that continue to adapt to new cultural contexts, technologies, and production techniques. Appel (2017) argues, for

example, that rather than viewing the singer-songwriter as a fixed genre or identity, it is more productive to understand it as a 'mode'—a set of aesthetic and rhetorical choices that can be applied across a wide range of musical styles and contexts (p. 8). He identifies four central principles of the "Singer-Songwriter Disposition" (p. 9):

(1) foregrounding of the ordinary human voice over virtuosity
(2) intimate sonic environments
(3) direct listener address
(4) an air of sincerity

This framework allows for a more inclusive and contemporary understanding of the singer-songwriter identity while reflecting a broader cultural evolution in how authenticity is perceived in popular music. The association between "realness" and unmediated, acoustic vocal performance has given way to a more expansive understanding—one that includes both raw, lo-fi aesthetics and highly constructed, digitally enhanced voices as valid forms of emotional expression. As Appel (2017) notes, "the association between acoustic instruments and a sense of intimacy and proximity that was widely accepted in the 1960s and 1970s seems to have faded in the last fifteen years. This is probably a result of the increasing ease of use and access to electronic and digital means of music production, which inevitably led to rising audience familiarity with their sound" (p. 11).

Appel further affirms that technological mediation can *enhance* rather than compromise authenticity (p.13), challenging traditional assumptions that associate authenticity exclusively with acoustic performance or unprocessed vocals.
While Laura Busby (2017) adds that the traditional concept of the singer-songwriter often overlooks the digital and post-digital methods of music-making that have become increasingly common today (pp.113 f.), Appel observes that "cyborgian vocalities" have become widespread among artists like Bon Iver and James Blake, who use technology to convey emotional honesty and intimacy. Similarly, Williams and Williams (2017) note that contemporary artists employ technology and diverse influences to create "new forms of intimate expression" (p. 1).

This understanding is particularly useful for my own research, which aims to explore vocal production choices and emotional resonance within my self-produced singer-songwriter work. By treating the singer-songwriter as an ethos rather than a genre label, I am able to experiment with aesthetics of shaping intimacy and narrative—such as doubling, reverb, layered harmonies, and signal processing such as vocoder or distortion—without abandoning the authentic core that characterises the singer-songwriter tradition.

The following section explores a specific selection of vocal production techniques that, from my own observation as both listener and artist, have become especially prominent in popular music over the past six to seven years (roughly 2018–2025), and how they can function expressively within the singer-songwriter 'mode'.

2.2 Contemporary Pop Vocal Production Aesthetics and Techniques

Within contemporary popular music[26], the voice remains the primary vehicle for emotional expression and narrative delivery. As Harding (2020) states, "the most important part of any commercial pop record in Western music is the vocal" (p. 307). Or as Lacasse (2000) puts it, the "voice is generally recognised as the most important sound source in popular music" (p.9). This centrality is also reflected in its treatment within the mix: "the normative arrangement of vocals within the mix places the voice centrally, and louder than other sounds, creating what Philip Tagg calls the 'monocentric mix'—the singer is central, foregrounded and often close to the listener" (Dibben, 2013, p. 113). This is reinforced by the intensive treatment of vocals in contemporary pop production practices that are multi-layered and highly detailed. As Thomas (2022) notes in her PhD thesis, "the sound of the commercial pop singing voice is, in its totality, produced through a host of cumulative digital processes" (p. 14). These processes typically unfold across three main stages: recording, editing, and mixing. In the recording phase, close-mic'ing techniques and acoustically treated environments are used to capture clean, isolated vocal performances, often across multiple takes (p. 144). The editing stage is arguably the most laborious as it involves "painstakingly working through vocal tracks at a microscopic level of detail" (p.141). This includes *comping* (assembling the best parts from several takes), timing and pitch correction using tools like *Auto-Tune* or *Melodyne* (p. 148 f.), clip gain adjustment, and spectral repair or noise removal using software such as *iZotope RX* (p. 161).[27]

---

[26] Throughout this thesis, terms such as 'contemporary popular music' and 'commercial pop' are used rather interchangeably in reference to the dominant aesthetics, production practices, and vocal treatments that characterise mainstream Western popular music in the 21st century.

[27] These tools exemplify what Thomas (2022) identifies as a visual turn in music production, whereby vocal editing becomes "as much a visual as it is a sonic process" (p. 157). Through pitch curves, spectrograms, and waveform displays, the act of shaping the voice increasingly relies on visual feedback, heavily influencing aesthetic decision-making. In this context, producer Kieran Hebden (aka Four Tet) comments "people who make music on computers don't realise how powerful the visual element is. Whether you like it or not, your mind starts to think in terms of patterns, because it's a natural human way to do things […]", and adds caution that, "it's important just to close your eyes and use your ears, and trust what's coming out of the speakers more than anything" (Marrington, 2017, p. 80).

Finally, in the mixing stage, these micro-edits are integrated into a cohesive sonic identity through a "vocal chain" of EQ, compression, de-essing, saturation, reverb, and delay—often including "multiple rounds of tuning as well as sometimes 7 to 10 different EQ and compression plug-ins in iteration, each enacting a very subtle modification to the voice." (p. 29).

Figure 6: An example of my lead "vocal chain" in "Forever Vow"
Description: Even though this song has a more natural-realistic aesthetic, it still makes extensive use of audio processing tools and plugins.

While many of these sonic treatments remain largely imperceptible (cf. *transparent mediation*, 1.2.3.4) to listeners untrained in music production, they nonetheless play a crucial role in shaping how the voice is perceived. While a lot of commercial vocal editing tends to favour transparency (subtly refining pitch, timing, and dynamics), *opaque* mediation (cf. 1.2.3.4) is also commonly employed, such as when Auto-Tune or vocoder effects are used in an exaggerated way to emphasise a more artificial, processed quality of the voice. That said, the boundary between these modes is not fixed. Even within a single track, vocal settings often shift every few seconds, reflecting a micro-level aesthetic that treats the voice as a dynamic, sculptable element. As listeners have grown increasingly accustomed to this shifting aesthetic over time, what was once experienced as overtly synthetic may now be perceived as part of the expressive norm in contemporary pop.

In today's popular music landscape, the tools available for vocal production are more precise, flexible, and accessible than ever before. At the same time, this accessibility has contributed to what may be described as a *preset culture* (Frith & Zagorski-Thomas, 2012, p. 5; Shields et al, 2024, p. 187*)*, in which ready-made vocal chains, processing templates, and

effect presets circulate widely through online platforms and commercial marketplaces. While this enables a broader range of creators to achieve professional sounding results, it also raises important questions about creative authorship, originality, and intention. These debates are further fuelled by automated processes powered by AI and machine learning being increasingly incorporated into vocal processing, pitch correction, and mixing. As Théberge (2020) notes, while these tools prioritise "efficiency, ease of use and low cost," they also "emphasize notions of control, conformity and perfection over creativity and experimentation" (p. 84). In this context, the artistic rationale behind vocal production choices—*why* a particular effect or technique is used—emerges as a critical and often overlooked variable.

Against this backdrop, the following section highlights four central techniques that illustrate the expressive range of today's vocal production:

(1) *close-mic'ing and compression* as tools of hyper-intimacy;
(2) *double-tracking and layering* as tools for emotional complexity;
(3) *reverb and echo* as emotive spatial and temporal staging devices; and
(4) *pitch-modulation, vocoder, and distortion*, which, if applied overtly, can transform the perceived authenticity, gender, or even humanity of a voice; or simply add emotional depth.

These four techniques have been chosen for their significant influence on the construction of the vocal sound and its emotional reception within contemporary pop music.
I demonstrate this by including examples of released works.[28] In addition to tracing key techniques in contemporary vocal production, this chapter adds to the conceptual framework for the case studies that follow (Chapter 4), establishing a shared vocabulary through which I will examine the production choices in my own original songs.

---

[28] Unless otherwise indicated, the interpretations of song lyrics and vocal production techniques discussed in relation to released works are my own. This is largely due to the limited availability of verified information regarding the creative intentions of artists and producers, which are often treated as proprietary knowledge or closely guarded trade secrets. While resources such as the YouTube Channel "Genius"'s series *Deconstructed* (example) and podcasts like *Song Exploder* are increasingly popular and offer valuable insights into artistic processes—for instance, the *Song Exploder* episode featuring Phoebe Bridgers (listen here)—obtaining reliable and comprehensive accounts through research remains a difficult and often frustrating task.

2.2.1 Close-Mic'ing[29], Compression and EQ as Tools for Hyper-Intimacy

The technique of *close-mic'ing* (placing a microphone in immediate proximity to the singer's mouth) has long been used to foreground vocal nuance and intimate detail. While it plays a crucial role in the sonic construction of emotional proximity today, its origins can be traced back to the 1920s, when the advent of electric microphones gave rise to the *crooning*[30] vocal style. Freed from the need to project their voices across a live ensemble, singers could now perform in hushed, conversational tones. The resulting sound, shaped by the *proximity effect* (the increase in low frequencies captured by directional microphones at close range), introduced an unprecedented warmth and closeness to the singer's voice.

Paralinguistic Sounds

Close-mic'ing enables (or risks) the capture of *paralinguistic effects* — lip smacks, tongue noise, breath, *vocal fry*[31]. These non-linguistic sounds, if used subtly, can contribute to an increased sense of intimacy, while used overtly, can create a sense of unease, particularly when experienced through headphones, where sound is perceived to originate *inside* the head — a psychoacoustic phenomenon known as *"in head localisation (IHL)"*. As Görne (2015) explains, when using headphones, spatial cues from the outer ear and room reflections are bypassed, leading to a lack of externalisation: the auditory image collapses into the space between the ears, giving the sensation that the voice is positioned *within* the listener's skull (p. 131).

---

[29] I prefer the spelling of 'close-mic'ing'—as a contraction of 'close-mic(rophon)ing' — rather than the more common variations 'close-miking' or even 'close-micing'. While the latter is frequently used, it risks unintended associations with rodents rather than recording techniques. That being said, it is admittedly quite fitting: close-mic'ing would indeed be an excellent method for capturing the rather quiet sound of a mouse.

[30] Famous figures from the crooning era include: Bing Crosby, Rudy Vallee, Fred Astair, Frank Sinatra, Nat King Cole and Andy Williams. The term 'crooner' was also applied to female singers, especially those with lower alto voices, for example Ruth Eating and Annette Hanshaw and Nina Simone.

[31] Vocal fry, also referred to as "creaky voice" is the "bubbling of air through closed vocal folds" at a slow enough rate, that "individual impulses are heard as separate ticks" (Heidemann, 2016, as cited in Malawey, 2020, p. 52)

This sonic staging contributes to what has become a hyper-intimate aesthetic in contemporary pop music. The rise of ASMR[32] culture — emerging around 2010 — has amplified this trend, drawing attention to the affective potential of whispering, breath, and tactile noise. Accornero (2022) argues that such effects are particularly potent because they invade the listener's *intimate space*, referencing Edward T. Hall's (1969) proxemic theory, in which the intimate zone (0–45 cm) is typically reserved for close personal relationships. In today's individualised, headphone-based listening culture, soft vocalisms and low-level detail therefore carry greater emotional impact.

The amplification of paralinguistic sound is often intensified by the use of *compression[33],* a standard audio processing technique that reduces the dynamic range of a signal by lowering the level of louder sounds and raising quieter ones. In pop music production, compression is widely used on vocals to ensure a consistent and present vocal level that sits clearly above the instrumental backing. However, this dynamic flattening also amplifies low-level sonic details such as breaths, tongue noise, or lip smacks, rendering them more audible than they would be in a natural listening environment. As a result, when these details are *not* part of the intended aesthetic, they must be painstakingly edited out, as mentioned in this sub-chapter's introduction — a tedious process that highlights the double-edged nature of the close-mic'ing technique.

Breathy Vocals

In addition to capturing paralinguistic sounds, close-mic'ing also enables the recording of breathy vocal timbres. *Breathiness (*a quality often associated with vulnerability, femininity, or

---

32 Short for "Autonomous Sensory Meridian Response", where whispering, tapping and slow hand movements trigger relaxing "brain tingles" experienced by some people as "shivers emanating from the back of the head and down the spine [inducing] a state of deep relaxation." (Holmes, 2023)

33 Tagg, 2012, p.312: "COMPRESSION basically makes loud sounds weaker and weak sounds louder, thereby compressing the audio signal's dynamic range. An audio track can be compressed to make it sound fuller and 'tighter' so that it stands out from other input sources. Compression is also often applied to the complete mix, to an entire song or album, even to the entire output of a radio station. Overall compression is useful if the music is to be heard in spaces containing a lot of extramusical sound, for example when driving a vehicle."
In other words, COMPRESSION decreases the dynamics of a track. So *perceived* dynamics are often achieved by the alteration of vocal performance intensity and timbre (is the singer whispering or shouting?) and its placement in the mix.

sensuality) results from a combination of airflow and low vocal cord tension resulting in partial vocal fold closure, creating a distinct airy texture (cf. Malawey, 2022, p.109). This quality is characterised by a dominant fundamental frequency and a relative absence of higher harmonics. In place of these overtones, we hear the sound of airflow, giving the vocal a softer, more diffuse texture. *Large-diaphragm condenser microphones*, most commonly used in studio vocal recording, naturally accentuate this effect due to their heightened sensitivity in the higher frequency region. Their ability to capture low-level detail makes breathiness more present and texturally rich. *Compression* further reinforces this by amplifying the quieter breath components and smoothing out dynamic variation, ensuring that the breathy quality remains consistent and audible throughout the performance. *Equalisation (EQ)* may be used to enhance this aesthetic even further — for instance, by boosting high frequencies to enhance breath noise or subtly cutting midrange frequencies to soften the vocal 'body'. Breathy vocals are particularly evident in genres like *bedroom pop* and *whisper pop*, which embrace low-dynamic and low-intensity, close-mic'ed vocal performances as core stylistic markers. These genres, emerging from prosumer DIY context, often rely on minimal production and inexpensive home recording setups where the close-mic'ed vocal is not only a practical result of spatial and technical limitations, but a deliberate aesthetic choice, reinforcing intimacy and introspection.

Together, close-mic'ing, EQ and compression form a key component of the hyper-intimate vocal aesthetic in much of today's pop music. In the following section I provide a few examples.

Examples of *Paralinguistic Effects*

Adele's *Love in The Dark (2015)* — vocal fry, e.g. from 0.20 "*Take your-eyes off of me, so I can leave / I'm far too ashamed*"

Billie Eilish *bad guy* (2019) — breath, e.g. 1.18 and 1.32

Billie Eilish i *love you (2019)* — mouth clicks, breath and vocal fry throughout

FKA Twigs *cellophane (2019)* — breath

King Princess *Winter is Hopeful* (2022)[34] — vocal fry, e.g. 0.14 *"you worry 'bout girls that I know"*

Olivia Rodrigo *drivers license* (2021) — vocal fry, e.g. 0.47 *"how could I ever leave someone else"*

Dodie *She* (2019) — raspiness at 0.39 and mouth clicks from 1.10

Examples of *Breathy Vocals*

A compelling case is Billie Eilish's i *love you* (2019), in which her voice is but a gentle whisper. Each vocal detail (breath, consonants, sibilance) becomes part of the track's emotional architecture. A similarly skilled, though more subtle, use of breathy vocals appears in Maro's *Saudade*[35] (2022), where her voice feels fragile and airy, underscoring the song's melancholic tone. Other examples of current artists that use this technique include: Lizzie McAlpine *reckless driving* (2022) — chorus from 0.44; Dodie *She* (2019); Freya Ridings *Lost Without You* (2019) — as heard in the first verse; Matilda Mann *Glass Ceiling* (2021); King Princess *Cursed* (2022); Gracie Abrams *Where Do We Go Now* (2023).

While many artists use vocal layering or heavier processing (discussed in the following subchapters) to help a breathy vocal cut through a dense or busy mix, Lucy Rose's *Shiver* (2012) provides a striking example of how a single breathy lead vocal can occupy the foreground of the mix without heavy processing. The track demonstrates how vocal intimacy can also be obtained by sparse instrumentation and deliberate microphone technique.

2.2.2 Double-Tracking and Layering as Tools for Emotional Complexity

Multitracking is the overarching term used to describe the process of "playing more than one […] track" simultaneously (Malawey, 2020, p. 131). *Live-multitracking* involves recording multiple instruments and/or vocal tracks simultaneously. Naturally, a lead vocalist who wants

---

[34] This track and artist in general can be considered representative of a contemporary singer-songwriter who self-produces and uses of the full array of digital production tools (close-mic'ing, echo, vocoder, distortion, reverb)

[35] *From the Oxford Languages Dictionary*: Saudade, n. (especially with reference to songs or poetry) a feeling of longing, melancholy, or nostalgia that is supposedly characteristic of the Portuguese or Brazilian temperament.

to multitrack their vocal (whether in unison, harmony or counterpoint) has to record individual takes one after the other over the existing track(s), which is called *overdubbing*.

Double-Tracking

Double-Tracking, also referred to as Doubling or Dubbing (Tagg, 2012), refers to the recording and layering of two (almost) identical vocal performances in unison, typically by the same singer, to enrich vocal texture and create a fuller sound. According to Tagg, the effect "can be used to flesh out a thin voice or to make a single voice sound like two or more of the same vocal persona" (p. 311).

While digital plugins can simulate this effect through detuning and delay[36], many contemporary artists continue to record additional takes manually — a technique that allows for subtle expressive differences to emerge. As Lacasse (2000) explains, we can distinguish between two primary modes of double-tracking in stereo production:

(1) *mono double-tracking*, where both vocal takes are panned to the same point in the stereo field (most commonly in the centre) to thicken the vocal sound, and
(2) *bilateral double-tracking*, where they are panned left and right to produce a sense of width and spatial separation (p. 129).

Mono double-tracking usually results in a *phasing effect*[37], caused by "micro-variations" in timing, pitch, and articulation between the two takes. These slight inconsistencies, while unintentional in early uses, are now valued for their aesthetic. The phasing effect can create a subtle "shimmering" or instability in the vocal sound that can be harnessed expressively, for example signifying nostalgia and emotional ambiguity.

Example: José Gonzales' *Heartbeats* (2003)

---

[36] see 2.2.3

[37] Phasing effects result from micro-variations between two sound sources with nearly identical timbre and amplitude. These variations cause patterns of *constructive* and *destructive interference*, where overlapping waveforms either amplify or cancel out certain frequencies. This leads to characteristic fluctuations in timbre and perceived spatial depth.

Bilateral double-tracking, on the other hand, emphasises spatialisation: when panned left and right, the voice appears to occupy a broader field, producing a "bigger" sound without increasing the volume (Lacasse, 2000, p. 130). As Philip Tagg (2012) similarly notes, doubling can "make audio input sound bigger […] and create the impression that there is 'more of the sound occupying more space', especially if the original and dubbed tracks are assigned different positions on the aural stage "(p.311).

Example: Elliot Smith's *Between The Bars* (1997)

Double-Tracking has a long history in popular music, from early experiments by Les Paul[38], Buddy Holly and Phil Spector to its famous implementation by The Beatles with the invention of *Automated Double-Tracking (ADT)* at Abbey Road Studios. As a result of John Lennon wanting his vocal to sound thicker and cut through the mix, engineer Ken Townsend invented the ADT effect in 1966, using a second tape machine to apply a delayed version of the lead vocal with a slightly varying offset (typically 24–30 milliseconds), effectively creating a form of automatic bilateral double-tracking. As Lacasse notes, ADT became a hallmark of *Revolver* (1966), appearing on thirteen of its fourteen tracks (p. 131).

