Artificial Authenticity - Real Authority: 你們是蟲子] (You are bugs), an AI-assisted playbite on Luigi Nono's ...*sofferte onde serene...* for piano and tape

Pavlos Antoniadis
University of Ioannina & IRCAM
p.antoniadis@uoi.gr
antoniadis@ircam.fr

ABSTRACT

On the occasion of Luigi Nono's birth centenary, I respond to his classic work for piano and tape, "..sofferte onde serene..", with a new work by the title 你們是蟲子 (You are bugs). This work extends Nono's techniques through live AI tools and interactive multimedia. It reflects on problems of authorship, authenticity and authority in musical AI, addressing the disturbing potential of human inferiority – the fact that we all are potentially "bugs" (with its multiple connotations), and we might have to accept a lower status in a new (or not so new) hierarchy. To do so, the work assumes an AI, as if alien, invasion of sorts, inspired by the "dark forest hypothesis", term coined by the Chinese sci-fi author Liu Cixin as a response to the Fermi paradox. The musical score consists of a fixed video, featuring borrowed cinematic materials of authoritarian figures, which are conducting, instructing, ordering, insulting and abusing the performer; augmented interactive tablatures; music notation and video synths. The score is projected in real-time and drives a combination of improvised, comprovised and fixed musical materials. These include the author's previous recordings of works by Luigi Nono, Iannis Xenakis and Maurice Ravel, which are remixed through a combination of SOMAX2 and GesTCom and are live controlled through inertial sensors and an MPE controller. Thus, both human and machine are trained and constrained by each other, and sparks of resistance to the machine are not missing.

1. INTRODUCTION

The advent of democratized AI tools based on generative architectures (GAN, RNN, CNN, Transformer, VAE and others) triggers new questions about the relation between authenticity and creativity [1], [2]. So do shallow learning approaches to generative AI, including methods such as probabilistic models, rule-based systems, Markov chains, genetic algorithms, and classical statistical techniques [3]. These methods generate content based on predefined rules, statistical relationships, or evolutionary processes, rather than the neural network architectures characteristic of deep learning.

But how novel are these questions? In the domain of musical performance and composition, the relation between authenticity and creativity has always been tensed, throughout both the history of contemporary music performance practice for traditional instruments, and the history of computer music performance practice for fixed or real-time interactive media. The emergent questions today are of a completely different order, but they still relate to the original aporias in often surprising ways to be explored in this paper.

Following up from this, I start by reviewing a classic problematic case, Luigi Nono's *..sofferte onde serene..*[4] for piano and tape, whose authentic performance by the standards of the so-called "high modernist model of performance practice" has been rendered impossible, due to parameters both endogenous and exogenous to the work. I subsequently address the relation between authenticity and creativity by reviewing three concepts, including Brian Massumi's notion of play as metacommunication - the conceptual background for the notion of "playbite" that features in the title - as well as my notions of Artificial Authenticity and Real Authority. I review the notion of Artificial Authenticity to address the creative and (auto)poietic aspects of any reconstruction of this type, before switching registers to link Artificial Authenticity to generative Artificial Intelligence.

To address this link, I present an analysis of my own piece, *你們是蟲子 (You are bugs),* which extends Nono's techniques through generative AI tools and multimedia based on a motion-following architecture, namely *GesTCom*[1][5], [6], [7] and *SOMAX2*[2] [8]. The poetic concept behind this creation is Liu Cixin's *dark forest hypothesis* applied to an imaginary AI invasion, which creates the narrative basis for the notion of "Real

[1] https://medias.ircam.fr/en/x2253e1

[2] https://forum.ircam.fr/projects/detail/somax-2/

Authority" as a counterpart to Artificial Authenticity. Real Authority is expressed both in the Human-Machine relationship and in the machine musicianship aspects. I highlight the tension between a fixed electronic medium, namely a video score for *你們是蟲子 (You are bugs)* and the latent free spaces of live Human-Machine interaction, ranging from free improv on Nono's materials, to drones and generative improvisations introduced through the *SOMAX2* and *GesTCom.* Finally, this analysis, which is accompanied by the video score with audio realization, and by a video performance documentation, serves as a trigger for further discussing issues relating to the semantic blind spots of generative AI and to ways of creating meaning through cyber-human musicianship.

2. *…SOFFERTE ONDE SERENE…* AND THE HIGH MODERNIST MODEL OF PERFORMANCE PRACTICE

In [9] Sebastian Berweck provides an overview of the challenges and remedies in contemporary music performance with electronics. One of his case studies is Luigi Nono's *..sofferte onde serene..*, (suffered, serene waves) for piano and tape. In assessing the problems for performance, Berweck points to both audio quality and synchronization. Both types of problems are related to the original mono tape and the denoising technique used during its digitization. The problems leave no realistic remedy options to the performer other than a complete restoration project, such as the one undertaken by Paulo de Assis [10].

Unlike Berweck, whose "findings are almost tragic" for performance authenticity, De Assis considers both the poetics of the work's creation and re-creation as breaking into new ground. Through meticulous analysis of "several new elements [..], namely a new understanding of the use of vertical sound-aggregates ("chords"), the exploration of complex variational and canonical procedures, and, crucially, new modes of organizing "multi-temporalities," with the piano and the tape [..]" [11, p.206], De Assis asserts that "its concert rendering involves various degrees of uncertainty and unpredictability of sonic combinations" [11, p.206].

