

Entangled Voices: AI, Intra-action, and the Body Multiple in Creative Practice

Fá Maria

Department of Computing
Goldsmiths, University of London
fpere002@gold.ac.uk

Matthew Yee-King

Department of Computing
Goldsmiths, University of London
m.yee-king@gold.ac.uk

Jenn Kirby

Department of Music
University of Liverpool
jenn.kirby@liverpool.ac.uk

Abstract

This paper explores how AI may shape the artist's creative output not only in technical and aesthetic terms but also in how, by utilising these tools, questions of artistic control, agency, and ethics arise and shape the artistic process. These concerns are explored through *ERASURE* (2025), an audio-visual work that interrogates gender biases in AI-generated voices. Drawing on Karen Barad's concept of *intra-action* and Annemarie Mol's *the body multiple*, this paper examines how AI tools co-constitute creative processes, challenging binary distinctions between artist and instrument. Using a hybrid methodology combining autoethnography and these two concepts, the paper investigates how AI tools complicate notions of authenticity, consent, and representation. Ultimately, this paper argues that AI tools, far from being neutral, pull artists into a negotiation of power that extends beyond aesthetics into the social and political.

1. Introduction

In contemporary electronic and experimental music, the integration of artificial intelligence (AI) as a creative tool has become increasingly prevalent, reflecting broader shifts in computational aesthetics and human-machine collaboration (Dean and McLean, 2018; Bown, 2021). This trend resonates with a broader move in human-computer interaction and music technology research toward seeing digital musical tools as entangled, active participants rather than passive instruments (Mudd, 2019; Frauenberger, 2020; Morrison and McPherson, 2024; Reed *et al.*, 2024)

This paper explores how AI may shape the artist's creative output not only in technical and aesthetics terms but also in how, by utilising these tools, questions of artistic control, agency, and ethics arise and shape the artistic process. To achieve this, we reflect on the process of composing *ERASURE* (2025), an audio-visual piece that explores questions of representation and gender biases in AI voice synthesising technologies. The piece builds on the premise that the voices from gender non-confirming people are generally overlooked or misinterpreted by AI models, often programmed with voices shaped by cultural assumptions about "masculinity" and "femininity" (Sindoni, 2024). The artistic process began as a speculative and sonic inquiry into the limitations of AI's vocal representation. However, as the project evolved, it became clear that voice synthesis technologies do not merely reproduce sound; they can also reconfigure notions of identity, ethics, and agency (Crawford, 2021). Working with a curated dataset of queer and trans vocalists, the piece also confronted questions around consent, authenticity, representation, and the politics of vocal embodiment. All these concerns demanded a situated approach - one that not only attends to the output of AI systems, but also to the broader histories, bodies, and relations (ethical, political, societal, among others) these systems touch.

The analysis in this paper unfolds through a hybrid methodology that intersects autoethnography, reported in first person by the first author (working under the artist name Fá Maria), combined with the theoretical lenses of Karen Barad's concept of *intra-action* (Barad, 2007) and Annemarie Mol's theory of *the body multiple* (Mol, 2002). Barad's concept of *intra-action* deepens this analysis by

rejecting a model in which humans and technologies are separate entities that interact. Instead, it views them as mutually constituted through their entanglement. In the context of *ERASURE*, we argue that voice, artist, tools, and artwork co-shape and co-constitute one another, emerging through the artistic and technological processes of training, synthesis, and performance as well as through the political, social, economic, and ecological themes that AI tools touch. Annemarie Mol's concept of the body multiple is used here to, for example, conceptualise voice not as a singular, coherent identity but as something enacted differently across varied contexts: in the human voice, in the dataset, in synthetic outputs, in political discourse, and in audience perception. Each tool and setting renders a different "voice," complicating notions of authenticity, ownership, and embodiment. Through this lens, the voice is understood as multiple, situated (context-related), and contingent, not fixed but continuously reconstituted.

Rather than treating technologies as passive tools, this approach foregrounds the mutual constitution of artist and system, voice and machine, self and data. The theoretical frameworks of *intra-action* and *the body multiple* offer a way to highlight the non-neutrality of the tools we use and recognise how AI and human bodies co-constitute one another in artistic practice.