Tagg (2012) notes that manual doubling is especially powerful when a radically different tone is used in the second take (for instance, when a singer whispers a line that was previously sung) so the listener hears the same words "both out loud and inside their heads." (p. 311). Here, double-tracking works in close tandem with close-mic'ing, a technique I presented in the previous subchapter (2.2.1).

In today's popular music landscape, double-tracking is used not only for a fuller sound but as a narrative tool to communicate emotional complexity, vulnerability, and layered identity, as shown through the following short analyses:

Analysis Example 1:

A key example is Phoebe Bridgers, who frequently uses (manual) double-tracking to create a diffused, ghostly presence in songs like *Motion Sickness* (2018) and *Scott Street* (2018). Her

[38] "[T]ape multi-tracking—[…] developed by Les Paul around 1949 […] [paved] the way to a number of multi-tracking techniques, including double-tracking" (Lacasse, 2000, p. 128).

doubled vocals suggest a kind of internal diffusion, nostalgia and melancholy, making the effect not just spatial but narrative: her voice sounds both present and distant, unified yet layered.

Analysis Example 2:

Packing it Up (2024) by Gracie Abrams employs a fairly obvious use of mono-double tracking. The two lead vocals appear to be similar in level, making the timing differences obvious and slightly distracting, especially on the sibilances and plosives.

As opposed to using the effect uniformly across a track, modern producers and artists also often make use of double-tracking selectively on specific phrases, choruses, or emotional peaks. This micro-level control, combined with stereo placement, has transformed double-tracking from a tool of clarity and presence into one of nuance and expressivity. What once served to thicken or polish a vocal line now operates as a narrative device, allowing artists to stage *multiplicity*, fragility, or even emotional dissociation.

Analysis Example 3:

An example of multiplicity, is Gracie Abram's *That's so True (*2024) where she uses a doubled vocal from the first verse, expanding the effect through panning and layering in the chorus. The second verse then features a double with a mocking character on the line *"she's so fun - wait I think I hate her'"* (from 0:57), enforcing the theme of jealousy that unfolds throughout the track despite attempts of the narrator to temper herself.

Layering and Harmonies

Layering, also referred to as *stacking*, involves recording and combining multiple vocal takes (whether lead vocals or harmonies) and is often enhanced through panning these across the stereo image. While double-tracking typically uses two nearly identical vocal performances

to reinforce a part, layering extends this concept further by adding three or more layers, which may differ in tone, phrasing, or harmony[39].

Layering is a particularly effective way of emphasising key phrases and words, as well as adding colour, texture and complexity to a song's narrative and has become a go-to technique in contemporary music production. It is common to double each harmony in order to create a balanced stereo image — sometimes adding multiple harmony layers and panning them at different points across the stereo field to fill out the space. Paralinguistic effects (cf. 2.2.1) become particularly pronounced when vocals are panned to the far sides of the stereo field. This is because the spatial separation isolates the vocals from the mix's central elements, making details such as plosives (rich in lower frequencies) and sibilants (rich in higher frequencies) more audible, as they are less likely to be masked by other instruments on the outer edges of the stereo image. For this reason, backing harmonies often leave out the end consonants of a word.

Analysis Example 4

Paris Paloma's *labour* (2023), particularly from 3:28 onwards, layers contrasting lyrics, overlapping phrasings, melodies, and harmonies to create a dense, textured sound. This deliberate 'busyness' sonically mirrors the song's lyrical focus on women's emotional labour and mental overload, effectively embodying the chaos and weight of these often-invisible burdens:

"All day, every day, therapist, mother, maid
Nymph, then virgin, nurse, then a servant
Just an appendage, live to attend him
So that he never lifts a finger

[39] Different types of harmonies can be categorised as: *parallel, counterpoint, pedal point, open* and *close*. While *parallel* harmony (shifting all notes up or down by the same amount within the scale, such as a 3rd) is often used to encourage the listener to sing along, the *counterpoint* approach, consisting of two completely different melodies, creates interesting textures, while the *pedal point* harmony can create tension *"because the root note never moves at all"*. The *open* harmony involves using wider intervals between the root note and the harmony, giving the vocal a much more open and expansive feel. The opposite, the *close* melody creates a rich, lush sound (Songs by Spencer, 2024).

24/7 baby machine

So he can live out his picket-fence dreams

It's not an act of love if you make her

You make me do too much labour

The capillaries in my eyes (all day, every day)

Are bursting (therapist, mother, maid)

If our love died (nymph, then virgin)

Would that be the worst thing? (Nurse, then a servant)

For somebody (just an appendage)

I thought was my saviour (live to attend him)

You sure make me do (so that)

A whole lot of labour (he never lifts a finger)"

Gang Vocals and Choirs

What once was limited by analog track counts (as in *Blue* by Joni Mitchell, recorded on a 4-track recorder) is now constrained only by the producer's vision and their computer's CPU[40]. A particularly expansive form of vocal layering is found in *gang vocals* and *choir-style arrangements,* both of which have become increasingly common in contemporary music production. These techniques involve stacking numerous vocal tracks — often performed by the lead singer, backing vocalists, or even a group of non-singers — to create the illusion of a crowd or communal voice. With *gang vocals*, the goal is typically raw energy and emotional immediacy: timing and tuning are deliberately imperfect, producing a sense of spontaneity and shared 'in the moment' emotion.

Examples:

Lumineers *Ho Hey* (2012) - e.g. intro

Mitski *Bug Like an Angel* (2023), e.g. at 0.50 "*family*" and *1.42 "they break you right back, break you right back, uh'"* - an interesting use because the sound is more choir-like, but used very selectively, e.g on just one single world (like gang vocals often)

---

[40] However, more isn't necessarily always better; Artists like Adele often (are able to) rely on a single, lead vocal to carry the full weight of a song. Her voice cuts through the mix beautifully through vocal performance, mic technique, arrangement, and dynamic control. Example: Adele *Easy on Me* (2021)

By contrast, *choir vocals* aim for a more unified, harmonic texture — ranging from lush gospel-inspired arrangements to ethereal soundscapes:

Examples:

Adele *Love in The Dark (2015)* chorus choir e.g. from 1.11 *"I can't love you in the dark"*

Cat Burns *People Pleaser* (2024) from 0.51 *"pleaser"* x3, and more from 1.53)

Alex Warren *Ordinary* (2025) e.g. *0.52 "Ordinary […] Sanctuary" and 1.35 "Ordinary, I want you layin' me down till we're dead and buried".*

Harmonisers and Octavers

The widespread use of DAWs has made these layered group effects easier than ever to produce. With virtually unlimited track counts and access to tools like pitch correction, stereo imaging, and modulation plugins, producers can simulate a full choir or crowd from just a few voices. Some *harmoniser* and *octaver plugins* further extend these possibilities, generating additional voices "in real time" and allowing for stereo spread, pitch shifting, or timbral manipulation — essentially automating what once required multiple takes and performers.

Example: Ed Sheeran *Afterglow* (2020) uses a harmoniser[41] effect all the way through.

In sum, double-tracking and layering in contemporary pop is not merely a technique for thickening and enhancing vocal sound; it is a symbolic tool for constructing emotional depth and spatial presence. One that draws on decades of recording history but has been reimagined to suit the expressive demands of today's contemporary music culture.

[41] Why is this not a vocoder? Because it sounds more natural, and doesn't sound as if it is re-synthesised through a different instrument. From observing Sheeran's live performances, he is able to recreate this effect live by sending his lead vocal through a harmoniser pedal.

2.2.3 Reverb and Delay as Spatial, Temporal and Emotional Staging Devices

Reverb and *delay* are among the most influential tools in vocal production, shaping both the spatial and emotional character of the recorded voice.

In acoustic terms, reverb is the persistence of sound following the impulse of a sound source, caused by the reflection of the source's sound waves off surfaces in its environment. When we hear a sound in a space, we receive it in three stages: the *direct sound,* which travels straight from the source to our ears; the *early reflections*, which bounce off nearby surfaces and arrive shortly after; and the *late reflections*, which diffuse and blend into a continuous decay. These spatial cues help us identify the size, shape, and material of a room —something the human brain is highly attuned to (cf. Görne, 2017).

While *delay*, in short, refers to the repetition of a sound, *echo* occurs when the delay time between the original sound and its repetition exceeds approximately 50 milliseconds (Lacasse, 2000, p. 116). When the gap becomes even longer and the repetitions are perceived as independent from the original sound, Lacasse refers to this as *reiteration* (p. 209). While an *echo* involves gradual decaying repetitions of the original phrase, *reiteration* involves the deliberate sampling and rhythmic redeployment of vocal material, creating a sense of disembodiment (cf. 1.2.2.3).

The staging of the voice through reverb and echo is not a modern phenomenon. Lacasse maps out how the origins of vocal staging can be traced back to prehistoric times. Quoting Murray Schafer (1977, as cited in Lacasse), he writes "primitive man was fascinated by the special acoustic properties of the caves he inhabited. [...] One imagines sacred rites being performed in these dark reverberant spaces" (p. 32). Lacasse refers to such practices not only as early examples of acoustic manipulation, but as intentional acts of sonic enhancement, where reverberant environments amplified the presence, reach and affective power of the human voice. In some cases, ancient cave drawings have been found at acoustically resonant points in cave systems, suggesting a deliberate alignment between voice, sound, and ritual (p. 33).

Similarly, in antiquity, vocal staging played a vital role in the design of Greek and Roman amphitheatres, where architectural features (such as curved walls and stone tiers) were crafted to project the voice across large outdoor spaces. Though microphones did not yet exist, these spaces acted as natural amplifiers, reinforcing vocal clarity and spatial presence. The voice, even then, was a staged phenomenon: shaped by and embedded within an acoustical environment to enhance its communicative impact.

Fast forward to the 20th century and reverb and echo, once products of physical spaces, became artificial tools; first used to approximate the acoustics of real-world performance environments, and soon reimagined as creative devices in their own right.[42]

William Moylan (1992, p.50) provides a helpful framework for understanding the broader function of reverb applied in music production today:

> "The coupling of a sound source—in our case, voice—with a given environment may give rise to a number of effects:
>
> - Have a significant impact on the meaning of the music, text, or sound source;
> - Supply dramatic effect;
> - Segregate sound sources, musical ideas, or groups of instruments; and
> - Enhance the function and effectiveness of a musical idea."

To better understand how these spatial choices operate aesthetically, Lacasse draws on Moylan's (1992) model of recording aesthetics. This framework maps a continuum between two extremes: on one end lies the *natural-realistic aesthetic*, which aims to faithfully capture a live performance 'as is' in its natural environment. This approach is typically associated with classical recordings which are usually captured in concert halls or churches, allowing for a coherent natural reverberant sound. On the other end of the spectrum is the *full-technological aesthetic,* common in electronic genres where sounds are synthesised,

---

[42] For instance, slapback echo (used extensively in Elvis Presley's Sun Studio recordings) gave his vocals a sense of distance, loneliness, and drama. This sonic treatment was not merely technical, but deeply aesthetic: "a stark, lonesome and incisive" identity emerged through the echo's spatial cues (Lacasse, p. 123).

layered, and spatially manipulated in ways that bear little resemblance to natural environments.

Most of today's popular music occupies the space between these poles, a zone Lacasse describes as the *artistic-technological aesthetic* (p. 113). The standard practice in contemporary pop music is to record vocals with as few reflections as possible. This *dry* sound (a reverberant sound is described as *wet)*, or what Lacasse refers to as a *flat voice* gives producers maximum control over aesthetic choices in post-production. The dry voice can then either be placed into a more natural sounding environment using artificial reverb, "as an extension of the *natural-realistic* approach" (Moylan, 1992); or, the vocal can be placed in a more unnatural sounding environment, leaning more into the full-technological side of the scale. This interplay between naturalism and artificiality resonates with Moorefield's (2010) concept of the "illusion of reality" and the "reality of illusion" in music production—terms that highlight the producer's ability to either mimic natural acoustics or to intentionally construct an imagined, surreal sonic environment.

Mike Senior (2011) provides a more technical lens on this practice by identifying five core functions of reverb in a mix: *Blend, Size, Tone, Sustain, and Spread* (p. 231). These categories describe how different parameters of a reverb can be shaped to influence how a vocal sits in the mix and is perceived by the listener. For example, "the biggest difference between reverbs designed for blend and size is that where the former is best provided by the earlier reverb reflections, the latter is best created by focusing the effect sound on the remainder of the reverb tail" (p. 242). These distinctions show how sonic space can be carefully constructed to serve artistic intention, not just technical polish.

In contemporary pop music, reverb is often used to create a sense of closeness or vulnerability. Short, soft reverbs with low diffusion can make a vocal feel intimate without overwhelming it. As Anderton (2018) puts it, low diffusion "blends in with the vocals rather than sounding like a separate effect". This kind of subtle, transparent reverb is common in genres like *bedroom pop* or *whisper pop*, where intimacy is a key aesthetic.

Lacasse further points out that, beyond its link to a perceived authenticity of live performance, listeners tend to connect reverberant sound with certain emotional or contextual qualities, for example evoking religious, sacred, or spiritual connotations (Lacasse, 2000, p. 51). Additionally, he observes that reverb can evoke temporal connotations, such as a sense of the past or present (2000, p. 161).

Another common use of reverb in contemporary vocal production is its reversal. *Reverse reverb sweeps* are often placed just before a word or phrase to create suspense and highlight emotional or lyrical significance. Due to the fact that the reversal of reverb can not exist in a natural space, the swelling effect can also signal a break from reality or an entry into memory or an internal world.

Beyond emphasising a word simply by repeating it, the *echo* effect also plays an important emotive role, and its use is often associated spatial distance, or nostalgia (Lacasse, p. 19). Producer Rick Beato[43] points out: In order for a delay to be effective, there has to be space for the vocal to "breathe", meaning there has to be space in the arrangement after the source material for the delay to be audible and make an impact.

In recent pop production, the mentioned effects taken together are often applied with extreme subtlety and precision, facilitated by automation techniques that allow for millisecond-level control of parameters like volume, wet/dry balance, and stereo imaging. This kind of "surgical" precision in production is especially evident in the work of artists like Billie Eilish (e.g., w*hen the party's over* (2019) or e*verything i wanted* (2019)), where delay and reverb are often automated in and out to create moments of hyper-intimacy or emotional dissociation.

Analysis Examples:

Freya Ridings' *Lost Without You* (2019) exemplifies the expressive use of a longer tailed reverb, enveloping her voice in a vast ambient field that amplifies the song's sense of emotional absence and longing. e.g. "*Let you go*" at 2.04; and "*I think I'm lost*" from 2.17.

---

[43] Rick Beato. (2017, March 2). *Vocal mixing for pros - using EQ, compression and FX | Featuring Michael Johns* [Video]. YouTube. https://www.youtube.com/watch?v=G_PIGop64Gs

In contrast, Lorde's *Liability* (2017) employs reverb with restraint. In the line "*until you're gone*" (1.46), a delicate reverb is added only on the last word "*gone*", as if shouting into a void—capturing a sense of loneliness. In the second chorus (2:08), a more subtle reverb is added beneath her voice, reflecting a quiet emotional opening rather than melodrama. This nuanced spatial shift helps underscore the track's vulnerability without overwhelming its rawness.

Sia's *Unstoppable* (2016) uses a huge reverb on the backing vocals of the choruses ("*I'm unstoppable, […] I'm so confident, Yeah I'm unstoppable today"*). The impact is best heard when listening from 0.38, as a contrast to the final line of the verse which is flatter. Here, the reverb employed in the chorus gives the impression of power and authority, taking over a large volume of space, turning the backing vocals into more of a powerful choir—supporting the song's lyrical message.

Dua Lipa's *Pretty Please* (2020) showcases a more stylised, rhythmic application of delay—specifically, tempo-synced vocal delays in the chorus that create bounce and momentum. At 1:48 and 2.06, a reverse reverb-like effect after the phrase "*sweet relief*" adds drama by 'pulling' the listener toward the next word "*pretty*". The track also features *reiterations*—cut-up, re-sampled fragments of the vocal woven into the beat (e.g. 2.12 and 2.21 and "*pretty").*

It is important to note that the emotional meaning of these effects always emerges in dialogue with the larger arrangement and above all, lyrical context. In *Pretty Please,* the lyric "t*rickle down my spine*" is preceded by a literal 'trickling' synth pattern at 1.51, emulating the sensation. While this thesis focuses on the voice, it recognises that narrative meaning and intent cannot be taken out of context from its surrounding musical landscape, but should rather be examined in relation to it.

Summary

In summary, reverb and delay are powerful tools for shaping the emotional and spatial perception of the voice. Reverb can enhance clarity, depth, and presence, while also invoking a sense of liveness or vulnerability. Used more heavily, it can carry sacred or

spiritual undertones. Delay, with its distinct repetitions, often evokes memory, reflection, or temporal disorientation. Both effects can also subtly communicate solitude, loneliness or emotional distance. Additionally, the *lack* of reverb, as Altman (1992, as cited in Lacasse) describes in his work on film sound, and presence of direct sound (dry sound) give the impression that the voice is speaking a message directly to the listener alone, creating a sense of *"for-me-ness"* (p.195).

Examples:

Charlie Puth *Attention* (2018) — works with the drastic reduction of reverb to focus the listener, fittingly on the line *"You just want attention"* at 0.46 — very effective when heard in context with the previous reverb-laden chorus section (0.29 - 0.46).

Alex Warren *Ordinary* (2025) — at *"I take one look at you"* (00:32), the dry vocal is very impactful.

Gracie Abrams *I miss you, I'm sorry* (2020) —*"I miss you - I'm sorry"* at 1.30 makes use of the same effect.

Before diving into heavier signal processing tools in the next chapter, I would like to present an example of contemporary vocal music production that showcases many of the points discussed so far: Lizzie McAlpine's *ceilings* (2022) combines close-mic'd fragile vocals with a gentle reverb from the start. During the first verse her plosives (*'**p**laster, '**c**an't', 'fas**t**er'* etc) are very audible in the mix. She also uses 'spot-layers' throughout the first verse, underscoring some of the words such as *"raining", "harder", "water"*. Her lead vocal overall is breathy and its presence is most likely boosted by a high shelf EQ, perceptible in the first chorus (from 0.59). The singular lead vocal comes across here as particularly intimate, as you would expect the performance intensity to increase along with the pitch, but it doesn't, rather the breathiness seems increase in the chorus. Verse 2 from 1:28 on features layered vocals (lower harmonies placed behind the lead vocal in the center of the mix) all the way through. From 2:07 (*"But it's over"* etc.) her vocals are doubled with whispered layers left and right, elevated in the high end - panning them wide helps elevate their presence over the instrumental buildup and adds intensity. It is noteworthy, that the buildup featuring new textures is introduced by the word *"chaos"* (2.03). Finally, the last word of the song — *"before"* — has an echo on it, trailing off towards the end of the song, evoking sentiments of the past.

2.2.4 Pitch-Modulation, Vocoder and Distortion as Transformative Tools

Whereas tools such as doubling and reverb tend to enhance the natural singing voice, other vocal processing techniques can alter the signal more dramatically, sometimes to the point of abstraction. These more radical forms of signal processing, including pitch-shifting, the vocoder, and distortion, shape not just the sonic texture but the very *identity* of the voice.

Middleton (1990, p. 262) describes the voice as "the profoundest mark of the human". He further underscores how alterations to the voice cut closer to the listener's emotional perception than the treatment of instruments:

> "[...] vocalising is the most intimate, flexible and complex mode of articulation of the body, and also is closely connected with the breath (continuity of life; periodicity of organic processes). Significantly, technological distortion of voice-sound (through use of a vocoder, for example) is far more disturbing than similar treatment of instrumental playing (which is regarded usually as a logical extension of manual performance)."