But the elements of uncertainty and unpredictability are not features of a re-creation strategy alone: They had already infused the very fabric of the work's creation. The piece was indeed based on a creative collaborative process between its dedicatee, Italian pianist Maurizio Pollini, and the composer. The process consisted of recording Pollini's improvisations and using them as input to be processed and remixed by the composer and sound engineer Alvise Vidolin, in ways that promote (con)fusion in the sonic and temporal realms:

> "This (con)fusion is enhanced by relatively free time relations between piano (live) and tape, allowing the performer on the piano and the performer controlling the sound-projection to intertwine a great variety of sonic relationships". [11, p. 206]

The score was composed as a reflection on the recorded improvisation, reversing the usual arrangement, whereby a performance follows the composition of the score and a recording follows a performance:

> "It was the concrete recorded sounds that slowly, in a constructive way, defined more and more precisely the sequence of sonic events—that is, the score for the pianist playing the piece's piano part. And if it is very clear that the score and its writing are the complete responsibility of Nono (who remains "the composer" in an orthodox sense), it is also true that the concrete sonic input produced by Pollini was of the utmost importance for the definition of the music". [11, p.207]

This flexible and unorthodox relationship between improvisation, score, recording and live piano performance manifests in the phenomenology of performance as a dialogue of echoes and resonances, which forms the essence of Luigi Nono's late compositional style, both aesthetically and politically.

De Assis's consideration of the multiplicity of relations between improvisation input, recorded materials, notated scores and final performances is crucial for the project described in the next sections. They particularly inform the sort of relationship sought after between AI agents and human performer. Before touching upon that, I would like to utter another word on the implications of Nono's approach for the wider context defined by Franklin Cox as the "High-Modernist Model of Performance Practice" (from now on abbreviated as "hmmpp").

In [12], Franklin Cox defines hmmpp as "a "noisefree", transparent relationship between all elements of the communicative chain [between conception, notation, performance and reception]" [12, p. 71], The hmmpp elaborates on the idea of a fully deterministic score-image. Such notation is considered able to enclose every possible characteristic of the sonic phenomenon it represents. As a result, it is supposed to assign concrete tasks to the performer and well-defined perceptual tropes to the listener. Cox suggests that complex music has brought about a "fundamental paradigm shift" away from that model. The reason is its purely quantitative characteristics, namely:

> "[..] extreme degrees of both density and fine detail, and [..] coalescence of highly rationalized materials, notated challenges and organization with an extreme physicality and almost irrationality of results" [12, p. 70].

The paradigm shift consists in the transformation of the above-mentioned communicative chain into "[..] an overlapping series of volatile conflicts between incompatibles. Thus, notation is treated as an essentially opaque medium and such notation demands less reading than decipherment" [12, p. 76].

From the discussion above, it becomes clear that the discrete existential domains of improvisation, recording, score and reproduction in Nono may be considered as a transposed instance of Cox's "overlapping series of volatile conflicts between incompatibles", yielding as a result the unpredictability and creativity that De Assis

considers a cornerstone for performing …*sofferte onde serene*… today.

But what if this notion of recreation as volatile conflict between incompatibles is further upgraded to the register of Human-Machine Interaction? How does one notate and represent a construction that potentially undermines representability? Could authenticity be redefined as Artificial Authenticity, when attempting the reproduction of Nono's "non-linear" techniques through machine musicianship? And how are these questions to be treated at the age of Generative AI? To answer these, we resort to the concept of play as metacommunication, in the instance of a 'playbite', developed by Massumi after Bateson.

3. PLAY AS METACOMMUNICATION

In [13] Brian Massumi offers an interesting interpretation on Bateson's definition of animal play.

"The play bite that says "it is not a bite" has the value of the analogous action without its force or function. The wolf cub says through his teeth: "this is not a bite; this is not a fight; this is a game; I am hereby placing myself on a different register of existence, which nevertheless stands for its suspended analogue." [13, p.4]

This take points towards play as some sort of analogous metacommunication of mutually inclusive (rather than exclusive) registers of existence: Fight includes play and play includes fight, or rather they are parts of a continuum that is differentiated by intensity and function. At the same time, this notion of play carries a semantic layer, thus the term "metacommunication", in the sense that through its analogy to fight it makes a non-verbal commentary, a statement, affirming the nature of a game rather than the nature of a fight.

Could we possibly think of the relationships between improvisation and composition, authenticity and authority, humans and machines, along these lines? What if, accepting that an authentic performance of …*sofferte onde serene*… today is irrelevant, we switch gears to a mutually inclusive register of play between what Cox perceives as incompatible existential domains in volatile conflict? This mutually inclusive register of play would denote: "this is not …*sofferte onde serene*…, nor is it a fight for authenticity, but a playbite on artificial authenticity in the face of generative AI". How is such a dynamic of metacommunication representable, both as musical material and as interaction schemes? And what eludes representation?

In that vein, I used …*sofferte onde serene*… as the basic material to create an "Artificially Authentic" double, whereby the "Real Authority" of the composer is outsourced to the score and to software agents, *SOMAX2* and *GesTCom*, trained on …*sofferte onde serene*…, filtered through other music and controlled through sensors, piano and an MPE controller in real-time.

4. ANALYSIS

你們是蟲子] (You are bugs) is conceived as a playbite on Nono's …*sofferte onde serene*…. The work is based on the dialectics between a static form manifested in the fixed video score (Figure 1) and a dynamic interactive process featuring media, human and artificial agents, and outputs (Figure 2).

4.1 Fixed form

The video score consists of a tripartite form. How do we represent a fixed video score? In an affirmation of Rodney Brooks' imperative "to use the world as its own model" [14], I employ a screenshot of the *iMovie* video editor used to create the score, further annotated in *Adobe Illustrator*. For the interaction part to be described in the next subsection, I employ a free adaptation of the visual language described in [15], designed in *Adobe Illustrator*.

The first part (*A*, 00:00 – 03:18) introduces the notion of Real Authority, associated with a video avatar of the author (Figure 3). The audio material is rooted on …*sofferte onde serene*…, including both a track of the original tape [16] and a track of a live recording of the work performed by the author [17]**.** The avatar visuals are based on video synthesized materials employing *Lumen*[3] and *OBS*[4]. Materials of feature films borrowed from *YouTube* are gradually introduced, presenting various authoritarian figures who instruct, order, insult and abuse the performer. The latter treats the recorded materials on stage according to the orders, remixing them live through *GesTCom* and *SOMAX2*.