This paper is structured as follows: In section 2, I describe the conceptual aims of *ERASURE*. In section 3, I introduce the theoretical frameworks of *intra-action* and *the body multiple*, which will be used to provide a lens for understanding the dynamic relationship between artists and AI tools. In section 4, I describe the specific AI techniques used in *ERASURE*, such as training AI models through RAVE architecture, running local audio models using the Neutone plug-in, and voice cloning in the ElevenLabs online platform, highlighting their impact on the creative process. In section 5, I analyse how these tools shape the artistic process as well as raise ethical questions about consent, transparency, and ownership. Finally, I explore how using concepts such as intra-action and the body multiple can help us navigate the complexities of using AI tools in artistic practices, reflecting on questions such as: What does it mean to compose music when the very tools of creation introduce unexpected transformations? How do ethics, consent, and transparency become part of the artistic process rather than external considerations? Ultimately, this paper argues that AI tools, far from being neutral, pull artists into a negotiation of power that extends beyond aesthetics into the social and political.

2. *ERASURE*: AI, bias, and queer vocality

The voice plays a crucial role in both the physical and symbolic aspects of identity and political representation (Cavarero, 2005; McGarry, 2024). For the queer and trans community, the voice is an important vehicle of self-expression but also a source of marginalisation (Stryker, 2008; Eguchi, 2018; Muchitsch, 2023) - from the "gay voice", to speech dysphoria, or even the changes in one's voice from the result of hormone therapy. Yet these ways of talking and sounding, which are outside normative norms, find expression and have deeply influenced many music genres from Hyperpop to House, Techno, Folk, and experimental music (Maus, 2022; Goldfine, 2023; Muchitsch, 2023).

AI technologies can particularly amplify gender-specific, racist, classist, and other biases (Kartal, 2022; Ferrara, 2023; Çırtlık and Cosar, 2024). At the heart of this issue lies the idea that AI systems, including digital voice assistants (DVAs) like Siri, Alexa, and Google Assistant, are often programmed with voices that mimic human gender characteristics. These choices are not neutral (Sindoni, 2024); AI voices, created and curated by designers and engineers, are influenced by historical and cultural assumptions about masculinity and femininity and perform gender in ways that perpetuate the notion that gender is fixed and binary (2024). This also echoes Judith Butler's theory of *performativity* (Butler, 2015), which suggests that gender is not an innate quality but a performance shaped by societal expectations. This *performance* in digital spaces replicates offline systems of oppression, such as sexism, homophobia, and transphobia, by excluding voices that don't conform to traditional gender norms (Çırtlık and Cosar, 2024). Female-coded voices are frequently designed to sound nurturing, patient, or subservient, mirroring traditional expectations of women's roles in society (2024). In contrast, male-coded voices are often imbued with authority, strength, or assertiveness, reinforcing patriarchal views that equate masculinity with leadership and competence (2024). Some attempts have been made to raise awareness of gender stereotypes built into these

tools, such as the world's first voice assistant service called *Project Q*. The voice of *Project Q* provides a “gender-neutral” voice that does not belong to either male or female (Carpenter, 2019).

Video 1: Trailer for [ERASURE](#).

ERASURE was developed as part of an open call by the University of the Arts London (UAL) and Responsible AI UK (RAI UK). As stated on RAI UK’s website, “Music RAI funded 3 mini-projects to create impact and interest in Responsible AI (RAI) concerns of bias in AI models. These mini-projects use AI tools such as low-resource AI models with small datasets to showcase the challenges of bias in AI and how RAI techniques can be used to address them.”

To create *ERASURE*, I assembled a group of queer and trans vocalists (including myself), featuring the artists Eli Huehuentro, Lyra Pramuk, megouem, Nadia Marcus, Odete, and trained bespoke AI models on our voices to compose the piece. I wanted to create speculative gender-fluid sonic identities that bridge human and machine, revealing the expressive potential of AI when applied to marginalised voices. To achieve this, I composed the music by interlacing both the original recordings and their AI-generated voices. The AI-generated voices range from being almost indistinguishable from the original recordings to machine-like. In some cases, I mixed different voices together to create new ones. In most of the piece, the words being heard were never actually spoken by the human vocalists, yet in some cases, the AI tools allowed me to replicate their voices so precisely that it becomes very hard to discern the difference, which I will talk about in more detail in section 4. If, on the one hand, these tools can extend our capabilities, they also raise profound questions about agency and the ethical implications of being able to put words into someone else’s mouth. Who controls a voice? What does it mean to replicate or alter someone’s identity through technology? Through the process of making this work, these questions emerge - from notions of representation, I was confronted to interrogate the boundaries of authorship and consent in the age of AI.

