DiMMI

Dictionary for Multidisciplinary Music Integration: Contamination

Trento, November 29-30, 2024

PROCEEDINGS

UNIVERSITÀ DI TRENTO

DiMMI. Dictionary for Multidisciplinary Music Integration

Vol. 2 - Contamination

The Dictionary for Multidisciplinary Music Integration (DiMMI) is a proceedings series about the event organized by the University of Trento and the Conservatory "F. A. Bonporti" of Trento and Riva del Garda, in which musicians and representatives of the academic world are called to reflect together on a word of common interest, each from the perspective of their own discipline.

Silvia Sacchetti, Nicola Conci (eds.)

CONTAMINATION

Proceedings of the 4th DiMMI conference, held in Trento, 29-30 November 2024

Università degli Studi di Trento

UNIVERSITÀ
DI TRENTO

Published by
Università degli Studi di Trento
via Calepina, 14 - I-38122 Trento
casaeditrice@unitn.it
www.unitn.it

Dictionary for Multidisciplinary Music Integration (DiMMI)
Vol. 2 - Contamination

Cover Design: Grafica > UniTrento | IMG > stock.adobe.com

ISBN 978-88-5541-141-7
DOI 10.15168/11572_469393

Table of Contents

Dictionary for Multidisciplinary Music Integration (DiMMI) 2024: Contamination

Silvia Sacchetti[1] and Nicola Conci[2]

[1] *Dept. of Sociology and Social Research, University of Trento*
[2] *Dept. of Information Engineering and Computer Science, University of Trento*

Editorial

The Integrated Multidisciplinary Music Dictionary (DiMMI) continues its tradition also in 2024, asking researchers, musicians and practitioners to interpret a single keyword within the perspective of their own discipline. Such an approach has been driving DiMMI since its inception, aiming to promote the exchange of ideas at the crossroad with other research communities. Following up to the 2022 edition, also for this year DiMMI releases the volume containing the proceedings of the event collecting the extended abstracts presented by all authors at the conference held in Trento, on the 29-30 November, 2024, jointly organized by the University of Trento and the Conservatory "F. A. Bonporti" of Trento and Riva del Garda.

Similarly to the conference, this manuscript collection is also very multidisciplinary.

After "dissonance" in 2020, "rules" in 2021, and "interaction" in 2022, the keyword chosen for the 2024 edition was "contamination". Contamination is the process in which a pure element is perturbed in its state by an external agent. While in chemical terms this has generally a negative connotation, blending knowledge from different research areas, meeting at the crossroads of hard and soft sciences, is among the most sought after and ultimate goals of scholars and practitioners whose aim is to broaden their research spectrum towards in a multidisciplinary fashion.

Overview of the Volume

The conference has seen the participation of researchers and practitioners from different countries and academic contexts, turning the event into a unique opportunity to broaden everyone's research perspectives, shaping the road towards the development of a different way of conducting academic research in a transversal fashion, fostering the cross-domain potential, a rarely explored territory.

The volume organization follows the conference program which, also for this year, consisted in different types of contributions, including oral presentation sessions, poster sessions, and demonstration sessions.

All contributions have undergone a meticulous peer-review process, which has lead to the selection of the final technical program of the event. In particular, oral presentation sessions have been selected according to the novelty of the theoretical and practical frameworks which demonstrated a significant contribution towards the keyword, "contamination", in different domains, also exploring the different sensing modalities of humans, spanning across social sciences, music, philosophy, and technology. The contributions include algorithms, teaching frameworks, multimodal and multidisciplinary paradigms.

The poster and demonstration sessions are instead aimed at presenting more practical implementations and in-progress works, such as software libraries, user studies, novel sensing/perception tools, complementing the technical program with tangible experiences that stimulated a fruitful discussion among all the participants.

We hope the reader will enjoy the contributions included in this volume, which, in the perspective of DiMMI, aim at offering a novel research paradigm, and a unique space in the vast panorama of conferences and workshops, where researchers, scholars, and musicians can benefit from the mutual exchange of knowledge.

Crossmodal Correspondence between Pitch Height and Olfactory Categories

Michelle Caldana and Elena Rusconi

Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy

Abstract

To date, not all crossmodal correspondences can be fully explained by statistical, semantic, linguistic or hedonic accounts. It has been hypothesised that the crossmodal correspondence between odours and auditory pitch may subtend a structural basis and originate in neural mechanisms intrinsic to the general organisation of sensory experience. Here we show that the same olfactory categories that have been shown to give rise to a vertical stimulus-response compatibility effect (SRC) are systematically associated with high and low tones, suggesting access to a common vertical spatial representation. The investigation of this type of crossmodal correspondence could help further our understanding of how we experience music and more generally of the architecture of our cognitive system, by identifying possible structural determinants underlying crossmodal correspondences.

1 Introduction

Crossmodal correspondences are consistent associations between certain attributes and dimensions in different sensory modalities [Spence, 2011]. Verbalizing olfactory perceptions is notoriously challenging and descriptors are frequently borrowed from other sensory modalities, which makes olfaction of particular interest in the field of crossmodal correspondences. This also suggests that the unisensory experience as we know it may in fact be the result of simultaneous contamination by other senses. To date, numerous correspondences between olfaction and other senses have been documented (e.g. [Gilbert et al., 1996, Belkin et al., 1997, Dematte et al., 2006, Hanson-Vaux et al., 2013]). Certain correspondences are easily ascribable to a specific mechanism. For example, those between odours and colours can be explained in terms of semantic mediation, and those between odours and thermal attributes through internalisation of environmental statistics. However, some peculiar correspondences between odours and contingent features, such as auditory pitch, are more difficult to explain. For example, odours and tones hardly co-occur in the environment, which may rule out a statistical account. The use of a lexicon referring to a common semantic field (high and low in the case of pitch; top and base in the case of olfactory notes) is unlikely to be the primary cause of this association since for the olfactory domain it occurs almost exclusively within the perfume industry. Finally, pitch choice for a given odour is not guided by hedonic values [Stevenson et al., 2012, Crisinel & Spence, 2012, Crisinel et al., 2013, Ward et al., 2022]. Deroy et al. [Deroy et al., 2013] proposed that this could be a case of structural correspondence, that is mediated by a common neural substrate or mechanism involving a structural feature. For instance, access to a common spatial representation. In support of this hypothesis, Caldana and Rusconi [Caldana & Rusconi, 2023] reported that, similarly to the auditory domain, where pitch evokes a subjective experience of spatial elevation and can also influence spatial response selection (e.g., [Pratt, 1930, Rusconi et al., 2006]), olfactory notes that across the literature have been individually associated with high tones (lemon, bergamot, orange) and low tones (coffee, caramel, vanilla), and that can be grouped into categories referred to as examples of top and base notes by professionals (e.g. fruity and gourmand), can also give rise to a vertical SRC effect. This effect does not correlate with the degree of perceived pleasantness, intensity, perceived lightness/heaviness, or correct odour recognition. However, in Caldana and Rusconi's study [Caldana & Rusconi, 2023], the relationship between such olfactory categories and tonal pitch was not explicitly investigated. It is therefore unknown whether such categories of stimuli would indeed give rise to crossmodal correspondences with auditory pitch. In this study, we investigate the presence of crossmodal associations between auditory pitch and the olfactory categories that in Caldana and Rusconi have originated a spatial SRC effect, by testing participants with an explicit association task, as part of a larger study. More precisely, we predict that odours in the fruity category will be associated with tones that are significantly higher in frequency than those with which odours in the gourmand category will be associated.

2 Methods

2.1 Stimuli and Procedure

We used the same six odorants as used by Caldana and Rusconi [Caldana & Rusconi, 2023]: bergamot, lemon and orange for the fruity category, caramel, coffee and vanilla for the gourmand category. Each odorant was delivered for 1,200 ms, with an inter-stimulus-interval of at least 11s, via a nosepiece connected to an olfactometer. During the task, an odorant could be reassessed for a maximum of two times by pressing an ad hoc key. Crossmodal correspondences were probed with pure tones having a range of frequencies from 65 Hz to 1,046 Hz; participants could play and evaluate as many tones as desired by moving the cursor on a slider, where pitch gradually changed in a horizontal direction, and pressing the "listen" button with a mouse click. Each tone lasted 1,500 ms and was presented over closed-ear headphones (HP, gaming OMEN blast) at a loudness of 70 dB [Crisinel & Spence, 2012, Crisinel et al., 2013, Ward et al., 2022]. The slider, subtending 26° degrees of visual angle in width, was centered on a PC screen positioned in front of the participant. The direction of pitch increase was counterbalanced (for half of the participants it was from right to left, for the rest from left to right) and responses were given by pressing with a mouse click the on-screen "confirm" button. Participants were then asked to score the degree of perceived heaviness/lightness of each of the 6 odorants, presented in random order, by selecting a number from 1 (extremely heavy) to 7 (extremely light) on a keyboard attached to the computer. The same procedure was used to assess the degree of perceived pleasantness (1 for extremely unpleasant; 7 for extremely pleasant), and intensity (1 for extremely weak; 7 for extremely strong). Finally, the 6 odorants were presented individually in random order and participants were required to locate in a list of 12 (lemon, pine, orange, cherry, green apple, bergamot, caramel, musk, coffee, lavender, banana, vanilla) the odour just delivered.

3 Results

3.1 Participants

One hundred and fourteen participants were tested. Four (3M, 1F; all right-handed; age: M = 25.5, SD = 8.5) were excluded due to failure to reach the cut-off (10 correct identifications out of 12) at the Sniffin' Sticks Screening Test (ODOFINTM). The rest (N = 110; 25M, 85F; Age: M = 24, SD = 7) reached or exceeded the cutoff and did not report any professional experience related to perfumes.

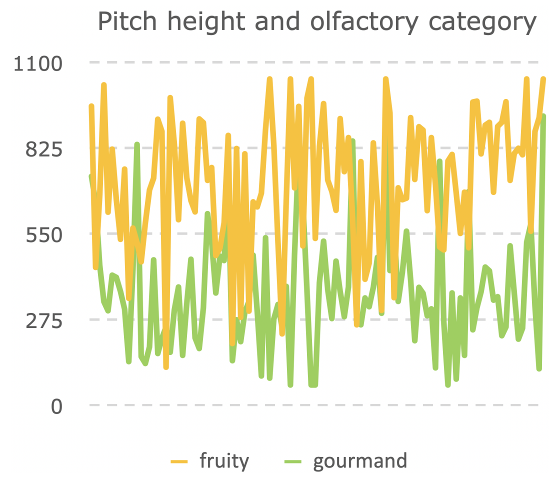


Figure 1: *The plot shows the median frequency of tones associated to fruity odours (in orange) and to gourmand odours (in green) for each of the 110 participants in the study (each corresponding to a point on the x-axis).*

3.2 Crossmodal Correspondence between Pitch Height and Olfactory Categories

Participants associated tones of higher frequency to fruity odors (M of medians = 717, SD = 212) than to gourmand odours (M of medians = 364, SD = 188; Figure 1). Because the distribution of frequencies associated to fruity odours did not significantly deviate from normality ($p = .055$) but the distribution of frequencies associated to gourmand odours did ($p < .001$), a Wilcoxon Signed-Rank test was performed. The test detected a significant difference between fruity and gourmand odours ($z = 8.277$, $p < .001$, matched rank biserial $r = .91$, $SE = .11$). A Bayesian version of the test, with data augmented using 5 chains of 5000 iterations, returned a BF10 superior to 5, indicating very strong evidence in favour of the alternative hypothesis (BF10 = 9.095 × 109, W = 5828, Rhat = 1.786). For only 12 out of 110 participants the difference between the median frequency associated with fruity odours and the median frequency associated with gourmand odours was a negative number. A series of Spearman correlations did not detect any significant associations between such difference and the difference between fruity and gourmand odours in their median scores of pleasantness, heaviness/lightness, intensity or in their percentage of correct recognition (all ps > .127), even though fruity odours were perceived as significantly more pleasant, lighter and less intense than gourmand odours (all ps < .009) and were better recognized than gourmand odours (p = .013).