The second part (*B*, 03:18 – 06:21) features a climactic opposition to Real Authority. The audio material introduces recordings of a work composed by Iannis Xenakis and performed by the author [18], remixed live through *GesTCom* and *SOMAX2*; as well as the invasion of electronic sounds, mostly pink noise, generated by the sound engine *Equator2* and controlled through the MPE controller *ROLI Seaboard RISE2*. The visual materials recede behind purple noise, featuring a single return of a cinematic snippet of yet another authoritarian figure.

The third part (*C*, 06:21 – end) features the return of the Nono material, this time filtered via *SOMAX2* through a recording of music by Maurice Ravel performed by the author. An improvisation on the acoustic grand piano, again based partly on Nono and partly on Ravel, is the only engagement of the author as a pianist performing live (as opposed to remixing). The visuals feature the authoritarian video avatar of the beginning, infused with the purple noise background and yet another snippet from a different film.

4.2 Developing variation

In this section, I explore the dynamic interplay of each part's materials, zooming in to identify detailed parameters and zooming out to identify interaction patterns.

Part *A* consists of two rather distinct sections: A static subsection *a*, which is based on an exposition of the fivefold varied repetition of a loop, consisting of the first 32s of …*sofferte onde serene*…'s tape, and on a recurring video-synthesized avatar of the author, instructing the performer; as well as a subsection *b*, initiating the dynamic