3. Theoretical Framework: intra-action and the body multiple

To understand the different ways in which AI co-shapes creative work, it is necessary to rethink traditional notions of the artist as an autonomous creator and artistic tools as passive (Partch, 1979), neutral instruments. In recent years, I have been increasingly integrating AI voice synthesis in my work. Through this process, I found that binary distinctions between artist and tool, or between human and machine, no longer adequately encapsulated the complexities of my creative process. These dualisms proved insufficient for describing a workflow in which agency, decision-making, and meaning emerged through entangled processes. In these contexts, the AI system did not merely function as a tool to be controlled; rather, it participated in shaping and was shaped by the artistic process. To better account for this dynamic, this paper adopts a hybrid methodology that weaves together autoethnography, a self-reflexive qualitative research method which foregrounds the researcher's subjectivity, and two theoretical frameworks: Karen Barad’s concept of *intra-action* and Annemarie Mol’s theory of the *body multiple*. These frameworks, I argue, allow for a more nuanced understanding of creativity as a relational and distributed phenomenon, where roles are not fixed and pre-defined but emerge through ongoing interactions - more accurately, *intra-actions* - between composer and AI.

First- and second-person methods enable artists and researchers to articulate experiences of the design practice or concept from within, using themselves and their practices as the subject of study (Ellis, Adams and Bochner, 2011; Rapp, 2018; Devendorf, Andersen and Kelliher, 2020). These methods allow for articulating insights that are typically overlooked, from design narratives (Howell, Desjardins and Fox, 2021), discourses surrounding machine learning (ML) and AI (Jourdan et al., 2024) to performance and music composition (Brown and Vasquez, 2020; Mainsbridge, 2022; McMillan, 2022; Tapparo and Zappi, 2022). At the core of this paper is an autoethnographic approach, drawing on my personal experience and engagement with voice synthesis tools to interrogate and reflect on how the tools artists use co-shape the artistic process.

Barad's work (Barad, 2007) is influenced by Judith Butler's concept of *performativity* (Butler, 1993). Barad engages with this concept through a diffractive reading of the scientific insights of Niels Bohr, specifically his "proto-performative account of scientific practices." (Barad, 2007: 31). According to Barad, Bohr's perspective suggests that scientific instruments and methodologies don't merely describe or measure pre-existing objects, but actively contribute to the formation and definition of those objects. She calls this *agential realism*. This term replaces the traditional notion of interaction, where pre-existing elements interact, with *intra-action*, where elements emerge and co-constitute one another (Blackman, 2020). In the context of music performance, thinking through the aforementioned concepts, and shifting from the notion of interaction to intra-action can offer some unique insights into the relationship between performers and musical instruments: traditional thinking might consider the instrument as a passive object that the performer "uses" to create music (Partch, 1979; Emmerson, 2007). Intra-action suggests that both the performer and the instrument come into their roles through their relationship with each other - *co-constitution*; It allows for a broader notion of agency that can be distributed across the performer and the instrument. For example, the affordances of a specific musical tool can direct the kind of music that is created. Intra-action allows for a complex understanding of how agency in music-making is not solely a human attribute - *agency*; rather than pre-existing categories of "performer" and "instrument," intra-action suggests that these categories emerge through their relationship. Barad's perspective can also help us pay close attention to how material conditions (like the design or state of an instrument) and discursive practices (like the way we talk about or conceptualise music and instruments) are entangled. This allows us to see how both human and non-human elements contribute to the practice and theory of music.

Complementing Barad's framework is Annemarie Mol's concept of the *Body Multiple*. I first encountered Mol's work through her book *The Body Multiple: Ontology in Medical Practice*. Using atherosclerosis (when arteries become narrowed), she illustrates how medical ontologies—how medical practitioners understand and categorise the body—are not fixed or universally agreed upon. Instead, they are constructed through interactions, negotiations, and collaborations among various actors (health professionals, patients, caretakers) in the medical field. Instead of following the concepts found in textbooks, Mol looks into how atherosclerosis is "done," "practised," or "enacted" (Jensen and Winthereik, 2005) in a Dutch hospital. The body multiple brings together the concepts of multiplicity, enactment and practices in engaging and radical ways (Jensen and Winthereik, 2005). In social sciences is common practice to refer to different views on the same object, what Mol calls "perspectivalism" - a single object that is looked upon by all the concerned parties. (2005) But for Mol an object, (a disease in this case) can not be seen as one fixed entity but as a "texture of partially coherent and partially coordinated enactments" (2005). In Mol's account, the ontology of an object is thus decentred to a multitude of practices. Objects do not exist in and of themselves but only through "multiple situated practices." (2005)

