UC Riverside
Conference Papers

Title
From Vocal Body to Vocal Network: AI and the Reconfiguration of Musical Co-Creativity

Permalink
https://escholarship.org/uc/item/4k64p1fz

Author
Paradiso, Paolo

Publication Date
2026-06-01

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

From Vocal Body to Vocal Network: AI and the Reconfiguration of Musical Co-Creativity

Paolo Paradiso
Free University of Bozen-Bolzano
Paolo.Paradiso@student.unibz.it

Abstract

Within the Euro-American art-music tradition, creativity has largely been framed through humanist paradigms privileging individual authorship, intentionality, and stylistic innovation. The growing presence of generative AI and machine learning (ML) in composition and performance unsettles these assumptions and calls for a rethinking of creative agency. This paper asks how the integration of artificial neural networks into live vocal improvisation reshapes musicological concepts of authorship, performance, and agency. Drawing on Donna Haraway's situated, hybrid subjectivities, Rosi Braidotti's relational posthuman subject, Deleuze and Guattari's theory of assemblage, and Dominic Pettman's account of sonic intimacy, it develops a posthumanist framework that is tested against two case studies: Tomomibot, by Tomomi Adachi, Andreas Dzialocha, and Marcello Lussana, and ULTRACHUNK, by Jennifer Walshe and Memo Akten. Combining philosophical inquiry with musicological analysis of vocal technique, performer-interface interaction, and improvisational structure, the paper examines how agency is distributed across performers, programmers, neural networks, training data, and technological infrastructure in live contexts. It argues that these performances do not demonstrate autonomous machine creativity; rather, they reveal creativity as a relational, materially embedded process in which the voice operates as a privileged bridge between biology and technology. Ultimately, the performances are read as a way of dwelling in the extimacy that constitutes voice and subjectivity alike, offering musicology new conceptual tools for understanding AI-mediated composition and performance as distributed yet asymmetrical co-creation.

Keywords: Distributed Agency; Human-AI Interaction; Posthumanism; Vocal Extimacy; Vocal Improvisation.

1. Introduction

The integration of generative artificial intelligence (GenAI) and ML into contemporary musical practice has moved rapidly from speculative experiment to established performance reality. This paper claims that for musicology, this development is not merely technological but conceptual: it destabilizes categories through which the Euro-American art-music tradition has long understood creative work, namely individual authorship, expressive intentionality, and stylistic innovation. When an artificial neural network (ANN) improvises with a human vocalist in real time, familiar questions lose their self-evidence. Who is the author of the resulting music? What does performance mean when part of the performing entity is a statistical model trained on the human performer's own past behavior? And where, exactly, is creative agency located?

A common way of framing these questions, whether AI can be creative in the way humans are, is, this paper argues, the wrong starting point. Posed in those terms, the question presupposes precisely the humanist model of the autonomous creative subject that AI-mediated practice puts in crisis. The more productive question is what happens to the concepts of creativity, authorship, voice, improvisation, and agency when the creative process itself becomes distributed across a socio-technical network of human and non-human participants. Debates about authorship are not idle: the inclusion of AI in creative processes

forces us to ask whether the author is the human who provides inputs, the AI as co-author, the programmers of the model, or the creators of the material in the training database, questions with direct implications for how copyright frameworks will need to be updated. The present study, however, is concerned not with legal authorship but with co-creation as a musicological and philosophical problem.

The paper asks: how does the integration of GenAI into contemporary musical practice reshape musicological concepts of authorship, performance, and agency? It approaches this question through a posthumanist framework drawing on Donna Haraway's situated, hybrid subjectivities, Rosi Braidotti's relational posthuman subject, and Deleuze and Guattari's notion of assemblage, with Dominic Pettman's theorization of vocal relationality informing the analysis of the voice as a site where species, technology, and affective proximity intersect.