[3] https://lumen-app.com/

[4] https://obsproject.com/

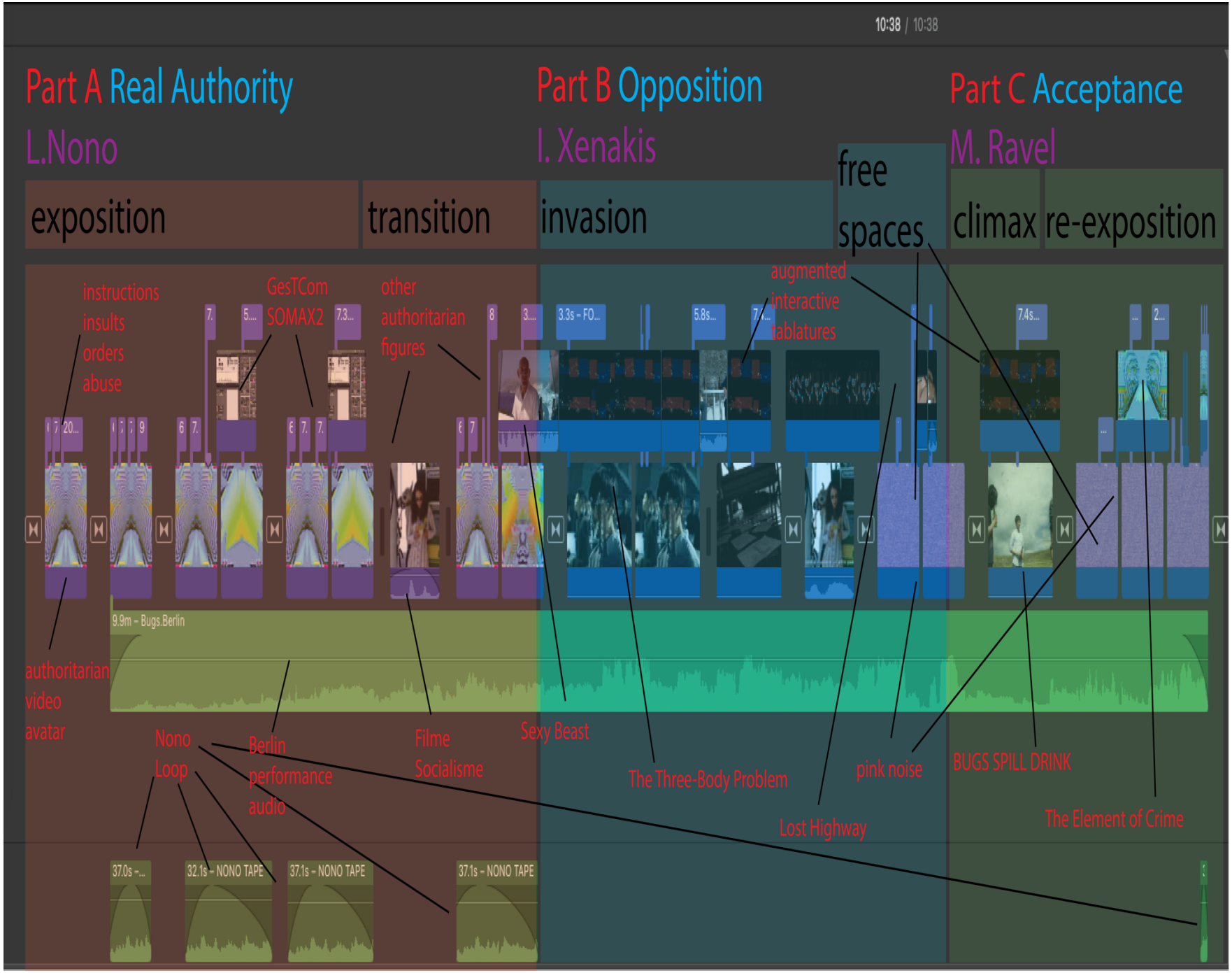

Figure 1: Overview of video score analysis as presented in section 4

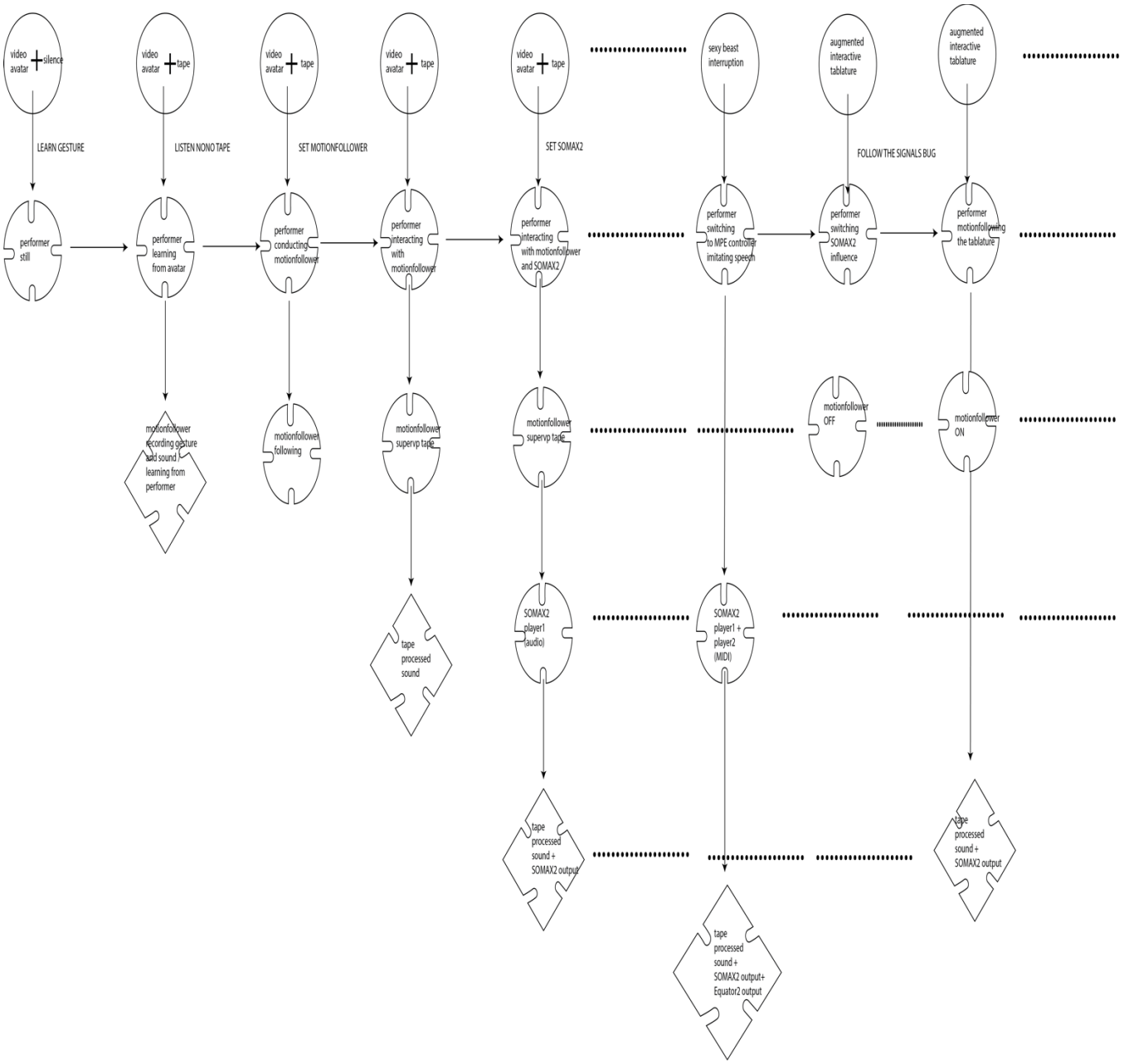

Figure 2: Interaction patterns with media, agents and outputs for part A, as described in sections 4 and 5. New media, agents and outputs, are indicated by circles, edged circles and edged rhombuses respectively. Please notice the proliferation and over-layering of agents and outputs over time.

transition to the next part, which consists of two interruptions of the *a* loop by cinematic materials: a snippet from *Film Socialisme* (2010) by Jean-Luc Godard [19], whereby a teenage girl in a rural gas station in France (played by Agatha Couture) ignores a couple of German tourists asking for instruction (ACH, DEUTSCHLAND, 02:37), and a snippet from *Sexy Beast* (2000) by Jonathan Glazer [20], whereby Ben Kingsley as Don Logan attempts to recruit a retired British gangster for a coming gig (03:12).

The piece starts with an authoritarian video avatar gesticulating in silence and assuming the role of the composer. Inscriptions that function as exclamations and/or instructions appear in the projected video score, for example: "你們是蟲子 (YOU ARE BUGS)". The scene evolves as if performing hypnosis and the inscriptions indicate actions to be performed in real time, driving the interaction. They unfold in ever-expanding cycles, that introduce new media, agents and outputs, as indicated in Figure 2 by circles, edged circles and edged rhombuses respectively. Thus, in the first cycle the inscription LEARN GESTURE orders the performer to mentally imitate the silent gestures that the pseudo-composer is performing. In the second cycle, the beginning of Luigi Nono's original tape for *…sofferte onde serene…* is introduced (LISTEN TO NONO TAPE) and the performer is sarcastically instructed to suffer (SUFFER) while imitating the gestures s/he has learned silently; at the same time, the performer's gestures train a motion follower system recording sound and movement as part of the *GesTCom* system. Thus, the power structure between the video score and the performer expands towards the motion follower that assumes the status of agent in the next cycle. In the third cycle, the performer conducts the motion follower that follows obediently, or varies her movements to obstruct the following, thus over-layering the output sound with frozen drones and spectral artifacts. A significant degree of unpredictability is introduced. The fourth cycle introduces the AI agent, *SOMAX2*, which filters a recording of a performance of the author through the Nono tape, adding to the previous mix. The cycle is abruptly interrupted by the introduction of the first snippet of cinematic materials described above, whereas the fifth cycle, wondering HOW LONG IS THIS GONNA TAKE, warns OUMUAMUA INCOMING, a reference to the first interstellar object formally designated 1I/2017 U1 and detected in 2017, and introducing the second cinematic clip described above.