4 Discussion

The association between odorants and tones appears to be one of the most intuitive crossmodal correspondences. In the middle of the 19th century, Septimus Piesse created the "Gamut of odors", one of the first and most notorious examples of spontaneous association between olfactory and auditory notes. He suggested that scent creation has analogies with music composition, as they both entail a harmonious combination of notes [Di Stefano et al., 2021, Spence, 2021]. In fact, to date, numerous examples of contamination between olfaction and music in the arts and entertainment scene have been proposed, to enhance multisensory processing without additional cognitive load, improve the quality of the experience and user enjoyment [Spence, 2021, Murray et al., 2016]. This study demonstrates that olfactory notes (fruity and gourmand), which are mapped by professionals onto higher and lower positions of the olfactory pyramid and can give rise to a vertical SRC effect [Caldana & Rusconi, 2023], are spontaneously associated to high- and low-pitch sounds in a consistent way across participants. This differential association is not correlated with differences in perceived pleasantness, intensity and lightness/heaviness of the two odour categories, or in the accuracy of their recognition. This rules out explanations in terms of statistical, language/semantic or hedonic accounts but is consistent with the hypothesis of a common (or analogous) spatial representation elicited by olfactory and auditory stimuli [Deroy et al., 2013]. To localize stimuli in physical space, both the olfactory and the auditory systems employ receptors with a fixed position. Localization of olfactory and auditory stimuli on the horizontal axis reflects a birhinal or binaural comparison of onset/duration and intensities, that are subsequently analysed and integrated centrally to construct a representation of auditory and olfactory spaces [Bao et al., 2019]. Localization on the vertical axis is generally poor in physical space (e.g. [Yost & Nielsen, 2000, Porter et al., 2007]) and it appears that certain stimulus characteristics are spontaneously mapped on a vertical direction instead. In the auditory domain, pitch height can evoke an experience of spatial elevation and consistently influence spatial response selection [Pratt, 1930, Rusconi et al., 2006]. In the olfactory domain, preliminary evidence suggests that certain categories of olfactory stimuli (which happen to be crossmodally associated high and low tones) may also influence spatial response selection [Caldana & Rusconi, 2023]. This line of investigation looks promising and may lead to a better understanding of the mechanism subtending this peculiar type of crossmodal correspondence, In addition to fostering contaminations in music, sound and art.

References

[Bao et al., 2019] Bao, X., Gjorgieva, E., Shanahan, L. K., Howard, J. D., Kahnt, T., & Gottfried, J. A. (2019). Grid-like neural representations support olfactory navigation of a two-dimensional odor space. *Neuron*, 102(5), 1066–1075.

[Belkin et al., 1997] Belkin, K., Martin, R., Kemp, S. E., & Gilbert, A. N. (1997). Auditory pitch as a perceptual analogue to odor quality. *Psychological science*, 8(4), 340–342.

[Caldana & Rusconi, 2023] Caldana, M. & Rusconi, E. (2023). Lo spazio mentale degli odori: evidenza di compatibilità spaziale stimolo-risposta nel dominio olfattivo. *Giornale italiano di psicologia*, 50(3), 523–540.

[Crisinel et al., 2013] Crisinel, A.-S., Jacquier, C., Deroy, O., & Spence, C. (2013). Composing with cross-modal correspondences: Music and odors in concert. *Chemosensory Perception*, 6, 45–52.

[Crisinel & Spence, 2012] Crisinel, A.-S. & Spence, C. (2012). A fruity note: crossmodal associations between odors and musical notes. *Chemical Senses*, 37(2), 151–158.

[Dematte et al., 2006] Dematte, M. L., Sanabria, D., Sugarman, R., & Spence, C. (2006). Cross-modal interactions between olfaction and touch. *Chemical senses*, 31(4), 291–300.

[Deroy et al., 2013] Deroy, O., Crisinel, A.-S., & Spence, C. (2013). Crossmodal correspondences between odors and contingent features: odors, musical notes, and geometrical shapes. *Psychonomic bulletin & review*, 20, 878–896.

[Di Stefano et al., 2021] Di Stefano, N., Murari, M., & Spence, C. (2021). Crossmodal correspondences in art and science: Odours, poetry, and music. In *Olfaction: An interdisciplinary perspective from philosophy to life sciences* (pp. 155–189). Springer.

[Gilbert et al., 1996] Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: The color of smells. *The American journal of psychology*, (pp. 335–351).

[Hanson-Vaux et al., 2013] Hanson-Vaux, G., Crisinel, A.-S., & Spence, C. (2013). Smelling shapes: Crossmodal correspondences between odors and shapes. *Chemical senses*, 38(2), 161–166.

[Murray et al., 2016] Murray, N., Lee, B., Qiao, Y., & Muntean, G.-M. (2016). Olfaction-enhanced multimedia: A survey of application domains, displays, and research challenges. *ACM Computing Surveys (CSUR)*, 48(4), 1–34.

[Porter et al., 2007] Porter, J., Craven, B., Khan, R. M., Chang, S.-J., Kang, I., Judkewitz, B., Volpe, J., Settles, G., & Sobel, N. (2007). Mechanisms of scent-tracking in humans. *Nature neuroscience*, 10(1), 27–29.

[Pratt, 1930] Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, 13(3), 278.

[Rusconi et al., 2006] Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the smarc effect. *Cognition*, 99(2), 113–129.

[Spence, 2011] Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995.

[Spence, 2021] Spence, C. (2021). Musical scents: On the surprising absence of scented musical/auditory events, entertainments, and experiences. *i-Perception*, 12(5).

[Stevenson et al., 2012] Stevenson, R. J., Rich, A., & Russell, A. (2012). The nature and origin of crossmodal associations to odours. *Perception*, 41(5), 606–619.

[Ward et al., 2022] Ward, R. J., Wuerger, S., & Marshall, A. (2022). Smelling sensations: Olfactory crossmodal correspondences (jpi-first). *Electronic Imaging*, 34.

[Yost & Nielsen, 2000] Yost, W. A. & Nielsen, D. W. (2000). *Fundamentals of hearing*, volume 3. Academic Press New York.

How Music Can Contaminate Research on Consciousness

Giulio Velardo[1,2] and Nicola De Pisapia[2]

[1]*Conservatory "F. A. Bonporti" of Trento and Riva del Garda, Italy*
[2]*Department of Psychology and Cognitive Science, University of Trento, Italy*

Abstract

Altered states of consciousness (ASC) are defined as a qualitative alteration in the overall patterns of mental functioning so that the experiencer feels that his/her operations of consciousness are radically different from ordinary functioning. The current literature on the topic predominantly focuses on psychedelic-induced or meditation-induced ASCs, although music is gaining progressively more traction as a research tool on the topic. In order to better understand and explain the specificities of music, psychedelics and meditation on ASC induction, we propose a classification framework which takes into account the qualitative differences between the ASCs induced by said stimuli, based on the neuroanatomical and phenomenological evidence available. In light of the proposed network, we argue that music might serve as a more flexible alternative to other ASC-inducing methods. Musical stimuli such as shamanic drumming are understood to induce a trance state quickly and without any previous training, unlike meditation. Although psychedelic drugs might still be preferred in trials where ASCs are studied in conjunction with cognitive or behavioural tasks, music might serve as an adequate alternative in contexts where access to psychedelics is limited or its use might be dangerous to subjects. Moreover, the literature suggests that music seems to occasionally present opposite activation patterns to the other aforementioned mechanisms in ASCs induction, thus making it a powerful tool to better understand how consciousness is structured at an anatomical level. Although music has the potential to greatly contribute to our understanding of consciousness, it is necessary that we allow our research paradigms to get contaminated by musical and acoustical analysis.

1 Introduction and Theoretical Interpretations of ASC

Psychedelics, meditation and music are all Altered States of Consciousness (ASC)-inducing methods that have received growing interest from neuroscientists in the past few decades [De Pisapia, 2024]. While they have all been proven effective in treating specific psychopathologies in clinical settings (e.g., depression or trauma), it is also possible to employ such practices as a tool to investigate the neuroanatomical and phenomenological bases of consciousness, as it is possible to infer how consciousness is structured by examining the reversible alterations caused by such stimuli. For brevity's sake, we will not take into account emotion-based ASCs, such as the ones induced by empathogens or entactogens psychedelics (e.g., psylocibine), perspective taking forms of meditation or absorptive states of consciousness induced by music. It would also be possible to classify said states inside the proposed classification network, but we will be hereby focusing on just attention and perception related ASCs. Among the few theoretical interpretations of the neural mechanisms underlying ASCs [Freeman, 1998, Dietrich, 2003, Schäfer et al., 2013], Carhart-Harris' entropic brain hypothesis is the most accredited one [Carhart-Harris, 2018]. According to Carhart-Harris, the brain is characterised at any given time by a certain degree of entropy (i.e., a system's degree of uncertainty). A normal level of entropy would correspond to the regular functioning state of consciousness. Higher entropy levels mean more complex, diverse, and unpredictable neural activity in the brain [Carhart-Harris, 2018], which lead to disorganisation and a more diverse range of dynamic neural states [Holas & Kamińska, 2023]. The subjective manifestations of ASCs are thought to be linked to such irregular activation of the cerebral cortex, and it is possible to temporarily elicit such irregularities by engaging in various practices, such as psychedelics ingestion, meditation and possibly (though it has not been verified experimentally yet) music listening/practising.

2 Classification Framework of ASC-inducing Stimuli

We propose that music-, meditation- and psychedelic-induced ASCs could be interpreted inside a framework composed of two main dichotomous dimensions: 1. Exogenous - Endogenous (i.e., whether the person's attention is directed toward the outer or inner environment respectively) and 2. Heterogenic - Autogenic (i.e. whether the ASC is induced by an external stimulus or by a strategy/behaviour employed by the subject). Classic psychedelics (e.g., mescaline, LSD) would be accordingly classified as an exogenous, heterogenic means of inducing ASCs, while meditation would be considered an exogenous, autogenic method. Music is more complicated to classify, as we found examples of both endogenous and exogenous ASCs induced by music, although classifying music as a heterogenic stimulus seems not up to debate. Additionally, it is possible to distinguish between chemical and physiological ASC induction, the former referring to psychedelic-induced ASCs (PIASC) and the latter to meditation-induced or music-induced ASCs (MEDI-ASC and MIASC respectively).

2.1 Psychedelics-induced ASCs (PIASC)

Psychedelics interact with serotonin receptors, namely the 2A subtype of serotonin receptors (5-HT2A) [Kelmendi et al., 2022], which are distributed near most of the serotonergic terminal rich areas the neocortex, thalamus, locus coeruleus, ventral tegmental area, and claustrum. More specifically, psychedelics reduce activities in the medial frontal cortex (mPFC) and posterior cingulate cortex (PCC), which compose part of the default mode network (DMN), a set of brain regions that are more active during wakeful rest and less so when engaging the external world [Kelmendi et al., 2022]. Moreover, psychedelics reduce communication among brain areas involved in planning and decision-making (such as the prefrontal cortex) but increase communication between areas involved in sensation and movement (i.e., the occipital cortex, the superior temporal gyrus, and the postcentral gyrus, as well as the precuneus) [Preller et al., 2018, Moujaes et al., 2024]. These findings are relevant, as the activation patterns mentioned elicit network disintegration and desegregation and increase brain entropy (see 1.) [Timmermann et al., 2023], leading to subjective alterations of consciousness which line up to the definition of trance (see glossary), such as increased feelings of unity and transcendence of time and space, visual hallucinations, illusions, synesthesia [Kelmendi et al., 2022, Meck, 2005] and ego-dissolution [Timmermann et al., 2023]. Corroborating this, it appears that more vivid subjective experiences are linked to a stronger hyper-connectivity pattern in somatosensory regions [Preller et al., 2018, Moujaes et al., 2024]. Other evidences supporting Carhart-Harris' theory come from the fact that psychedelic-induced ego-dissolution is related to greater reductions in DMN metabolism, oscillatory power, and connectivity between its main hubs (mPFC, PCC, and medial temporal areas) as well as decoupling of its activity from other brain networks.[Carhart-Harris et al., 2014, Millière et al., 2018, Kometer et al., 2015, Mason et al., 2020, Lebedev et al., 2015]. We argue that the subjective experiences associated with DMN deactivation and decoupling are sufficient to classify the related ASCs as "exogenous".

2.2 Meditation-induced ASCs (MEDI-ASC)

The term "meditation" encompasses a variety of practices, such as Mindfulness, Mantra Recitation and Open Monitoring, associated with different cognitive tasks [Millière et al., 2018]. These cognitive tasks are related to specific cortical area's activation patterns (e.g., Mantra Recitation is related to activations in Broca's area [Millière et al., 2018]), but there is no evidence these patterns support ASC induction, as they seem correlated exclusively to the action that distinguish them. Due to this fact, the aforementioned practices will still be hereby grouped under the umbrella term "meditation", as they share the cognitive mechanisms of meta-awareness, focusing of attention and cognitive defusion [Lutz et al., 2015]. These cognitive mechanisms are related to attenuated activation of the DMN (especially in the mPFC and PCC) [Simonsson et al., 2023, Gu et al., 2015, Visted, 2015], much like psychedelic drugs. More specifically, the act of focusing attention (which is a commonality in virtually all meditation practices) is correlated to significant activation clusters in executive brain areas such as the premotor cortex (PMC) and the dorsal anterior cingulate cortex (dACC), which may underlie top-down regulation of attention and monitoring of spontaneous thoughts during meditation (thereby supporting the "autogenic" classification term), and to the deactivation of the PCC and inferior parietal lobule (both part of the DMN), involved in introspection related functions (e.g. mind-wandering, self-reflection, etc.) [Millière et al., 2018], thereby supporting the "exogenous" classification criterion.