Methodologically, the study is a qualitative, theoretically informed musicological analysis of two paradigmatic case studies. It combines philosophical inquiry with musicological attention to vocal technique, performer-interface interaction, improvisational structures, and the role of ANNs in shaping musical form and timbre, with particular attention to how agency is distributed across composers, programmers, performers, neural networks, and technological infrastructure in live contexts. The two case studies, Tomomibot (2019), by Tomomi Adachi, Andreas Dzialocha, and Marcello Lussana, and ULTRACHUNK (2018), by Jennifer Walshe and Memo Akten, are structurally similar: both stage a live improvisation between a highly experienced performer of experimental extended vocality and an artificial counterpart generated by neural networks trained on pre-existing material from that same performer. Yet their aesthetic outcomes and technical architectures differ significantly, and these differences are analytically productive rather than incidental.

The central argument is that these performances do not simply demonstrate machine creativity as the simulation of human creativity. Rather, they reveal creativity as a materially embedded process of distributed agency, in which improvisational structures take shape precisely from the human-AI interaction. Within this process the voice plays the central role: it is the bridge between the organic and the artificial, the key that enables the hybridization of the two dimensions, which reconfigure themselves into a mixed assemblage in which nature and culture are not opposing poles but rhizomatically connected elements within a shared environment. What these performances ultimately explore, the paper proposes, is the extimate structure that voice and subjectivity share: a threshold at which inside and outside are not opposed but mutually constituted, and at which the other is always already present. Precisely for this reason the encounter with an artificial vocal double, far from being a mere technical curiosity, can be read as a way of testing new forms of relation and coexistence.

The paper proceeds as follows. Section 2 reconstructs the concept of the posthuman subject in Haraway and Braidotti; Section 3 develops the theory of assemblage as the ontological hinge between posthuman subjectivity and musical analysis; Section 4 analyzes Tomomibot and ULTRACHUNK individually and comparatively, proposing the notion of a vocal network; Section 5 interprets the results through Pettman's theory of sonic intimacy and the concept of extimacy; Section 6 concludes.

2. The Posthuman Subject

Why discuss subjectivity at all when the central topic is musical improvisation? Because including AI in the creative process forces a rethinking of authorship and expressive intent, and because the humanist model of the author presupposes a specific model of the subject:

unitary, autonomous, rational, and self-transparent. Posthumanist philosophy offers an alternative, and it is important to stress at the outset that posthumanism is not a homogeneous doctrine. Francesca Ferrando distinguishes its three constitutive gestures: post-humanism, the recognition that the human is plural rather than universal; post-anthropocentrism, the decentring of the human in relation to the non-human; and post-dualism, the critique of rigid identity-defining dichotomies such as nature/culture and human/machine (Ferrando, 2019). Within this constellation, Haraway and Braidotti reject the idea of a universal and autonomous human subject in different, complementary ways.

Haraway's intervention is primarily epistemological. Knowledge, for Haraway, always derives from embodied, partial, and localized perspectives; subjects are formed through the crossing of boundaries, notably between human and machine (Ferrando, 2019; Haraway, 1991). Her cyborg names this condition: "By the late twentieth century, our time, a mythic time, we are all chimeras, theorized and fabricated hybrids of machine and organism; in short, we are cyborgs" (Haraway, 1991, p. 50). Crucially, the cyborg is not a celebration of technology but an argument that the relation between organism and machine has been a "border war", and a plea for "pleasure in the confusion of boundaries and for *responsibility* in their construction" (ibid.). Subjectivity, on this account, arises from interactions among materials, technologies, bodies, histories, and power structures.

Haraway's later work sharpens this position into an explicit rejection of autopoiesis, the self-organization of the living theorized by Maturana and Varela. "Nothing makes itself; nothing is really autopoietic or self-organizing" (Haraway, 2016, p. 58). Her counter-concept, *sympoiesis*, from the Greek *sún* (together) and *poiesis* (making), designates "complex, dynamic, responsive, situated, historical systems. It is a word for worlding-with, in company" (ibid.). The relevance of this move to a discussion of co-creation is direct: if nothing makes itself, then making, including music-making, is always making-with.