Part *B* (03:18) introduces new material in both sound and image, capitalizing on an audio recording of Iannis Xenakis's *Evryali* (1973) [21] remixed through *SOMAX2* and *GesTCom*, filtered through Nono's work in the second version of the author's performance, as well as on augmented interactive tablatures based on Xenakis's sketches for *Evryali* and functioning as graphic scores for the performer – another *GesTCom* application. The tablatures are combined with materials from the last episode of the Chinese series adaptation [22] of Liu Cixin's *The Three-Body Problem* [23], as a recurrent threatening and awe-inducing inscription in the retinas of

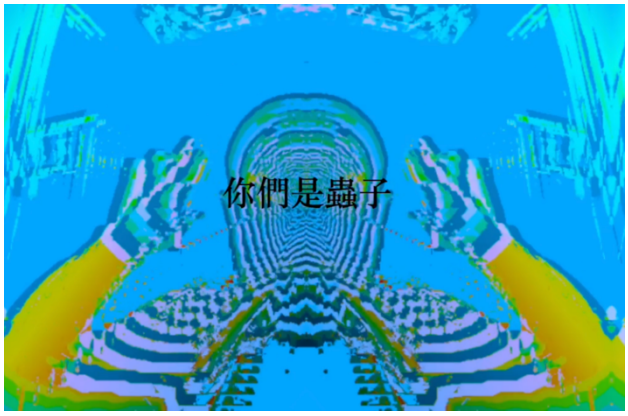

Figure 3: Avatar figure representing Real Authority, Part A

higher human officials (04:03), originating in an alien and superior civilization, the *San Ti*, who invade the Earth after a signal sent to them by a Chinese scientist in the years of Mao's Cultural Revolution. This development is interrupted yet again by a clip from *Sexy Beast,* with Don Logan in ever more abusive appearances, as well as a clip from *Film Socialisme*, with the teenage girl exclaiming that the tourists should "go invade another country". This moment signals an exodus of the score into the "purple noise" segment, introducing a free improvisation on *ROLI Seaboard RISE 2* MPE controller - a sort of tone (or rather pink noise) painting of the various invasions alluded to in the score, produced by the *Equator2* synthesizer (Figure 4). Despite the specification NOTHING TO SEE HERE, as an indication and invitation to free and unrepresentable spaces, the climax recedes fast to give its place to a hidden drone based on *...sofferte onde serene…*, and introducing a clip from David Lynch's *Lost Highway* (1997) [24], and yet another abusive figure, Alice (played by Patricia Arquette): Pete and Alice arrive at a remote desert cabin to fence stolen goods. During sex outside, Alice taunts Pete, saying, "You'll never have me," before disappearing into the cabin. This short episode marks the return of the climactic material comprised by a combo of Xenakis' *Evryali* filtered through Nono and pink noise on *ROLI*, before the introduction-premonition (06:08) of the Ravel material: an unpublished recording of Ravel's *Oiseaux Tristes* (nr. 2 in the collection *Miroirs*, 1904-1905) at IRCAM [25].

Part *C* starts at 06:23 with a stunning moment from the *Three-Body* series, whereby, in a gesture of acceptance of inferiority and animalistic-bug survival in the face of the alien invasion, the main characters perform a unique gesture of spilling their drinks (the equivalent of an ancient Greek *sponde*) with the epigram BUGS SPILL DRINK. The gesture marks both the reprise of the initial avatar material, as well as the first introduction of live piano playing which continues to the end, an improvisation on *…sofferte onde serene…* materials aimed to "zombify" the Pollini improvisation as transmitted through the Nono tape. The initial epigrams-instructions are slightly varied, through the inscription LEARN COMPOSURE AS WELL

and there is only a short interruption-reminiscence of the climactic part, in the form of a clip by Lars von Trier's *The Element of Crime* (1984) [26] - an image of a horse (yet another animal) accompanying the SUFFER instruction. This Nono tribute ends with a final loop of the tape, over the inscriptions SUFFER SERENE WAVES (the translation in English of *sofferte onde serene*).

Figure 4: Pink noise patch in *Equator2*

5. INTERACTION WITH AI AGENTS

In this section, more technical details are provided on the interaction patterns between the AI agents and the human performer as diagrammed in Figure 2.

5.1 LEARN GESTURE: The use of *GesTCom*

By virtue of a predefined syntax of movement and *a* probabilistic motion-following methodology employing Hidden Markov Models [27], *GesTCom* can be trained to follow (or unfollow) the performer during variations of the initial performance. This process involves a *recording phase* and a *following phase*. In the recording phase, the user-performer follows the movements of the avatar figure in Part A to train *GesTCom*'s *motionfollower*. Alternatively, performers can also follow any mobile element of the augmented interactive tablatures (in part B 03:20 – 03:43), which is set to move at a desired speed, like a classic metronome would do.

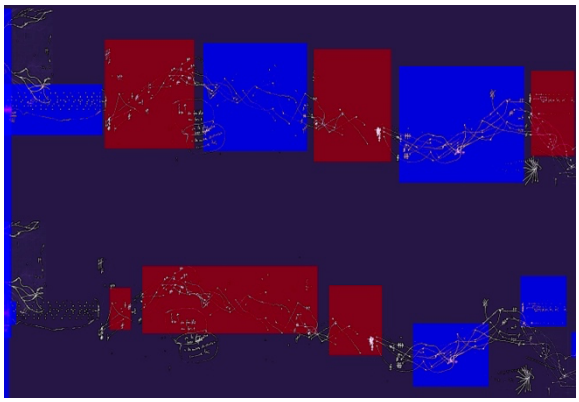

Figure 5: explicit score mapping on a Xenakis sketch for *Evryali* for motion following, Part *B*, realized with *INScore*. Sketch transcription by Aurélien Duval.

The musical sketch of Figure 5 has already been graphically segmented through the INScore and assigned a duration according to a space-time formalism (explicit mapping). In this phase, the motion follower "learns", so to speak, the mapping from the performer's gesture captured by R-IoT IMUs (implicit mapping), while s/he follows the video avatar or the explicit mapping of the video or sketch. In the next phase, defined as "following", the performer can pursue highly varied performances, ranging from heterophonic re-interpretations of the original to the introduction of completely novel material that shares the same gestural segmentation. This time, it is not the performer that follows the system, but rather the system that follows the performer, given that the segmentation is correct and common in all these varied performances. Thus, the performer may control the mobile elements of the tablature. The feedback of the follower has been extended to score compound representations. The gesture-following has been turned into score-following.