In this light, heavy signal processing can be understood not merely as an aesthetic choice, but as a transformative and emotive tool which can rob a voice of its humanness, or add unsettling, expressive, ironic or even intimate qualities, depending on the context and intention behind its use.

Digital Pitch-Modulation: From Pitch Correction to Stylistic Marker

Originally designed as a pitch correction tool, *Auto-Tune* by Antares (released in 1997) was intended to subtly fix off-pitch notes without noticeable artefacts. However, within just a year, it was famously repurposed as a stylistic effect: Cher's 1998 hit *Believe* marked one of the first mainstream uses of Auto-Tune in its exaggerated, robotic form. This sound (often referred to as the "*Cher effect*"[44]), achieved by turning the effect's correctional speed down to zero, quickly permeated pop and hip-hop music, becoming a hallmark of digital-era vocal

[44] In the Podcast "Switched on Pop: Auto-Tune always and forever", Charlie Harding points out that despite the effect being named after Cher, Kid Rock had used Auto-Tune in a similar way on *Only God Knows Why* (1998), released a few moths before Cher's *Believe*.

aesthetics. **Examples:** T-Pain *Buy U a Drank* (2007); Kanye West *Heartless* (2008) and *Love Lockdown* (2008)

Whereas Cher's *Believe* was received as an innovation and intriguing vocal effect, Maroon 5's release of *She Will Be Loved* (2003) a few years later set off a wave of criticism, with the broad view that pitch correction tools were stripping the music of its artistry and authenticity (Harding, 2024). Olsson (2016) points out, however, that this backlash often reveals a bias: "Is the hate only towards people who are put on the spot as being 'unable to sing'? Looking at the majority of articles and blogs the evidence points toward just that" (p. 16).

According to Charlie Harding of the *Switched on Pop* podcast, around 90% of pop vocals today undergo some form of pitch correction, underscoring how embedded these tools have become in contemporary vocal production. Today, "Auto-Tune" functions as an umbrella term encompassing a range of pitch-modulation tools, including *Melodyne, Logic's Flex Pitch,* and many more. These tools are used both subtly to gently smooth imperfections, and overtly, to create hyperreal, stylised textures.

As Brøvig-Hanssen and Danielsen (2016) argue, Auto-Tune's staying power lies in its dual function: "both its ability to perfect the pitch of the human voice and [its ability] to add genuinely new sounds to the human musical repertoire" (p. 132). The effects of this aesthetic are far-reaching. In their words, such tools can "block the access to the 'real' self of the singer," allowing for a transgression of traditional notions of identity—including "man and woman, black and white, culture and nature"—and enable new expressions of emotional states such as "distance, numbness, and the absence of presence in one's life" (p. 132). The result, they argue, is that Auto-Tune has become nothing less than a "signature sound of contemporary pop music."

This influence spans genres and artist types. Even artists rooted in acoustic or singer-songwriter traditions have embraced them. Ed Sheeran, for example, whose career began with minimalistic, loop-based live arrangements, has increasingly adopted vocal effects and Auto-Tune in his more recent work. On tracks like *Bad Habits* (2021), his vocals are subtly

processed and stylised to align with an electronic pop aesthetic. The effect is likely perceived as transparent by most listeners, with gentle pitch correction and added reverb creating a smooth vocal sheen that blends seamlessly into the track's synthetic production, maintaining vocal authenticity while adapting to contemporary standards. In contrast, in Tate McRae's *greedy* (2023), Auto-Tune is applied more noticeably, especially on the lead vocal in the chorus, where it lends the voice a glassy, synthetic edge. Yet even here, the processing doesn't obscure McRae's vocal identity—instead, it becomes part of the track's expressive signature. These examples illustrate how pitch correction has become a standard and flexible tool, shaping genre aesthetics without necessarily compromising individuality or emotional resonance.

Bon Iver's *Woods* (2009)—a haunting a cappella track that stands in stark contrast to the otherwise stripped-down, acoustic-folk aesthetic of the *Blood Bank* EP—demonstrates how vocal processing techniques like pitch correction and vocoder (the latter discussed in the next section) can radically transform the perceived authenticity, gender, or even humanity of a voice. The song consists of a single stanza repeated eleven times ("*I'm up in the woods, I'm down on my mind / I'm building a still to slow down the time")* creating a meditative loop that echoes emotional overload and the disorientation of isolation. Written during a period of self-imposed solitude in a cabin in northern Wisconsin, the track reflects Justin Vernon's inner turmoil and introspection. The overt processing of the vocals, while never disguising Vernon's identity, adds a layer of emotional distance and ambiguity: the voice sounds both human and otherworldly, intimate yet estranged. This layered, processed vocal becomes a kind of performative persona—a vocal "costume" that blurs the line between raw feeling and stylised expression. Rather than diminishing the song's emotional weight, the effects heighten it, enveloping the listener in a sonic atmosphere of claustrophobia and vulnerability. In this way, *Woods* becomes a compelling example of how Auto-Tune and vocal manipulation can be used not just for pitch correction, but as expressive tools that deepen narrative and emotional resonance.

Before digital pitch-shifting existed, producers used *varispeed* on tape machines to alter pitch and formants organically. By speeding up the tape recording, the vocal formants would be shifted upwards, resulting in higher, squeakier, younger-sounding, or even

chipmunk-like voice. Or vice versa: slowed down, the vocal formants would be lower, resulting in a deeper, darker, slightly ghostly or menacing sound.
Today, DAWs like Logic Pro X are able to emulate this process digitally. The effects described above can be achieved by slowing down project tempo, recording a lower vocal line, and then returning to the normal project speed; or likewise, recoding at a higher pitch and tempo and then returning to the original project speed.
Examples:
Jacob Collier *Time Alone With You* (2019) — e.g. from 0.42 *"It's a beautiful Sunday morning, I see the sun up in the sky…It's gonna shine for you"*
The Beatles *Lucy In the Sky With Diamonds* (1967) — from first verse.

A popular effect in contemporary pop vocal production is the overt use of pitch-modulation on short phrases or even words. Overtly lowered and darkened vocals can appear to be "monster-like". For instance, Billie Eilish's *bury a friend* (2019) features overtly darkened vocals between 1:12–1:26 ("*what do you want from me… where do we go*"), creating a monster-like quality that enhances the track's eerie, dissociative mood. More subtle examples appear in Dermot Kennedy's *Glory* (2017), where brief pitch manipulations on lines like "*and a hopeful rhythm woke within him / had some letters written, 'course she's in 'em / tried to tell her, throat was linen*" (0.48-0.50; 0.54-0.56; 1.01-1.03) add texture and urgency.

Due to gendered coding and connotations of voice in relation to the fundamental frequency (and therefore perceived pitch) (cf. Malawey 2020, pp. 59, 61), pitch shifting can also change the perceived gender of a voice, suggest ambiguity, or subvert its meaning entirely. In some social contexts, "individuals may alter [the perceived pitch of their voice] to achieve a desired social status". In this context, a lower pitch maybe seen to portray power and authority, whereas a higher voice might be adopted to show kindness or submission (p. 61).
Example: Sia *The Greatest* (2016) — e.g. at 00.23 and 00.28 *"I got stamina"*

As Brøvig-Hanssen and Danielsen (2016) note, such vocal manipulations often appear in genres like electropop and R&B, particularly among female artists, where extreme Auto-Tune is combined with imagery of "monstrous femininity and/or hyperembodiment"—

constructing a vocal body that is "cultivated beyond the human" (p. 138). These aesthetic choices challenge essentialist readings of gender, voice, and emotional authenticity.

The Vocoder[45]

"The vocoder manipulates frequencies in an audio signal to produce a non-human, robotic sort of sound" (Tagg, 2012, p.310).

The *vocoder*, short for "voice encoder," invented by Homer Dudley at Bell Labs in 1939, dates back to military communications technology, developed to compress, encode and re-synthesise the speaking voice in order to save bandwidth in transmission and secure messages against interception (Malawey, p. 132 f.). In musical contexts, the vocoder was first adopted by avant-garde and electronic artists in the 1970s and '80s, most famously by Kraftwerk and Laurie Anderson. Anderson's *O Superman* (1981) is a landmark example, using a vocoded voice to unsettling effect: breathy, intimate, yet alien, especially juxtaposed against the repetitive dry vocal samples and the ambient landscape featuring tweeting birds (audible for example from 1.23 - 1.46 and beyond).

One of the most prominent examples of the vocoder's use in the 2000s is Imogen Heap's *Hide and Seek* (2005); her solo voice, entirely processed through a vocoder, delivers stark emotionality within a disembodied, synthetic soundscape. The lack of instrumental accompaniment heightens the effect, giving the voice full emotive weight despite its digital mediation.

A more recent example of such overt use is Bon Iver, whose 2016 album *22, A Million* is an experimental fusion of folk songwriting and electronic manipulation. On tracks like *715 - CRΣΣKS*, Justin Vernon sings through a modified vocoder effect, bending harmonies and formants throughout. The result is a fractured but deeply emotive/touching/hypnotic vocal presence.

In King Princess' *If You Think This is Love* (2019) her lead vocal features a vocoder from the start. While the effect is obvious, the narrator is still perceived as being human and

---

[45] The *Vocoder* effect is sometimes confused with the *Talkbox* effect, which is featured on the lead vocal in the opening lines of Bruno Mars' *24K Magic* (2016).

importantly, as King Princess herself. Here the use of the vocoder provides a sense emotional depth and inner fragmentation, enforcing the song's narrative, which, according to *Genius*[46], depicts "the end of a relationship and the regret that comes with it."

See also Rose Betts *Sober* (2023), where the lead vocal is vocoded from the start. The harmonies give them a melancholy touch. The effect becomes particularly significant with the lyric *"But you never kiss me when you're sober"* (i.e. "only when you're drunk") in the chorus (from 0.39) as the effect really enforces the sentiment of feeling drunk, filtered, feeling removed from reality.

In recent years, the vocoder has become a standard tool used in contemporary pop music, for example for giving the lead vocal a subtle lift in choruses whilst providing more width (if steep panned as usual), or adding bass frequencies that aren't normally occupied by the vocals in a mix. As Zedd[47] points out, "the frequency spectrum you can reach with just the vocal is limited" (Beato, 2025), so bass vocoders can be used to give the voice more presence in the lower frequencies, ensuring more vocal dominance in the overall mix. The effect is fully accepted by listeners and often goes unnoticed if not paying close attention.

Examples:

King Princess *1950* (2018) — subtle vocoder from 0.29 "*So tell me why my gods [look like you] / And tell me why it's wrong*"

Taylor Swift *Vigilante Shit* (2022) — subtle vocoder from 0.30 "*I don't start shit but I can tell you how it [ends] - don't get sad, get even*"; from 0.53 "*She needed cold hard proof so I gave her some*" etc.

Zedd, Marren Morris, Grey *Middle* (2018) — e.g. 0.30 *"why don't you just meet me in the middle […] so why don't you just meet me in the middle".*

Bruno Mars *Treasure* (2012) — uses a vocoder on the final Chorus from minute 2.00 on the repetition of the lyrics "*you are my treasure*" to add colour and momentum. The vocoder has a subtle flanging effect on it as well in this case.

---

[46] https://genius.com/King-princess-if-you-think-its-love-lyrics

[47] Rick Beato. (2025, May 24). *Zedd: "Clarity" to "The Middle" unlocking his hit song production SECRETS* [Video]. YouTube. https://www.youtube.com/watch?v=UM2zJ9NNjho

Distortion: Sonic Tension and Overload

"Distortion effects (a.k.a. overdrive, saturation) radically alter the character of overtones in a sound's frequency spectrum to create timbres that have been variously described as rough, gritty, harsh, rich and full-bodied" (Tagg, 2012, p.309).

Distortion (meaning *harmonic distortion),* long associated with electric guitars and rock music, can also be applied to vocals to introduce qualities such as graininess, aggression, or emotional rawness. Harmonic distortion is caused by clipping due to signal overload, whereby the waveform is cut off at its peaks, producing additional harmonic content in the signal's frequency spectrum (cf. Görne, 2015, p. 230). Similarly, the human voice, when strained or "overloaded" — through what Malawey calls "pressed phonation" — can produce a certain roughness, "in which we can hear things beside the tone of the voice itself", including sounds of friction, rasp and hoarseness (p. 102-103). Consequently, digital vocal distortion can be associated strongly with corporal strain, rendering its use highly emotive. Lastly, in his reception test with 128 participants, Lacasse (2000) found that a distorted vocal sound evoked connotations of malevolence and unnaturalness (p. 161).

Intelligibility

While distortion can thicken a sound, elevating its presence in the track, the intelligibility of the lyrics can sometimes suffer through the application of these more stylised effects. However, poor intelligibility itself can be used as a narrative and creative tool that can aid storytelling or emotional expression by deliberately clouding or coding the message. This can evoke a sense of censorship**,** or a blurred sense of self in the narrator.

In the singer-songwriter realm, *distortion* on vocals may seem counterintuitive, but when used purposefully, it can convey feelings like urgency, alienation, or resistance.

Analysis Examples:

Artists such as girl in red use this technique to powerful effect: in *i wanna be your girlfriend* (2018) and *bad idea!* (2019), her lead vocals are subtly distorted and doubled, underscoring themes of adolescent longing, frustration, and queer identity. The roughness, at times,

contrasts with the softness of her delivery, creating a tension between intimacy, desperation and rage.

Billie Eilish employs distorted vocals selectively, often on specific phrases or backing layers. In *you should see me in a crown* (2019), gritty vocal processing adds menace to her whisper-like delivery, highlighting the push-and-pull between vulnerability and power. These choices blur the lines between genre conventions—merging singer-songwriter introspection with punk, grunge, or industrial aesthetics.

Summary

While Auto-Tune, vocoder, and distortion were once seen as revolutionary or genre-specific, they are now part of the broader pop production toolkit—readily available and creatively reinterpreted by self-producing artists. For the singer-songwriter, these tools offer new expressive possibilities: they can enhance vulnerability, suggest alternate identities, or introduce friction into the narrative arc of a song. When used consciously, they do not obscure the voice's message, but instead transform it, adding layers of meaning by design.

3 — Methodology

This chapter outlines the artistic research methodology underpinning the project, with a focus on the practice-based inquiry into vocal production aesthetics. Rather than generating new theoretical insights through practice alone, the creative work was informed and shaped by the conceptual and theoretical frameworks established in Chapters 1 and 2. These frameworks were applied directly to the songwriting and production process, using each song as a site to explore, test, and reflect on specific vocal production techniques within a real-world creative workflow.

While the production process occasionally yielded additional insights, the primary aim was not to discover new theory but to critically engage with existing discourse through artistic practice. The chapter details the tools, workflows, and documentation strategies employed throughout the project, with a particular emphasis on the vocal recording and production process. It also situates the project within current scholarly debates surrounding the analysis of the voice in popular music, addressing the methodological gap between technical production practice and musicological analysis.

3.1 Artistic Research as Practice-Based Inquiry

This thesis is situated within the field of artistic research and follows a practice-based case study approach, in alignment with Frayling's (1993) model of "research through art," where the production of three original songs serves as both artistic outcome and site of inquiry. As Borgdorff (2012) notes, theory and practice in the arts are co-constitutive, and reflection is embedded in the act of making. Each song thus functions as a case study, allowing for a deeper understanding of how production choices contribute to emotional expression and narrative without undermining the authenticity and intimacy of the singer-songwriter format.

My role as researcher and subject positions this project within an auto-ethnographic tradition. As Malawey (2020) notes, the relationship between the voice and identity is deeply intertwined, and my own positionality as a female, self-producing artist inevitably shapes both the practice and its interpretation. This lens allows for a nuanced understanding of how personal history, artistic intention, and cultural context intersect in the vocal production process.

Self-producing in a home studio brought both challenges and freedoms. Limited resources required creative problem-solving, while the autonomy allowed for a deeply personal exploration of sonic identity. This iterative preparation stage helped me refine the artistic direction without overcommitting to fixed sonic ideas too early, fostering a more fluid and responsive production process. My work contributes to a growing field of self-producing singer-songwriters who assume both creative authorship and technical control. This reflects broader shifts in the landscape of "democratised" music production and expands the expressive possibilities of the singer-songwriter tradition through technological mediation.

3.2 Bridging Theory and Practice

As pointed out in the introduction, there remains a persistent gap in scholarly discourse regarding the intersection of production technology and musical meaning. Music theory often overlooks technological mediation, despite its central role in contemporary popular music. As Steinbrecher (2021) writes, "music-theoretical approaches actually reveal a significant blind spot" in addressing "production and technology topics" (p. 416), echoing Brøvig-Hanssen and Danielsen's (2016) call for a deeper examination of digitisation's aesthetic implications on popular music sound. This methodological gap has led to a lack of conceptual frameworks for discussing the emotional and narrative roles of production techniques.

The methodological gap is even more critical with regards to the analysis of the singing voice in popular music. Even in writings where vocal production is highlighted, such as in practical handbooks like Craig Anderton's (2018) *How to Record and Mix Great Vocals*, the focus remains on technique rather than expressive meaning or narrative function[48].

In response to Exarchos and Zagorski-Thomas (2020) observation that "there is very little connect between the technical theory and a musicological theory of production aesthetics", this project seeks to bridge that divide through creative practice. By analysing my own songs from narrative and interpretive perspectives, I explore how techniques such as close-

[48] Granted, handbooks are typically designed for wide applicability, and musical meaning is highly context-dependent. However, since the voice is the primary carrier of meaning in popular music, I argue that efforts to understand and articulate its expressive impact through production are not only worthwhile but necessary.

mic'ing, vocal layering, reverb, vocoder, and pitch modulation work together to construct vocal presence, shape emotional tone, and contribute to the overall narrative arc of each song.

3.3 Creative Process and Tools

3.3.1 Tools and Environment

All vocal recordings and most mixing were carried out in a home studio setup, using a Neumann U87 Ai microphone in cardioid mode, Logic Pro X, and a minimal selection of plugins and monitoring systems. The setup reflects the realities of independent, self-producing artists today, and this constraint played a key role in shaping the aesthetic of the work.

3.3.2 Workflow and Production Strategies

Pre-production began with a larger pool of self-written songs, from which three were selected for further development during the thesis. This phase enabled me to explore arrangement ideas alongside initial vocal drafts, laying an analytical foundation for later refinement. For the thesis, I returned to the lyrical core of each piece, rebuilding the songs from just voice and guitar to realign my approach with their emotional centre. Using phone recordings and annotated lyric sheets, I sketched early production concepts, retaining elements from pre-production that supported the songs' intent and discarding those that felt emotionally misaligned.

Each song was approached as a self-contained narrative, prompting me to conduct a brief analysis of its lyrics based on Randle and Evans' (2013) typology for distinguishing open and closed readings and narratives (see 1.2.2.6; see appendix for documentation). I analysed earlier demos, cleaned up Logic sessions, created tailored to-do lists and outlined an initial production concept for each song (see appendix).

Recording vocals followed a "comp-as-you-go" method, supported by my collaborator Rami Olsen. His input functioned as what Burgess and Toulson (2017) describe as a "practical facilitator" and "experienced consultant" — offering targeted input on

arrangement, sonic detail, or structure without overriding my artistic vision. Importantly, Rami handled technical monitoring, allowing me to focus on performance.

My mixing workflow followed a "bottom-up" approach (Harding, 2020, p. 303), beginning with instrumental balance before layering working on the vocals. However, mixing is always an iterative process: individual sounds and elements were continuously adjusted throughout to shape an overall coherent and balanced sonic image. The vocal mix received particular attention, guided by techniques outlined in practical handbooks (e.g. Mike Senior, Craig Anderton). Mixes were tested for translation across diverse playback environments (studio speakers, headphones, 'cheap' bluetooth boxes). I worked entirely in the box (ITB), beginning with rough mixes that focused on establishing the vocal character and overall mood early on.