While for Haraway the issue is epistemological, for Braidotti it also becomes ontological. Braidotti develops the theme of subjectivity within a systematic neo-materialist theory that conceives the posthuman subject as a relational, transversal, and post-anthropocentric form of life connecting humans, animals, technology, and the environment (Ferrando, 2019; Braidotti, 2013). The subject is not pure consciousness dominating a body from above; it emerges from a network of material, biological, technological, social, and environmental relations (Braidotti, 2019). The posthuman subject's relational capacity is not confined to the human species but includes non-anthropomorphic elements within a nature-culture continuum (Braidotti, 2013). Braidotti therefore describes the subject as embodied and embedded, incarnated and immersed in the world: we are always already caught in a web of relations with other humans, animals, environments, and technologies (Braidotti, 2019). Posthuman subjectivity is, in her formulation, "a transversal alliance that nowadays involves non-human agents", relating simultaneously to the earth and to technological agents such as codes and algorithms (Braidotti, 2019, p. 58).

Two features of Braidotti's account matter especially for what follows. First, the definition of the subject through affect rather than rationality: "what defines us is not primarily rationality, but the capacity to affect and be affected – that is, to act upon others and to be transformed by relationships" (Braidotti, 2019, p. 57). Second, the consequence that the subject does not precede its relations but emerges within them, which leads directly to the concept of assemblage.

3. Theory of Assemblage

If the subject emerges from relations rather than preceding them, an ontological vocabulary is needed for relational configurations. Deleuze and Guattari's concept of assemblage (*agencement*) provides it. An assemblage is a dynamic composition of heterogeneous elements that function together (Deleuze and Guattari, 1987). Its elements are of several kinds at once: "human, social, and technical machines, organized molar machines; molecular machines with their particles of becoming-inhuman" (Deleuze and Guattari, 1987, p. 37), and, decisively for the question of authorship, "there are no individual statements, only statement-producing machinic assemblages" (ibid.). What Deleuze and Guattari call consistency, the "holding together" of heterogeneous elements, is not imposed from outside but achieved in the functioning of the assemblage itself (Deleuze and Guattari, 1987, p. 323). The subject, on this account, is not a self-sufficient 'I' that subsequently enters into relation with the world, but a provisional node of material, affective, linguistic, social, and technical relations (Deleuze and Guattari, 1987).

Braidotti explicitly grafts her theory of the posthuman subject onto this ontology. The posthuman subject must be conceived as a collective assemblage, grounded in a process ontology that refuses to identify subjectivity solely with rational consciousness (Braidotti, 2013); she also speaks of a technoassemblage to emphasize a composition that includes life and technology. The subject thus coincides not with a separate individual but with a relational configuration made up of body, environment, technology, affect, institutions, and languages (Braidotti, 2019); it is "a moveable assemblage within a common life-space that the subject never masters nor possesses but merely inhabits" (Braidotti, 2013, p. 193). Conceiving subjectivity as an assemblage that includes non-human agents implies that subjectivity is not the exclusive prerogative of *anthropos*, is not linked to transcendental reason, and is based on the immanence of relations (Braidotti, 2013).

For musical analysis, the concept does three kinds of work. First, it dissolves the question 'who is the author?' into the more tractable question 'how is the statement, here, the musical event, produced, and by which configuration of elements?' Second, it directs attention to relations rather than essences: what matters analytically is not what the performer or the algorithm is, but what each does to and with the others, their ways of affecting and being affected. Third, it warns against symmetry: an assemblage is composed of heterogeneous elements whose contributions are not equivalent, so describing a performance as an assemblage does not mean declaring humans and machines interchangeable. The musical experiments examined below provide a practical and aesthetic illustration of these concepts, a sonic and performative translation of them, but they also test them: the analysis must show how, where, and through which relations the assemblage is constituted, not merely assert that it is one.

4. Posthuman Music

Tomomibot and ULTRACHUNK are two experimental musical performances with very distinct aesthetic outcomes, yet structurally similar: in both, a human vocalist improvises with an artificial counterpart generated by neural networks trained on aspects of that vocalist's own practice. Each is first analyzed individually; their comparison then motivates the concept of the vocal network.