Although focused on bodies, techniques and practices that are multiplied, however, *the body multiple* does not stand for fragmentation. Mol observes that a mystery to be solved is how divergent objects align so that they "hang together". Not necessarily as a coherent body, but instead aligned in a way that their "contradictions and tensions are made to matter in very particular ways" (125). According to Mol, objects do not exist as entities in and of themselves, but only through multiple situated practices, and are, as such, *body multiples*. From this position, we cannot develop an understanding of body multiples by just describing and defining an object. Rather, we should describe the multiple activities, practices and enactments in which body multiples are encountered, what Mol coins as "ontological multiplicity". This idea of enactment implies looking at the body not as a singular, bonded and stable entity, but as a process. "The focus on process", scholar Lisa Blackman argues, "is on composing rather than composed, pre-formed entities. The focus on composing looks at how bodies become assembled in particular ways through their coupling or conjoining with particular objects, practices, techniques and artefacts such that they are always bodies in the making rather than being ready-made" (Blackman 2009:107).

The concepts of intra-action and the body multiple deal with the complexities of identity, agency, and the relationships between human and non-human entities. They arise from feminist theory, science and technology studies, and posthumanist philosophy. They emerge from a set of relations

rather than possessing an intrinsic, immutable identity. These entanglements suggest that no entity exists in isolation but is constantly being made and remade through its relationships. They have implications for how we understand agency, responsibility, and ethical action, challenging us to think beyond individualistic and anthropocentric frameworks and acknowledging the complexity of the systems in which we're embedded. By integrating these concepts, we can recognise the performer-musical tool relationship as a dynamic network of human and non-human agents that are constantly affecting each other and being affected in return. This view, I argue, allows for a richer, more nuanced understanding of the performer-musical tool relationship, one that accommodates the complexities and multiplicities involved in musical or artistic production.

By integrating these concepts, we can reconceptualise the performer-musical tool relationship, not as a hierarchical exchange between human and instrument, but as a dynamic, reciprocal network of material-discursive intra-actions. This framing is especially pertinent when considering AI systems that “learn” from our bodies and voices, and in turn, can, for example, alter how we perceive ourselves. As such, theoretical perspectives like Barad’s and Mol’s are not abstract; they offer a language through which to describe the lived complexities of artistic practice in a posthuman landscape. They allow me to trace how my work is embedded in and shaped by entangled systems of embodiment, technology, and identity.

4. AI Compositional Techniques

In this section I will describe the steps and some of the composition techniques I used to create *ERASURE*.

Dataset: I assembled a group of queer and trans vocalists from Berlin’s experimental music scene. The goal was to have two sets of recordings from each vocalist (including my own): one file using the speaking voice, in which they read poems (some of them included in the piece), and the other file a singing voice exploring different vocal textures, including sustained tones, breaths, and non-verbal utterances. The goal was to capture the grain and idiosyncrasies of individual voices for later synthesis and transformation. Each file had between 30-60 minutes in duration.

AI training: Different types of AI tools were employed to explore how varying training methods affect vocal output. The approaches included:

- **RAVE:** The *RAVE* (Realtime Audio Variational autoEncoder) (Caillon and Esling, 2021) architecture was trained via Google Colab. The training process was lengthy and technically demanding, often requiring multiple restarts due to errors. In some cases, training a single model could take up to ten days. Once trained, the resulting models were deployed using the *Neutone FX* plug-in, which was integrated into *Logic Pro X* for compositional work.
- **Neutone:** On *Neutone*’s own platform (<https://neutone.ai/>), a custom model was trained by uploading audio files directly to their online training dashboard. This process was significantly faster, usually taking under three days. The resulting AI model could then be used within their dedicated *Morpho* plug-in.
- **ElevenLabs:** Additionally, the *ElevenLabs* online platform (<https://elevenlabs.io/>) was used to clone participants’ voices. This approach generated vocal outputs ranging from nearly indistinguishable replicas of the original speakers to highly synthetic blends and morphings of multiple voices.

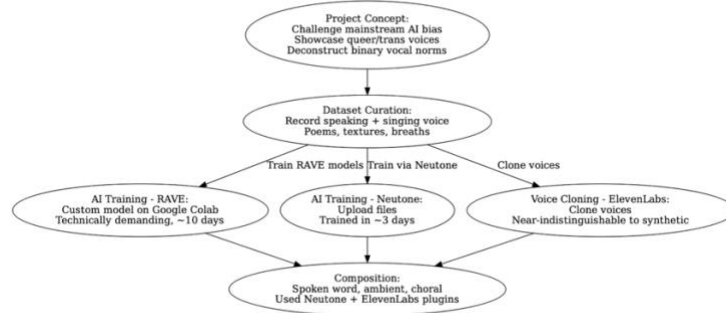


Figure 1: Diagram of *ERASURE*'s compositional process.