Moreover, meditation-induced ASCs share the ego-dissolution phenomenon with the psychedelics induced ones, which is related to the same cortical activation patterns. [Buckner et al., 2008].

2.3 Music-induced ASCs (MIASC)

First of all, it is important to clarify that the term "music" refers to a vast array of acoustical stimuli, ranging from full fletched recordings of orchestral performances to monotonous drumming or chanting. This being said, the scientific community seems to agree that a few characteristics qualify musical stimuli as better than others at inducing specific ASCs. As it pertains to trance states, repetition and simplicity are regarded as the most important aspects of the stimulus [Meck, 2005, Grill-Spector et al., 2006, Kjellgren & Eriksson, 2010], while other altered states of consciousness such as absorption or flow seem to seem to be more heavily influenced by the emotional valence of the music heard or played [McNamara & Ballard, 1999, Nater et al., 2005, Diaz, 2011, Lange et al., 2017, Lange & Fink, 2023, Hall et al., 2016, Vroegh, 2019, Schäfer & Fachner, 2014, Kreutz et al., 2007].

While the neural correlates of flow and absorption induced by music have not been systematically investigated as of yet, there is some evidence of the effect of trance-inducing acoustical stimuli on the brain [Rogerson et al., 2021, Konopacki & Madison, 2018, Huels et al., 2021, Hove et al., 2015, Dvorak & Hernandez-Ruiz, 2019].

In order to better compare the effects of music listening to psychedelics and meditation, we will only be focusing on the neuroimaging findings of trance states induced by music listening, as functional magnetic resonance imaging (fMRI) and connectivity maps are the predominant methods employed in studies regarding PIASC and MEDIASC [Timmermann et al., 2023, Holas & Kamińska, 2023, Preller et al., 2018, Kelmendi et al., 2022, Millière et al., 2018, Yu et al., 2024, Zhigalov et al., 2019].

[Hove et al., 2015] found that eigenvector centrality increased during shamanic drumming-induced trance in the PCC, dACC, and insula, all part of the DMN. Increased PCC centrality during trance supports its importance in internally directed mental states, and suggests that trance involves increased large-scale neural integration in this key hub [Hove et al., 2015]. Moreover, higher eigenvector centrality of the dACC, insula, and the PCC suggests increased coupling among these three regions, which could help maintain an internally directed cognitive state [Hove et al., 2015]. These results are interesting, as it appears that trance states induced by music present opposite activation patterns of psychedelic- and meditation-induced trance states (characterised by reduced centrality if these regions). This difference seemingly compels us to distinguish between MIASC's trance from PIASC and MEDIASC's trance: the first one being classified as endogenous and the latter two as exogenous. However, Dvorak et al. present differing results [Dvorak & Hernandez-Ruiz, 2019].

In contrast with Hove's account, Dvorak found that predictable music further deactivates the DMN in subjects undergoing mindful meditation practices [Dvorak & Hernandez-Ruiz, 2019]. These findings are consistent with the ones discussed in 2.1 and 2.2, and would lead us to define music-induced trance states as exogenous.

Before affirming that music can work both as an endogenous and exogenous stimulus, it is important to note a few criticalities. Hove's subjects were experienced shamanic practitioners, unlike the naif ones in Dvorak's work. Moreover, Hove's subjects deliberately tried to undergo a trance state during the experimental conditions, contrary to Dvorak's, which were instructed to just meditate. Lastly, Hove's study made use of a shamanic drumming recording, while Dvorak composed the stimulus ad hoc, which differed from Hove's on various acoustical parameters: it is possible that, although both stimuli were "repetitive" and "predictable", acoustical characteristics such as sound wave form, amplitude and frequency might have had an impact on the results.

3 The Aim of the Proposed Classification Framework

We argue that classifying ASCs based on their qualitative and phenomenological differences might help researchers gain insight into the characteristics of the stimuli that can elicit them. It appears that, although trance states are usually considered as a whole, the neural dynamics associated with specific stimuli are not the same. This is important, although rarely addressed by the current literature; it seems to be possible to elicit similar subjective experiences even if the ASCs associated present opposite cortical activation patterns.

It has recently been proposed that there might be functional convergence between overstimulation and inhibition of cortical areas, as it pertains to other cognitive processes [Chen et al., 2024]: we argue that a similar concept could be applied to the phenomenon of brain entropy, too. It is possible that consciousness might operate "regularly" at an optimal level of entropy, which might increase and therefore alter consciousness not only when decoupling and deactivation occurs in certain key brain hubs, but also when these hubs are overstimulated. Employing the ex/endogenous - hetero/autogenic framework researchers might be better equipped to investigate how consciousness is mediated by cerebral activity and connectivity, as it would take into account the differences between "overstimulating" ASCs (e.g. as it pertains to the DMN, overstimulation this hub would lead these ASCs be defined as endogenous) and "inhibitory" ASCs (e.g. these would be defined exogenous in the case of inhibition occurring the DMN).

4 Methodological Issues of Current Music-induced ASC's Paradigms

Although the MIASC approach seems promising, a few methodological issues are to be addressed. First of all, music listening is rarely used as an isolated stimulus during procedures of ASC induction: music is most commonly employed in conjunction to shamanic rituals, explicit instructions, meditation practices or psychedelics use [Perry et al., 2021, Diaz, 2011, Becker-Blease, 2004, Alali-Morlevy & Goldfarb, 2022, Maurer et al., 1997, Fachner et al., 2019, Rogerson et al., 2021, Woodside et al., 1997]. In order to investigate whether music is capable of inducing ASCs by itself, we need more studies employing music stimuli on their own. Secondly, the musical stimuli employed in the existing studies are too varied and their acoustical characteristics are rarely addressed. In order to establish more accurately the variables responsible of the observed effects a more cautious approach is to be followed: we propose that by allowing the ASC research field to be more extensively contaminated with a more precise acoustical and theoretical analysis of the musical stimuli employed to induce ASCs it would be possible to establish a methodological approach better suited to study the phenomenological and behavioural aspects of ASCs.

5 Conclusion and Suggestions for Future Research

The most interesting aspect of the data and hypotheses here presented refers to the apparent ambivalence of musical stimuli. We already proposed possible causes of said incongruence, but it might be possible that certain musical stimuli such as the ones employed in Hove's study could elicit trance states inner of itself, even without the explicit intention of participating in a shamanic trance state. This hypothesis is somewhat forwarded by evidence provided in studies on auditory driving [Hove et al., 2015, Rouget, 1985], although the results are mixed [Konopacki & Madison, 2018, Kjellgren & Eriksson, 2010, Grill-Spector et al., 2006]. We propose that music-induced trance states might be linked to the polymodal activation elicited by music listening, meaning that listening or playing music leads to simultaneous activation of cortical areas that are not usually communicating [Chorna et al., 2019, Sihvonen et al., 2021, Chatterjee et al., 2021]. So far, polymodal activation has been employed to explore cortical plasticity [Schlaug, 2001] and to rehabilitate neurological patients, but, due to the fact that this phenomenon partially resembles the one of brain entropy, we are compelled to propose a few hypotheses. First, that expert musicians might be more resistant to ASCs, as their years of musical training might have raised the baseline for brain entropy to be significant, although it is also possible that expert musicians might be more sensible to ASCs, due to the fact that the effects of brain entropy might be easier to obtain. Secondly, it is possible that music could serve as a more flexible alternative to other ASC-inducing methods. Musical stimuli such as shamanic drumming are understood to induce a trance state after just a few minutes of listening and do not require any training in order to experience their effects, unlike meditation. This would make similar musical stimuli better suited to induce ASCs in naif individuals. Although psychedelic drugs might still be preferred in trials where ASCs are studied in conjunction with cognitive or behavioural tasks (as the ASC induced chemically are more resistant to distractions and alteration in the environment), music might serve as adequate alternative in contexts where access to psychedelics is limited or its use might be dangerous to subjects. In order to explore both hypotheses we suggest using stimuli similar to the one employed by Hove [Hove et al., 2015]. Shamanic drumming appears to be the best candidate to study music-induced ASCs for a few reasons. 1. Monotonous drumming isolates the main variables which make an acoustical stimulus effective (predictability) while allowing researchers to control other aspects of the stimulus; 2. Monotonous drumming better lends itself to be acoustically analysed and 3. the specific acoustical variables which might aid or impede ASC induction are more easily manipulated compared to more complex stimuli; 4. It does not carry along emotional, semantic or syntactic significance that might confuse the results and 5. It is anthropologically significant, as monotonous drumming has been used transculturally in ritualistic settings with the specific aim of aiding transcendental, trance-like experiences on the participants [Perry et al., 2021].

We briefly discussed how music, or acoustical stimuli more in general, present unique interactions with the brain compared to other ASC-inducing stimuli. While these interactions appear at times contradictory, we argue that a more musically insightful approach would lead to a better understanding of the specificities of music, thereby allowing for a clearer classification of MIASCs. We are convinced that understanding whether music is unique (and, given the case, to what degree) in the field of ASC research, would entrust the scientific community with a powerful tool to deepen our knowledge of how consciousness is structured and interpreted, as it can shed light on both the neuroanatomical sites associated and the phenomenological perception of consciousness. Of course, said "musically insightful approach" is made possible by allowing the current, well established paradigms of consciousness research to be contaminated with the ones pertaining to musical research.

GLOSSARY

Heterogenic ASCs: ASCs caused by the system's interaction with external stimuli.
Autogenic ASCs: ASCs induced by the strategies directed by the subject himself/herself.
Exogenous ASCs: ASCs for which a person's locus of attention is directed towards the external environment.
Endogenous ASC: ASCs for which a person's locus of attention is directed towards his/her internal world.
Absorption: personality trait associated with higher abilities to focus attention on a specific object and monitor one's internal state [Perry, 2021] OR ASC characterised by a state of heightened attention directed towards one's internal state. It is considered an endogenous ASC.
Trance: ASC characterised by alterations or discontinuity in consciousness, awareness, personality, or other aspects of physical functioning [Cohen, 2007] . It is considered an exogenous ASC.
Flow: ASC characterised by is a state of full engagement, control, concentration and action awareness, occurring during an activity perceived as highly self-rewarding and characterised by clear goals, unambiguous feedback, distortion of time perception, loss of self-consciousness and a balance between challenges and skills required to best perform it [Csikszentmihalyi, 1975,1990]. It is considered an exogenous ASC.
Psychedelic-induced ASCs (PISASC): ASCs which are due to psychedelic drugs intake. They are considered heterogenic and chemical and lead to exogenous ASCs.
Meditation-induced ASCs (MEDIASC): ASCs induced by meditation practices. They are considered autogenic and physiological and lead either to exogenous or endogenous ASCs depending on the chosen practice.
Music-induced ASCs (MIASC): ASCs induced by either music playing or music listening. They are both considered heterogenic and physiological and can lead either to exogenous ASCs in the case of music playing, or endogenous ASCs as it pertains to music listening.
Entropic brain hypothesis: entropy is defined as a system's uncertainty. Higher entropy levels mean more complex, diverse, and unpredictable neural activity in the brain, which lead to disorganisation and a more diverse range of dynamic neural states [Carhart-Harris, 2018].
Default mode network (DMN): cortical network composed of the medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC)/precuneus (PCu), inferior parietal lobe, lateral temporal cortex and hippocampal formation. The DMN is typically deactivated during tasks requiring externally-oriented attention and activated during passive rest states or internally-oriented mental processes [Xu, 2016].
Ego dissolution: attenuation/dissolution of the subject–object distinction that structures ordinary experience [Timmermann, 2023].
Shamanic drumming: ceremony where a drum is played at a steady, fast beat (usually around three beats per minute) for a variable period of time. Alterations in the normal flow of consciousness are experienced after five to seven minutes.

References

[Alali-Morlevy & Goldfarb, 2022] Alali-Morlevy, K. & Goldfarb, L. (2022). The effect of background sounds on mind wandering. *Psychological Research*, 87(6), 1848–1861.