4.1. Tomomibot

Tomomibot is a 2019 experimental music project by Tomomi Adachi, Andreas Dzialocha, and Marcello Lussana. The name designates both the project and the software, based on ML algorithms, that progressively learns from the voice and musical behaviors of the Japanese sound artist Tomomi Adachi, allowing him to perform in dialogue with an artificial version of himself. The project's objective is not only musical but also critical and cultural: to make computer culture 'audible' and to interrogate the relationship between computational logic and human culture (Adachi, n.d.).

The system's technical architecture is worth specifying, because it shapes the musical interaction. Tomomibot relies on LSTM (Long Short-Term Memory) recurrent neural networks, a sequential architecture that determines which sound to reproduce based on the sounds heard previously during the performance. The software was designed by Andreas Dzialocha, while Marcello Lussana created, using AI-based speech-synthesis algorithms (WaveNet, WaveRNN, and FFTNet), a vast archive of sounds mimicking Adachi's voice; this repertoire constitutes the sonic vocabulary with which Tomomibot improvises and interacts with the human performer (Adachi, n.d.). The division of labor is thus already a distribution of creative agency: the performer supplies the vocal practice from which the system learns; one collaborator designs the listening-and-response logic; another constructs the very sound-world the system can speak.

Musically, the decisive factor is Adachi's improvisational style, characterized by fragmentation, noise, and vocal experimentation, which causes Tomomibot to behave more like a sound machine than a singing system. The system does not model 'song' as a musical category; it models Adachi's extended vocality, and thanks to ML it can learn patterns from any type of sound material and use them as a basis for improvisation. In performance, the LSTM's sequential prediction establishes a real-time loop of listening, response, and anticipation between performer and system.

Adachi's artistic research explicitly targets 'machine creativity' and the non-anthropomorphic possibilities of vocal synthesis, a third way beyond generic voice technology and voice cloning, in which technological alterity is explored aesthetically rather than concealed (Napolitano, 2022). He states that "the machine's incomprehensible expressions reveal the deepest part of its mind" (Napolitano, 2022, p. 307), a formulation whose anthropomorphism should be read critically, as part of the work's poetics rather than a technical description. More precise, and more analytically consequential, is his account of the resulting voice: "Tomomibot's voice is both my voice and the machine's voice. Through neural networks, the machine 'performs' my voice as if it were its own, autonomously. Ultimately, my identity as an artist is expressed precisely at this boundary" (Napolitano, 2022, p. 308). Adachi here locates his own artistic identity not in the vocal body, nor in the software, but at the boundary between them, an emblematic formulation of the situated, hybrid subjectivity theorized by Haraway, in which the artist aims at synthetic identities that are personal yet deliberately non-realistic, non-human.

4.2. ULTRACHUNK

ULTRACHUNK is a 2018 collaboration between vocal performer and composer Jennifer Walshe and artist and computer scientist Memo Akten. Here too the work involves a vocal improvisation between the human performer and her artificial alter ego, generated by the ANN GRANNMA MagNet (Granular Neural Music & Audio with Magnitude Networks).

GRANNMA is a system of generative neural networks integrating various machine learning architectures, including GANs (Generative Adversarial Networks), VAEs (Variational Autoencoders), and other custom architectures, for the real-time generation of audiovisual material derived from Walshe's vocal improvisations (Akten, 2018). Training required hours of audio and video documentation of Walshe's improvising, accumulated over the course of a year, to build the machine-learning-based live performance system (ibid.).

Two features distinguish ULTRACHUNK from Tomomibot. First, its output is audiovisual: in the live performance, digital video and audio are generated in real time and interactively, so that Walshe confronts not only a sonic double but a visible one. Second, its generative logic differs. Where Tomomibot recombines a pre-synthesized sonic vocabulary according to LSTM sequence prediction, GRANNMA synthesizes its material continuously. As Akten describes:

> During the performance, the video and audio output from the machine are neither recordings nor processed — every frame and sound is synthesized live in realtime, navigating the latent hypersphere, constructed from the fragments of memories in the depths of multiple neural networks. The original and virtual Walshe inhabit the Uncanny Valley together, singing in duet, improvising, listening and responding to each other (Akten, 2018).