Although final mixing and mastering will take place after thesis submission, external feedback has already proven invaluable in shaping the emotional impact of the songs. To complete this final stage, I intend to work with a dedicated mixing engineer whose fresh perspective can reveal sonic blind spots that often arise through prolonged close listening and creative isolation, ensuring the finished tracks achieve their full expressive potential.

3.3.3 Analytical Listening

Close listening was integral to both the analytical and creative phases of this project. I practised informal reverse-engineering of vocal production aesthetics from artists such as Billie Eilish, Phoebe Bridgers, and Lizzy McAlpine, among others, to better understand how specific techniques contribute to narrative delivery and emotional resonance. These insights helped shape my own aesthetic decisions. While some of these observations are reflected in the contextual discussion in Chapter 2, they are not explicitly cited within the case studies.

3.4 Documentation and Reflective Journal

Throughout the project, I maintained a reflective journal that documented production decisions, emotional responses, challenges, and evolving artistic concepts. The creative process further documented and analysed in detail within the individual song case studies in

Chapter 4. Selected DAW screenshots and mixing drafts supplement this documentation and serve as visual evidence of the evolving creative process.[49]

[49] See appendix for select DAW screenshots; some audio examples are included on the USB drive.

4 — Case Study: Original Song Production Process

4.1 Overview

This chapter presents the three produced original songs as individual case studies, each including artistic intention, technical process, and reflective analysis. Each case study starts with a short analysis and summary of the song's lyrical narrative.[50]

Importantly, by sharing my artistic intentions and technical decisions, I aim to offer insight into the creative process rather than *impose* specific interpretations. The richness of songwriting and production often lies in its ambiguity — in the space it leaves open for listeners to find their own meanings.

Before delving into the individual case studies in detail, the following overview (Table 2) serves to orient the reader by summarising the conceptual focus, vocal aesthetics, production techniques, and narrative function of each track, offering a comparative overview of how each song's sonic choices support its lyrical themes and affective intentions.

[50] To distinguish between open and closed readings and narratives, I conducted a brief analysis of each song's lyrics based on Randle and Evans' (2013) typology for analysing pop song lyrics (see 1.2.2.6). See appendix for documentation (Example: *Forever Vow*).

Table 2: Overview of Case Studies: Vocal Concepts, Aesthetics & Narrative Function

Title	Core Emotion & Lyrical Theme	Reading of Narrative	Overall Vocal Aesthetics	Key Concepts applied	Key Production Techniques	Narrative & Emotional Impact	Vocal Aesthetic Position
Forever Vow	Vulnerability, doubt, weariness, sincerity, nostalgia, past vs present, broken relationship, perseverance, renewal of vows, reciprocity	Closed Reading/ Undefined	natural, intimate, warm, varying between close-up/ breathy fragility and power; layered	• Performance Intensity • Embodied Voice • Register & Timbre Considerations • Spatial and Temporal Manipulation • Hall's Proxemic Zones	• Close mic'ing • low to high performance intensity • ceremonial reverb and subtle echo • selective layering and panning • call-and-response panning	▶ reflects narrator's state of mind, perceived as the 'unfiltered truth' ▶ builds narrative arc through increasing intensity ▶ reverb places core memory into a perceived past, echo ties together past and present ▶ layers add emphasis, warmth and and urgency ▶ underlines reciprocity	natural-realistic, transparent mediation
What Do You Expect	emotional diffusion, physical desire, powerlessness, emotional fragmentation, moral struggle, ambivalence, rational restraint, denial	Closed Reading/ Defined	Dual-layered vocals; doubled phrases; close, direct timbre; subtle edge	• Embodied vs Disembodied Voice • Hall's Proxemic Zones • Spatial, Temporal and Timbral Manipulation • Opaque Mediation	• Double-tracking, layering • close mic'ing • distortion • light vocoder • Vocal sampling (and hyper-close"mmhhs")	▶ Creates emotional ambiguity and inner dialogue effect, supports lyrical tension ▶ Emotional overload and tension ▶ voice remains embodied but begins to split into a more complex vocal persona	artistic-technological, opaque-leaning mediation
I Don't Mind	Reflective introspection, memory, fractured identity, regret, disembodiment, surrender, haunted	Open Reading/ undefined	Warped, off-kilter, ghostly; from embodied to synthetic; heavy processing	• Disembodied voice • Register and Timbre Considerations • Identity/Persona • Timbral Manipulation	• Varispeed formant shifts on lead vocal • Vocoder • Auto-Tune-style pitch correction • pitch-shifted 'moster'doubles; • reverb	▶ Sonically reflects haunted emotional state ▶ distances vocal persona ▶ manipulations mirror internal breakdown; vocals move from physical memory to digital ghost ▶ Emotional override ▶ Warped sense of time and space	artistic-technological, opaque mediation

4.2 'Forever Vow' – Bronty[51]

4.2.1 Summary of Lyrical Narrative

Forever Vow tells the story of a tired narrator addressing a growing emotional distance in a long-term relationship or marriage. The first verse ("*Well I don't think you're feeling it right now")* opens as an intimate, resigned admission that sets the tone for a song grounded in both vulnerability and perseverance. The lyrics trace the narrator's repeated efforts to rekindle a fading connection — an honest attempt to hold onto something that is slipping away. Woven throughout are allusions to wedding vows, used not only as lyrical anchors but as emotional cues — reminders of a shared past that still holds meaning.

According to Randle and Evans' (2013) typology of pop song lyrics, *Forever Vow* aligns with a closed reading and undefined narrative, characterised by its repetitive phrasing and focus on mood and emotional state rather than a clearly defined, linear story.

The structure of the song itself mirrors this emotional arc. Each verse returns to the same phrasing *("I don't think you're feeling it right now […] trying my best to fix this, but I don't know how'"*, reflecting a loop of effort and stagnation. The pre-chorus ("*Promised that I'd remind you […] that you said forever and forever's not over yet"*) gently bridges the verses and choruses, nudging the other person back into a shared memory and urging reconnection.

The chorus — evoking the language of vows — introduces a temporal shift. Here, memory and longing intersect with hope, contrasting the worn-out present with the promises of the past. ("*All that I am I give to you" […] "all that you are you give to me"*). The idealised sentiment of 'forever' is broken down into small, everyday acts, with the line *'we'll take turns with cooking and making pots of tea'* evoking the lived reality behind long-term romantic commitment. In this way, the song highlights how emotional labour in long-term relationships often lies not in grand gestures, but in the quiet persistence of showing up—again and again—for the small, shared rituals that sustain connection. As the song

---

[51] While the original songs in this project are grounded in emotional sincerity, they are not strictly autobiographical. The lyrics weave together fragments of personal experience and observations, always grounded in authentic emotional states. The decision to release this work under the artist name *Bronty* reinforces this approach, allowing for a degree of separation between myself and the narrator, while providing a creative frame through which emotional material can be explored with greater expressive freedom.

progresses, the narrator's will to fight is reignited, transforming from emotional fatigue to new-found resolve. The bridge functions as the emotional and structural peak of the song. While it echoes the language of the original vows, the line *"No matter the tide, I'll be by your side"* articulates a conscious decision to remain committed in the now. This shift from memory to agency underscores a turning point in the narrator's emotional trajectory — not only continuing to endure but actively choosing to reengage with their partner. Musically and lyrically, the bridge encapsulates the thematic core of the song: that love is not only a romantic gesture to be remembered but a decision to be remade time and time again, especially in the face of difficulty. In the third and final verse, the line shifts from *"but I don't know how"* to *"could you help me out?,"* marking another crucial turning point — the narrator moves from passive helplessness to agency, inviting their partner into the resolution and acknowledging that emotional repair is a shared responsibility.

Though the lyrics remain simple and repetitive throughout, it is in the performance intensity and production where the emotional transformation unfolds.

4.2.2 Artistic Intention and Production Concept

With *Forever Vow*, my artistic intention was to sonically embody the emotional weight of endurance in long-lasting love—specifically, the quiet kind of love that persists through disconnection, fatigue, and doubt. I wanted to create a listening experience that feels close, vulnerable, and at times nostalgic, mirroring the inner world of someone who is emotionally worn out but still holding on. Just as much as the lyrics serve as a reminder to the partner, they function as a reminder to the narrator herself. While the lyrics are intentionally sparse and repetitive, the production was designed to carry the nuances between the lines, as well as support the overall emotional arc of the song, tracing the narrator's journey from weariness and resignation to renewed agency.

A central part of this concept was the use of spatial and temporal manipulation (cf. Lacasse, 2000) such as reverb and echo—placing the narrator's voice in subtly different spatial and temporal dimensions. I imagined the verses as anchored in the present: dry, intimate, and closely mic'ed. In contrast, I envisaged treating the choruses with more spatial depth, evoking distant memories of the ceremony and vows and echos to connect the past and the present.
I further wanted to explore in how close-mi'cing and performance intensity, closely tied to considerations on register and timbre and embodiment, could help shape the narrative arc from vulnerability and weariness at the outset, to a renewed sense of agency towards the end of the track. I also intended to experiment with the physical and emotional disconnect between the narrator and her partner, working with the idea of Hall's proxemic zones.

4.2.3 Technical Considerations and Reflective Process

The case study on *Forever Vow* is somewhat more detailed than the others, as the song's emotional arc unfolds gradually and relies heavily on subtle shifts in vocal delivery, spatial treatment, and layering. As a result, a more section-by-section analysis was necessary to trace how these understated production choices accumulate meaning over time and support the song's evolving emotional landscape.

To realise the emotional and narrative arc of *Forever Vow,* I drew on a range of vocal production techniques, often adapting my approach in response to what I discovered during the process.

Verse 1

For the lead vocal in the first verse, I recorded at low vocal intensity, very close to the microphone — approx. 5-8cm — to capture an embodied sense of exhaustion, weariness, and intimacy. The proximity to the microphone enabled me to capture details in the high frequency region. I chose to counter the proximity effect — the boosting of low frequency signals — by using the low cut switch on the microphone, as I wanted the vocal to sound more fragile and thin at the beginning. Despite moving the microphone to a 45° angle, this approach created a lot of issues with mouth noises and other artefacts, which took careful comping and editing to clean up (see appendix, *Figure A*). I avoided using the pencil tool or plug-ins such as iZotope, as these can often lead to unwanted artefacts. Similarly, I avoided overt compression, as this exacerbated the problem. These decisions were driven by a desire to preserve the rawness and authenticity of the lead vocal, foregrounding emotional truth over aesthetic perfection. The lower register coupled with weak articulation (sung through a more closed jaw) resulted in the capture of vocal fry on the words *'you're', 'been', 'but'* and 'that you said' serving as indication of fatigue.

In order to create space for the fragile lead vocal to come through in the mix without overt processing, I double tracked the guitar and panned the accompaniment to left and right. Furthermore, I decided to use only one lead vocal for clarity (as opposed to an increasingly popular choice of doubling the vocal for density, timbre and style). The low performance

intensity coupled with an unpolished appeal and lack of reverb, suggest that the narrator is speaking a weary admission, perhaps not even directly to the partner, but to herself.

Pre-Chorus 1

While the vocal performance here is still captured in close proximity to the microphone, in a breathy, restrained voice, I treated the vocal with subtle compression, still keeping it raw, but giving it a gentle lift over the growing instrumental arrangement. Rather than increasing the performance intensity as one would naturally do in order to reach a higher note, the '*I'd re[mind you]*' was sung in a gentle head voice. The end of the first pre-chorus introduces a first temporal shift: this transition was carefully crafted using subtle adjustments in reverb on the last phrase *'[forever's] not over yet__*', opening up the emotional space without fully stepping outside the present. I was unsure at first, whether to use reverb from the start of the pre-chorus, but decided to save it for the final few words as it worked better as a transition into the chorus and kept the pre-chorus grounded in the present, and made the temporal effect in the chorus more impactful.

Chorus 1

In the first chorus, reverb and echo serve a clear narrative function. Instead of using reverb for polish or lift—as is common in pop—I applied a symbolic, church-like reverb to evoke the memory of a wedding ceremony. Echoes of key phrases act like lingering vows, suggesting a nostalgic, temporal distance. While the lyrics only hint at this emotional shift, the production clarifies it: the ethereal, reflective echoes convey the sense of promises fading, mirroring the narrator's emotional detachment. Automation helped fine-tune these nuances. As Rick Beato[52] notes, echo requires space to breathe—something I leaned into by removing an early brass section (later featured in the second chorus) to allow the vocal echoes to unfold fully.

Verse 2

Verse 2 draws the listener back into the present, stripping away all the reverb from the lead vocal. The rather abrupt change of environment focuses the listeners attention on the vocal

[52] see footnote 43

performance, which is more determined and less fragile this time than in the first verse, mirroring an emotional shift from helplessness to instilled perseverance. Here, I engaged the diaphragm more for a higher, but still audibly restrained performance energy. To avoid capturing plosives, I backed away from the microphone a bit further this time, to a distance of around 12 cm, singing 'past' the microphone at a 45° angle. Though the lyrics are a repetition of the first verse, the difference in performance intensity here influences the narrative, underlining the emotional momentum building beneath the lyrical surface. I found that the contrast achieved by taking away most of the reverb worked really well, suggesting a return to the present.

Pre-Chorus 2

With the return of the Pre-Chorus, the lead vocal is joined by low, breathy backing vocals panned left and right on the lines "*I'd remind you should you forget*" and "*forever, and forever's not over yet.*" Positioned on either side of the the lead vocal in the mix, they offer emotional support without pulling focus. Inspired by Hall's intimate proxemic zone concept, these close, breathy, low-register layers add warmth and intimacy, reinforcing the vulnerability of the moment. Meanwhile, the lead vocal gains intensity —"*I'd [remind you]*" is delivered in a sharper head voice this time—signalling the narrator's growing resolve. The stereo spread of the backing vocals subtly expands the emotional space, suggesting an opening-up or turning point. Light reverb blends the voices without distancing them, keeping the moment grounded in the "now", with a slight increase in reverb on the last phrase "*and forever's not over yet*" leading into the second chorus.

I briefly considered placing the lead vocal itself in Hall's intimate zone, but pairing that level of closeness with the vocal's rising urgency felt too aggressive. Instead, experimenting with the backing vocal arrangement allowed me to preserve a sense of emotional closeness while letting the lead retain clarity and momentum.

Chorus 2

Instead of leaning into the reverberant memory of the past, like in Chorus 1, the lead vocal in the second chorus is noticeably drier and more grounded. This shift was a significant turning point in the production process. Initially, I had planned to keep both choruses in the

same reverberant, distant and echoing space. However, I realised that the emotional arc required contrast: the second chorus had to feel present, like a renewal of vows rather than a recollection of memory. While the dryness of the lead vocal brings the listener closer, the backing vocals take on a more distant choral texture (created through multiple harmonies of 'aaaah's, in both chest and head voice registers). This contrast forms a metaphorical portal between past and present and reflects the narrator's growing conviction.

Interlude

The interlude opens into an even wider stereo field, immersing the listener in a morphing soundscape of the past and present, contrasting more reverberant elements such as ethereal reverberant guitar plucks with the more organic, dry drums (now widened across the stereo field) - representing a transformation of the narrator's inner emotional world. The choral texture of backing vocals ("*Aaaahh"*), reappears in the second half of the interlude, now making use of the full stereo width, heightening emotive effect. The widening of the stereo image in the interlude allows for the off-centre lead vocals — panned slightly to the left and to the right — in the following section (bridge) to feel less out of place/more organic.

Bridge

The bridge marks the emotional and dynamic peak of the song. Its opening lines "*All that's mine is yours" and "All that's yours is mine"* were panned alternately to the right and left, mirroring the lyrical theme of reciprocity. The lead vocal here is delivered with a high intensity in the chest voice and recorded at a greater distance from the microphone, creating a spatial contrast to the more intimate, close-mic'ed verses. This change in spatial treatment reflects the narrator's emotional shift — from quiet, internal reflection to an outward-facing moment of emotional conviction. The next line "*For better or for worse"* is panned in the centre again, underlining its contextual relevance. The lead vocal in the penultimate line *'No matter the tide'* is doubled and spread left and right, adding richness to the vocal as well as creating width, underlining the omnipresence of its lyrical theme. Finally, the centred lead vocal "*I'll be by your side"* is supported by layered harmonies on either side of the stereo image. The bridge was treated with less reverb to emphasise its immediacy and resolve. Here, the narrator is no longer reflecting, but confronting the

moment with everything they have—throwing their full emotional weight into the statement. Pulling back from the microphone during this section also helped accommodate the higher vocal intensity and diaphragm-supported chest voice delivery without harshness or distortion, preserving clarity while amplifying expressive power.

Stereo-panned drum rolls were introduced here as a call-and-response device, mirroring the narrator's emotional intensity, as well as provoking a possible a reaction from the partner. The combination of rhythmic momentum, stereo vocal spread, and harmonic layering transforms this section into a moment of heartfelt declaration and determination.

Verse 3

In contrast, the final verse opens with a 1-bar break, featuring only the backing vocal stack *"well I don't think"*, setting up another turning point: The lead vocal here enters mid-line on "*you're feeling it right now"*, creating a deliberate sense of ambiguity against the backing vocal. By omitting the narrator's voice from the fist half of the phrase, this leaves room for the possibility that connection to the partner has been re-established, suggesting they *are,* in fact, "*feeling **it** right now"*. Interestingly, in this context, the word *"it"* shifts in meaning—from the relationship itself to the narrator's attempt to reconnect.

The next line "*think I'm out of ways to fix this, could you help me out?"*, confirms this shift, marking a change from previous wordings of the verses: the narrator moves from passive helplessness in previous iterations to a place of shared agency, extending an invitation to their partner to collaborate in the emotional repair.

What surprised me during this process was how small structural decisions in the arrangement (specifically vocal layering) and omitting words from the lead vocal could unlock a narrative shift that I hadn't originally envisioned - letting production not just support the story, but subtly rewrite it.

Ending (Pre-Chorus)

The final iteration of the pre-chorus lead vocal is sung in the chest voice, against a thick layer of backing vocals, for which I recorded two harmonies, doubled and panned either side of the stereo image. While the lead vocal carries a sense of determination and grounded presence, this backing vocal arrangement amplifies the narrator's newly found emotional resolve.

Comparing all three pre-chorus sections throughout the production demonstrates how vocal tone, performance intensity, and microphone placement can dramatically shape the perceived fragility of a vocal delivery thereby supporting the emotional arc of a song. In particular, the opening word "*I'd*" in each version reveals subtle but significant emotional shifts—from tentative head voice to resolute chest voice (see audio example 1 for comparisons on USB drive).

For the final lead vocal line of the song, "*forever's not over yet*", I wanted the voice to convey a soft invitation rather than a determined plea. I leaned in very closely (approx. 5 cm) to the microphone again, letting the vocal return to a more intimate, gentle tone which places the voice firmly within the intimate proxemic zone (Hall, 1969).
The final phrase—"*forever's not over yet*"—is delivered a cappella and left completely dry. I had initially considered adding reverb to suggest a merging of past, present, and future, evoking a suspended, unresolved state. However, the harmonically suspended chord in the backing vocals already fulfilled that function. In the end, I found it more powerful to close the song in the direct, intimate zone without reverb.
This final moment demonstrates how the absence of reverb can be just as expressive as its presence. By resisting the impulse to sonically "lift" or blur the ending, the production remains grounded in raw presence, letting the uncertainty linger without being romanticised.