Sound Manipulation and Composition: To compose *ERASURE*, I combined spoken word, choral, ambient, and rhythmic elements. ElevenLabs was primarily employed for the synthesis of spoken word and poetic segments, as its voice cloning capabilities allowed for a high degree of manipulation and transformation of the recorded voices. The Neutone plugins (Neutone FX and Morpho) were used to generate new vocal textures and timbres, focusing on more abstract, ambient, or hybrid sonic outputs. For example, in the third song of the piece, I morphed my vocals with Lyra Pramuk's AI voice model. In two audio tracks (Fig.2), I recorded myself singing *a cappella* non-verbal vocalisations. On one of these tracks, I applied the Morpho plug-in using Pramuk's AI model. This plug-in enabled me, through various morphing parameters, to blend my voice with Pramuk's synthetic voice, producing a new hybrid vocal layer. Achieving a result I was satisfied with required multiple iterations and careful tweaking of the plug-in's parameters. This reflects my overall experience of using these AI tools throughout the piece: achieving a sound I was satisfied with required multiple iterations and careful tweaking of different parameters, highlighting the exploratory and hands-on nature that working with AI tools in a compositional process demands.

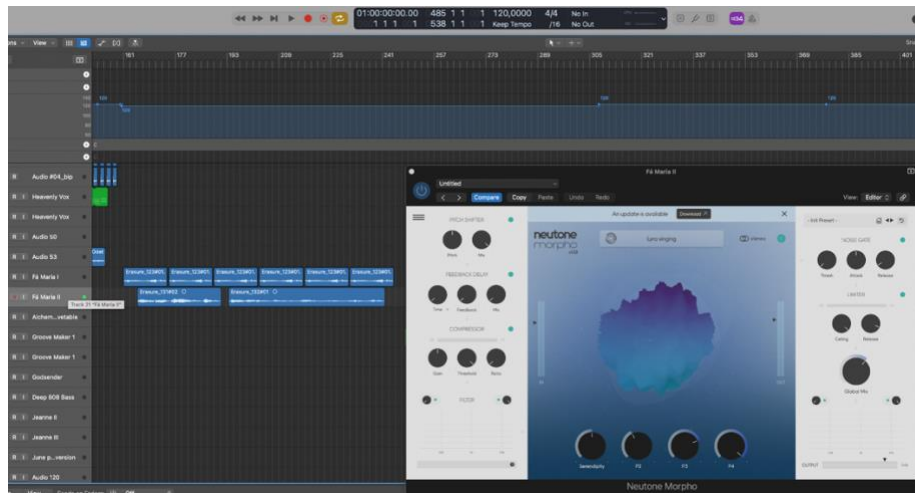


Figure 2: Screenshot of Logic Pro X

5. Artistic Practice and AI Techniques: the emergence of new questions

The initial goal of the project was to create music that explores how mainstream AI models usually overlook or misrepresent non-normative voices such as the ones from queer and trans people by using low-resource AI models and small datasets. The idea was to showcase different types of voices and explore how, by combining analogue and digital voices, we could deconstruct established

notions of how “male” or “female” voices should sound and challenge gender binaries and instead represent a gender spectrum.

If initially my goal with this project was to challenge questions of representation and gender norms in AI-generated voices, by *intra-acting* with these tools, other questions emerged, such as questions of consent. Tools like ElevenLabs allow the cloning of voices that, in some cases, are nearly indistinguishable from the original voices. This meant I could “put” words in the vocalists' mouths that they did not say. Two illustrative examples of voice cloning involve the artist Eli Huehuentro, who, when presented with poems in English, shared that he did not feel comfortable pronouncing some of the words due to some of the pronunciation being challenging for him. Therefore, he translated parts of the poems into Spanish. By cloning his voice, practically indistinguishable from his, I could create a version in which Eli read those poems in English (that I later used in one of the songs); the other example is with artist Lyra Pramuk's voice, which can be heard reading a poem she never read in the first place. This made me realise the power of these tools and the importance of not only having the consent of the artists but also contextualising the communities from which these voices come.