Walshe's own account of the improvisational relationship is strikingly concrete:

> Improvising with ULTRACHUNK was like improvising with someone in an altered state of consciousness, if that person was also me, high on a drug I've never taken. My job was to trust that ULTRACHUNK was going to contribute very strongly, if chaotically, at times. I should perform as normal, while also managing the things that people who are high tend to neglect, like long-term structure, and thanking the sound engineer (Walshe, 2023).

The statement is analytically rich. It registers genuine machine contribution while precisely delimiting it: responsibility for long-term musical form, and for the social frame of performance, remains with the human. Distributed agency, in other words, is not symmetrical agency. The machine participates in the moment-to-moment texture of the improvisation; the human holds the temporal architecture. Walshe's phrase "if that person was also me" (ibid.) simultaneously names the uncanny dimension of the encounter: the alterity she improvises with is a transformation of her own past vocal behavior, returned to her as an interlocutor.

4.3. Vocal Network

The two cases are excellent examples in which the voice expands to become fully extracorporeal: the human voice comes into contact with its digital translations, and what results is neither human nor artificial, because the two poles intersect in a way that evokes the posthumanist continuum of nature-culture and nature-technology; the shared environment constructs the resonant space that accommodates an other, hybrid, mixed voice (Di Scipio, 2023). But the comparison must be kept precise, because the two systems participate in improvisation differently.

In Tomomibot, the machine's contribution is bounded by a pre-constructed sonic vocabulary (Lussana's synthesized archive) and organized by sequential prediction: the LSTM selects what to play on the basis of what it has just heard. Its agency is that of a listening-and-answering interlocutor operating within a finite lexicon. In ULTRACHUNK, by contrast,

nothing is pre-recorded: GRANNMA's generative networks synthesize every frame and every sound live, by navigating a latent space learned from a year of Walshe's audiovisual material. Its agency is that of a continuous transformational double, whose outputs are constrained not by a lexicon but by the statistical structure of the training data. What is 'learned' also differs: Tomomibot learns patterns of sonic behavior to be recombined; GRANNMA learns a compressed representation of Walshe's voice and body from which new, never-recorded material is generated. Correspondingly, authorship is redistributed differently: in Tomomibot the programmer's and sound-designer's contributions are individuated and audible as such (response logic, vocabulary); in ULTRACHUNK the designer's contribution is absorbed into the architecture and training regime of a single generative system.

What the cases share is the structure of mutual influence. The musical choices of the human performers are influenced and co-defined by the outputs generated by the ANNs, and vice versa; creativity is distributed among the collaborators, vocalist, ANN, programmer, and in this division of labor technological agency enters the performative sphere as a new type of active agent, enriching and complicating the paradigm of distributed creativity. Specifically, the AI contributes to the formation of a vocal frame, plausibly corresponding to the aesthetic pursuit of the performers and programmers involved: the result of the combined action of an intention of empathic communication exceeding the semantic sphere with the technological mediation of the voice, capable of shaping the empathic perception of the vocal persona's identity and the context of voice perception in a given spatial situation (Occhio, 2023).

This mutual influence gives concrete musical form to Braidotti's affective definition of subjectivity (Braidotti, 2019): performer and system are characterized not by any rational faculty but by their openness to one another, each shaping the other's sonic behavior while being reshaped by it in turn. The relationship is rhizomatic, multilevel, corporeal, material, and collective, shifting the focus away from the isolated self and placing interdependence at the center (Braidotti, 2013). The musical examples thus display the transformation announced in this paper's title: from a bodily, biological voice to a network voice, in which the nodes of the network co-participate in musical creation in real time, integrating and hybridizing into a voice that inhabits the space situated at the intersection of the nodes themselves.