4.2.4 Findings

Forever Vow deepened my understanding of how subtle vocal production techniques such as mic placement, reverb, echo, panning, and performance intensity can shape emotional storytelling, especially when working with sparse, repetitive lyrics. Rather than functioning as surface-level polish, these elements became structural tools that supported the song's narrative arc, sometimes revealing emotional layers the lyrics only implied.
The production process reinforced a key learning: that when lyrical development is limited, sonic choices carry even more narrative weight. In *Forever Vow*, where each verse revisits familiar phrases, it was these shifts in space, tone, and vocal delivery that carried the emotional momentum forward. Without them, the movement from vulnerability to

commitment might remain too subtle to register. Through techniques like stereo-panned layering in the bridge, or the shift from reverberant distance to grounded presence across the choruses, production offered a kind of emotional counterpoint—a way to let the unspoken rise to the surface.

A key turning point was the realisation that the two choruses couldn't live in the same spatial world. Initially, I had planned for both to be soaked in reverb, but this flattened the emotional arc (in a "back where we started" kind of way). The shift to a drier second chorus (atypical for a contemporary/pop production) was subtle, but crucial as it helped underscore the narrator's transition from nostalgia to active recommitment in the present.

One of the most valuable discoveries was the expressive potential of spatial contrast. A dry vocal could immediately anchor the listener in the present, while reverb or echo hinted at memory, distance, or unresolved feeling.

Viewed through Randle and Evans' (2013) typology of pop lyrics, *Forever Vow* resides in the open/undefined narrative quadrant: the lyrics are emotionally evocative and metaphorical, but lack a defined fabula in terms of event, time, or setting. This made production decisions even more central to the storytelling. Spatialisation, vocal texture, and proximity filled in the narrative gaps, nudging the listener toward an emotional reading while still leaving space for personal interpretation. In this sense, vocal production became not just a supporting layer, but a primary means of constructing and shaping the perceived narrative space.

Ultimately, *Forever Vow* became a blueprint for how I want to approach vocal production more broadly: not as decoration, but as a form of emotional architecture. This case study shows how even minimalist, singer-songwriter-informed production can benefit from pop aesthetics by simply amplifying what's already present between and within the lines.

4.3 'What Do You Expect' – Bronty

4.3.1 Summary of Lyrical Narrative

This song explores the internal conflict between emotional disarray, physical desire and rational restraint, captured through a first-person narrative voice. The lyrics convey a sense of vulnerability, powerlessness, longing and temptation, where the narrator is acutely aware of crossing a moral boundary yet finds themself unable to resist. Based on Randle and Evans' (2013) typology of pop song lyrics, *What Do You Expect* reflects a closed reading with a more defined narrative, marked by explicit events and a clear emotional conflict, even though a resolution remains elusive.

The song's structure reinforces this emotional tension: intimate verses grounded in embodied sensation using physical imagery (e.g. *"shivers run right through me when you touch me like that")* contrast with choruses that attempt to restore reason and distance through repetition and declaration ("*this shouldn't be happening, no"*). The bridge functions as a final attempt at emotional distance ("*I'll just pretend that we're only good friends"*), yet the lyrics never arrive at resolution, reflecting the ongoing emotional duality of the narrator ("*I hope my thoughts will follow soon"*).

This fluctuation between physical desire and emotional restraint becomes the central narrative and sonic tension of the song.

4.3.2 Artistic Intention and Production Concept

This song was an experiment in emotional duality, echoing the narrative tension between physical desire and emotional restraint. From the lyrical narrative, three implicit psychological layers emerged: the physical world, the emotional world, and the realm of reason and restraint. Each of these layers provided grounds for their own distinct vocal aesthetic. I was thus interested in exploring how production choices could articulate the subtle shifts between the three states.

Several key concepts shaped my initial intentions:

(1) Edward T. Hall's theory of proxemic zones;
(2) the distinction between embodied and disembodied voice;
(3) the expressive potential of spatial, temporal, and timbral manipulation; and
(4) the idea of *opaque mediation*, where sonic treatment becomes a veil rather than a window.

These ideas translated into specific techniques: close mic'ing and low performance intensity to place the voice in the intimate zone and capture paralinguistic detail (breath, vocal fry, sibilance); mono vs. bilateral double-tracking to explore emotional immediacy; contrasting wet vs. dry processing to indicate shifts in space and affect; vocal reiterations to reveal internal thoughts and fragment temporal flow; distortion to convey overload or rupture; and vocoder as a tool for emotional distancing and self-censorship, reflecting the narrator's turn toward rationality and restraint.

4.3.3 Technical Considerations and Reflective Process

During the production process, a number of aesthetic and technical insights shaped the final version. Again, these are best presented in a section-by-section approach.

Verse 1

In earlier versions, the track opened with a single, intimate lead vocal line, recorded at close proximity to the microphone. This was however revised in favour of a mono double-tracked vocal, placing more emphasis on the narrator's emotional conflict form the outset. This choice allowed the lyrics to take on an explicit narrative role, while the production shaped the underlying emotional subtext, setting the tone for the ambiguity and tension that permeate the song. I placed the lead vocals in a subtle airy reverberant space only to take the reverb away again on the ending phrases *"when you look at me like that"* and *"when you touch me like that"* which helped draw focus to the lyrics and emphasise their explicitly intimate nature. The reverberant inhale audible at the transition into the second verse was consciously placed for intimate affect, drawing further attention to the physicality of the lyric and associated embodied sensations.

Verse 2

The mono-doubled lead vocal continues throughout the second verse. Here, the narrator turns more inwards with the lyrics becoming more introspective and mentally turbulent, characterised by lyrical repetition that reflect obsessive thinking *("messing with my mind, messing with me").* The production emphasises this state of mind, creating a distant and slightly distorted reiteration of the phrase "*messing with my mind"*. The distortion, a theme that reappears throughout, symbolises internalised thoughts and emotional "system overload".

Chorus 1

The chorus marks an attempt to break free from the moral dilemma and retreat into reason. This is mirrored in the production by widening the stereo image and placing bilaterally double-tracked lead vocals within lush, layered backing vocals (each recorded twice and panned to the sides). Leaving the centre of the stereo field empty was a deliberate choice, symbolising a departure from the emotional core established earlier. The lightly distorted echo on "*shouldn't be"* reinforces the internal struggle, acting as a moment of self-persuasion within the narrator's psyche.

Interlude

The interlude features non-linguistic vocalisation (*"uhhh")* that is pushed back in the mix, rendered airy and detached through a close-mic'ed, breathy delivery and reverb. These sounds function as faded memories or emotional echoes, providing contrast for the intensity that follows.

Verse 3

The attempt to restore reason and restraint in the chorus collapses in Verse 3, as the narrative returns to a space of physical desire. This verse contains the most lyrically explicit section of the song, and the vocal production reflects this shift: the lead vocal is presented completely dry and closely mic'd, enhancing intimacy and directness. This decision contrasts sharply with the layered, spatialised treatment of the previous chorus and interlude, drawing the listener into a more intimate scene. Initially, I experimented with bilateral double-tracking of the lead vocal, panning the doubles hard left and right, to envelop the listener in

emotional entanglement while leaving the centre open for the prolonged distorted scream (*"ahhh"*), evoking a sense of system overload. However, this approach diminished the sense of direct address essential to the verse's explicit lyrics. I ultimately opted for mono double-tracking, with the scream panned slightly off centre to the left to heighten its sense of intrusion (see audio example 2 USB drive for comparison). The distorted scream acts as an embodiment of inner turmoil and a momentary loss of control.

Chorus 2

Chorus 2 is a repeat of the first Chorus, with an added vocoder layer, subtly shifting the vocal timbre towards a more disembodied, synthetic quality. This production choice reflects the narrator's narrator's attempt to regain control in a turn toward rationality and restraint, using the vocoder as a tool for emotional distancing, self-censorship, and the suppression of desire.

Interlude

The second interlude mirrors the first, with its breathy, reverberant vocalisations, though this time it is interrupted by a vocal sample (*"I'll just pretend").* Functioning as a sonic foreshadowing of the bridge, this inserted fragment acts like a musical motif, subtly reintroducing the narrator's internal monologue.

Bridge

The bridge pulls back slightly in performance intensity, marked by a more restrained vocal delivery. Unlike the verses, however, the voice here is not entirely stripped of reverb, placing it within what Hall defines as the personal rather than intimate proxemic zone. This slight distance helps convey the narrator's vulnerability as something carefully guarded and deliberately contained. Meanwhile, the reiterated line *"I'll just pretend"* functions like an inner monologue or mantra, as if the narrator is trying to persuade themself to stay on this chosen path of restraint. Yet at the very end, the final word *"soon"* is uttered by a dry and closely mic'ed lead vocal, pushing the voice forward into the intimate zone and ending the section with a lingering sense of ambiguity.

Outro

The outro contrasts, hyper close and breathy humming with a reiterated, distorted *"I'll just pretend"*. This sonifies the internal conflict between emotional truth and mental denial, reinforcing the underlying contradiction even in the final moments of the song.

General comments

A common technical challenge encountered was the prominence of sibilants, plosives, and mouth clicks in stereo vocal stacking, especially with the very closely mic'd vocals. The process of manually locating the clicks and using a pencil tool to correct the source information proved to be very labour-intensive and unrealistic given the sheer amount of mouth noises (probably made worse by the remnants of a cold when recording!). Tools such as iZotope can be used to remove mouth noises more time effectively, however they can sometimes compromise the overall quality of the recording. Drawing on choral techniques (such as omitting end consonants on doubles), these issues were smoothed out without compromising clarity. Mouth clicks can be masked by the percussive elements (e.g. drums or strums) and thus are negligible, so it is worth checking the vocal against the rest of the backing track to find the parts where the mouth noises are audible in context of the whole mix. Sometimes mixing can solve these issues by masking them, and other times exacerbate them, so I learned to make time to edit out remaining 'annoyances' after the mixing stage. The use of mono-double tracking was more forgiving in terms of mouth noises and clicks—as the diffusion allowed the vocal to sit a bit further back in the mix, making the high frequency noises less irritating and diffusing them.

4.3.4 Findings

This song demonstrated how vocal production can communicate emotional and narrative complexity through intentional spatial, temporal, and timbral manipulation. Processed fragments—reiterations of existing lyrics—not only heightened emotional tension but also introduced a secondary layer of narration, revealing the narrator's subliminal struggle beneath the more controlled surface of the lead vocal.

Guided by Edward T. Hall's theory of proxemic zones, I explored how shifting vocal proximity—from intimate close-mic'd verses to more distant choruses—could reflect

vulnerability and emotional guardedness. Combining natural, embodied vocal qualities with digitally processed, disembodied textures provided a means to explore psychological tension and emotional ambivalence. The use of a vocoder sound softened and obscured the voice, creating a sense of emotional detachment, while distorted reiterations introduced harsher timbral elements that conveyed overstimulation and internal conflict.

However, the production process of *What Do You Expect* also highlighted that over-processing can dilute a vocal's emotional impact and weaken the clarity of the song's narrative voice. Not all production ideas can coexist effectively—for instance, a single, lead vocal might traditionally signal intimacy, but in this case, it felt redundant, as the lyrics themselves already used physical imagery. Instead, mono double-tracking allowed for an emotionally complex vocal presence that reflected the narrator's internal duality from the outset. The evocative breath samples proved to be more effective and contrasting when paired with the emotionally conflicted double-tracked lead vocal, compared to how they would have functioned against a single, clear, and intimate vocal. Sometimes production is at its most impactful when it supports the emotional subtext rather than reinforcing the literal content.

Overall, this song exemplifies how hybrid artistic-technological mediation can negotiate authenticity and affective complexity. It shows that vocal production is not simply a technical process but an integral narrative device capable of expressing layered emotional states and articulating the tension between embodied desire and rational restraint.

4.4 'I Don't Mind' – Bronty

4.4.1 Summary of Lyrical Narrative

The song unfolds through a first-person narrative, exploring themes of somatic memory[53], regret, surrender and penance. From the outset, it remains unclear who exactly the narrator is addressing—whether the "you" in the verses and the "you" in the pre-chorus and chorus even refer to the same figure, or whether the song's emotional core is projected onto multiple sources. It is equally ambiguous whether the emotional impact described is self-inflicted or caused by the unspecified other. This deliberate ambiguity establishes a blurred, unstable emotional landscape, inviting multiple interpretations.

According to Randle and Evans' (2013) typology, *I Don't Mind* exemplifies an open reading with an undefined narrative, characterised by its ambiguous actors, setting, and temporality, allowing the listener to inhabit the song's unresolved emotional space.

In the first verse, the narrator evokes a visceral sense of haunting connection: "*I remember you, I still feel you / In my bones, in my body / In my blood, under my skin.*" These lines root the memory in the physical body, suggesting a past relationship or experience that has left a permanent imprint. However, the tone is neutral, which reinforces the song's emotional ambiguity. The image of something—or someone—residing "*under*" the narrator's "*skin*" introduces a quiet unease, a suggestion that this lingering presence may not be entirely welcome.

The pre-chorus introduces a pivotal shift, both thematically and tonally: "*And I don't mind if you take my body / No I don't mind, but please leave my soul.*" This line establishes a fragile boundary between physical surrender and the preservation of something more profound. It suggests a power dynamic, a plea, or a form of protection, but again, its intent remains ambiguous. Is this a willing act of surrender, a protest, or a strategy for survival?

The chorus continues to navigate internal contradiction: "*Cause I'm no good, but I'm not evil.*" This repeated phrasing articulates a refusal to be categorised in simplistic moral terms where the narrator admits fallibility but does not accept full condemnation. In the second chorus, however, a subtle shift occurs: "*I may be no good…*" introduces a tone of hesitation or internalised doubt. This variation could suggest a softening of the narrator's stance,

[53] Somatic memory refers to the way our bodies hold on to past traumas and stresses.

perhaps even a move toward self-forgiveness. Alternatively, it may reveal that the narrator has internalised an accusation—that the label "*no good*" may not have originated from within, but from an unspecified other. In this reading, the narrator is not only wrestling with self-image, but also with the echo of an external judgement that continues to shape their emotional landscape. The second verse deepens the sense of internal struggle with the line: *"I'm filled with regret, and to this day, I feel it."* This *"it"* remains unnamed—perhaps a specific event, a feeling, a choice, or an entire past—underscoring the ineffability of the emotion. The repetition of the bodily imagery from verse one links this emotional burden back to the physical realm, suggesting that the sense of regret is not just emotionally but somatically encoded.

The post-chorus section consists entirely of the repeated line: *"I don't mind, no I don't."* This mantra-like repetition, stretched across multiple lines, oscillates between numbness and surrender. It remains unclear whether the narrator's repetition signifies genuine or feigned indifference, emotional detachment, or surrender. Whereas the lack of variation in the line suggests emotional stasis, its persistent return may also be read as a strategy of self-soothing or dissociation.

Throughout the song, the narrator never clarifies the nature of the relationship at stake, nor do they attempt to resolve their own emotional contradictions. Instead, the lyrics construct a deliberately ambiguous and ominous emotional terrain in which memory, guilt, and identity blur together. The refusal to name the lyrical "*you*", the oscillation between physicality and metaphysical plea, and the repeated motifs of regret and moral ambivalence create a narrative space that is open, unstable, and resonant. In doing so, the song invites the listener not to decode a fixed story, but to inhabit its unresolved emotional state.

4.4.2 Artistic Intention and Production Concept

I Don't Mind began as a sonic exploration of identity and persona. Rooted in a lyrically open and narratively ambiguous text that fluctuates between corporality and spirituality, the song centres on a speaker haunted by memory, disoriented by regret, and detached from the self.

I was drawn to the idea of a voice that was already altered, marked by the aftermath of its own story, before the listener arrives. This shaped the production concept: Inspired by Bon Iver's harmoniser- and Auto-Tune- sculpted textures and Laurie Anderson's vocoder-inflected speech and vocal sampling in *O Superman* (1981), I imagined the vocal as emotionally resonant yet technologically alienated — a site of ambiguity. The idea was to use layering, sampling, pitch-shifting, and spectral effects as narrative devices, the production itself becoming the story.
One key concept I explored was the disembodied voice (cf. 1.2.2.3), and how opaque technological mediation can dislocate the voice from the body without diminishing its emotional impact. Timbre and register considerations were central to the aesthetic design. I was reminded of Zedd's (cf. 2.2.4 *The Vocoder*) observation that adding low-end frequencies to the voice—a range that humans, and particularly females, cannot naturally produce—holds especial affective power. This insight affirmed my intent to use processed, pitch-shifted doubles to create tension and emotional unease.

Through this hybrid vocal aesthetic, I set out to explore how tools typically associated with artificiality impact the perceived identity of the vocal persona, while also unlocking unexpected depths of expression and affect.

4.4.3 Technical Considerations and Reflective Process

Intro

The song opens with a layered, reverberant a cappella choral hum — a decision made at a late stage in the production process. Originally, the track opened with guitar and warped flageolets (as featured later in the arrangement), but I found that beginning with the choral vocal instantly established the emotional tone I was seeking, creating a suspended, almost sacred atmosphere, which aligns with the song's central themes of reflection, judgement, and penance. The choral texture was achieved by first recording and layering 30 vocal takes — including different vowel sounds and pitches at various distances from the microphone — and then panning them to different points across the stereo image (see appendix, *Figure B*). This approach was intended to emulate the spatial and tonal variety of a real choir, suggesting different singers, vocal registers, and physical positions within a resonant space, such as a church.

For the lead vocal that enters from the first verse, I used the *varispeed* method — recording my voice at a faster speed and higher pitch, then slowing the recording to match the project's tempo. This process resulted in darker vocal formants, causing the vocal persona to sound slightly "off" from the start. I chose to keep the effect subtle, applying only a 5% speed in-/decrease respectively. While the choral arrangement is placed in a lush, reverberant space, evoking an ambiguous sense of time and place, the lead vocal has less reverb on it, but enough still to make the vocal feel haunted from the start. Lastly, I applied a long-tailed reverb on the final word of the verse — *"skin"* — to mirror the sense of lingering feeling mentioned in the lyric.

Chorus

The Chorus is preceded by an atmospheric *reverse reverb sweep,* further distorting a sense of space and time. I originally planned for this to be done using a vocal take, however I already had recorded some atmospheric flageolets from the guitar that were well suited. During the first Chorus, a vocoder doubles the lead vocal (stereo spread for more clarity), adding layers of colour and complexity to the vocal sound — and more depth vocal persona. Additionally, select lines (*"but please leave my soul"* and *"cause I'm no good"*,

“evil" and "but I'm not evil”) are doubled with AutoTuned, marking the beginning of the de-humanising process.

Verse 2

Verse 2 continues in the same fashion as the chorus until the line *“under my skin”*, which marks a key turning point in the song. Here, an overt pitch-shifting effect is introduced, creating a monster-like vocal tone that reveals an even darker, more unsettling layer beneath the vocal persona’s surface. The lead vocal rises in pitch and intensity — as if straining, resisting, or losing control — precisely the moment at which the “monster voice” begins to emerge. From this point onward, the warped vocal layers become more present and consistent, giving the impression that they are gradually taking over. This subtle shift is reinforced by the removal of all breath sounds from the varispeeded lead vocal, contributing to a sense of growing unease and disembodiment (see appendix, *Figure C*).

Chorus 2

The second chorus features a lyrical shift from the line *"I'm just no good”* to *“I may be no good”* suggesting internalised doubt, or perhaps an echo of an accusation. This sense of external judgement folding into internal conflict is supported by increasingly present vocal artefacts, pushing the voice further into a digitally unstable dimension. At the same time, breathy, ethereal vocal layers enter underneath the fragmented vocal persona, adding to the ambiguity and tension. I recorded these using the close-mic’ing technique and low performance intensity to capture as much breathiness as possible (not a 30 person choir this time). The end of the Chorus introduces a rhythmic shift and with it, a sense of release.