What initially started as a sonic investigation into gendered biases embedded within AI tools evolved into a broader critical engagement with structural themes underpinning these technologies. As the work unfolded, it became apparent that interrogating the absence—or misrepresentation—of queer and trans voices in mainstream AI systems could not remain at the level of aesthetics alone. The project demanded a contextual and situated approach, one that accounted for the lived realities and political histories entangled with voice and vocal representation. In response, I integrated textual elements that adopted a more informative tone and also included poems by Nadia Marcus, megouem, and myself, whose speak to embodied trans and queer experience, affect, and resistance.

Throughout the process, the concepts of *intra-action* and *body multiple* operated not merely as theoretical reference points but as methodological guides. These perspectives facilitated a shift from thinking of the voice as a stable, individualised entity to understanding it as something enacted differently across personal, relational, technological, and sociopolitical contexts. Retrospectively, these frameworks supported the ethical reframing of the work, helping moving it beyond sonic inquiry into a broader investigation of how identity is co-constructed through physical, digital, and meaning entanglements. This lens ultimately repositioned my role as a musician: not just as a composer of sound, but as a co-navigator of the complex *intra-actions* between self, system, and society, supporting my exploration and expanding my work into the conceptual and fine arts. This shift was aligned with a broader evolution in my practice, as I began to explore ways of expanding my work beyond the role of composer and music producer to incorporate visual, textual, and conceptual dimensions. Engaging with these questions through multiple media has given me a framework and subtract to not only expand the formal language of my work but also reposition my role as an artist by bringing together sound, theory, ethics, and technology.

Crucially, the affordances and constraints of the tools I worked with also played an active role in shaping both the process and the final form of the piece. Many of the AI tools—such as RAVE, Neutone, and ElevenLabs—were new to me. This unfamiliarity required a posture of openness and experimentation, inviting me to “trust the process” and treat the language and idiosyncrasies of these tools as intrinsic to the aesthetic language of the work. Although I began the project with a general sense of how I wanted the piece to sound, this vision was inevitably transformed through encounters with the tools themselves. One might say that this is the case with any musical instrument. An instrument, a tool interface and its limitations and affordances become not just technical features but co-constitutive elements of the compositional process and final result. In the next section, I will discuss in more detail how using the concept of *intra-action* can help us better understand this.

Important to mention that the visual component was created with Runway AI, blending manually shot digital footage (later processed using AI) and fully AI-generated visuals prompted through text. However, due to space limitations, this paper will focus on the audio component of the piece.

5.1. Connecting Intra-Action and the Body Multiple to *ERASURE*

Intra-Action

Karen Barad's concept of intra-action can help us understand the role of AI in developing *ERASURE*. Barad's theory emphasises the interdependence of agents, in this case, human and AI, both in the technical and creative aspects of the work. Rather than viewing the artist and AI as separate entities interacting with each other, intra-action posits that both the artist and AI are co-constituted through their engagement. The boundaries between the artist's intentions and the AI's responses become porous, and meaning is not solely generated by the artist but is developed in response to the affordances and outputs of the tools they use.

In the context of *ERASURE*, AI techniques such as speech synthesis, voice morphing, and sound manipulation do not simply act as extensions of my artistic will. They bring forth unexpected transformations, shifts in sound, and new configurations that I could not have predicted beforehand or had full control over. This intra-action is integral to understanding how the project evolves and reflects Barad's notion of the co-constitutive entanglement between artist and technology.

The Body Multiple

Annemarie Mol's concept of the body multiple offers a compelling lens for exploring the embodied and fragmented nature of artistic practice in *ERASURE*. Mol's theory challenges the traditional notion of a single, unified body and instead posits that the body (or object of study) is multiple—constructed through various practices and encounters depending on its different contexts of enactment.

From this perspective in *ERASURE*, each of its parts, whether the vocals, the tools used (plug-ins, AI tools), or the final piece itself, need to be looked at in its different enactments. For example, let's take the case of my own voice: there is my "a cappella" analogue voice, my AI-cloned voice, and my AI-resynthesized voice through plug-ins. Each one of these different bodies has its own characteristics depending on its different enactments, revealing different bodies, properties, challenges, and possibilities. All of these different voices, these different bodies become this way, my voice - a body multiple. This lens can be applied to any "object" of study.