Two caveats guard the argument against overstatement. First, distribution does not imply equivalence. The performer's embodied training, the programmers' design decisions, the datasets' histories, and the algorithms' statistical operations are heterogeneous kinds of contribution; the assemblage holds them together precisely as heterogeneous (Deleuze and Guattari, 1987). Walshe's retained responsibility for long-term structure is a musical index of this asymmetry. Second, none of this licenses claims about machine intentionality. The systems exercise agency in the minimal, relational sense of making a difference to the unfolding of the performance and eliciting responses; they do not thereby acquire expressive intent. It is exactly this distinction, between agency and intention, generation and authorship, interaction and autonomy, that the posthumanist framework allows us to keep in view without collapsing.

A more detailed examination of interpreting the role of the voice in AI-driven music through a posthumanist lens can be found in Paradiso (2025).

5. Posthuman Sound Intimacy

What makes this hybridization possible is, first and foremost, the voice, and the intimacy that a 'posthuman listening' allows to be created between biological and digital voices. Here the

analysis draws on Dominic Pettman's theory of sonic intimacy, which aims to overcome the sharp division between the human voice and the non-human voice, including the machine voice (Pettman, 2017). Pettman proposes considering 'natural' and 'artificial' sounds as being in continuous dialogue, without erasing the distinction between the origins of the various voices (human, animal, artificial, environmental) (ibid.); this attention to non-human vocal sources leads to a reconsideration of human specificity itself, from a posthuman and post-anthropocentric perspective.

Pettman argues that "in being open to vocal solicitation from nonhuman sources, we might find ourselves being re-interpellated, and in ways that are empowering, not just for ourselves but also for the entity hailing us" (Pettman, 2019, p. 7). He proposes the notion of a *vox mundi*, the voice of the world, as the sum of the cacophonous, heterogeneous sounds of the postnatural world, within which to situate the voices "of creatures, agents, entities, objects, and phenomena" (Pettman, 2019, p. 79). In this planetary chorus, even the cybernetic voice finds its agency in communication, as speaker or listener, bearing witness to how sociality itself has long been forged "through the practice of listening intently to voices that do not necessarily have human bodies as their source" (Pettman 2019, p. 19).

Pettman's account thus operates on a dual register: on one side, the personal interiority of the intimate experience with the artificial voice; on the other, an extended and heterogeneous voice of the world. This duality echoes the extimate quality of the voice, to use the Lacanian neologism: belonging completely neither to the body nor to language, neither to the subject nor to the Other, the voice is situated at the intersection of these elements, as a remnant produced in the subject's relationship with the Other (Lacan, 2004). Pettman himself invokes *extimité* to name the way "the internal, psychic world is created by the infolding of the wider world" (Pettman, 2019, p. 113). The voice of the world, in order to be perceived as a voice at all, must relate to the subject in an intimate manner.

What is required, therefore, is an attentive, environment-oriented listening that illuminates the sonic intimacy of voices otherwise "dormant" in the world, a posthuman, post-anthropocentric listening that allows one to perceive the aural punctum even in non-biological agents (Pettman, 2019, p. 85). Pettman adapts the term punctum from Barthes's account of photography, where it names "that fatality which, within it, pricks me (but also wounds me, seizes me)" (Barthes, 2003, p. 28). Transposed to sound, the aural punctum names that peculiar voice effect, resistant to verbal explanation, found even in artificial voices: it "conveys the improper within the proper, the plural within the singular, the contingent within the essential, the alien within the human – and vice versa" (Pettman, 2019, p. 48).

It is in this kind of posthuman listening that sonic intimacy comes to life, transforming the voice from a static, traceable property of a sound into an emergent property, a relational event, regardless of whether the entities involved are human or non-human (Napolitano, 2022). Through the medium of the voice a "vector of connection or *sympathy*" is created between producer and listener: we become subjects of the call of other entities (Pettman, 2017, p. 53).

Read through this lens, the two case studies acquire their full significance. When Adachi locates his artistic identity precisely at this boundary between his voice and the machine's, and when Walshe improvises with a chaotic interlocutor who "was also me" (Walshe, 2023), both performers are practicing exactly the posthuman listening Pettman describes: they allow themselves to be re-interpellated by a non-human vocal source, one that is, uncannily, derived from their own embodied vocal history. The voice here is at once embodied and technological, situated yet transferable, a marker of identity that is detached, modeled,

transformed, and returned to the performer as an other. The performances stage and invite the audience into the voice's extimate position: an inner-outer space that is constitutively shared and traversed by others.