Post-Chorus

The aforementioned release is present in the lead vocal as well, which has now gone fully digital, creating a sense of emotional override. The Post-Chorus thus marks the narrator’s final descent — or ascent — into de-humanisation. The lead vocal is doubled, bilaterally panned, and heavily pitch-corrected with a fast correction speed to reveal glitches and artefacts, creating the recognisable "AutoTune sound”. A formant-shifted “monster-double” (detuned -1200 cents) adds weight and distortion. The space becomes more open through increasing (automated) reverberation, yet also more distant and disembodied. Due to the

ambiguous nature of the lyrics (if you have to repeat the same sentence 12 times in a mantra-like fashion, do you really believe it?), I wanted to add another layer of tension to this final section. Through experimenting, I recorded some grainy "*ah-ah-ah*" vocal samples layered over ethereal breathy background vocals that grow in complexity until the very end. Within this push-and-pull between the corporal and the spiritual sphere, remorse and submission blur into an emotionally ambiguous, estranged vocal persona.

4.4.4 Findings

This case study demonstrates how vocal production can not only shape emotion and narrative, but ultimately transform the vocal persona itself. The progressive stripping of breath and natural vocal tone culminated in a highly processed, post-human vocal persona —an aesthetic choice aligned with the song's underlying themes of submission and emotional detachment.

Influenced by my research into vocal mediation, I deliberately embraced overt digital manipulation, something I would not have attempted prior to this project. Working with a limited set of tools—primarily Logic Pro's built-in plugins—I explored the expressive potential of artefacts, glitches, formant-shifting, and spatial widening to construct a voice that feels broken and steeped in sorrow.

Inspired by Moorefield's (2010) observation that the recording aesthetic has shifted from "the illusion of reality" to embracing "the reality of illusion," the spatial design of *I Don't Mind* reflects bolder choices, where multiple reverberant textures coexist within the same mix, creating a morphed sense of space.

Interestingly, despite the heavy use of Auto-Tune and manipulation at the end of the track, the voice still conveys a sense of remorse. This could possibly be linked to Brøvig-Hanssen and Danielsen's (2016) concept of the *naturalisation* of technological mediation in music (cf. 1.2.3.4). In today's listening culture, our perception of authenticity has adapted to accept (even expect) such radical forms of vocal mediation. Initially, listeners may focus on the audible traces of technology rather than the musical content itself; over time, these features become "naturalised," no longer perceived as intrusive but embraced as part of a new aesthetic norm.

It is difficult to imagine how abstract sonic ideas will behave in context until they are realised. Allowing space for trial and error was crucial to this process—underscoring the value of self-production as a flexible, iterative creative mode. Ultimately, I welcome external feedback, as fresh ears can offer insights that I, from within the work, can no longer access.

5 — Conclusion

This thesis set out to explore how contemporary vocal production aesthetics can enrich narrative and emotional expression in the self-produced singer-songwriter format, with particular attention to the mediated female voice. The central hypothesis — that *integrating contemporary pop vocal aesthetics into the traditionally sparse singer-songwriter genre can enhance emotional depth and narrative clarity without compromising authenticity* — was tested through both theoretical exploration and practice-based research. The process of producing and analysing three original songs allowed me to critically engage with key concepts from the literature whilst discovering how they resonate in real-world creative practice. This final chapter thus synthesises the findings that emerged across both the contextual and practice-based components of this project.

My research began with a hesitation: I feared that overt vocal processing and "opaque" mediation might be perceived as inauthentic within the singer-songwriter tradition. Traditionally celebrated for its intimacy, minimalism, and "raw" expression, this genre seemed to prescribe a limited sonic palette. However, tracing the historical evolution of the singer-songwriter and interrogating shifting ideas of authenticity in pop production (Chapter 2) helped reframe this anxiety. I began to see the singer-songwriter not as confined to acoustic purity, but as a narrative-driven artist — someone who can, and perhaps must, adapt their sonic language to reflect emotional complexity and contemporary listening habits. Contemporary music audiences are attuned to a highly mediated sound world; using production as part of the storytelling toolkit enables singer-songwriters to remain artistically relevant without compromising their core values of intimacy and lyrical focus.

Throughout this journey, theoretical frameworks from Lacasse (2000), Malawey (2020), (Hall, 1969), and others became not just analytical lenses but creative tools. Concepts such as proxemics, vocal staging, and the embodied/disembodied voice actively shaped how I approached production choices, from microphone technique and spatialisation to vocal layering and tuning artefacts. These frameworks gave me a language to understand how production mediates affect, identity, and storytelling — and how contrast, hybridity, and context can be strategically deployed to heighten emotional resonance.

Each of the three case study songs explored a different point along the spectrum of vocal mediation. *Forever Vow* relied on a purer vocal production aesthetic — using minimalist methods like close-mic'ing, layering, panning, reverb and echo — to preserve a sense of emotional vulnerability and narrative immediacy. Here, I discovered that reverb, when used with intention, can serve not only as an aesthetic embellishment or "glue" of a track, but as a dramaturgical device in its own right. Subtle shifts in spatialisation can convey different psychological states, and even even impact the temporal backdrop of a narrative, implying past events or lingering memories. Another key discovery was how small structural decisions in the vocal arrangement, such as omitting words from the lead vocal and letting the backing layers stand alone momentarily, could unlock a narrative shift that I hadn't originally envisioned - letting production not just support the story, but subtly rewrite it.

In *What Do You Expect*, I leaned into contrast and hybridity, juxtaposing raw lead vocals with processed textures, vocoders, and layered doubles to sonically reflect themes of inner conflict and emotional ambivalence. This track revealed how vocal production can embody contradiction by combining both "real" and "synthetic" sounds to reflect complex emotional landscapes. A key takeaway form this case study was how production unlocked *a whole other layer of narration*, existing in parallel with the lead vocal, simply by repurposing and manipulating original lyrical fragments.

I Don't Mind delved the furthest into technological mediation: pitch and formant shifting, varispeed, and subtle glitch artefacts were used to construct an unstable and fragmented vocal persona that conveyed dissociation and internal dissonance. These tools didn't just decorate the song — they *were* the story.

Across the production process of all three songs a key insight emerged: that when lyrical development is limited, sonic choices carry even more narrative weight. Furthermore, these case studies reaffirmed that vocal production is far more than a technical postscript to songwriting. It is a compositional act in its own right — a site where narrative, emotion, and identity are sculpted. This realisation echoed and reinforced what I had begun to uncover through my theoretical research: that authenticity is not a by-product of minimal intervention, but the result of intentional, meaningful creative choices. Sometimes the most emotionally truthful and resonant voice is one that has been digitally mediated, transformed, or distorted. The tools of contemporary production are not inherently inauthentic; they become authentic when used with purpose.

I became more confident experimenting with these possibilities once I found an artist identity (Bronty) that gave me room to explore — a name and creative space under which I could house diverse vocal expressions. Contextualising my position as a self-producing singer-songwriter not only clarified my aesthetic direction, but expanded it. The options, I realised, are vast, even within a simple home-studio environment with limited tools.

My basic home-studio setup did not hinder me; it proved to be a catalyst. Working within constraints pushed me to be more resourceful, more intentional, and ultimately more curious. Had I used automated vocal alignment tools or vocal chain presets, I might never have discovered the expressive potential of imperfection — the slightly off harmonies, the (at times) rough edges of double-tracked vocals that shimmer and feel like two conflicting sides of the same vocal persona. These "flaws" often became the moments of greatest emotional impact. Furthermore, had I recorded in a professional analog studio, I may have leaned more into conventional reverb aesthetics; instead, in-the-box automation and digital plugins led me to craft spatial shifts on a near-phoneme (to use a linguistic term) level — "micro-staging" the voice to match lyrical or emotional nuance.

Importantly, self-producing gave me full control over how I wanted to represent myself as an artist. But it also taught me that production is not always best done alone. Creative agency includes the freedom to invite others into the process — to collaborate, co-produce, or simply seek a second opinion. Working with Rami during this project reminded me that collaboration can offer new perspectives without undermining artistic authorship. Deciding when and how to bring others in is itself an act of creative autonomy.

Another valuable habit I developed during this project was regularly looking up songwriting and production credits when I found a track interesting — paying closer attention to which artists were involved in their own production processes, and how that shaped their sound. This informed my thinking around live playability. I initially struggled with the question: Should my recordings reflect what I can perform live? I've come to see that the recorded and live versions of a song need not be identical. Rather, they can reveal different facets of the same story. The studio affords a different kind of expressivity — one where personas, textures, and sonic illusions can flourish. The stage offers immediacy, connection, and

reinterpretation. Both are valid. Some artists even release multiple versions to showcase that evolution which is an idea I now embrace.

Though this project focused specifically on the vocal with regard to its production, staging, and emotional and narrative impact, it was always developed in context with the instrumental arrangement. While I did not explicitly analyse instrumental production in this thesis, I was constantly aware of how timbre, rhythm, and spatial relationships between voice and instruments shaped the overall aesthetic. In hybrid arrangements especially, the voice must find its place within a dynamic, shifting sonic landscape. Vocal decisions are never made in isolation — especially following a "bottom up" approach, they respond to the instrumental backdrop.

My exploration further affirmed that the historical underrepresentation of women in music production makes self-production a crucial space for female singer-songwriters to assert creative agency, challenge persisting industry norms, and shape authentic artistic identities. While the term "singer-songwriter" traditionally emphasises writing and performing, the "self-producing singer-songwriter" — or "artist-producer" (cf. Wolfe, 2019) — reflects a more integrated creative identity, where composition and production processes are inseparable. That said, this holistic approach isn't always efficient. The freedom to endlessly revise, re-record, and reshape can be creatively rich but also exhausting, resulting in decision fatigue and creative burnout or overwhelm. Without collaborators or external deadlines, it becomes easy to lose perspective. Moving forward, I will embrace both approaches: solo self-production when the vision is clear and and the process intuitive; co-production when a fresh perspective is needed, or when efficiency matters.

Ultimately, this thesis has shown me that contemporary vocal production is not a threat to authenticity but a powerful extension of artistic expression. It allows singer-songwriters to stay emotionally honest while also staying sonically relevant[54]. For the self-producing female

[54] To quote Music Producer Mike Senior (2011), "If you're producing seriously chart-oriented music, then another useful guideline is to make sure that there's some interesting little feature happening every three to five seconds […] Chart music needs to command attention continuously if it's to cater for the attention span of young, media-bombarded music fans" (pp. 114/115).

singer-songwriter navigating today's popular music culture, this expanded toolkit is both a creative asset and a form of empowerment.

Above all, this project reminded me that process is everything. Each song taught me something new — not only about production, but about myself as an artist. There is no final version, no perfect take. There is only the evolving relationship between the voice, the story, and sound; and often it is the "messy middle" of the process where transformation happens.

That, to me, is what it means to find your voice.

Final Outlook

As we move deeper into an era shaped by automation and algorithmic logic, the rise of generative AI in music production raises pressing questions about artistic identity, human expression, and what it means to be truly "authentic." Projects like *Velvet Sundown*, a fully AI-generated band, exemplify a new frontier that simultaneously fascinates and unsettles. But rather than replacing the human voice, such developments may in fact amplify our collective craving for intimacy, imperfection, and real emotional presence. The turn toward a hyper-intimate vocal aesthetic over the past few years is evidence of this. In a scroll-heavy, sonically saturated music culture, the self-produced singer-songwriter's voice (literal and metaphorical) remains a vital counterpoint: flawed, embodied, emotionally raw.

While AI may support the more strategic or logistical aspects of being an artist today—suggesting social media strategies, reading platform algorithms, or even helping with time management, it cannot replicate the inward-facing, self-therapeutic process that underpins songwriting or art in general. The act of crafting a song, of shaping raw emotion into melody and texture, is not just a creative output; it is a form of emotional processing. For many artist-producers, making music is not simply about the product—it is a way of making sense of the world and of oneself.

References

Accornero, G. (2022). What Does ASMR Sound Like? Composing the Proxemic Intimate Zone in Contemporary Music. *Contemporary Music Review*, *41*(4), 337–357. https://doi.org/10.1080/07494467.2022.2087377

Alberhasky, M., & Durkee, P. K. (2024). Songs tell a story: The Arc of narrative for music. Plos one, 19(5), e0303188.

Alleyne, M. (2020). Authenticity in Music Production. In Zagorski-Thomas, S., & Bourbon, A. (Eds.). *The Bloomsbury Handbook of Music Production*, 19-32. London: Bloomsbury Academic.

Anderton, C. (2018). *How to Record and Mix Great Vocals.* Hal Leonard Books.

Apolloni, A. (2017). Vocal authorship: Marianne Faithfull and "Sister Morphine". In Williams, J., & Williams, K. (Eds.). *The Singer-Songwriter Handbook*, 213-222. Bloomsbury Publishing UK.

Appel, N. (2017). The Singer-Songwriter Disposition. In Williams, J., & Williams, K. (Eds.). *The Singer-Songwriter Handbook*. Bloomsbury Publishing UK.

Auvinen, T. (2020). Creative communities of practice: Role delineation in record production in different eras and across different genres and production settings. In Zagorski-Thomas, S., & Bourbon, A. (Eds.). *The Bloomsbury Handbook of Music Production, 161-176.* London: Bloomsbury Academic.

Bal, M. (1997). *Narratology Introduction to the Theory of Narrative*. (2nd ed.) University of Toronto Press.

Barthes, R. (1974) *S/Z*. Translated by Richard Miller. New York: Hill and Wang.

Biddle, I. (2007). "The Singsong of Undead Labor": Gender Nostalgia and the Vocal Fantasy of Intimacy in the 'New' Male Singer/Songwriter. In Jarman-Ivens, F. (Ed.). *Oh Boy!: Masculinities and Popular Music,* 125. New York and London: Routledge.

Bell, A. P. (2018). *Dawn of the DAW: The studio as musical instrument*. Oxford University Press.

Bennett, S. (2016.) Behind the magical mystery door: history, mythology and the aura of Abbey Road Studios. *Popular Music,* Volume 35/3, 396–417. Cambridge University Press.

Boak, S. (2016) The female singer-songwriter in the 1990s. In Williams, K., & Williams, J. (Eds.). *The Cambridge Companion to the Singer-Songwriter,* 257-264. Cambridge University Press.

Borgdorff, H. (2012). *The conflict of the faculties: Perspectives on artistic research and academia*. Leiden University Press.

Brøvig-Hanssen, R. (2018). 'Listening to or through technology: Opaque and transparent mediation in popular music'. In Bates, E. & Bennett, S. (Eds.) *Critical Approaches to the Production of Music and Sound,* 195-210. London: Bloomsbury Academic. https://doi.org/10.5040/9781501332074.0021

Brøvig-Hanssen, R., and Danielsen, A. (2016). *Digital Signatures. The Impact of Digitization on Popular Music Sound.* Cambridge, MA, The MIT Press.

Burgess, R. J. (2013). *The Art of Music Production: The Theory and Practice*. (4th ed). Oxford University Press.

Burgess, R., & Toulson, R. (2017). Singer-Songwriter Meets Music Production. In Williams, J. & Williams, K. (Eds.). *The Singer-Songwriter Handbook*, 91-111. Bloomsbury Publishing UK.

Busby, L. (2017). Collage, cut up, and pop mantras: Post-digital approaches to songwriting. In Williams, J., & Williams, K. (Eds.). *The Singer-Songwriter Handbook*, 113-133. Bloomsbury Publishing UK.

Dibben, N. (2013). The Intimate Singing Voice: Auditory Spatial Perception and Emotion in Pop Recordings. In Zakharine, D. & Meise, N. (Eds.). *Electrified Voices. Medial, Socio-Historical and Cultural Aspects of Voice Transfer.* Vandenhoeck & Ruprecht Verlag (V&R Unipress).

Frayling, C. (1993). Research in Art and Design. *Royal College of Art Research Papers series,* 1(1). London: Royal College of Art.

Frith, S., & Zagorski-Thomas, S. (Eds.) (2012). *The Art of Record Production: An Introductory Reader for a New Academic Field*. Routledge.

Gaston-Bird, L. (2019). *Women in audio*. Focal Press.

Görne, T. (2015). *Tontechnik*. (4th ed). Carl Hanser Verlag München.

Görne, T. (2017). *Sound Design: Klang, Wahrnehmung, Emotion.* Carl Hanser Verlag München.

Harding, P. (2020) Desktop Production and Commerciality. In Zagorski-Thomas, S., & Bourbon, A. (Eds.). *The Bloomsbury Handbook of Music Production,* 307. Bloomsbury Publishing USA.

Hall, E. T. (1969). *The hidden Dimension*. Anchor.

Harding, C. (Host). (24th December 2024) Auto-Tune always and forever. [Audio podcast]. In *Switched on Pop.* Spotify. https://open.spotify.com/episode/2SDgEyDL0MMNvtEhs9sP3h?si=bd343db4334846da

Holmes, J. A. (2023) Billie Eilish and the Feminist Aesthetics of Depression: White Femininity, Generation Z, and Whisper Singing. *Journal of the American Musicological Society* 1 December 2023; 76 (3): 785–829. https://doi.org/10.1525/jams.2023.76.3.785 [last accessed on 10th July 2025]

Kealy, E. R. (1979). From Craft to Art: The Case of Sound Mixers and Popular Music. *Sociology of Work and Occupations*, *6*(1), 3-29. https://doi.org/10.1177/009392857961001 [last accessed on 10th July 2025]

Lacasse, S. (2000). "*Listen to My Voice": The Evocative Power of Vocal Staging in Recorded Rock Music and Other Forms or Vocal Expression.* PhD thesis, University of Liverpool.

Leyshon, A. (2009). The Software Slump?: Digital Music, the Democratisation of Technology, and the Decline of the Recording Studio Sector within the Musical Economy. *Environment and Planning a Economy and Space*, *41*(6), 1309–1331. https://doi.org/10.1068/a40352 [last accessed on 10th July 2025]

Malawey, V. (2020). *A Blaze of Light in Every Word: Analyzing the Popular Singing Voice.* Oxford University Press.

Marc, I., Green, S. (Eds.) (2016). *The Singer-Songwriter in Europe: Paradigms, Politics and Place*. Routledge.

Marrington, A. (2017). Composing with the digital audio workstation. In Williams, J., & Williams, K. (Eds.). *The Singer-Songwriter Handbook*, 77-89. Bloomsbury Publishing UK.

Middleton, R. (1990). *Studying Popular Music*. Milton Keynes: Open University Press.

Moore, A. (2002). Authenticity as Authentication. *Popular Music*, *21*(2), 209–223. http://www.jstor.org/stable/853683 [last accessed on 10th July 2025]

Moorefield, V. (2010). *The Producer as Composer. Shaping the Sound of Popular Music.* Cambridge, Mass.: MIT Press.

Moylan, W. (1992). *The Art of Recording: The Creative Resources of Music Production and Audio*. Van Nostrand Reinhold.

Muchitsch, V. (2023). "Genrefluid" Spotify Playlists and Mediations of Genre and Identity in Music Streaming. *IASPM Journal*, *13*(3), 48-65.

Negus, K. (2012). Narrative, interpretation, and the popular song. *The Musical Quarterly*, *95*(2-3), 368-395.

Olsson, C. (2016). *Out of Touch? A Study of the Technology and Framework That Is Supposedly Killing Music.* Bachelor's Thesis, Lund University. https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8626616&fileOId=8626629 [last accessed on 18th July 2025]

Petrusich, A. (2021). Genre is Disappearing. What Comes Next? *The New Yorker, 8th March.* https://www.newyorker.com/magazine/2021/03/15/genre-is-disappearing-what-comes-next. [last accessed on 15th July 2025]

Rada, M., Mason, R. D., & De Sena, E. (2025). Immersive Music Production Workflows: An Ethnographic Study of Current Practices. *158th Audio Engineering Society Convention*. Audio Engineering Society.

