UC Riverside
Conference Papers

Title
“Duo Squared” and “Mirage”: Reflections on playing with AI versions of ourselves

Permalink
https://escholarship.org/uc/item/2hz056zd

Authors
Albert, Jeffrey
Cammarota, Anthony

Publication Date
2026-06-01

Peer reviewed

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

"Duo Squared" and "Mirage": Reflections on playing with AI versions of ourselves

Jeffrey Albert
Georgia Institute of Technology
jeff.albert@music.gatech.edu

Anthony Cammarota
Georgia Institute of Technology
acammarota3@gatech.edu

Abstract

This article reflects on "Duo Squared" and "Mirage," two improvisational works that pair human musicians with AI-generated versions of themselves. Both works use Somax2, an interactive music system that recombines analyzed slices of audio corpora in response to live performers. Their bespoke corpora are drawn from the H2H Music Improv Dataset, specifically recordings of the participating musicians improvising together. Consequently, the systems reproduce not only each performer's characteristic sounds and gestures, but also musical responses originally created in dialogue with a particular partner.

Through practice-based reflection, the authors examine whether musical AI should resemble human collaborators, generate unfamiliar behavior, or operate as a reflective tool through which improvisers can interrogate their own creative practices. Performance experiences reveal moments in which distinctions between human and AI partners become difficult to hear, complicating perceived boundaries between performer, corpus, and system.

The article argues that the value of bespoke-corpus improvisation lies not solely in successful human imitation, but in its capacity to expose what musicians recognize, expect, and value in interaction. It concludes by identifying further questions concerning corpus temporality, creative advancement, and the social dimensions that may be lost when AI substitutes for human collaborators.

Keywords: improvisation; Somax2; interactive systems; bespoke corpus

1. Introduction

Musical improvisation with interactive computer systems allows us to explore several questions. Do we want the computer system to behave musically like our human improvisation partners? If yes, why? If no, what are we looking for from the computer system that are absent from human interaction? Is compelling musical performance the goal of this endeavor, or are we really using these interactive systems to interrogate our own musical practice? How do we get generative systems to make low level musical gestures that "feel" human/musical? These are some of the questions that arose or re-emerged while making the pieces "Duo Squared" and "Mirage." This essay will not answer all these questions, and it may end up posing more questions than it answers, but it will examine these ideas through the lens of these two pieces that were a part of the Co-Creativity in Music, Sound and AI Conference at University of California Riverside in June of 2026.

"Duo Squared" is an improvised duet between two human musicians and AI versions of themselves that is part of the virtual exhibition in CCMSAI 2026 (https://audiovisualmusic.ucr.edu/co-creativity-exhibition/duo-squared/). The piece uses the Somax2 system, developed at IRCAM, (Borg et al., 2025) which uses a corpus of audio or MIDI data that is sliced and analyzed such that the slices of the corpus can be reassembled in reaction to input from a human improviser. The corpora used in "Duo Squared" are recordings

of the performers created in dialog with each other for the H2H Music Improv Dataset (Ma et al., 2026). The existence of these recordings allows for an exploration of how bespoke corpora may increase the performers' sense of interaction with Somax, as well as our sense of how human-like the Somax output can be.

"Mirage" is a similar piece that was performed on the Saturday night concert at CCMSAI and adds an additional AI musical voice to the two human improvisers and the AI versions of themselves. This additional voice also uses audio corpora that were recorded in duet with the human improvisers as part of the creation of the H2H Music Improv Dataset.

2. How "human" should it be?

If a computer system simply makes music like a human, why do we need the computer system? One answer is that we may not. Another common answer is that the computer system can surprise us in ways that humans do not or usually do not. Should surprise be a goal when improvising? George Lewis reframes the quest for surprise saying, "rather than making a benchmark from the production of surprise, my goal in making musical AI is to avoid the kind of familiarity that prompts contempt." (Basica & Zhu, 2026)

The similarity of the Somax output to what Jeff would normally respond with provoked new reflective questions: Is it possible that the times the computer system really does act like a human is where the surprise lives? There is a moment, starting around 4:44 and going through 5:03 in the video of "Duo Squared," in which Jeff smiles a couple of times while the Somax trombone (a corpus of his own playing) played with human Anthony. After the take ended, he commented to Anthony, "I almost laughed out loud because that was completely something I would have played there."

If regular human-like interaction is not the goal per se, and we are not chasing surprise, then why are we improvising with computer systems? Could these systems serve as mirrors that allow us to interrogate our own human creative praxis? Playing with the system and analyzing the results, then tweaking the system or corpus, and repeating the process lets us consider the aspects of these creative acts that compel or inspire us. Will working with the system help us see when we are nearing the "familiarity that prompts contempt?" We come to know our own playing and what we value in improvisation partners by designing, playing with, and analyzing these systems and their output. As Ritwik Banerji writes about musicians interacting with his system called Maxine:

> Commenting on the system's resemblance to or difference from human musicians enabled improvisers to narrate their experience of co-presence and interaction with other performers with a level of detail that is rarely afforded by their typical modes of sociality. (Banerji, 2021)