6. Conclusion

This paper asked how the integration of GenAI into contemporary musical practice reshapes musicological concepts of authorship, performance, and agency. The analysis of Tomomibot and ULTRACHUNK supports the following answer. In these works, creativity is not the property of an autonomous subject, human or artificial, but a relational and distributed process emerging from a heterogeneous assemblage: vocalists, neural networks of specific and distinct architectures, programmers and system designers, training datasets sedimenting years of vocal practice, interfaces, and the shared performance environment. 'Artificial creativity' is therefore best understood not as simulation of human creativity but as a materially embedded process of distributed agency, in which improvisational structures take shape from the human-AI interaction itself.

At the same time, the case studies show that distribution is asymmetrical. The systems differ in what they learn, generate, and contribute, sequence prediction over a designed vocabulary in Tomomibot, live latent-space synthesis of audiovisual material in ULTRACHUNK, and in both, decisive dimensions of the music, long-term form above all, remain with the human performer. A posthumanist account decenters the autonomous human subject without declaring humans and machines equivalent; it keeps distinct agency and intention, generation and authorship, interaction and autonomy.

The contribution of this study is to propose a posthuman redefinition of musical creativity that integrates philosophical theory with close analysis of contemporary experimental vocal practice, offering new conceptual tools for understanding AI-mediated composition and performance within musicology. Its specifically musicological claim concerns the voice: as a privileged bridge between biology and technology, the voice enables a posthuman listening in which the biological and the computational appear not as opposing poles but as elements situated within a continuum The performances examined here represent an attempt to inhabit the extimacy constitutive of both voice and subjectivity: an inner-outer space within us that is constantly shared and traversed by others, with whom we interact to rethink and shape new ways of being together.

7. References

Adachi, T. (n.d.). *AI improvisation*. ADACHI Tomomi. https://www.adachitomomi.com/ai-improvisation/.

Akten, M. (2018). *Ultrachunk*. https://www.memo.tv/projects/2018/ultrachunk/.

Barthes, R. (2003). *La camera chiara: Nota sulla fotografia*. Einaudi.

Braidotti, R. (2013). *The Posthuman*. Polity Press.

Braidotti, R. (2019). *Posthuman Knowledge*. Polity Press.

Deleuze, G., & Guattari, F. (1987). *A Thousand Plateaus: Capitalism and Schizophrenia*. University of Minnesota Press.

Di Scipio, A. (2023). In viva voce. In M. Garda (Ed.), *La mediazione tecnologica della voce* (pp. 155–190). NeoClassica.

Ferrando, F. (2019). *Philosophical Posthumanism*. Bloomsbury Academic.

Haraway, D. J. (1991). A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century. In *Simians, Cyborgs, and Women: The Reinvention of Nature* (pp. 149–181). Routledge.

Haraway, D. J. (2016). *Staying with the Trouble: Making Kin in the Chthulucene*. Duke University Press.

Napolitano, D. (2022). *La voce artificiale. Un'indagine media-archeologica sul computer parlante*. Editoriale Scientifica.

Occhio, M. (2023). La voce come frame. Caratteri e spazializzazione del vocalico. In M. Garda (Ed.), *La mediazione tecnologica della voce* (pp. 83–94). NeoClassica.

Paradiso, P. (2025). Voice contamination in contemporary music: Towards a posthuman interpretation of music through AI. In S. Sacchetti & N. Conci (Eds.), *Contamination: Proceedings of the 4th DiMMI conference, held in Trento, 29–30 November 2024* (Vol. 2, pp. 32–35). Università degli Studi di Trento. https://hdl.handle.net/11572/469393.

Pettman, D. (2017). *Sonic Intimacy: Voice, Species, Technics (or, How to Listen to the World)*. Stanford University Press.

Walshe, J. (2023, December 15). *Unsound Dispatch: 13 ways of looking at AI, art & music*. Unsound. https://unsoundfestival.substack.com/p/unsound-dispatch-13-ways-of-looking.