Intra-action, the body multiple, and gender non-conforming voices

AI technologies for voice synthesis tend to treat the voice as a disembodied signal, something to be captured, replicated, or optimised, often neglecting the complex interplay between voice, identity, and lived experience (Bendel, 2019; Chadwick, 2020; Bergner, Hildebrand and Häubl, 2023). However, the voice is never neutral. It emerges through the body, through language, embedded in a particular culture, social and political contexts, and personal experiences - this takes great importance in the case of minorities such as queer and trans people. To reduce these voices to spectral elements is to show just a small part of them. *ERASURE* becomes a space not just of sonic experimentation but where the lived experiences of queer and trans vocalists are centered rather than abstracted (also multiplied). Similarly, Mol's concept of *the body multiple* invites us to take into account the different enactments of which an object is part of. Mol's theory suggests that rather than a single, stable entity, the body (or here, the voice) is performed differently depending on the context. In *ERASURE*, the voice appears: as a technological artefact, processed and reconfigured by AI; as a queer/trans testimony, shaped by lived experience and embodiment; as a political subject, revealing bias and silences embedded in datasets and technical design; as a sonic aesthetic, used compositionally to express emotion, distortion, fragmentation, and resistance. Each of these is a different enactment of "voice," shaped by different practices — technical, artistic, embodied, and discursive.

Mol argues that the way bodies are enacted is also political. Likewise, in *ERASURE*, the piece does not just present different voices; it intervenes in how AI enacts what a voice *is*. It questions the assumed neutrality of datasets by including the queer/trans vocal experience.

6. Conclusion: A Holistic View of AI in Artistic Practice

Through the frameworks of intra-action and the body multiple, this paper discusses how AI does more than expand musical possibilities—it reconfigures the artist’s agency, while also introducing new ethical, political, and social entanglements.

Key takeaways:

1. AI tools do not just assist artists; they transform their role, shifting from control to negotiation.
2. What begins as a focus on bias can lead to larger ethical concerns, such as consent, transparency, and ownership.
3. AI tools are not just about technical and artistic output, but they also force us to engage with the socio-political dimensions of these tools.

By employing Karen Barad’s concept of intra-action and Annemarie Mol’s theory of the body multiple, this paper explores the dynamic, unpredictable, and entangled relationship between the artist and AI in *ERASURE*. These theoretical frameworks help articulate how AI tools, far from being neutral instruments, actively participate in the creation process, shifting the artist’s role and raising questions about control, agency, and ethics. By acknowledging the co-constitutive nature of human-AI interactions, this paper provides insights into how the integration of AI into artistic practice challenges traditional notions of authorship, ethics, and identity, offering new ways to understand the complexities of using AI in creative processes.

Acknowledgments

I acknowledge the support provided for the first author’s PhD studentship by the FCT: Foundation for Science and Technology in Portugal and the Music RAI project, RAI UK International Partnerships (UKRI EPSRC grant reference EP/Y009800/1) musicrai.org that funded the project *ERASURE*. I also gratefully acknowledge the contribution of the co-authors for their invaluable assistance with editing and their guidance and support in their roles as PhD supervisors.

References

- Barad, K. (2007) *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham, NC: Duke University Press.
- Bendel, O. (2019) ‘The synthetization of human voices’, *AI & SOCIETY*, 34(1), pp. 83–89. Available at: <https://doi.org/10.1007/s00146-017-0748-x>.
- Bergner, A.S., Hildebrand, C. and Häubl, G. (2023) ‘Machine Talk: How Verbal Embodiment in Conversational AI Shapes Consumer–Brand Relationships’, *Journal of Consumer Research*. Edited by J.J. Inman, R.J. Lutz, and E.J. Kyung, 50(4), pp. 742–764. Available at: <https://doi.org/10.1093/jcr/ucad014>.
- Bown, O. (2021) ‘Sociocultural and Design Perspectives on AI-Based Music Production: Why Do We Make Music and What Changes if AI Makes It for Us?’, in E.R. Miranda (ed.) *Handbook of Artificial Intelligence for Music*. Cham: Springer International Publishing, pp. 1–20. Available at: https://doi.org/10.1007/978-3-030-72116-9_1.
- Butler, J. (2015) *Gender Trouble: Feminism and the Subversion of Identity*. First issued in hardback. New York London: Routledge, Taylor & Francis Group (Routledge classics).
- Caillon, A. and Esling, P. (2021) ‘RAVE: A variational autoencoder for fast and high-quality neural audio synthesis’. arXiv. Available at: <https://doi.org/10.48550/ARXIV.2111.05011>.
- Carpenter, J. (2019) ‘Why project Q is more than the world’s first nonbinary voice for technology’, *Interactions*, 26(6), pp. 56–59. Available at: <https://doi.org/10.1145/3358912>.
- Cavarero, A. (2005) *For more than one voice: toward a philosophy of vocal expression*. Stanford, Calif: Stanford university press.
- Chadwick, R. (2020) ‘Methodologies of voice: Towards posthuman voice analytics’, *Methods in Psychology*, 2, p. 100021. Available at: <https://doi.org/10.1016/j.metip.2020.100021>.
- Çırtlık, B. and Cosar, S. (2024) ‘Gender Bias in AI’, *Feminist Asylum: A Journal of Critical Interventions*, 2. Available at: <https://doi.org/10.5195/faci.2024.124>.