In Figure 6, the grey signal represents the implicit mapping - gesture that the augmented multimodal tablature "learns" along the explicit mapping of Figure 5 in the *recording phase*. The green signal represents the new incoming signal that controls the tablature in the *following phase*, the signal that the tablature follows during variations of the initial performance.

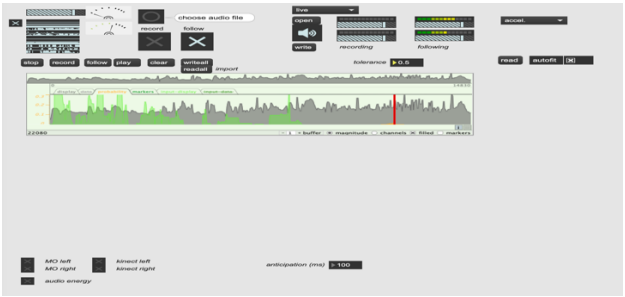

Figure 6: interface for motion following

More than following, *GesTCom* here has a function of unfollowing: Learned gestures are intentionally obstructed, or the tolerance of the motion follower algorithm is set extremely low, so that the system cannot follow, thus sabotaging its initial intent**.** As a result, various drones occur, as the *supervp.scrub~* object (advanced phase vocoder position controlled player module) allows resynthesized audio output to be time-stretched, transposed in pitch and undergo several spectral envelope transformations (Figure 7). More importantly, these effects can be unpredictable. The triggering from the performer's conducting gestures is not hard-wired: It partly depends on the complexity of the input recording and the performer's knowledge of it, on the degree of accuracy of the performance during both the *recording* and the *following* phase, and on the algorithm tolerance threshold. In this respect, it acts as the meta-commentary or playbite evoked in the introduction, and it eludes the usual mapping strategies in precisely controlling parameters with hand gestures. As such, it allows for a

very flexible and often surprising re-shaping of the input recording.

The motion following software, alone or combined with the sketch-based tablature was used in combination with audio tracks by Luigi Nono, Maurice Ravel and Iannis Xenakis as described above. The first introduction of the system's following function is featured in 01:42 – 01:53 of the video score and in 0:41 onwards on the performance video.

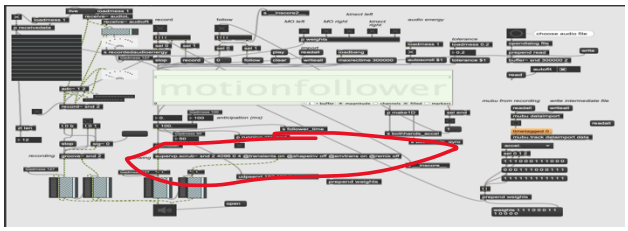

Figure 7: supervp.scrub~object

5.2. Training the AI Agent for comprovisation: *SOMAX2*

The motion following output, alone or combined with the sketch-based tablature, was used in combination with audio tracks by Luigi Nono, Maurice Ravel and Iannis Xenakis, as well as with MIDI input from the *ROLI Seaboard RISE 2* as training and influencer material for *SOMAX2*. *SOMAX2* is a multi-agent interactive system performing live machine comprovisation with musicians, based on machine-listening, machine-learning, and generative units. The current version [8] is a recent development and algorithms' improvement from the former *SOMAX* version. Agents provide stylistically coherent improvisations based on learned musical knowledge while continuously listening to and adapting to input from musicians or other agents in real time. The system is trained on any musical materials chosen by the user, effectively constructing a generative model (called a *corpus*), from which it draws its musical knowledge and improvisation skills. Corpora, inputs and outputs can be MIDI as well as audio, and inputs can be live or streamed from MIDI or audio files. *SOMAX2* is one of the improvisation systems descending from the *Omax* software [28], presented here in a totally new implementation. As such it shares with its siblings, the general loop [listen/learn/model/generate], using some form of statistical modeling that ends up in creating a highly organized memory structure from which it can navigate into new musical organizations, while keeping style coherence, rather than generating unheard sounds as other Machine Learning systems do. The open-ended nature of the resulting system allowed me to experiment with different kinds of interaction and materials in a setting of structured comprovisation based on Nono's original tape, a recording of …*sofferte onde serene*…, as well as Xenakis' *Evryali* and Ravel's *Oiseaux Tristes*, and finally Xenakis's *Evryali* sketches. Capitalizing on the chroma affinities between Nono, Xenakis and Ravel, I was able to construct composite corpora including harmonically similar language.

In Part *A*, an audio influencer treating sequentially an audio file of the Nono tape part (for loop 2), and an audio file of a complete performance of …*sofferte onde serene*… by the author (for loops 3-5) was used in pitch detection mode (loops 2-3) and, as the density increased, in onset detection mode (loops 4-5). The function of influencing equals a type of stylistic filtering of the Nono original tape. In Part *B*, an audio influencer playing the piece's complete performance, as well as a MIDI influencer capitalizing on the live input from *ROLI*, filtered a recording of Iannis Xenakis' *Evryali*. Finally, in Part *C* the materials of reference in the Player oscillated between Ravel and the tape by Nono, filtered by the audio of the live piano performance and the MIDI input by *ROLI*. Please note that the employment of up to four Players is possible, which would probably explode the complexity of interaction described above.

Several degrees of distancing (same, similar and alien material) of the human pianist from the original material used to train *SOMAX2* result in a variety of interactions, reflecting the echoes and resonances of Nono's own aesthetics. In a second instance, the pianist performer can shape the output of the *SOMAX2* through the motion follower, allowing for the supervp.scrub~ object (advanced phase vocoder position controlled player module) audio output to be sent to *SOMAX2*.