[Becker-Blease, 2004] Becker-Blease, K. A. (2004). Dissociative states through new age and electronic trance music. *Journal of Trauma & Dissociation*, 5(2), 89–100.

[Buckner et al., 2008] Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124, 1–38.

[Carhart-Harris, 2018] Carhart-Harris, R. L. (2018). The entropic brain—revisited. *Neuropharmacology*, 142, 167–178.

[Carhart-Harris et al., 2014] Carhart-Harris, R. L. et al. (2014). The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in Human Neuroscience*, 8, 20.

[Chatterjee et al., 2021] Chatterjee, D., Hegde, S., & Thaut, M. (2021). Neural plasticity: The substratum of music-based interventions in neurorehabilitation. *Neurorehabilitation*, 48(2), 155–166.

[Chen et al., 2024] Chen, S., Tsou, M., Chen, K., et al. (2024). Impact of repetitive transcranial magnetic stimulation on cortical activity: a systematic review and meta-analysis utilising functional near-infrared spectroscopy evaluation. *Journal of NeuroEngineering and Rehabilitation*, 21(108).

[Chorna et al., 2019] Chorna, O. et al. (2019). Neuroprocessing mechanisms of music during fetal and neonatal development: A role in neuroplasticity and neurodevelopment. *Neural Plasticity*, 2019, 1–9.

[De Pisapia, 2024] De Pisapia, N. (2024). *La Coscienza*. Il Mulino.

[Diaz, 2011] Diaz, F. M. (2011). Mindfulness, attention, and flow during music listening: An empirical investigation. *Psychology of Music*, 41(1), 42–58.

[Dietrich, 2003] Dietrich, A. (2003). Functional neuroanatomy of altered states of consciousness: The transient hypofrontality hypothesis. *Consciousness and Cognition*, 12(2), 231–256.

[Dvorak & Hernandez-Ruiz, 2019] Dvorak, A. L. & Hernandez-Ruiz, E. (2019). Comparison of music stimuli to support mindfulness meditation. *Psychology of Music*, 49(3), 498–512.

[Fachner et al., 2019] Fachner, J. C. et al. (2019). "telling me not to worry..." hyperscanning and neural dynamics of emotion processing during guided imagery and music. *Frontiers in Psychology*, 10, 1–13.

[Freeman, 1998] Freeman, W. J. (1998). A neurobiological role of music in social bonding. https://escholarship.org/uc/item/9025x8rt.

[Grill-Spector et al., 2006] Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14–23.

[Gu et al., 2015] Gu, J. et al. (2015). How do mindfulness based cognitive therapy and mindfulness-based stress reduction improve mental health and well-being? a systematic review and meta-analysis of mediation studies. *Clinical Psychology Review*, 37, 1–12.

[Hall et al., 2016] Hall, S. E., Schubert, E., & Wilson, S. J. (2016). The role of trait and state absorption in the enjoyment of music. *PloS One*, 11(11), e0164029.

[Holas & Kamińska, 2023] Holas, P. & Kamińska, J. (2023). Mindfulness meditation and psychedelics: potential synergies and commonalities. *Pharmacological Reports*, 75, 1398–1409.

[Hove et al., 2015] Hove, M. J. et al. (2015). Brain network reconfiguration and perceptual decoupling during an absorptive state of consciousness. *Cerebral Cortex*, 26(7), 3116–3124.

[Huels et al., 2021] Huels, E. et al. (2021). Neural correlates of the shamanic state of consciousness. *Frontiers in Human Neuroscience*, 15.

[Kelmendi et al., 2022] Kelmendi, B. et al. (2022). Psychedelics. *Current Biology*, 32(2), R63–R67.

[Kjellgren & Eriksson, 2010] Kjellgren, A. & Eriksson, A. (2010). Altered states during shamanic drumming: A phenomenological study. *International Journal of Transpersonal Studies*, 29(2), 1–10.

[Kometer et al., 2015] Kometer, M. et al. (2015). Psilocybin-induced spiritual experiences and insightfulness are associated with synchronisation of neuronal oscillations. *Psychopharmacology*, 232, 3663–3676.

[Konopacki & Madison, 2018] Konopacki, M. & Madison, G. (2018). Eeg responses to shamanic drumming. does the suggestion of trance state moderate the strength of frequency components? *Journal of Sleep and Sleep Disorders Research*, (pp. 16–25).

[Kreutz et al., 2007] Kreutz, G. et al. (2007). Using music to induce emotions: Influences of musical preference and absorption. *Psychology of Music*, 36(1), 101–126.

[Lange & Fink, 2023] Lange, E. B. & Fink, L. K. (2023). Eye blinking, musical processing, and subjective states—a methods account. *Psychophysiology*, 60(10).

[Lange et al., 2017] Lange, E. B., Zweck, F., & Sinn, P. (2017). Microsaccade-rate indicates absorption by music listening. *Consciousness and Cognition*, 55, 59–78.

[Lebedev et al., 2015] Lebedev, A. et al. (2015). Finding the self by losing the self: neural correlates of ego-dissolution under psilocybin. *Human Brain Mapping*, 36, 3137–3153.

[Lutz et al., 2015] Lutz, A. et al. (2015). Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. *American Psychologist*, 70, 632–658.

[Mason et al., 2020] Mason, N. et al. (2020). Me, myself, bye: regional alterations in glutamate and the experience of ego dissolution with psilocybin. *Neuropsychopharmacology*, 45, 2003–2011.

[Maurer et al., 1997] Maurer, R. et al. (1997). Phenomenological experience in response to monotonous drumming and hypnotizability. *American Journal of Clinical Hypnosis*, 40(2), 130–145.

[McNamara & Ballard, 1999] McNamara, L. & Ballard, M. (1999). Resting arousal, sensation seeking, and musical preference. *Genetic, Social and General Psychological Monographs*, 125, 229–250.

[Meck, 2005] Meck, W. H. (2005). Neuropsychology of timing and time perception. *Brain and Cognition*, 58(1), 1–8.

[Millière et al., 2018] Millière, R. et al. (2018). Psychedelics, meditation, and self consciousness. *Frontiers in Psychology*, 9, 1475.

[Moujaes et al., 2024] Moujaes, F. et al. (2024). Comparing neural correlates of consciousness: From psychedelics to hypnosis and meditation. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 9(5), 533–543.

[Nater et al., 2005] Nater, U., Krebs, M., & Ehlert, U. (2005). Sensation seeking, music preference, and psychophysiological reactivity to music. *Musicae Scientiae*, 9(2), 239–254.

[Perry et al., 2021] Perry, G., Polito, V., & Thompson, W. (2021). Rhythmic chanting and mystical states across traditions. *Brain Sciences*, 11(1), 101.

[Preller et al., 2018] Preller, K. et al. (2018). Changes in global and thalamic brain connectivity in lsd-induced altered states of consciousness are attributable to the 5-ht2a receptor. *eLife*, 7, e35082.

[Rogerson et al., 2021] Rogerson, R., Barnstaple, R., & DeSouza, J. (2021). Neural correlates of a trance process and alternative states of consciousness in a traditional healer. *Brain Sciences*, 11(4), 497.

[Rouget, 1985] Rouget, G. (1985). A theory of the relations between music and possession. *Journal of Ethnological and Historical-Religious Studies*.

[Schäfer & Fachner, 2014] Schäfer, T. & Fachner, J. (2014). Listening to music reduces eye movements. *Attention, Perception, & Psychophysics*, 77(2), 551–559.

[Schäfer et al., 2013] Schäfer, T., Fachner, J., & Smukalla, M. (2013). Changes in the representation of space and time while listening to music. *Frontiers in Psychology*, 4.

[Schlaug, 2001] Schlaug, G. (2001). The brain of musicians: A model for functional and structural adaptation. *Annals of the New York Academy of Sciences*, 930, 281–289.

[Sihvonen et al., 2021] Sihvonen, A. J. et al. (2021). Resting-state language network neuroplasticity in post-stroke music listening: A randomized controlled trial. *European Journal of Neuroscience*, 54(11), 7886–7898.

[Simonsson et al., 2023] Simonsson, O. et al. (2023). Altered states of leadership: mindfulness meditation, psychedelic use, and leadership development. *Frontiers in Psychology*, 14, 1151626.

[Timmermann et al., 2023] Timmermann, C. et al. (2023). A neurophenomenological approach to non-ordinary states of consciousness: hypnosis, meditation, and psychedelics. *Trends in Cognitive Sciences*, 27(2), 139–159.

[Visted, 2015] Visted, E. (2015). Emotion regulation in current and remitted depression: a systematic review and meta-analysis. *Frontiers in Psychology*, 6, 1892.

[Vroegh, 2019] Vroegh, T. (2019). Zoning-in or tuning-in? identifying distinct absorption states in response to music. *Psychomusicology*, 29(2–3), 156–170.

[Woodside et al., 1997] Woodside, L. N., Kumar, V. K., & Pekala, R. J. (1997). Monotonous percussion drumming and trance postures: a controlled evaluation of phenomenological effects. *Anthropology of Consciousness*, 8(2–3), 69–87.

[Yu et al., 2024] Yu, Z. et al. (2024). Alterations in brain network connectivity and subjective experience induced by psychedelics: a scoping review. *Frontiers in Psychiatry*, 15, 1386321.

[Zhigalov et al., 2019] Zhigalov, A. et al. (2019). Decoding attentional states for neurofeedback: Mindfulness vs. wandering thoughts. *NeuroImage*, 185, 565–574.

Expanded Authorship in Digital Art

Enrico Dorigatti
University of Portsmouth, School of Creative Technologies, Portsmouth, UK

Abstract

Traditionally, authorship and agency over the realisation of an artwork have been a prerogative of the artist. However, more recently, the advent of digital technologies and their exploration and use within the artistic context have challenged this axiom.

Focusing specifically on the broader context of digital art and some of its subsets, this paper introduces an expanded concept of authorship and shared agency, facilitated and mediated through digital technologies. While the first part of this paper presents the artistic domains considered, the second part reflects over the concepts of agency and authorship.

1 Digital Art

Digital art, popularised in the 90s by the first museum exhibitions showcasing artworks described as belonging to this category, can be defined as an artistic domain heavily based on computers and digital technologies [Marcos, 2007]. *Digital art* is a term developed after earlier definitions, either too generic or too specific to effectively convey all the nuances of digital art. Examples include the labels *media art* and *computer art*. Digital artworks are not only immaterial works made of data but, more generally, artworks relying on digital technologies as part of their production process or embedding them as fundamental features [Kwastek, 2023].

The domain of digital art has witnessed a surge in terms of popularity and enthusiasm shown by artists, critics, scholars, and the audience, to the extent that artistic forms of expression relying on digital technologies are nowadays widely popularised and not confined within avant-garde and specialised exhibition venues any longer. Digital art leverages technologies and techniques made available by rapid technological advancements. Sometimes, these are borrowed from fields different from the artistic one. Data visualisation [Rahlf, 2017] and data sonification [Bonebright et al., 2010] offer two examples of appropriation [Dix, 2007] exercised by artists to explore artistically what originally was meant to be used for different, non-artistic purposes [Viégas & Wattenberg, 2007]. The possibility of mixing various technologies and techniques makes the digital art context extremely fluid and broad, and open to virtually limitless artistic combinations.

2 Sound Art

As for the *digital art* label, sound art is an umbrella term used to identify a number of artistic practices. Often referred to with the even broader *sonic arts* concept proposed by Trevor Wishart [Wishart, 1996], the field of sound art seems too vast and rapidly changing for scholars and practitioners to define it through a shared and widely acknowledged definition, which therefore remains elusive [Licht, 2009]. Most authors writing about sound art acknowledge this field comprises a heterogeneous set of practices and aesthetic perspectives focusing on sound or, at least, employing it as a prominent artistic feature among others, with the most common being interactive interfaces, thus interactivity, and visual elements [Keylin, 2023, Knight-Hill & Margetson, 2024, Porcello & Patch, 2022]. Multidisciplinarity seems to be a key aspect of sound art, allowing artists and artworks to be inscribed within this artistic field while not necessarily dealing exclusively with sound. Yet, others criticise the seemingly endless openness of the sound art label, and the field it defines, as a shortcut to define 'art' what is not, with the consequence that "Much of what has been called 'Sound Art' has not much to do with either sound or art" [Neuhaus, 2000, p. 1]. While this might seem a pessimistic standpoint, it is undeniable that a relaxed or even non-existent framework for defining sound art may lead to the inclusion in this context of practices and artistic production (artefacts) dealing with sound but with poor artistic value, or vice versa.