- Crawford, K. (2021) *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*. New Haven London: Yale University Press.
- Dean, R.T. and McLean, A. (eds) (2018) *The Oxford handbook of algorithmic music*. New York, NY: Oxford University Press (Oxford handbooks).
- Eguchi, S. (2018) 'From Drag Queens to Leathermen: Language, Gender, and Gay Male Subcultures. Rusty Barrett. New York: Oxford University Press, 2017. 272 pp.', *American Ethnologist*, 45(2), pp. 297–298. Available at: <https://doi.org/10.1111/amet.12652>.
- ElevenLabs, no date. *ElevenLabs*. [online] Available at: <https://elevenlabs.io>.
- Emmerson, S. (2007) *Living electronic music*. Aldershot Burlington, Vt: Ashgate.
- Ferrara, E. (2023) 'Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies', *Sci*, 6(1), p. 3. Available at: <https://doi.org/10.3390/sci6010003>.
- Frauenberger, C. (2020) 'Entanglement HCI The Next Wave?', *ACM Transactions on Computer-Human Interaction*, 27(1), pp. 1–27. Available at: <https://doi.org/10.1145/3364998>.
- Goldfine, R. (2023) 'We're Here, We're Hyper, and We're Popping: A Queer Analysis of Hyperpop'. Available at: <https://www.bowdoin.edu/gender-women/news/2023/were-here-were-hyper-and-were-popping-a-queer-analysis-of-hyperpop.html>.
- Kartal, E. (2022) 'A Comprehensive Study on Bias in Artificial Intelligence Systems: Biased or Unbiased AI, That's the Question!', *International Journal of Intelligent Information Technologies*, 18(1), pp. 1–23. Available at: <https://doi.org/10.4018/IJIT.309582>.
- 'Material-Oriented Musical Interactions' (2019) in Mudd, T., *Springer Series on Cultural Computing*. Cham: Springer International Publishing, pp. 123–133. Available at: https://doi.org/10.1007/978-3-319-92069-6_8.
- Maus, F.E. (2022) 'Introduction', in F.E. Maus and S. Whiteley (eds) *The Oxford Handbook of Music and Queerness*. 1st edn. Oxford University Press, pp. 1–31. Available at: <https://doi.org/10.1093/oxfordhb/9780199793525.013.112>.
- McGarry, A. (2024) *Political Voice: Protest, Democracy, and Marginalised Groups*. 1st ed. Oxford: Oxford University Press, Incorporated (Oxford Studies in Culture and Politics Series).
- Mol, A. (2002) *The body multiple: ontology in medical practice*. Durham: Duke University Press (Science and cultural theory).
- Morrison, L. and McPherson, A. (2024) 'Entangling Entanglement: A Diffractive Dialogue on HCI and Musical Interactions', in *Proceedings of the CHI Conference on Human Factors in Computing Systems. CHI '24: CHI Conference on Human Factors in Computing Systems*, Honolulu HI USA: ACM, pp. 1–17. Available at: <https://doi.org/10.1145/3613904.3642171>.
- Muchitsch, V. (2023) 'Listening to Anohni's variously vibrating voice: studying transfeminine vocality in 21st-century popular music culture through the concept of vocal figurations', *Popular Music*, 42(1), pp. 59–78. Available at: <https://doi.org/10.1017/S0261143023000107>.
- Neutone, no date. *Neutone*. [online] Available at: <https://neutone.space>
- Partch, H. (1979) *Genesis of a music: an account of a creative work, its roots and its fulfillments*. 2. ed., enlarged, 1. paperback ed. New York, NY: Da Capo Press (Da Capo Paperback).
- Reed, C.N. et al. (2024) 'Sonic Entanglements with Electromyography: Between Bodies, Signals, and Representations', in *Designing Interactive Systems Conference. DIS '24: Designing Interactive Systems Conference*, Copenhagen Denmark: ACM, pp. 2691–2707. Available at: <https://doi.org/10.1145/3643834.3661572>.
- Stryker, S. (2008) *Transgender history*. Berkeley, CA: Seal Press : Distributed by Publishers Group West (Seal studies).