Randle, Q. & Evans, K. (2013). So what does" Set Fire To The Rain" really mean? A typology for analyzing pop song lyrics using narrative theory and semiotics. *MEIEA Journal,* 13(1), 125-147. https://doi.org/10.25101/13.7 [last accessed on 15th July 2025]

Reddington, H. (2021). *She's at the Controls: Sound Engineering, Production and Gender Ventriloquism in the 21st Century.* Equinox Publishing (UK).

Senior, M. (2011). *Mixing Secrets For The Small Studio.* Focal Press.

Shields, L., Moore, A., & Dewey, C. (2024). A Content Analysis Study of Distortion's Influence on Drum and Bass Music Production. *IASPM Journal*, *14*(2), 182-202.

Shuker, R. (2008). *Understanding Popular Music Culture (*3rd ed.), 53. Routledge.

Shumway, D. (2016). The emergence of the singer-songwriter. In Williams, K., & Williams, J. (Eds.). *The Cambridge Companion to the Singer-Songwriter,* 11-20. Cambridge University Press.

Smith, S. L., Pieper, K., Hernandez, K., & Wheeler, S. (2025). Inclusion in the Recording Studio?: Gender & Race/Ethnicity of Artists, Songwriters & Producers across 1,300 Popular Songs from 2012–2024. *USC Annenberg Inclusion Initiative.*

Stiegler, Z., & Campbell, T. (2023). *Musical Intimacy*. https://doi.org/10.5040/9781501372285 [last accessed on 10th June 2025]

Tagg, P. (2012). *Music's Meanings: a modern musicology for non-musos*. The Mass Media's Scholar's Press.

Théberge, P. (2020). Transitions: The History of Recording Technology from 1970 to the Present. In Zagorski-Thomas, S., & Bourbon, A. (Eds.). *The Bloomsbury Handbook of Music Production*, p. 84. Bloomsbury Publishing USA.

Thomas, A. (2022). *Digital vocal editing, plug-ins and the fetishism of voice and technology in 21st-century Western commercial popular music production* [PhD thesis]. University of Oxford.

Watson, A. (2015). *Cultural Production in and Beyond the Recording Studio*. New York: Routledge.

Williams, J., & Williams, K. (Eds.). (2017). *The Singer-Songwriter Handbook*. Bloomsbury Publishing UK.

Williams, K., & Williams, J. (Eds.). (2016). *The Cambridge Companion to the Singer-Songwriter.* Cambridge University Press.

Wolfe, P. (2019). *Women in the studio: Creativity, control and gender in popular music sound production*. Routledge.

Young, M. (2019). Proximity/Infinity: The Mediated Voice in Mobile Music. In Eidsheim, N. S., & Meizel, K. (Eds.). (2019). *The Oxford handbook of voice studies.* 403-418. Oxford University Press.

Zagorski-Thomas, S. (2014). Musical Meaning and the Musicology of Record Production. In Helms, D., & Phleps, T. (Eds.). *Black Box Pop.* 135-148. Bielefeld: transcript Verlag. https://doi.org/10.14361/transcript.9783839418789.135 [last accessed on 18th May 2025]

Appendices

Discography

Adele. (2008). Chasing pavements [Song].
Adele. (2011). Set fire to the rain [Song].
Adele. (2015). *Love in the dark* [Song].
Anderson, L. (1981). *O superman* [Song].
Bareilles, S. (2007). *Gravity* [Song].
Bareilles, S. (2007). *Lovsong* [Song].
Bareilles, S. (2010). *King of everything* [Song].
Betts, R. (2023). *Sober* [Song].
Blake, J. (2011). *Limit to your love* [Song].
Blake, J. (2011). *The Wilhelm scream* [Song].
Bridgers, P. (2018). *Motion sickness* [Song].
Bridgers, P. (2018). *Scott street* [Song].
Burns, C. (2024). *People pleaser* [Song].
Bush, K. (1985). *Running up that hill* [Song].
Burna Boy, & Sheeran, E. (2022). *For my hand* [Song].
Callum Scott. (2018). *Dancing on my own* [Song].
Cohen, L. (1967). *One of us cannot be wrong* [Song].
Collier, J. (2019). *Time alone with you* [Song].
DiFranco, A. (1995). *Not a pretty girl* [Song].
Dixon, E. (2022). *Green grass* [Song].
Dodie. (2019). *She* [Song].
Eilish, B. (2019). *Bad guy* [Song].
Eilish, B. (2019). *Bury a friend* [Song].
Eilish, B. (2019). *Everything I wanted* [Song].
Eilish, B. (2019). *I love you* [Song].
Eilish, B. (2019). *When the party's over* [Song].
Eilish, B. (2019). *You should see me in a crown* [Song].
Elliot Smith. (1997). *Between the bars* [Song].
Feist. (2011). *How come you never go there* [Song].
Fitzsimmons, W. (2010). *I kissed a girl* [Song].
FKA Twigs. (2019). *Cellophane* [Song].
González, J. (2003). *Heartbeats* [Song].
Grannis, K. (2014). *Fire* [Song].
Heap, I. (2005). *Hide and seek* [Song].
Iron & Wine. (2016). *Time after time* [Song].
Keane. (2004). *Somewhere only we know* [Song].
King, C. (1971). *So far away* [Song].
King Princess. (2018). *1950* [Song].
King Princess. (2019). *If you think this is love* [Song].
King Princess. (2022). *Cursed* [Song].
King Princess. (2022). *Winter is hopeful* [Song].
Lipa, D. (2020). *Pretty please* [Song].
LMFAO. (2011). *Party rock anthem* [Song].

Lorde. (2017). *Liability* [Song].
Lucy Rose. (2012). *Shiver* [Song].
Maroon 5. (2003). *She will be loved* [Song].
Maro. (2022). *Saudade* [Song].
Mars, B. (2012). *Treasure* [Song].
Matilda Mann. (2021). *Glass ceiling* [Song].
McAlpine, L. (2022). *Ceilings* [Song].
McAlpine, L. (2022). *Reckless driving* [Song].
Mitchell, J. (1971). *A case of you* [Song].
Mitski. (2023). *Bug like an angel* [Song].
Morissette, A. (1995). *You oughta know* [Song].
Obadiah Parker. (2008). *Hey ya* [Song].
Paloma, P. (2023). *Labour* [Song].
Puth, C. (2018). *Attention* [Song].
Radiohead. (1992). *Creep (Acoustic version)* [Song].
Ridings, F. (2019). *Lost without you* [Song].
Rodrigo, O. (2021). *Drivers license* [Song].
Run River North. (2017). *Mr Brightside* [Song].
Sara Bareilles. (2007). *Gravity* [Song].
Scott, C. (2018). *Dancing on my own* [Song].
Sheeran, E. (2017). *Shape of you* [Song].
Sheeran, E. (2020). *Afterglow* [Song].
Sheeran, E. (2021). *Bad habits* [Song].
Sheeran, E. (2025). *Sapphire* [Song].
Sia. (2016). *The greatest* [Song].
Sia. (2016). *Unstoppable* [Song].
Smith, E. (1997). *Between the bars* [Song].
Swift, T. (2008). *Love story* [Song].
Swift, T. (2020). *Cardigan* [Song].
Swift, T. (2022). *Vigilante shit* [Song].
T-Pain. (2007). *Buy U a drank* [Song].
Tate McRae. (2023). *Greedy* [Song].
The Beatles. (1966). *Revolver* [Album].
The Beatles. (1967). *Lucy in the sky with diamonds*
The Killers. (2004). *Mr Brightside* [Song].
Warren, A. (2025). *Ordinary* [Song].
Zedd, Morris, M., & Grey. (2018). *Middle* [Song].

Comparative Overview of Singer–Songwriter Categories

Table A: Comparative overview of Singer-Songwriter Categories as per Chapter 2.1

Category	Definition/ Characteristics	Examples	Values	Vocal Aesthetic	Technological Mediation
Genre	Musical tradition from the 1960s–70s, rooted in confessional lyricism, acoustic instrumentation, and emotional truth, often performed and recorded in with backing band.	Joni Mitchell, Bob Dylan, James Taylor, Carole King	Emphasis on personal storytelling, emotional honesty, and sincerity.	Simple, acoustic arrangements that support and foreground lyrics. Natural-realistic vocal sound.	Generally low tolerance for heavy processing; favors natural, unprocessed sound to preserve "authenticity." Bob Dylan turning electric caused an uproar of dismay amongst followers (at first)
Identity	Artist as self-authoring creator writing and performing own material, across genres and production styles.	Ed Sheeran, Imogen Heap, Phoebe Bridgers	Creative control, personal authorship, and expressive individuality prioritised.	Varied sonic approaches; may blend genres and production styles, but retain personal narrative focus/ coherent narrative persona.	Flexible acceptance of digital tools and effects, so long as they serve expressive goals and narrative clarity.
Style	Sonic approach emphasising voice and lyrics with often minimal or sparse arrangements that foster intimacy and direct listener engagement.	Acoustic cover songs or sparse, arrangements foregrounding a gentle voice, possible across different genres. Examples: Radiohead *Creep - Acoustic* Version (1992); William Fitzsimmons *I Kissed A Girl* (2010)	Values sparse arrangements, oftentimes just vocal and solo acccompaniment, intimacy, directness, and emotional connection with the listener.	Use of vocal techniques like doubling, reverb, layering, to enhance closeness and emotional depth.	Openness to subtle production effects that enhance vocal presence and intimacy without overshadowing content.

Original Song Lyrics

Forever Vow - Bronty

Verse 1
Well I don't think you're feeling it right now
Been trying my best to fix this
but I don't know how

Pre-Chorus
Promised that I'd remind you should you forget
That you said forever and forever's not over yet

Chorus
All that I am I give to you
From this day forth, together, we'll conquer the world anew
And all that you are you give to me
And we'll take turns with cooking and making pots of tea

Verse 2
But I don't think you're feeling it right now
Been trying my best to fix this
but I don't know how

Pre-Chorus
Promised that I'd remind you should you forget
That you said forever and forever's not over yet

Chorus
All that I am I give to you
From this day forth, together, we'll conquer the world anew
And all that you are you give to me
And we'll take turns with cooking and making pots of tea

Interlude

Bridge
All that's mine is yours
All that's yours is mine
For better or for worse
No matter the tide
I'll be by your side

Verse 3
Well I don't think you're feeling it right now
Think I'm out of ways to fix this
Could you help me out?

Pre-Chorus
Promised that I'd remind you should you forget
That you said forever and forever's not over yet

What Do You Expect - Bronty

Verse 1
Don't you know what you do to me
When you look at me like that
Shivers run right through me
when you touch me like that

Verse 2
I'm trying not to think
about the things you said
All the things you said, pictures in my head
Messing with my mind, messing with me, messing with me

Chorus
But you and I both know, this shouldn't be happening,
shouldn't be happening right now, no
But you and I both know, that this shouldn't be happening,
shouldn't be happening right now, no

Interlude
Uhh-Uhhh-Uhh-uh

Verse 3
What do you expect when you look at me like that,
you look at me like that?
What do you expect when you're touching me like that,
you're touching me like that?
I want you in my bed, want you in my bed

Chorus
But you and I both know, this shouldn't be happening,
shouldn't be happening right now, no
But you and I both know, that this shouldn't be happening,
shouldn't be happening right now, no

Interlude
Uhh-Uhhh-Uhh-uh

Bridge
So I'll just pretend that we're only good friends
and I hope my thoughts will follow too
So I'll just pretend that we're only good friends
and I hope my thoughts will follow, soon

Outro
Mmhm mhhh mhh mhhh

I Don’t Mind - Bronty

Verse 1
I remember you, I still feel you
In my bones, in my body
In my blood, under my skin

Chorus
And I don’t mind if you take my body
No I don’t mind, but please leave my soul
Cause I’m no good, but I’m not evil
I’m just no good, but I’m not evil

Verse 2
I’m filled with regret, and to this day, I feel it in
In my bones, in my body
In my blood, under my skin

Chorus
And I don’t mind if you take my body
No I don’t mind, but please leave my soul
Cause I’m no good, but I’m not evil
I may be no good, but I’m not evil

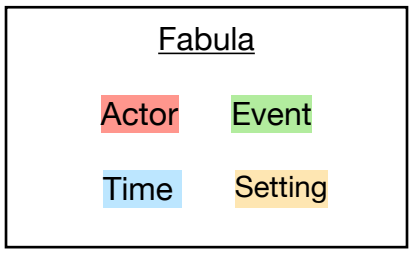
Post-Chorus
I don’t mind, no I don’t
I don’t mind, no I don’t
I don’t mind, no I don’t

I don’t mind, no I don’t
I don’t mind, no I don’t
I don’t mind, no I don’t, no I don’t

I don’t mind, no I don’t
I don’t mind, no I don’t
I don’t mind, no I don’t, no I don’t

I don’t mind, no I don’t
I don’t mind, no I don’t
I don’t mind, no I don’t

Fabula Analysis Example *Forever Vow* (part I)



Verse 1
Well I don't think - you're feeling it right now
Been trying my best to fix this
But I don't know how

Pre-Chorus 1
Promised that I'd remind you should you forget
That you said forever and forever's not over yet

Chorus 1
All that I am I give to you
From this day forth, together, we'll conquer the world anew
And all that you are you give to me
And we'll take turns with cooking and making pots of tea

Verse 2
But I don't think you're feeling it right now
Been trying my best to fix this
but I don't know how

Pre-Chorus 2
Promised that I'd remind you should you forget
That you said forever and forever's not over yet

Chorus 2
All that I am I give to you
From this day forth, together, we'll conquer the world anew
And all that you are you give to me
And we'll take turns with cooking and making pots of tea

Interlude

Bridge
All that's mine is yours
All that's yours is mine
For better or for worse
No matter the tide
I'll be by your side

Verse 3
Well I don't think you're feeling it right now
Think I'm out of ways to fix this
Could you help me out

Pre-Chorus (Outro)
Promised that I'd remind you should you forget
That you said forever and forever's not over yet

Fabula Analysis Example *Forever Vow* (part 2)

Actors:

- **Narrator** (the "I" voice): Someone committed to a relationship that feels like it's faltering. They're emotionally invested, trying to save the relationship, holding on to past memories, and initiating communication. They appear vulnerable and direct, yet patient and gentle.
- **Addressee** (the "you"): The romantic partner who appears emotionally distant or disengaged. They once made a vow ("you said forever") but are now withdrawn or emotionally unavailable, prompting the narrator's concern and repeated attempts to reconnect.

Event:
The core event is the slow emotional disconnection or detachment of the addressee in a long-term relationship. The narrator is actively trying to fix this, invoking shared memories, vows, and domestic routines to reignite connection and repair the emotional bond.

Setting:
There are two layers of setting woven into the lyrics:

- Domestic/Everyday Setting
    - Implied by lines like "cooking and making pots of tea," shared routines, and emotional repair attempts.
    - Suggests a lived-in, long-term relationship with physical and emotional closeness.

- Ceremonial/Implied Wedding Setting
    - Introduced by the chorus and bridge language resembling traditional wedding vows:
      "All that I am I give to you," "From this day forth," "For better or for worse"

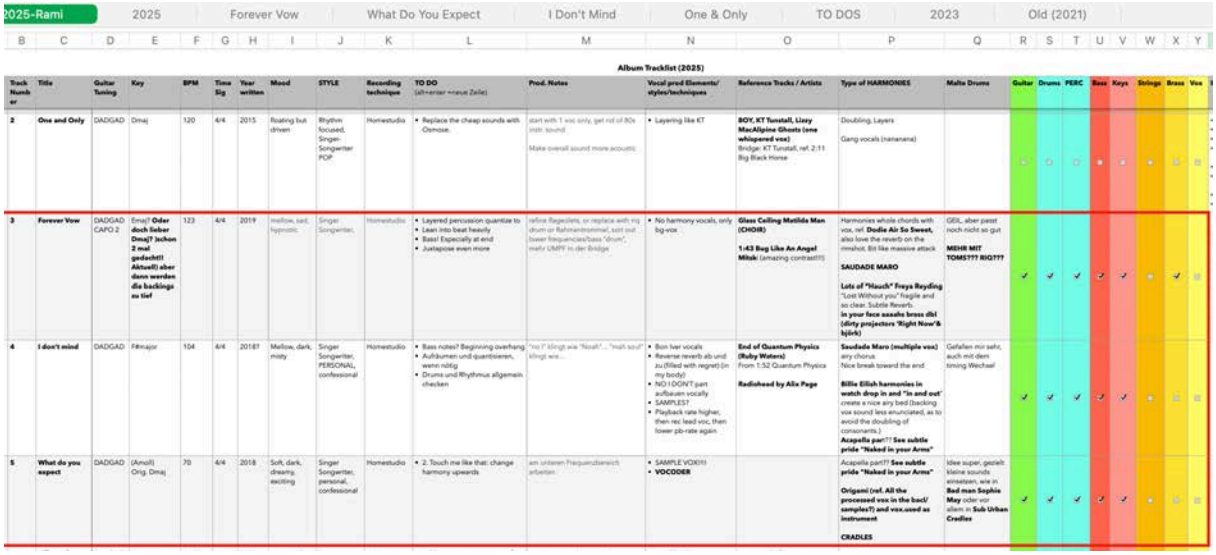
    - This invokes either:
        - A memory of a past wedding,
        - A metaphorical vow renewal, or
        - A contrast between past formal promises and present emotional distance.
    - This ceremonial space functions symbolically—reminding the addressee of shared ideals and lifelong commitment.

Time:

- The present dominates (repeated "I don't think you're feeling it right now").
- There are references to past promises ("you said forever") and an implied hopeful future ("we'll conquer the world anew"), showing a temporal tension between past commitment, present doubt, and future hope.