Within the context of this paper, a possible definition of sound art acknowledging the breadth of this field [Knight-Hill & Margetson, 2024, Licht, 2009] might define sound art as *a set of artistic practices piv-*

oting around the employment of sound possibly in conjunction with other artistic features, and aimed at realising artworks meant to be showcased in spaces and with modalities different from those traditional of music.

3 Generative Art

Generative art defines a broad and nuanced domain, yet easier to define when compared to sound art. Generative artworks leverage entities and processes to produce an artistic outcome [Dorin et al., 2012]. Although generative art is technology-agnostic *per se* [Galanter, 2016] and, thus, the entities involved and processes carried out may be of any nature, generative art often relies on digital technologies and computers [Boden & Edmonds, 2009]. The computational power offered by these machines, in fact, allows for complex operations to be carried out in short periods, consequently allowing for results otherwise impossible or hardly possible to achieve.

The *generative* adjective describes the way in which an artwork belonging to this category is realised and evolves, thus through processes (tasks) assigned to physical or virtual entities to perform. The outcome of such processes effectively leads to the ontological manifestation (realisation) of the artwork. As these processes can produce both audio and visuals, some generative artworks can be rightfully considered as sound art.

4 Interactive Art

Interactive art defines artworks as providing an environment within which various agents develop a network of relationships, in turn allowing the artwork itself to manifest [Harries, 2013]. The main agents involved within the network of interactions are the artist providing a system allowing the interaction to take place, the system itself (artefact), and the audience. Within an interactive context, the audience specifically has a prominent role and makes a fundamental contribution to the realisation of the artwork [Kluszczynski, 2010]. Indeed, it is worth highlighting that the artefact provided by the artist is not the final and complete artwork but rather a system allowing the artwork to manifest through the interaction performed by the audience. This statement highlights a shift in the role of the agents involved, to the extent that the notions of *participants* and *art system* were proposed to replace *audience* and *artwork* [Cornock & Edmonds, 1973], underlining the dynamism of the interactive exhibition setting.

Interactive art relies on the constraints and affordances [Magnusson, 2010] set by the artist through the artefact provided to the audience. The artefact is meant to be explored by the audience and, simultaneously, to guide and direct the interaction performed [Bongers, 2000, Edmonds et al., 2004]. By defining an interactive artistic space that the audience can, physically or metaphorically, freely explore, interactive art becomes an effective paradigm for a more personal, dynamic, and engaging artistic experience unfolding within the limits defined by the artist and embedded within the artefact.

As for generative art, the *interactive* adjective characterises the artwork in terms of affordances and functioning but does not suggest anything concerning its output, which, as in generative art, could include audio, visuals, and other features. Interactive artworks, therefore, may belong to the domain of sound art.

4.1 The Open Work

As a precursor of the modern, digitally-soaked concept of interactive art, it is worth mentioning the notion of *opera aperta* (open work) proposed by Umberto Eco. Laying the groundwork for the development of the field of interactive art, Eco acknowledges the artist as responsible for creating an artefact rather than a fixed, immutable, and finished product. The artefact, in turn, does not provide the final, complete artistic experience but rather sets interactive and perceptual boundaries, leading the audience through an intimate, personal, yet controlled interpretation of the concept proposed by the artist:

> Le suggestioni sono volute, stimolate, richiamate esplicitamente, ma entro i limiti preordinati dall'autore, o meglio dalla macchina estetica che egli ha messo in moto. La macchina estetica non ignora le capacità personali di reazione degli spettatori, anzi le chiama in gioco e ne fa condizione necessaria della sua sussistenza e del suo successo; ma le indirizza e signoreggia [Eco, 1971, pp. 81–82]

In the translation:

> The suggestions are intentional, provoked, and explicitly reiterated, but always within the limits fixed by the author, or, better, by the aesthetic machine that he has set in motion. This aesthetic machine does not ignore the audience's capacities for response; on the contrary, it brings them into play and turns them into the necessary condition for its subsistence and its success, while directing them and controlling them [Eco, 1989, p. 35].

It is worth noting that, according to Eco, even the traditionally most seemingly closed and fixed forms of artistic expression, such as music or fine arts, are, on the contrary, open and interactive at least from a conceptual level, as they are prone to the personal interpretation and assimilation performed by the audience.

5 Rationale for Categorisation

The artistic categories overviewed offer an increasing level of specificity regarding artworks. It is common for generative and interactive artworks to employ sound as a fundamental feature of their output, thus being correctly considered as belonging to the sound art domain. Likewise, thanks to the ease of access for artists to digital technologies, their pervasiveness and their immediate usefulness within these artistic contexts, most generative and interactive artworks are also digital artworks, as the bare usage of digital sound or images can be enough to categorise an artwork as such [Kwastek, 2023]. It is, therefore, very likely that generative or interactive artworks making use of sound, thus fitting the sound art context, in turn, also fit the overarching domain of digital art.

6 Expanded Authorship

The contexts of generative and interactive art offer opportunities for reflection on the concepts of authorship, artwork, and artist, as anticipated by the notion of open work proposed by Eco [Eco, 1971]. While generative art relies upon processes carried out by external entities, often of digital nature, interactive art leverages, as the name suggests, the interaction performed by the audience, which can be again considered as an external agent. In general, agents contributing to the realisation of generative or interactive artworks may be of various origins, from human beings to algorithmic systems acting as virtual agents [Dorigatti, 2022]. However, in all cases, these agents are external to a traditional conception of the creative process, which sees the artist as the only agent with agency over the artwork.

The external agents involved in the creative process acquire agency over the creative process itself and, consequently, the resulting artwork. In both the generative and interactive artistic domains, scholars investigated the issues arising from granting agency within a creative process and over an artwork to agents other than the artist (author). In the context of generative art, Galanter proposed a set of questions on the matter, aimed at those involved in generative art, namely artists, audience, and critics [Galanter, 2016, Galanter, 2019]. Likewise, within the context of interactive art, Dannenberg and Bates reflected upon the notions of agency and authorship [Dannenberg & Bates, 1995].

Among the specificities of the two domains considered, two points may be relevant in both generative and interactive art. In particular, these points prompt reflections on the role of the artist (what does the artist do and which role do they fulfil?) and authorship (who realises the artwork?). Answering these questions means drawing upon the characteristic traits that generative and interactive art share. In both fields, the artist does not provide the final artwork but rather an artefact, a sandbox in which other agents, whether the audience, digital entities, or both, are free to explore, manipulating the settings according to the constraints and rules set by the artist through the artefact itself. The changes made are subsequently manifested through the artwork.

The artist does not disappear or abdicate their role but consciously relinquishes some control over the artistic and creative process. This relinquishing of control can be seen as the *conditio sine qua non* for the realisation of generative and interactive artworks, without which they cannot happen and thus be considered as such. The artist, therefore, while retaining full authorship as the entity who designed and set in motion the creative process, at the same time expands authorship to encompass, within the creative process, external agents, granting them a variable degree of agency.

7 Discussion and Conclusions

The concept of expanded authorship is nuanced, given the virtually unlimited ontological scenarios in the digital art domain and specifically in its generative and interactive subsets, which in turn may be inscribed within the sound art context. This paper has reviewed the relevant background for the concept of expanded authorship, which, however, will require further elaboration and expansion. The discussion developed on the topic, drawing from Eco's open work concept, should be seen as a sketched starting point for future investigations, which may be furthered by the critical reflection performed on their own artworks by practice-based scholars and artists exploring the domains of generative and interactive art. Given the growing and ubiquitous use of AI within the arts, the development of this concept may be especially beneficial. In particular, it could contribute to providing a framework for the correct attribution of agency and authorship to all the agents involved in the creative process, to better describe their role not only in relation to the actual realisation of the artwork, but also concerning its design and conceptualisation.

References

[Boden & Edmonds, 2009] Boden, M. A. & Edmonds, E. A. (2009). What is generative art? *Digital Creativity*, 20(1-2), 21–46.

[Bonebright et al., 2010] Bonebright, T., Cook, P., Flowers, J., Miner, N., Neuhoff, J., et al. (2010). *Sonification report: status of the field and research agenda*. White paper, National Science Foundation and International Conference on Auditory Display (ICAD).

[Bongers, 2000] Bongers, B. (2000). Interaction in multimedia art. *Knowledge-Based Systems*, 13(7-8), 479–485.

[Cornock & Edmonds, 1973] Cornock, S. & Edmonds, E. (1973). The creative process where the artist is amplified or superseded by the computer. *Leonardo*, 6(1), 11–16.

[Dannenberg & Bates, 1995] Dannenberg, R. B. & Bates, J. (1995). A model for interactive art. In *Proceedings of the Fifth Biennial Symposium for Arts and Technology* (pp. 102–111). Connecticut College.

[Dix, 2007] Dix, A. (2007). Designing for appropriation. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but Not as We Know It - Volume 2*, BCS-HCI '07 (pp. 27–30). Swindon, GBR: BCS Learning & Development Ltd.

[Dorigatti, 2022] Dorigatti, E. (2022). Automating art: A case-study of cellular automata in generative multimedia art. In G. Torre (Ed.), *Proceedings of the International Computer Music Conference 2022* (pp. 175–181). Limerick: ICMC International Computer Music Association, Inc.

[Dorin et al., 2012] Dorin, A., McCabe, J., McCormack, J., Monro, G., & Whitelaw, M. (2012). A framework for understanding generative art. *Digital Creativity*, 23(3-4), 239–259.

[Eco, 1971] Eco, U. (1971). *Opera aperta*. Milano: Bompiani.

[Eco, 1989] Eco, U. (1989). *The Open Work*. Cambridge, Massachusetts: Harvard University Press. (A. Cancogni, Trans.).

[Edmonds et al., 2004] Edmonds, E., Turner, G., & Candy, L. (2004). Approaches to interactive art systems. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia* (pp. 113–117).

[Galanter, 2016] Galanter, P. (2016). Generative art theory. *A companion to digital art*, (pp. 146–180).

[Galanter, 2019] Galanter, P. (2019). Artificial intelligence and problems in generative art theory. In *Proceedings of EVA London 2019*: BCS Learning & Development.

[Harries, 2013] Harries, G. (2013). 'the open work': ecologies of participation. *Organised Sound*, 18(1), 3–13.

[Keylin, 2023] Keylin, V. (2023). *Participatory Sound Art: Technologies, Aesthetics, Politics*. Springer Nature.

[Kluszczynski, 2010] Kluszczynski, R. (2010). Strategies of interactive art. *Journal of Aesthetics & Culture*, 2(1).

[Knight-Hill & Margetson, 2024] Knight-Hill, A. & Margetson, E. (2024). *Art of sound: Creativity in film sound and Electroacoustic Music*. Routledge.

[Kwastek, 2023] Kwastek, K. (2023). *Aesthetics of Interaction in Digital Art*. Cambridge, Massachusetts; London, England: The MIT Press.

[Licht, 2009] Licht, A. (2009). Sound art: Origins, development and ambiguities. *Organised Sound*, 14(1), 3–10.

[Magnusson, 2010] Magnusson, T. (2010). Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal*, 34(4), 62–73.

[Marcos, 2007] Marcos, A. F. (2007). Digital art: when artistic and cultural muse merges with computer technology. *IEEE computer graphics and applications*, 27(5), 98–103.

[Neuhaus, 2000] Neuhaus, M. (2000). Sound art? *Volume: Bed of Sound*.

[Porcello & Patch, 2022] Porcello, T. & Patch, J. (2022). *Re-making Sound: An Experiential Approach to Sound Studies*. Bloomsbury Publishing USA.

[Rahlf, 2017] Rahlf, T. (2017). *Data visualisation with R: 100 examples*. Springer.

[Viégas & Wattenberg, 2007] Viégas, F. B. & Wattenberg, M. (2007). Artistic data visualization: Beyond visual analytics. In *2nd Online Communities and Social Computing (OCSC) as Part of HCI International* (pp. 182–191). Beijing, China: Springer.

[Wishart, 1996] Wishart, T. (1996). *On sonic art*, volume 12. Psychology Press.

Promenadologie: Dynamic Maps for Soundscape Composition

Oliver Turbitt[1], Jelle Akkerman[1], and Francesco Ardan Dal Rì[2,3]
[1] *Independent Researcher*
[2] *DISI - Department of Information Engineering and Computer Science, University of Trento, Italy*
[3] *Conservatory "F. A. Bonporti" of Trento and Riva del Garda, Italy*

Abstract

In this paper we present *Promenadologie*, a prototype of a novel browser-based interface for geo-tagged field recording reproduction. Building on the legacy of soundscape composition and inspired by the idea of *dérive*, we collected a series of audio recordings by freely exploring the urban soundscape of the city of Trento, Italy, along with geographical coordinates in real-time. These data are then used to create an interactive soundmap, allowing the user to listen to the soundscapes and follow the movements of the recorder as it unfolds throught the map.