Figure 8: SOMAX2 setup

5.3. New Virtuosity

The virtuosity involved in controlling *GesTCom* and *SOMAX2* in live settings is nontrivial, and so is the possibility to enter a dialogue with materials already learned and recorded, but also to surprise the systems as well as the human performer with alien materials not featured in the Nono score or in the recordings. Several types of action are employed, including: silent conducting; recording conducting gestures and sounds; imitating sound from fixed media; following and varying gestures and sound; launching programs and tweaking parameters; piano performance, including traditional score reading and free improvisation; controlled improvisation, including the performance of *GesTCom* tablatures and/or variations of the notated Nono score; and finally, performance of the MPE controller, including 5D touch such as note-on

velocity MIDI messages (strike), polyphonic aftertouch or channel pressure MIDI messages (Press), MIDI pitch bend value on a given MIDI channel (Glide), MIDI Control Change Message (Slide) and note-off velocity MIDI message (Lift)[5]. Employing a paradigm of interaction and musical context based on Rasmussen's model of human information processing [29], [30] we can here detect all three types of behavior: skill-, rule- and model-based behavior. Skill-based behavior involves real-time, ongoing responses to continuous signals. In contrast, rule-based behavior entails selecting and executing stored procedures based on cues extracted from the system. Finally, model-based behavior is even more abstract, with actions directed toward a conceptual goal, requiring active reasoning to determine the appropriate response—whether rule-based or skill-based. In the repertoire of actions described above, instances of skill-based and rule-based behavior are self-explanatory, whereas model-based behavior is detected both on the macroscopic level of the multimedia composition of the work, as well as on the microscopic but potentially very calculated responses to the semantic network established by the cinematic and video-synthesized materials. Thus, the work is inscribed in the general discussion on contemporary musical virtuosities as exposed in [31].

6. FROM SCORE INTERPRETATION TO HUMAN-AI COMPROVISATION AS PLAYBITE

The topic of algorithms inheriting human biases, both hidden and overt, has sparked heated discussions in AI research. Notably, epistemic biases [32], [33] often reflect preconceived ideas and models about human thought, learning, and judgment that are integrated into AI systems without much scrutiny. Consequently, these systems, along with their human components, are often stuck in traditional frameworks, missing out on modern advances in their areas.

The probabilistic systems used in *GesTCom* and *SOMAX2* reveal deep-seated epistemic biases, especially in how music knowledge is represented. Both rely heavily on the probabilistic scanning of audio—responding to gestures in the motion follower or analyzing pitch and chroma in *SOMAX2*. This reveals what some describe as AI's "semantic blind spot," [34] where states lacking memory swap syntactic roles without creating meaningful relationships or causal effects, unless genuine, unpredictable changes occur. Such change springs from the unknown, not from familiar, navigable data.

Looking back, one might argue that symbolic biases rooted in classic replication models—or their modern incarnations like high-modernist performance practices—are rigidly imposed. These include the unchanging nature of scores, their layered parameters, the symbol grounding challenge [35], [7]. These biases constrain innovative systems of improvisation, highlighting issues such as symbolic representation's limits when compared to real-world experiences, the preference for abstract sound patterns over rich social interactions, and simplifying complex dynamics into basic parameters [36].

However, the idea of creating diverse connections among distinct representation domains—like improvisations, recordings, videos, and scores—is promising. It allows for unexpected, personalized interactions within time, even if these domains merely rearrange fixed timelines. The potential for breaking one-to-one score associations, exploring one-to-many mappings, and creatively transforming them into video scores with both familiar and novel materials is enticing.

Finally, the question of a semantic blind spot is addressed through the inherent semanticity of the borrowed cinematic materials. A network of inherently meaningful materials is articulated into an overarching narrative form, counterbalancing the free play of semantically neutral permutations.

7. CONCLUSIONS

I have attempted to present a concise analysis of my piece 你們是蟲子] (*You are bugs*) reflecting on the relationship between authenticity and creativity in relation to generative AI. Reviewing Nono's late aesthetics, introduced by his piano and tape piece *..sofferte onde serene..* , I wonder if the poetics of transformation of an original improvisation by Maurizio Pollini into a fixed piano score accompanied by a tape echoing and ghostly reflecting the original could be more accurately rendered as an AI-generated reconstruction of the piece, whereby Artificial Authenticity is coupled to Real Authority. The term of choice in this case is the notion of "playbite", exhibited by the philosopher Brian Massumi after Bateson, as an analogical existential domain of Otherness, which also bears a semantic layer potentially invaluable, given the epistemic biases exposed in the last section as far as AI is concerned. After exposing my concepts, I offered an analysis of the video score and I elaborated on the tools and concerns they spark.

8. REFERENCES

[1] R. Mitra and I. Zualkernan, "Music Generation Using Deep Learning and Generative AI: A Systematic Review," in *IEEE Access*, vol.13, pp.18079-18106, 2025, doi:10.1109/ACCESS.2025.3531798 https://ieeexplore.ieee.org/abstract/document/10845168 accessed 06.10.2025

[2] Gioti A-M. "From Artificial to Extended Intelligence in Music Composition". *Organised Sound*. 2020;25(1):25-32. doi:10.1017/S1355771819000438

[3] Dubnov, S. and Greer, *Deep and Shallow. Machine Learning in Music and Audio.* New York: Chapman and Hall/CRC, 2024

[5] https://support.roli.com/support/solutions/articles/36000376229-roli-seaboard-2-creator-manual-previously-rise-2-

[4] Luigi Nono, *...sofferte onde serene...*, 1st ed. (Milan: Edizione Ricordi, 1977)

[5] Antoniadis, P., Bevilacqua, F. and Fober, D., "Gesture cutting through textual complexity: Towards a tool for online gestural analysis and control of complex piano notation processing", Proceedings of Conference on Interdisciplinary Musicology, 2014, Berlin, Germany

https://hal.science/hal-02158741/file/CIM14-final.pdf accessed 06.10.2025 accessed 06.10.2025

[6] Antoniadis, P. *Embodied navigation of complex piano notation: rethinking musical interaction from a performer's perspective.* PhD Thesis. Université de Strasbourg & IRCAM, 2018 https://theses.hal.science/tel-01861171/file/Antoniadis_Pavlos_2018_ED520.pdf accessed 06.10.2025