Using an interactive improvisation system to interrogate our own improvisatory practices is certainly a reason to develop and play with these systems. If we also perform with these systems, we must consider their musical output since we are now asking audiences to enjoy, or endure, it. In 2010, while Jeff was working on his "Interactive Musical Partner" project (Albert, 2013), Jeb Bishop told him, "you will never make it think like a human." Jeff's response was that he did not care if it "thought" like a human, he simply wanted it to be fun to play with and that it to make good music. Setting aside the likely impossible task of defining "good music," let's look at the idea of being fun to play with, or put another way, resembling human interaction. Banerji writes about "how the design of these systems is fundamentally oriented

towards creating the experience of interacting with one or more members of the human species." (Banerji, 2021) He unpacks this statement to reveal the, possibly obvious, idea that each of these systems are meant to be human-like in terms of a specific human, not some general ideal human. And there is certainly no general ideal musical improviser.

3. Bespoke Corpora

The audio corpus integration in Somax allows us to take the idea of modeling the interaction or behavior of a particular human to a new level of specificity. Somax is literally creating its responses from slices of music that were originally made by a particular human. Much of the creative act in playing Somax comes from the choice of the corpus. For these pieces, we used recordings from the H2H dataset that were real-time musical dialogs. This allows Somax to create not only an AI version of Jeff, but an AI version of Jeff based on his musical reactions to Anthony, and vice versa.

The H2H dataset is "six hours of audio-visual expert duo improvisations with clean per-player stems and per-player annotations of both their own intentions and their perception of their partner's intentions." (Ma et al., 2026) The dataset was created to help researchers develop ways of understanding communication between improvisers. An unintended beneficial by-product of the process was that Jeff and Anthony have an hour of recordings of their improvised musical interactions in separate isolated audio files which can be used to make bespoke Somax corpora. There is also an hour of trumpeter Alexandria Smith playing with Anthony, and another hour of her playing with Jeff. These recordings are the third AI voice in "Mirage." These bespoke corpora are notable in that they are not simply recordings of a trombonist improvising with a guitarist, but they are *this particular* trombonist improvising with *this particular* guitarist. That means that the specific ways that Jeff reacts musically to Anthony, and vice versa, are available to the AI. This allows the AI duets to be almost indistinguishable from the human duets, entangling the human and AI interactions in duo or trio formats. During live performance, the audience is able to see the difference, but possibly not hear the difference, challenging the perceived boundary between the human performers and their AI versions. According to Albert, "if I am playing with my eyes closed, there are times when I am not sure if it is human Anthony or AI Anthony playing with me. The sounds are coming from the same speaker, and the tone and gestures are so similar that I only really know which it is if I look."

4. (New?) Question(s)

Post hoc reflection led to an additional curiosity that the authors would like to explore further: when playing with each other, would Somax select time slices that directly align to the original temporal sequence of the corpora recordings? Presently, this project is planning the next phase of analysis to address this enquire and the authors intend to report their findings after completion. However, this brings us back to the question: do we want Somax to just do what we do? An interactive system that is programmed to exactly mimic human creativity may be an interesting computer science problem to solve, and the process of creating it certainly gives us a mirror on our own creative practice, but does it advance our creativity or give us more interesting music to hear?

Returning to Banerji's idea that while the goal of much AI research is to duplicate the cognitive capabilities of humans, the only way we can really judge the output is through the experience. If the experience is like interacting with a human, do we call it successful? The experience of

playing with these bespoke corpora in Somax is very similar to the experience of playing with the human who played the music that became the corpus. The computer and gear needed to play with a Somax version of Jeff or Anthony certainly takes up less space in the car than Jeff or Anthony and does not need its own plane ticket. It is also much less fun at dinner or the bar after the concert. Does saving space or airfare make it worth removing one, or both, of the humans from the performance? Can the music be the same without the extra-musical social aspects of music making? The answers to these questions may not be universal, and are certainly beyond the scope of this project, but they are nonetheless important to the future of human-AI system design.

5. References

Albert, J. V. (2013, April 5). *Interactive Musical Partner: A System for Human/Computer Duo Improvisations*. http://etd.lsu.edu/docs/available/etd-03272013-153428/

Banerji, R. (2021). *Artificial Intelligence and Phenomenological Ethnography*. https://academic.oup.com/edited-volume/38571/chapter/375437137

Basica, C., & Zhu, J. (Eds.). (2026). *Musicians Coding AI for Themselves*. Routledge. https://doi.org/10.4324/9781003615873

Borg, J., Fiorini, M., & Assayag, Gérard. (2025, March 20). *Somax2 User's Guide*. IRCAM. https://github.com/DYCI2/Somax2/blob/master/Somax2%20User's%20Guide.pdf

Ma, A. T., Cammarota, A., Wang, J., Smith, A., Huang, C.-Z. A., Albert, J., & Lerch, A. (2026). *H2H Music Improv: A Communication Model and Audio-Visual Dataset for Music Improvisation* (arXiv:2608.13957). arXiv. https://doi.org/10.48550/arXiv.2608.13957