1 Introduction

The idea of *soundscape* has long served as a mean to explore the acoustic environments we live in. Initially popularized by Schafer [Schafer, 1977] with a particular focus on natural, in-the-wild recordings, soundscapes soon started to incorporate other contexts, especially urban settings, thereby including a variety of sounds and noises produced by human activities [Jurková, 2012], and further refined by comprising concepts related to psychoacoustic and auditory display by scholars such as Neuhoff [Neuhoff, 2021].

The technique of *field recordings* played a central role in capturing such soundscapes, aiming at distilling the pristine essence of specific environment at specific time [Gallagher, 2015]. The traditional approach involves static and controlled audio takes, removing as many traces of the recorder's presence as possible. As such, footsteps, breathing, and other unintentional sounds are typically filtered out, resulting in a somewhat disembodied listening experience deprived from its original context.

In recent years, soundscapes have evolved from static representations into more dynamic and interactive experiences, particularly within the context of urban settings. For instance, the emergent practice of *soundmapping* [McMurray, 2018] entails capturing sounds from various locations and visually charting them on a digital map via geographical coordinates. Historically, sound maps are usually static, offering a limited interaction that often abstracts the temporal and spatial dimensions of the soundscape. While mostly developed in several research domains, such as archiving natural [Bruyninckx, 2011] or ethnomusicological [Seeger, 1986] sounds, promoting awareness on acoustic ecologies [Droumeva, 2017], or assessing urban noise levels [Raimbault et al., 2003], the practice of field recordings also developed in a thriving variety of artistic expressions, where in-the-wild recordings became proper works of sound art themselves.

Building on these concepts and inspired by the Situationist International's idea of *dérive* (see 2.1), in this paper we propose a novel interface for field recording listening, emphasizing mobility and interactivity. Reflecting on the keyword of this edition of the conference, we promote a **contamination** between the exploration of the urban soundscape, in this case of the city of Trento, Italy, with a compositional and artistic approach, in a dynamic and human-centered fashion. As such, by exploiting geographical real-time annotations, we created an interactive digital map and transformed the traditional linear playback bar into a mirror of the actual route taken by the recorder, or composer, during his/her geographical wandering: as listeners engage with the soundscape, they can follow the recorder's journey, discovering the sonic environment as it unfolds in time and space through an intuitive interface and fostering a closer connection with the exploratory experience.

2 Theoretical Framework

2.1 Psychogeography

The concept of psychogeography originated in the 1950s among the Situationist International [Wollen, 1989], a group of radical intellectuals and artists. In his 1955 essay *Introduction to a Critique of Urban Geography* [Debord et al., 1955], Debord first defined the term as the study of the "specific effects of the geographical environment on the emotions and behaviour of individuals". This involves a creative and critical exploration of urban spaces, with the end goal of (re)discovering one's surroundings. Debord's recommended method for engaging with psychogeography was the loosely structured practice of urban wandering, known to the Situationists as *dérive* (lit. "drift"), "a technique of rapid passage through varied ambiances" [Debord, 1958]. Walking in such a manner was envisaged as a way of widening an individual's horizons by deliberately breaking out of the routes they are familiar with (for instance, the route from one's home to one's workplace). The experience of doing so heightens one's attentiveness and may provide the spark for creative production.

Indeed, since Debord's era, psychogeography-derived methodologies have continued to be inspire myriad works in different artistic areas including, for example, the ambulant writings of British authors Iain Sinclair and Will Self [Brown, 2019], the films of Patrick Keiller [Mudie, 2016], the music of ambient composer Jon Hassell [Feehan, 2010], and the "locative installation" *Net Derive* (2006) by Atau Tanaka [Tanaka, 2006]).

2.2 Soundscape Composition

Soundscape composition aims at understanding, rethinking and expressing acoustic environments through creative audio works, implying recognizable field recordings as primary sonic elements to highlight the richness and complexity of real-world acoustic ecologies and invoke peculiar associations, memories, and imagination [Wrightson, 2000]. Such practice have been pioneered within the World Soundscape Project (WSP) from the late 1960s [Truax et al., 2013] and popularized by artists such as Hildegard Westerkamp, whose works promoting the enhancement of environmental awareness and a deep reconnection between the people and their auditory surroundings [Westerkamp, 2002] - relevant works including *A Walk through the City* (1981) and *Gently Penetrating beneath the Sounding Surfaces of Another Place* (1997), and Berry Truax, which creates imaginary soundscapes by blending layers of environmental recordings with electronic sounds - examples being *Pacific* (1979) and *La Sera di Benevento* (1999).

While the general trend is to further elaborate the compositions, organizing the material according to a narrative will, including synthetic sounds, or involving heavy processing of the raw recordings, other compositions follow a more documentary or *phonographic* approach [Drever, 2002], defined by Truax as *found sound*, in which recorded materials are proposed "with minimal or no alteration that can be listened to as if they were music" [Truax, 2008], pivotal work being WSP's *The Vancouver Soundscape* (1973-1996). Examples of such perspective include Annea Lockwood's *A Sound Map of the Hudson River* (1989), in which the author explores the sound textures of the river in different locations [Richardson, 2012], or Peter Cusack's *Sounds from Dangerous Places* (2022), where several abandoned places are creatively investigates through "sonic journalism" [Pisano, 2015]. The interest in listening to field recordings as actual compositions is further demonstrated by the presence of numerous dedicated labels, such as for instance Presque Tout[1], which regularly publishes contributions requesting practitioners to record the sonic environment from their open window.

2.3 Soundmaps

The term soundmap - or *cartophony* - refers to a representation of the acoustic environment of a particular location by spatially organizing audio recordings, typically on a digital and interactive map [Thulin, 2018].

From the early 2000s, a plethora of web-based interactive soundmaps and geo-tagged audio archives emerges, mostly as means of audio archivism or raising awareness on urban noise pollution and forms of sonic activism, often fostering a crowdsourcing approach to audio data collection [Droumeva, 2021]. For instance, the *Hush City* project, launched in 2017, provides a mobile app allowing users to identify and assess quiet areas in urban contexts [Radicchi, 2017]; the web app *soundinbetweenness*[2] allow users to upload soundwalks in public spaces and create manually-annotated "postcards" displaying their routes on an interactive map [Karahan, 2020]; or the *Radio Aporee* project, promoting contribution from practitioners from all around the world to collect locative audio data [Noll, 2013]. Creative application of soundmaps also emerged, for instance the *UrbanRemix* project, a collaborative web-based platform in which users can record geo-tagged sounds, remix them, or share them with DJs for live performances [Freeman et al., 2011], or *The Museum of Portable Sounds*, a collection of in-the-wild recordings collected in a single mobile phone and organized as an exhibition [Gates, 2021]. Finally, with particular regards of the city of Trento, Beozzo et al. recently proposed a soundmap of the Municipality inspired by the Kandinskyan essential geometries [Beozzo et al., 2023].

[1] `https://presquetout.bandcamp.com/`
[2] `https://soundinbetweenness.org`

3 Promenadologie

3.1 Project Concept

As composers interested in exploring urban soundscapes by the mean of field recordings, we reflected on our practice which, being strongly linked to the Situationist's idea of *dérive*, is based on the free investigation of *found* sound spaces, in which the compositional will is expressed through choices of movement and focus on extemporaneous sounds spontaneously emerging during the walk. This already constitutes a substantial difference compared to traditional field recording methodologies, where practitioners aim to remove their individual presence and agency. Furthermore, we believe that the geographical context of these recordings is of fundamental importance; in this sense, sound maps can help restore this relationship. However, being our experience centered on movement, we find that the existing literature suffer from a lack of effective kinetic representation.

Our project seeks to challenge these conventions by embracing the human element, fostering the preservation of spatial and temporal context of the recording process: as such, we propose an interactive web-based interface consisting of a soundmap in which the musician's route is transformed into a progress bar. By doing so, we aim to create a more immersive and authentic connection between the listener and the recorded environment. In this way, our soundmaps are not just a product of their own, but rather the final result - and expansive representation of - a much longer creative process which involves going outside, moving, recording and collecting data.

We called it *Promenadologie* in homage to the sociologist Lucius Burckhardt, creator of the critical-aesthetic methodology of the same name - or *Strollogy*, from the German *Spaziergangswissenschaft* - based on the practice of walking [Burckhardt, 2015].

3.2 Audio and Data Acquisition

We individually acquired a few paths around the city of Trento. Paths have been tracked arbitrarily over different days and times and last from ∼7 to ∼30 minutes, according to personal preferences.

Audio has been recorded using a wind-shielded Zoom H4n Pro, 44100Hz 16-bit, in stereo .wav format, with a 80Hz low-cut filter. A pair of closed AKG K-271 MKII headphones have been used for monitoring. No post-processing has been applied.

For location tracking, we initially considered several fitness-specific apps which allow to capture geographical coordinates in common formats includes GPS Exchange Format (.gpx), Training Centre XML (.tcx) and Flexible and Interoperable Data Transfer (.fit) [Bekker & Scholtz, 2024]. We choose the latter since it is also contains timestamps, and recorded the data relying on the free app Wahoo[3], which allowed us to store timestamped data for single individual sessions, and directly download the relative .fit files. By default, the app stores a datapoint every second, which is a suitable resolution for our purpose.

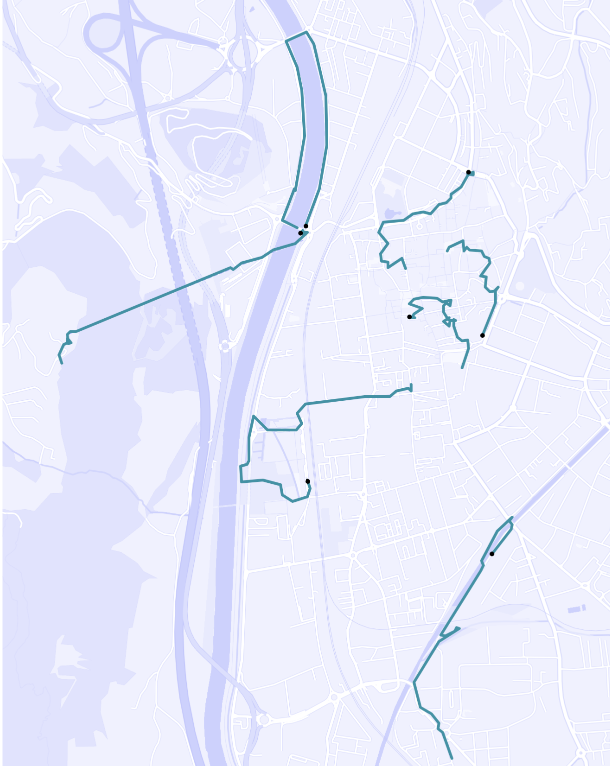


Figure 1: *An overall view of the interactive soundmap centered on the city of Trento, displaying several recodings which can be selected by the user.*

3.3 Implementation Details

At the current stage, our browser-based interface is developed to work locally.

First, in the backend, we iterate to every .fit file in the project folder and convert them to .json files, thus retrieving longitude and latitude for every timestamp, via fit-parser[4].

In the JavaScript frontend, we exploit Leaflet[5] [Cheng et al., 2024] to instantiate an interactive map centered around the city of Trento, and draw every path as .svg according to the respective GPS coordinates (Figure 1). Paths are smoothed in order to deal with slight drifts in the data. We managed audio playback trough event listeners, and update pointers every second along the corresponding paths, To further mimic the behaviour of typical media player interfaces, we use colors to highlight the past segment according to the current position of the pointer. More-

[3] https://it-eu.wahoofitness.com/fitness-apps
[4] https://github.com/jimmykane/fit-parser
[5] https://leafletjs.com

over, users can scrub through the recordings by moving the dots around the path. We finally add pop-up timestamp for the selected path, showing the current time and overall duration (Figure 2). Since there are multiple path displayed on the map, we also take care of automatic zooming, as in normal maps: when the user clicks on a path, the script zooms in to match the path area to the screen size. When a player is paused or the recording reaches its end, the script automatically zooms out to the original view.

We open-source the entire prototype along with audio files (converted to .mp3 for portability). The code is available at: `https://github.com/jellea/promenadologie`

Figure 2: *A closeup of a single recording while being played - a walk from Piazza Duomo to Piazza Fiera*

4 Conclusions and Future Work

In this paper, we proposed a project which aims to bridge the gaps between the existing disciplines of psychogeography, soundscape composition and soundmapping by way of mutual **contamination**, thus expanding their horizons with an emphasis on mobility and interactivity.