Reading: Closed Reading / Undefined Narrative

Production Plan - Work in Progress



Album Tracklist (2025)

Track Number	Title	Guitar Tuning	Key	BPM	Time Sig	Year written	Mood	STYLE	Recording technique	TO DO	Prod. Notes	Vocal prod Elements/ styles/techniques	Reference Tracks / Artists	Type of HARMONIES	Mahe Drums	Guitar	Drums	PERC	Bass	Keys	Strings	Brass	Vox
2	**One and Only**	DADGAD	Dmaj	120	4/4	2015	floating but driven	Rhythm focused, Singer-Songwriter POP	Homestudio	• Replace the cheap sounds with Osmose.	start with 1 voc only, get rid of 80s instr. sound Make overall sound more acoustic	• Layering like KT	**BOY, KT Tunstall, Lizzy MacAlpine Ghosts (one whispered vox)** Bridge: KT Tunstall, ref. 2:11 Big Black Horse	Doubling, Layers Gang vocals (nananana)								
3	**Forever Vow**	DADGAD CAPO 2	Emaj7 **Oder doch lieber Dmaj7 (schon 2 mal gedacht!! Aktuell) aber dann werden die backings zu tief**	123	4/4	2019	mellow, sad, hypnotic	Singer Songwriter,	Homestudio	• Layered percussion quantize to • Lean into beat heavily • Bass! Especially at end • Juxtapose even more	refine fingerstyle, or replace with rig drum or Rahmentrommel, sort out lower frequencies/bass "drum", meh/ UMPF in der Bridge	• No harmony vocals, only bg-vox	**Glass Ceiling Matilda Man (CHOIR)** **1:43 Bug Like An Angel Mitski** (amazing contrast!!!)	Harmonies whole chords with vox, ref. **Dodie Air So Sweet,** also love the reverb on the rimshot. Bit like massive attack **SAUDADE MARO** **Lots of "Hauch" Freya Reyding** "Lost Without you" Fragile and so clear. Subtle Reverb. **in your face aaaahs brass dbl (dirty projectors 'Right Now' & björk)**	GEIL, aber passt noch nicht so gut **MEHR MIT TOMS??? RIQ???**							
4	**I don't mind**	DADGAD	F#major	104	4/4	20187	Mellow, dark, misty	Singer Songwriter, PERSONAL, confessional	Homestudio	• Bass notes? Beginning overhang • Aufräumen und quantisieren, wenn nötig • Drums und Rhythmus allgemein checken	"no I" klingt wie "Noah"... "mah soul" klingt wie...	• Bon Iver vocals • Reverse reverb ab und zu (filled with regret) (in my body) • NO I DON'T part aufbauen vocally • SAMPLES? • Playback rate higher, then rec lead voc, then lower pb-rate again	**End of Quantum Physics (Ruby Waters)** From 1:52 Quantum Physics **Radiohead by Alix Page**	**Saudade Maro (multiple vox)** airy chorus Nice break toward the end **Billie Eilish harmonies in watch drop in and "in and out"** create a nice airy bed (backing vox sound less enunciated, as to avoid the doubling of consonants.) **Acapella part?? See subtle pride "Naked in your Arms"**	Gefallen mir sehr, auch mit dem timing.Wechsel							
5	**What do you expect**	DADGAD	(Amoll) Orig. Dmaj	70	4/4	2018	Soft, dark, dreamy, exciting	Singer Songwriter, personal, confessional	Homestudio	• 2. Touch me like that: change harmony upwards	am unteren Frequenzbereich arbeiten	• SAMPLE VOX!!! • **VOCODER**		Acapella part?? **See subtle pride "Naked in your Arms"** **Origami (ref. All the processed vox in the bac/ samples?) and vox.used as instrument** **CRADLES**	Idee super, gezielt kleine sounds einsetzen, wie in **Bed man Sophie May** oder vor allem in **Sub Urban Cradles**							

DAW Session Screenshots

Figure A: *Forever Vow* Lead Vocal Comp. (cf. p. 78)

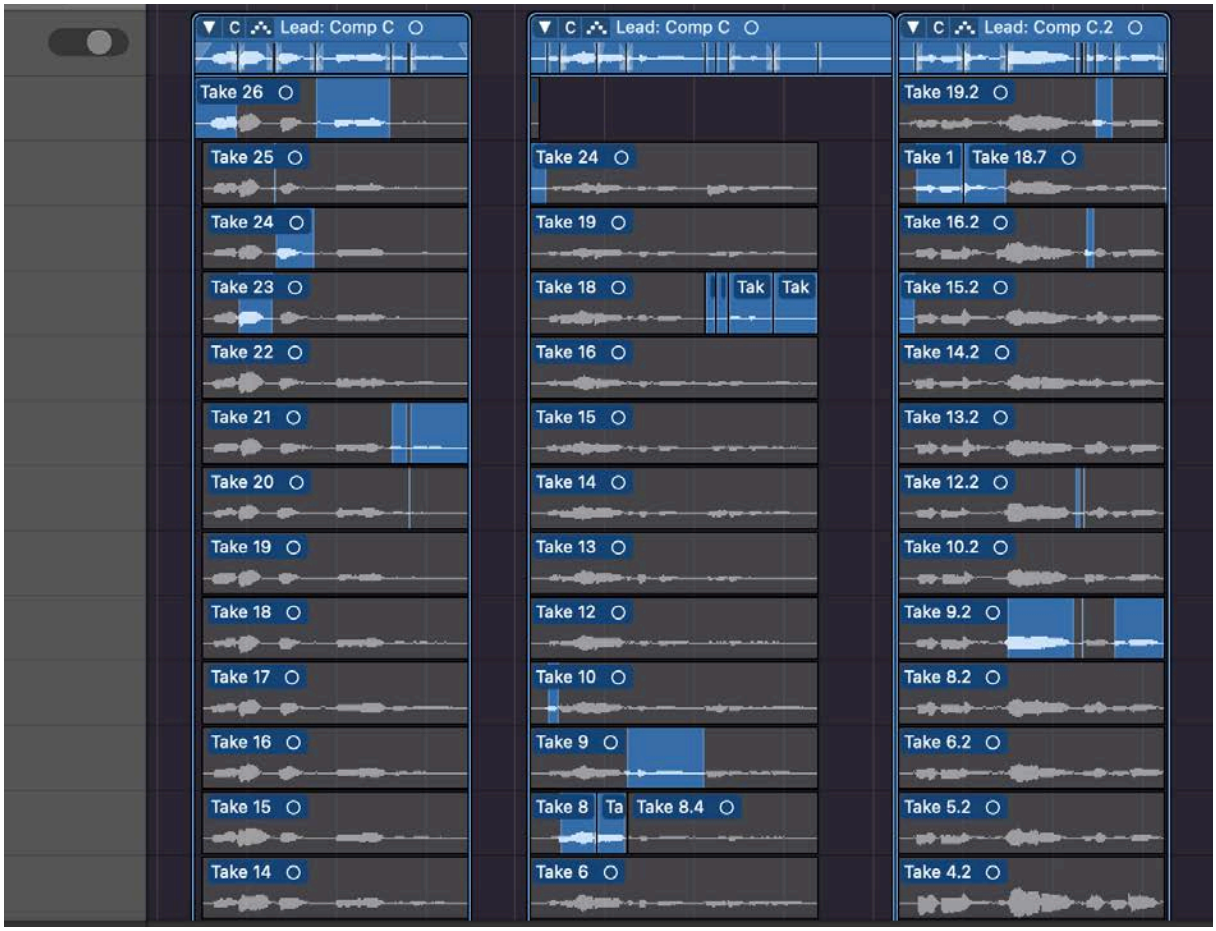


Figure B: *I Don't Mind* Intro Choral Layers (cf. p. 94)

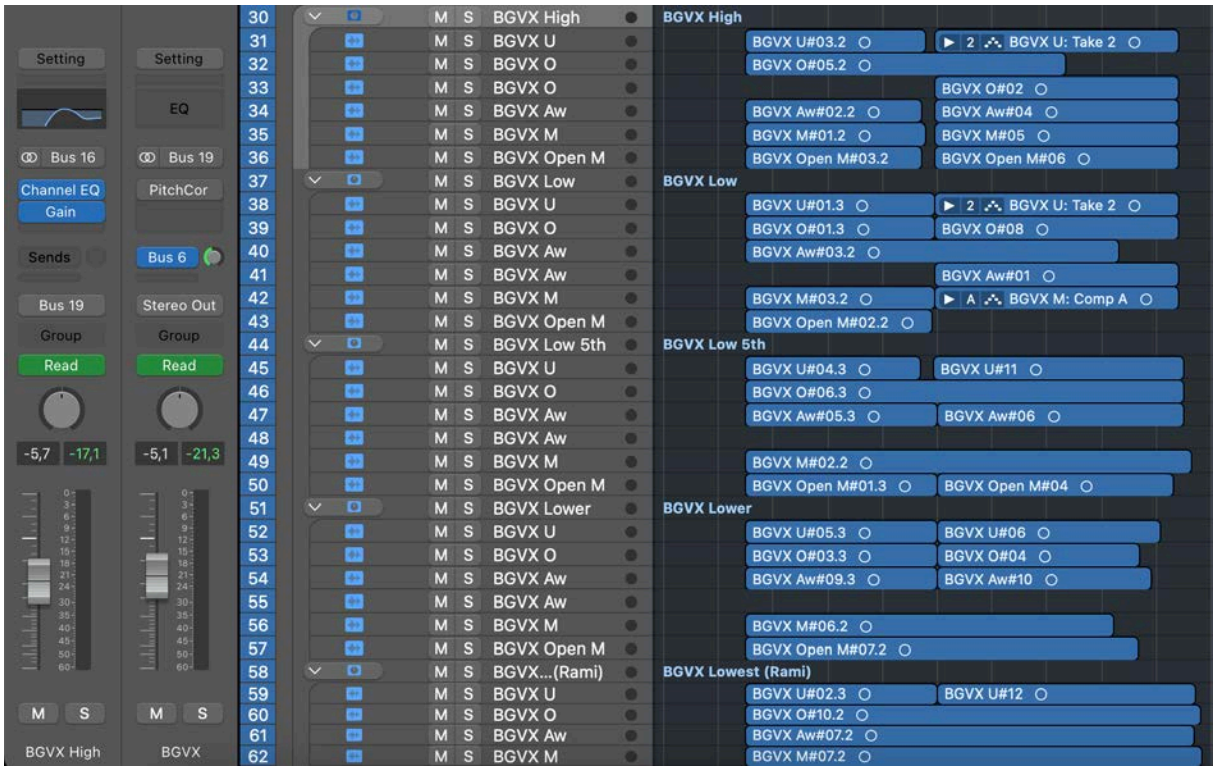


Figure C: *I Don't Mind* Breath removal and lead vocal effects (cf. p. 95)

Figure D: Vocal Tracking Session in Rami Olsen's home-studio.

Figure E: *I Don't Mind* DAW Session in progress - automation and note taking

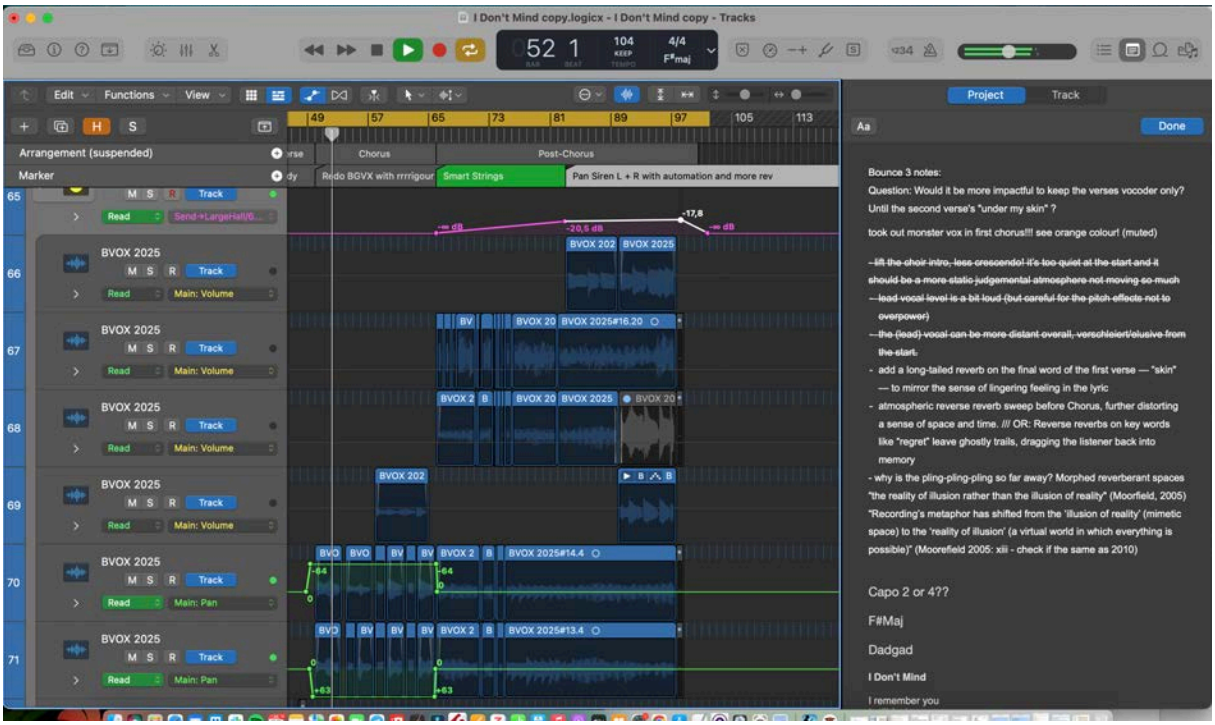


Figure F: Reflective Journaling

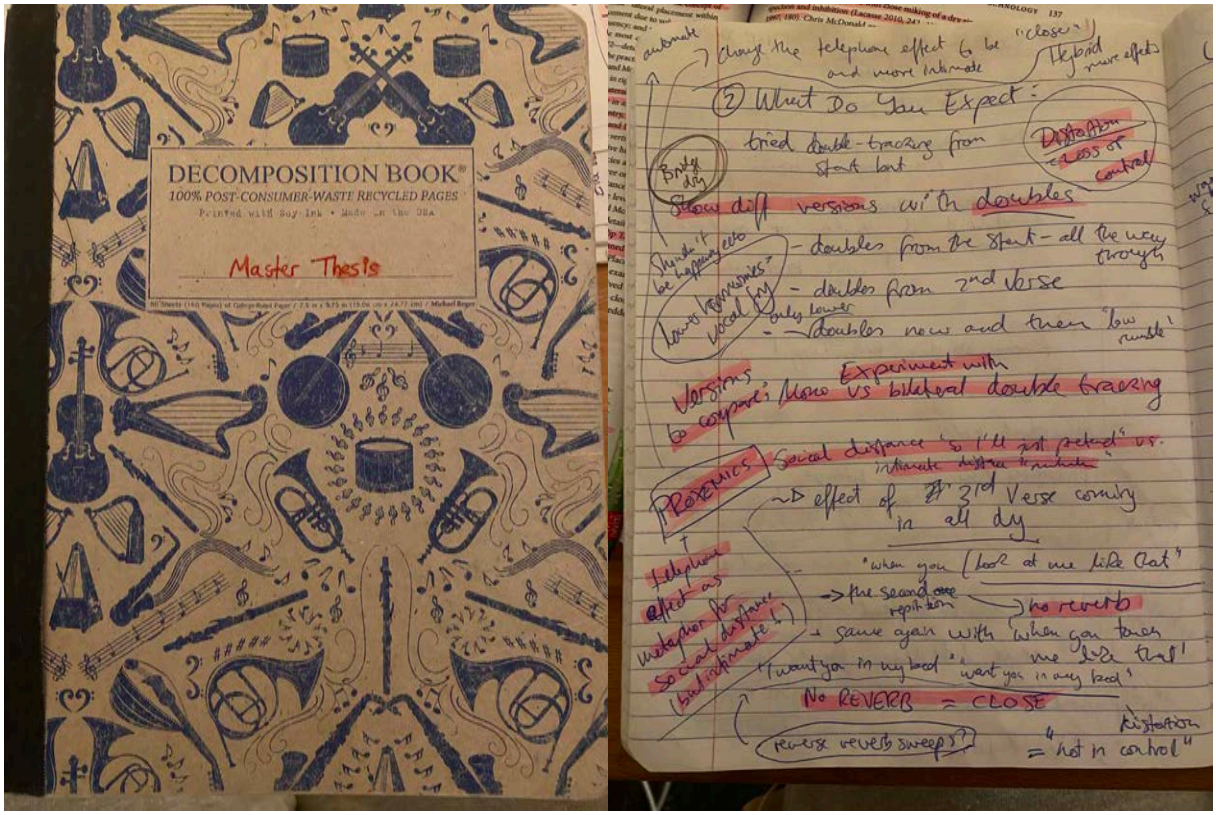
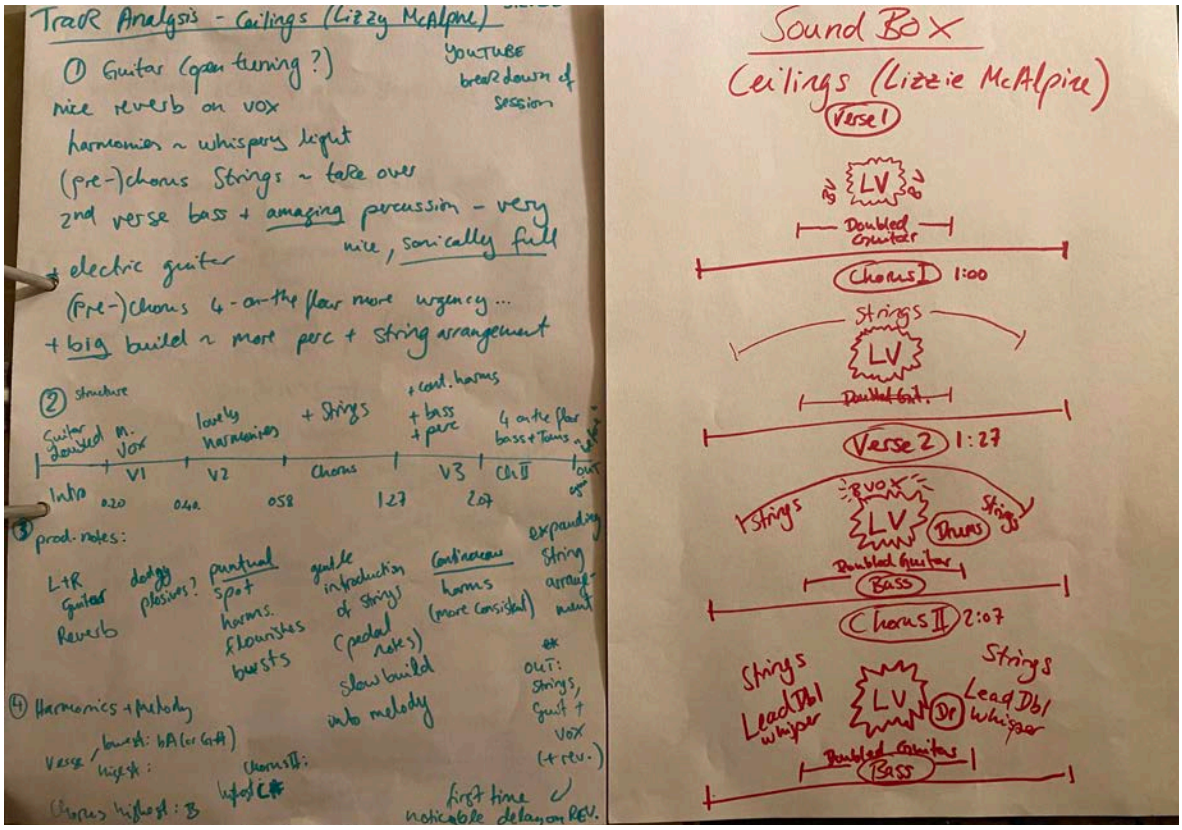


Figure G: Analytical listening practice, Lizzie McAlpine *Ceilings* (2022)



Note on the Use of AI Tools ('Hilfsmittel')

In the course of writing this Master's thesis, I made selective use of the generative AI tool *ChatGPT* (Model: GPT-4, https://chat.openai.com) as a digital assistant. This use was conducted independently and in accordance with the academic policies and ethical guidelines of HAW Hamburg regarding the responsible and fair use of AI tools in higher education.

ChatGPT was employed to support the writing process in areas such as language refinement and structural development. It did not serve as a source of academic knowledge or scientific citation. All intellectual content, analysis, and conclusions in this thesis are the result of my own independent research and creative decision-making. AI-generated outputs were critically evaluated, edited, and contextualised before integration.

Statement of Originality (‘Eigenständigkeitserklärung’)

Hiermit versichere ich, dass ich die vorliegende Masterarbeit mit dem Titel:

“Finding My Voice: A case study on vocal production aesthetics and techniques for the self-producing female singer-songwriter in today’s popular music culture”

selbständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Alle Passagen, die ich wörtlich aus der Literatur oder aus anderen Quellen wie z. B. Internetseiten übernommen habe, habe ich deutlich als Zitat mit Angabe der Quelle kenntlich gemacht. Zudem versichere ich, dass die eingereichten schriftlichen Ausfertigungen der elektronischen Fassung entsprechen.

Datum	Unterschrift

Elektronische Veröffentlichung

Mit der elektronischen Veröffentlichung der MA-Arbeit erkläre ich mich hiermit einverstanden.

Datum	Unterschrift