[7] Antoniadis, P. and Chemero, A. "Playing without mental representations: Embodied navigation and the *GesTCom* as a case study for radical embodied cognition in piano performance". *journal of interdisciplinary music studies* season 2020, volume 10, art. #20101207, pp. 126-174 https://unipub.uni-graz.at/jims/download/pdf/5758295?fbclid=IwY2xjawNQadJleHRuA2FlbQIxMABicmlkETFWOUhwVEpDeHJ3TjlpWWhQAR5uxtLEihsvyk6WjRATZ220hr08_s2etGWvhh7cRzEb2b1ldFUb_yfaM6HL9Q_aem_bLJSoZAhCp4QIXY9Yk_9OA accessed 06.10.2025

[8] Gérard Assayag, Laurent Bonnasse-Gahot, Joakim Borg, "Cocreative Interaction: Somax2 and the REACH Project". *Computer Music Journal* 2022; 46 (4):7–25. doi: https://doi.org/10.1162/comj_a_00662 accessed 06.10.2025

[9] Sebastian Berweck, *It worked yesterday: On re-performing electroacoustic music*, Doctoral Dissertation, University of Huddersfield, 2012, http://eprints.hud.ac.uk/17540/ accessed 06.10.2025

[10] Paulo de Assis, *Luigi Nonos Wende. Zwischen Como una ola fuerza y luz und .....sofferte onde serene...* (Hofheim: Wolke, 2006)

[11] Paulo de Assis, «Revisiting Luigi Nono's Suffered, Serene Waves », in: *Artistic Experimentation in Music: An Anthology* (eds. Crispin, D. and Gilmore, B.), Leuven University Press, 2014

[12] Franklin Cox, "Notes Toward a Performance Practice for Complex Music" in Claus-Steffen Mahnkopf, Frank Cox and Wolfram Schurig, eds.: *Polyphony and Complexity, New Music and Aesthetics in the 21st Century*, Hofheim: Wolke Verlag, pp. 70-132, 2002

[13] Massumi, B. *What Animals Teach Us About Politics*, Duke University Press, 2014

[14] Rodney A. Brooks, "Intelligence without representation", *Artificial Intelligence 47* (1991), 139–159.

[15] Baalman, M. *Composing Interactions. An Artist's Guide to Building Expressive Interactive System*, V2_Publishing: Rotterdam, 2022

[16] Luigi Nono, *...sofferte onde serene...*, 1st ed. (Milan: Edizione Ricordi, 1977)

[17] Antoniadis, P. Live recording at SARC, Queen's University, Belfast, 18.01.2018 of Luigi Nono, *...sofferte onde serene...*, 1st ed. (Milan: Edizione Ricordi, 1977)

https://on.soundcloud.com/2xWXgYYKXt3pIzTeFH accessed 06.10.2025

[18] Antoniadis P. Live recording at Amphi X, Université Paris 8, 29.03.2021 of Iannis Xenakis *Evryali* (1973) https://on.soundcloud.com/MVjIwbaZqJ4HaQ2IuN, accessed 06.10.2025

[19] Godard, J.L., *Filme Socialisme*, Vega Film, 2010

[20] Glazer, J., *Sexy Beast*, FilmFour/Kanzaman S.A./ Recorded Picture Company/Jeremy Thomas Productions, 2000

[21] Xenakis, I., *Evryali*, Salabert, 1973

[22] Tian Liangliang, *Three-Body* (*三体*), Tencent Video/CCTV/The Three-Body Universe/Linghe Cultural Media, 2023

[23] Cixin, L., *The Three-Body Problem* (三体), Chongqing Press, 2008

[24] Lynch, D., *Lost Highway*, Ciby 2000 / Asymmetrical Productions, 1997

[25] Ravel, M. « Oiseaux Tristes » (nr. 2 in the collection *Miroirs*, Edition Peters, 1904-1905

[26] Von Trier, L. *The Element of Crime*, Det Danske Filminstitut, 1984

[27] Bevilacqua, F. et al 'Online Gesture Analysis and Control of Audio Processing. In: Solis, J. & Ng. K.(eds.), *Musical Robots and Interactive Multimodal Systems*. Berlin Heidelberg: Springer-Verlag, 2011, 127-142

[28] G. Assayag, G. Bloch, M. Chemillier, A. Cont, S. Dubnov, "Omax brothers: a dynamic topology of agents for improvisation learning,", in *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*. Santa Barbara, United States, 2006, pp. 125–132.

[29] Rasmussen, J. *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. Elsevier Science Inc., New York (1986)

[30] Malloch, J., Sinclair, S., Hollinger, A. and Wanderley, M. M. "Input Devices and Music Interaction" in Solis, J. & Ng. K.(eds.), *Musical Robots and Interactive Multimodal Systems.* Berlin Heidelberg: Springer-Verlag, 2011

[31] Devenish, L., Hope, C. (eds.) *Contemporary Musical Virtuosities*, Routledge, 2025

[32] M. Shanahan, "Talking About Large Language Models," arXiv:2212.03551v2 [cs.CL], 2022. https://doi.org/10.48550/arXiv.2212.03551

[33] Antoniadis, P. "Challenging epistemic biases in musical AI: a guerrilla approach to human – machine comprovisation based on Xenakis's sketches for *Evryali*", in *Proceedings of TENOR24 - 9th International Conference on Technologies for Music Notation and Representation (TENO*R), pp. 65-72 https://www.tenor-conference.org/proceedings/2024/09_TENOR2024_Antoniadis.pdf accessed 06.10.2025

[34] AA. Cavia, *Logiciel. Six Seminars on Computational Reason*, Berlin: The New Centre for Research and Practice, 2022

[35] L. Barsalou, "Perceptual Symbol Systems," in *Behavioral and Brain Sciences*, 22: pp. 577–609, 1999

[36] D. Borgo, D. *Sync or Swarm. Improvising Music in a Complex Age*, Bloomsbury, 2022.