In this sense, our project's contribution is to encourage the kinetic production of soundscapes: the soundmap is created because of the soundscape, not vice-versa. The creative process becomes intertwined with the final product: our project's audiovisual reproduction of the route(s) taken allows users, in turn, to fully appreciate the composition itself. Such blending of sound mapping with field recordings also opens up new possibilities for engaging with our environments. By allowing listeners to remotely experience the essence of a sound walk beyond being mere "sound objects" — without the need for physical movement — we create a novel interface that democratizes access to these sonic experiences. The digital map does more than just chart sound: by reintroducing the dynamic factor, it serves as a medium for understanding the intricate relationship between sound and space, providing a platform where listeners can explore the acoustic ecology of both urban and natural environments.

Regarding the future of the project, we would like to elaborate more on the prototype, by adding images, details and contextual data on each recording (e.g. day, hour, name...) along with audio descriptors which can be useful to trace similarities between different soundscapes. The end outcome would be turning it into a website. We envision that possible use cases could span from tourist-oriented websites for local audio-based promotions, to an alternative player for field recording labels.

5 Acknowledgments

We would like to thank Emanuele Lapiana from *oSuonoMio* for providing support during the recordings.

References

[Bekker & Scholtz, 2024] Bekker, C. & Scholtz, B. (2024). An evaluation of health data accessibility and sharing features of smartwatches used by south africans. In *2024 IST-Africa Conference (IST-Africa)* (pp. 1–12).: IEEE.

[Beozzo et al., 2023] Beozzo, S., Chioni, C., Venturini, C., et al. (2023). Is landscape sound? a multidisciplinary approach to the soundscape of trento (italy). In *Interaction: proceedings of the 3rd DiMMI conference, held in Trento, 25-26 November 2022* (pp. 51–55).

[Brown, 2019] Brown, W. (2019). The essay and psychogeography: Negotiating marxism in the essays of iain sinclair and will self. *Art & the Public Sphere*, 8(1), 83–96.

[Bruyninckx, 2011] Bruyninckx, J. (2011). Sound sterile: Making scientific field recordings in ornithology. *The Oxford Handbook of Sound Studies The Oxford Handbook of Sound Studies*, (pp. 127–150).

[Burckhardt, 2015] Burckhardt, L. (2015). *Why is landscape beautiful?: the science of strollology*. Birkhäuser.

[Cheng et al., 2024] Cheng, J., Schloerke, B., Karambelkar, B., & Xie, Y. (2024). *leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library*. R package version 2.2.2.9000, https://github.com/rstudio/leaflet.

[Debord, 1958] Debord, G. (1958). Theory of the dérive. *Visual Culture: Spaces of visual culture*, 3, 77–82.

[Debord et al., 1955] Debord, G. et al. (1955). Introduction to a critique of urban geography. *Les lèvres nues*, 6(2).

[Drever, 2002] Drever, J. L. (2002). Soundscape composition: the convergence of ethnography and acousmatic music. *Organised Sound*, 7(1), 21–27.

[Droumeva, 2017] Droumeva, M. (2017). Soundmapping as critical cartography: Engaging publics in listening to the environment. *Communication and the Public*, 2(4), 335–351.

[Droumeva, 2021] Droumeva, M. (2021). The sound of the future: listening as data and the politics of soundscape assessment. *Sound Studies*, 7(2), 225–241.

[Feehan, 2010] Feehan, N. N. L. (2010). *Syncwalk: a framework for locative audio composition*. PhD thesis, Massachusetts Institute of Technology.

[Freeman et al., 2011] Freeman, J., Disalvo, C., Nitsche, M., & Garrett, S. (2011). Soundscape composition and field recording as a platform for collaborative creativity. *Organised Sound*, 16(3), 272–281.

[Gallagher, 2015] Gallagher, M. (2015). Field recording and the sounding of spaces. *Environment and planning D: society and space*, 33(3), 560–576.

[Gates, 2021] Gates, R. (2021). The museum of portable sound.

[Jurková, 2012] Jurková, Z. (2012). Environmental aesthetics and urban soundscapes. *Journal of Urban Culture Research*.

[Karahan, 2020] Karahan, S. (2020). Urban sonic research and sound map. soundinbetweenness. org. In *Proceedings of the 4th International Congress on Ambiances, Alloaesthesia: Senses, Inventions, Worlds*, volume 2 (pp. 148).: Réseau International Ambiances.

[McMurray, 2018] McMurray, P. (2018). Ephemeral cartography: on mapping sound. *Sound Studies*, 4(2), 110–142.

[Mudie, 2016] Mudie, E. (2016). 'reading the territory': Psychogeography and urban intertextuality. *Journal of Urban Cultural Studies*, 3(2), 275–284.

[Neuhoff, 2021] Neuhoff, J. (2021). *Ecological psychoacoustics*. Brill.

[Noll, 2013] Noll, U. (2013). radio aporee. *Berlin: Udo Noll*.

[Pisano, 2015] Pisano, L. (2015). The third soundscape: How a sonic exploration of abandoned spaces can lead to the discovery of unforeseen places and the geographies of a territory in the post-digital era. *Third Text*, 29(1-2), 75–87.

[Radicchi, 2017] Radicchi, A. (2017). Hush city: a new mobile application to crowdsource and assess' everyday quiet areas' in cities. In *Proceedings of invisible places: the international conference on sound, urbanism and the sense of place* (pp. 7–9).

[Raimbault et al., 2003] Raimbault, M., Lavandier, C., & Bérengier, M. (2003). Ambient sound assessment of urban environments: field studies in two french cities. *Applied Acoustics*, 64(12), 1241–1256.

[Richardson, 2012] Richardson, J. (2012). *River Soundscapes: Ecological Perspectives in the Music of Annea Lockwood, Eve Beglarian, and Leah Barclay*. Arizona State University.

[Schafer, 1977] Schafer, R. M. (1977). The tuning of the world: Toward a theory of soundscape design.

[Seeger, 1986] Seeger, A. (1986). The role of sound archives in ethnomusicology today. *Ethnomusicology*, 30(2), 261–276.

[Tanaka, 2006] Tanaka, A. (2006). Creating for mobile culture. *Intensive Science*.

[Thulin, 2018] Thulin, S. (2018). Sound maps matter: expanding cartophony. *Social & Cultural Geography*, 19(2), 192–210.

[Truax, 2008] Truax, B. (2008). Soundscape composition as global music: electroacoustic music as soundscape. *Organised Sound*, 13(2), 103–109.

[Truax et al., 2013] Truax, B., Westerkamp, H., Woog, A., & Kallmann, H. (2013). The world soundscape project. *WORLD SOUNDSCAPE PROJECT. Accessed May*, 15.

[Westerkamp, 2002] Westerkamp, H. (2002). Linking soundscape composition and acoustic ecology. *Organised Sound*, 7(1), 51–56.

[Wollen, 1989] Wollen, P. (1989). The situationist international. *New Left Review*, 174(1), 1.

[Wrightson, 2000] Wrightson, K. (2000). An introduction to acoustic ecology. *Soundscape: The journal of acoustic ecology*, 1(1), 10–13.

Contamination as Creative Process of the Intermedia Artwork: the Case of the Collaboration between Lucia Romualdi and Fausto Sebastiani

Angelica Speroni[1] and Alessandro Bile[2]

[1] *Paris 8, Paris, France*
[2] *Sapienza University, Roma, Italy*

Abstract

Lucia Romualdi has been a Roman artist active since the 1970s in conceptual art. In 1990, she moved on to the use of new technologies and the creation of 'light scores', installation interventions realised with video and diaprojections. She has collaborated with some of the most important contemporary composers: Franco Donatoni, Ivan Fedele, Fausto Sebastiani, or important performers such as Antonio Ballista, Claudio Jacomucci and Stefano Cardi, who have composed, transcribed and performed some pieces in dialogue with her scores. Since the 1990s, music has become an integral and indispensable part of her installations, and the relationship between this and the image is never didactic or descriptive, it does not arise from any transposition, but remains a dialogue open to mutual contamination, both on the compositional level and on that of the fruition. This does not present itself as visual or musical, nor simply multi-media, but creates an intermedia place, technically and conceptually, as a consequence of the coexistence of its various components.

This paper examines the collaboration between Lucia Romualdi and composer Fausto Sebastiani that lasted from 1995 to 2007. Over this period, they presented four works that testify to how the research of each was 'contaminated' by that of the other. Moreover, this contamination became increasingly stronger until the last work, *op. K6m33* (2007), in which the sound of the projection machines becomes part of a piece, *op. KM33*, that almost no longer has a musical autonomy of its own but only exists in dialogue with the visual part.

1 Introduction

After the experiences of the historical avant-gardes, only in recent decades has the dialogue between visual arts and music been rediscovered by artists. In this, in Italy, Lucia Romualdi was a pioneer. A Roman conceptual artist active since the 1970s, she has in fact explored this terrain, challenging the traditional boundaries between media. Concentrating since the 1990s on the creation of installations in which music and image meet in a continuous dialogue, she collaborated with important contemporary composers, such as Fausto Sebastiani, who played a crucial role in this evolution.

Between 1995 and 2007, Romualdi and Sebastiani created four works that represent a significant example of artistic contamination. In particular, the last of these, *op. K6m33* (2007), marks a high point, where the sound of the projection machines becomes an integral part of the musical composition, overcoming the distinction between the two languages.

This paper examines the collaboration between Romualdi and Sebastiani, analysing how in their common artistic journey contamination was at the basis of the creation process of the works themselves, redefining the concept of multimedia installation and creating an aesthetic experience that challenges the separation between sound and image.

2 Working with the Intermedia

The collaboration between Lucia Romualdi and Fausto Sebastiani lies in the field of intermediality. This section will therefore present the two artists and the concept of intermedia.

2.1 The Visual Artist

Active since the late 1970s in areas relevant to conceptual research, after an interesting and coherent experimentation with painting, Lucia Romualdi has been working in close symbiosis with musical languages since 1989. In 1990, she began collaborating with Franco Donatoni who, in dialogue with her work, composed the two well-known pieces *Cloches II* for two pianos (1990) and *Feria IV* two pieces for accordion (1997). She has collaborated with other contemporary composers; Ivan Fedele (*Canone Infinito* for magnetic tape, in 2002 - *Two Nocturnes with Figure* for piano and electronics, in 2010); Francesco De Gregori (*Cardiologia* for voice and kinematics, new transcription in 2014); and accordionist Claudio Jacomucci (he composed *Incantesimi* for spatialised accordion, in 2023) and Fausto Sebastiani.

A careful analysis of her work from the very beginning immediately reveals that intermediality is a working methodology for her, although she never uses this term. Her installations are not only on the threshold between music and the visual arts, but precisely in that in-between place that arises in the encounter of the vocabulary of the visual arts, the language of speech and other artistic media: processes typical of film editing, theatre, and of course music.

The chemistry with the composers has always been very important to her; only to them has she allowed to operate within a work that, as she conceives it, is not only the fruit of her individual genius, but also the result of a synergetic collaboration. Using figures, numerical diagrams and printouts, the artist creates scores of light (frames printed on acetate, digitally and manually edited, which are eventually presented as slides) that are transformed through projection into highly suggestive environmental installations. In fact, this approach gives the works a depth and complexity that goes beyond the narrative telling, and the contamination with the musical medium completely frees itself from 19th century experiments in transposition.

Her works have been installed in major museums on the contemporary circuit, in Rome at the Galleria Nazionale d'Arte Moderna (1999, 2010 and 2011) and at Palazzo delle Esposizioni (1993 and 2000), at the Museo Pecci in Prato (1998), in Antwerp at MuHKA (2002), and at MAXXI in Rome (2016), but the artist's favourite venues remain those outside the realm of art in the strict sense of the word, among them: Aquarium Neapolitanum, Naples 1977-1992; Serre Medicee di Rufina, 1981; Stanze d'acqua di Palazzo della Cancelleria, Rome 1992; Teatro di Palazzo delle Esposizioni, Rome 1993; Mitreo della Basilica di S. Clemente, Rome 1994; Clausura delle Suore Agostiniane della Basilica dei Santi Quattro Coronati, Rome 1994; Cinecittà 2, Rome 1994; Sala di lettura del Teatro di Rieti, Rieti 1995; Aquarium, Rome 1995; Scalo de Pinedo, Rome 1995; Cantina di Bomarzo, Bomarzo 1996; Osservatorio Astronomico, Rome 1998; Orestiadi, Gibellina 2000; Castel dell'Ovo, Naples 2003; Sinagoga, Šamorín 2011; Forte Michelangelo, Civitavecchia 2013, Sala Assoli 2023. Together with Sebastiani, she also experimented with music venues: the Centro Cultural Recoleta in Buenos Aires (1995) and the Acquario Romano as part of the event Progetto Musica '95.

2.2 The Composer

Fausto Sebastiani studied composition and electronic music at the Santa Cecilia Conservatory in Rome and attended courses with Sciarrino, Carter, Xenakis and Ferneyhough at the Ferienkurse in Darmstadt; he furthered his studies at the C.S.C in Padua and IRCAM in Paris. He realises orchestral and instrumental compositions with or without the aid of electronic media, melologues, art installations presented at national and international festivals and seasons. He has been invited to the LIPM in Buenos Aires, the C.R.M. in Rome, the Grame in Lyon and has had commissions from various institutions such as the Teatro dell'Opera in Rome, Accademia Filarmonica Romana, Biennale Musica in Venice, Festivali Heliloojad, Auditorium Parco della Musica, Sixty Fourth Aldeburgh Festival of Music and the Arts, Lisbon - Culturgest, International Contemporary Music in Akiyoshidai. He has participated in three recent editions of the Venice Biennale - Contemporary Music Festival 2013, 2016 and 2020. He teaches at the Frosinone Conservatory of Music and has held composition courses at Estonian Academy of Music and Theatre, Conservatorio Superior de Mùsica de Canarias, Liszt Ferenc Academy of Music, Conservatorio Superior de Mùsica de Sevilla. He was President of the Nuova Consonanza Association, is a Director of the Isabella Scelsi Foundation and a member of the Music Commission of the PSMSAD - INPS fund.

He met Lucia Romualdi in 1993 through the artist Luiz Allegretti and formed a professional relationship with her, but also a deep friendship. Until 2007, he composed four pieces for her, which will be discussed in detail in section 3. Generally speaking, in his professional activity Sebastiani has shown an openness to experimentation that goes beyond the strictly musical sphere, but it was in his collaboration with Romualdi that he truly and totally allowed himself to be 'contaminated' by the visual arts, offering himself for a contribution that led him to compose more and more, as will be seen, in function of the installation itself, as in the piece op.KM33. If for Romualdi it is therefore a given that the work cannot be separated from the music, which is indeed an integral part of it, in this case it can also be said that it is the visual part that is an integral part of Sebastiani's composition. This empathy between the two artists is at the same time what allowed the mutual contamination and the result of it.

2.3 Intermedia

The term 'intermedia' was coined by Dick Higgins in the context of the inclusive aesthetics of Fluxus to describe an experimental strategy of convergence of different artistic languages on both a technical and conceptual level [Higgins, 1966]. Unlike mixed-media, popular in the 1960s, intermedia emphasises fluidity rather than rigid categorisation. Higgins sees this practice as one of the hallmarks of modern society, and over time the concept has become central to contemporary aesthetics, especially since the digital revolution.

The historical roots of intermediality date back to the early 20th century, when artists began to rebel against the methodologies of institutionalised arts, leading to a variety of experimental experiences. In the first half of the century, artists such as Marcel Duchamp disrupted traditional artistic conventions, paving the way for new forms of spectatorship and participatory art. At the same time, there was a renewed interest in synaesthesia, as evidenced by works such as Schoenberg's *Die Glückliche Hand* and Skrjabin's *Prometheus*, which attempted to unite different arts into one balanced multi-sensory experience [Gelli, 2008].

These experimentations did not create a new genre, but gave rise to a multiplicity of different experiences. From the mid-1950s, and especially in the 1960s, we see the culmination of these intermedial experiences with sound installations and video art, particularly through the Fluxus movement. In Italy, despite a prolific artistic context, the absence of specific institutions and markets led to widespread but fragmented experimentalism. Before Romualdi, few visual artists oriented their research towards the intermedia, unlike composers such as Sylvano Bussotti, who explored this direction, particularly in the post-Webernian Nuova Musica [Torelli Landini, 1993].

In the contemporary panorama, intermediality, understood according to Dick Higgins' vision, is certainly characterising, and for this reason it is necessary to propose a specific theory of it, as a dialectic of convergence between different artistic languages, not as a fusion, but as a terrain of conflict and "threshold", also rediscovering its ethical value [Speroni, 2024].

3 The Contamination

In this section, the four works resulting from the collaboration between Lucia Romualdi and Fausto Sebastiani are analysed.

3.1 *Codice C*

Codice C is the title of the first work created in collaboration between Romualdi and Sebastiani. The fact that the title is unique for the two parts, visual and musical, is in itself a unicum. The piece was composed by Sebastiani during his residency at the LIPM in Buenos Aires (MAECI scholarship, July-September 1995).

During his stay in Argentina he actually also installed Romualdi's work entitled regard de l'etoile (for which she had entrusted him with the score) during a concert he organised (Musica Electroacustica en Italia y Argentina, 1995). While in this case one cannot speak of a true collaboration, at least in terms of the composition, it is interesting that Romualdi allowed him to bring one of his works into a space she had not seen and over which she would have no control. This case, unique in the artist's history, testifies to the great closeness and trust she had with Sebastiani.

Codice C, on the other hand, is a composition for acoustic guitar, electric guitar and electronic events, premiered on 27 October 1995 by guitarist Stefani Cardi. The total duration of the work is about 18 minutes, the entire duration of Romualdi's installation cycle. The composition is a journey from acoustic guitar to electric guitar; it begins with an initial section for acoustic guitar only, entitled Astrapen, which is divided into two parts: in the first, melodic fragments are presented, extremely minimal, almost like short melodies in the form of babble that represent Romualdi's luminous signs, minimal in character, abstract but in which the link with Leonardo's sign from which both artists draw inspiration is clear. Slowly, some of these sounds, especially those performed as a rapid, lightning-fast tremolo, are superimposed on similar electronically processed sounds, thus creating the entrance of double, out-of-phase figures, almost a shadow of those performed live. Gradually we come to a second part made up of Rasgueado chords alternating with longer sounds in the form of tremolos. This is followed by a section of 'electronic music only' that expresses the transition between the two instruments. In the third section, for electric guitar and electronics, the two musical figures from the first section are somehow echoed and electronic interventions, including rhythmic ones, creep in. A real dialogue is established, the electric guitar doubles up in the electronics and the total sound becomes broader and more involving, as if to completely envelop even the sounds of the installation's projectors, ever present, constant as the cicadas in a nocturnal meadow. It is like a musical journey that dialogues with the images, a gradual approach to the sound of Romualdi's 'visual machine', but one that does not accompany the scanning of his slides in the traditional way.

At the premiere *Codice C* is played in concert form at the Acquario Romano, as part of the event Progetto Musica '95, financed by Edi-Pan. Romualdi asks that long black sheets be lowered from the balcony above the stage, as screens for the projections. If the form is that of a concert, in reality one can already speak of an installation, although not in the strict sense of the term.

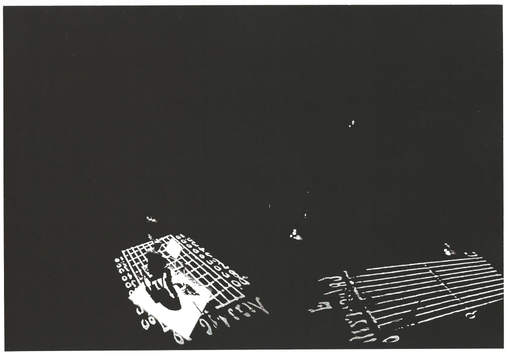


Figure 1: *Codice C, Progetto Musica 95, Acquario Romano, Rome 1995. Courtesy Lucia Romualdi Archives*

The work was then presented again in 1996 at Opera Paese, Festa della Musica, at the invitation of Paolo Pachini, in the former industrial spaces of Pietralata in Rome. The space was very different from the early 20th century Art Nouveau flavour of the Acquario. It was also smaller, more immersive, the projection also affected the spectators, Romualdi reduced the number of projectors, and here it had more the character of an installation.

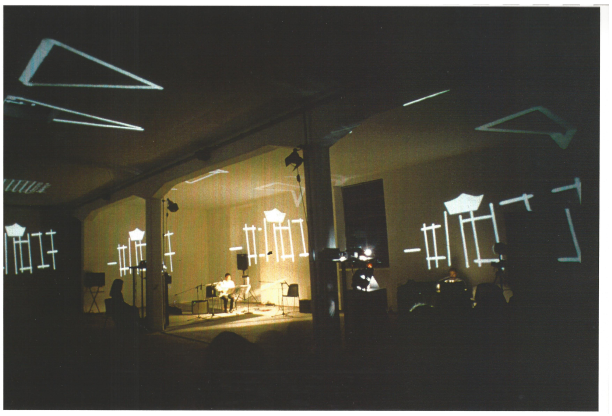


Figure 2: *Codice C, Festa della Musica, Opera Paese, Rome 1996. Courtesy Lucia Romualdi Archives*

3.2 *162 minuti Venezia - Cromatico*

The second collaborative work was *162 minuti Venezia*, a light score for 9 projectors, installed at the Studio d'Arte Contemporanea Pino Casagrande in Rome. The light score dialogues with *Cromatico*, a piece written in the late 1980s for accordion and revised for the occasion for accordion, to be played by Claudio Jacomucci. Until then, Romualdi had worked with piano and guitar, but she immediately enthusiastically agreed to dialogue with this new instrument. *Cromatico* is a very unusual piece, experimental in its own way because it makes the accordion emit extremely high-pitched sounds, known as harmonics, with a non-traditional technique of pressing the key, touching it and using the bellows in a very gradual way, and this is where the title comes from, not from a synaesthetic link with the visual art it nevertheless evokes. On this occasion again one can appreciate the contamination: Lucia fell in love with the accordion so much that she later asked also Franco Donatoni to compose for this instrument, which was to play a leading role in many of her later works.

3.3 *rasalhague la 55* α *- Melodie II*

rasalhague la 55 α is a light score for 21 projectors and rotating machines installed in 1999 at the MLAC in Sapienza, University of Rome and simultaneously in the dome of the Osservatorio Astronomico in Rome in dialogue with *Melodie II*. The electronic piece, also previously composed (1993) and reworked for the occasion, is a sound diagram for magnetic tape, written with the Csound software and the score shows the commands that the computer has to send to the synthesiser. Romualdi, who had already been working on tide tables for some time, became interested in the score graphically, visually as an object, and made it her own, inserting it into the work as a visual object. Here there is a further rapprochement in aesthetics between the two, a musical and artistic one, and the work, now properly an installation, is a great success. Also this work was performed again, at the Muspac in L'Aquila the following year, with the involvement of composer and perfomer Silvia Lanzalone who, by means of a computer, intervened with a real-time processing algorithm in counterpoint to the electronic sounds of the digital medium. In those days it was something futuristic or at least something that was being investigated. Lanzalone's performance was also visually addressed by Romualdi, who asked her for a 'geometric' gymnast's gesture.

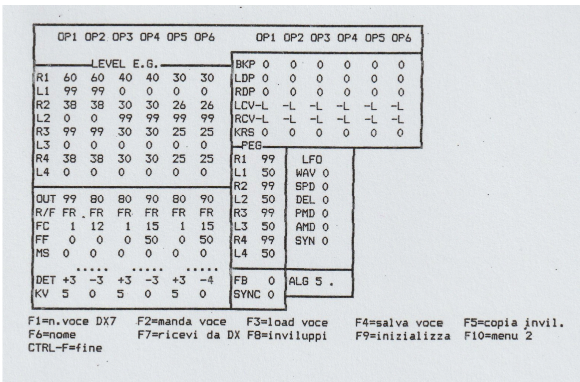


Figure 3: *Melodie II, detail of the score. Courtesy Lucia Romualdi Archives*

3.4 *op.K6m33 - op.KM33*

The last work, *Variazione op.K6m33*, is a light score for 11 optical machines, projectors and mechanical sounds also installed at Casagrande's. For this occasion, Sebastiani composes from scratch a piece based precisely on the recording of the sounds of the image projectors that become an integral part of it, inaugurating a methodology that will also be taken up by Francesco De Gregori for the rewriting of his work Cardiologia for soundings (Studio Trisorio, Naples 2014 - MAXXI, Rome 2016). Sebastiani understands that the installation has its own sound preceding the music, that of the projectors that make the slides move every few seconds. In writing the piece, he starts from this and then gradually departs from it, finding absolute harmony at the expense of the piece's musical autonomy.

Yet even here there remains a minimal gap, as can be seen in the titles (*op.K6m33* and *op.KM33*) that are 'almost' identical, which allows that friction from which the intermedia originates.

4 Conclusions

The article examines the creative process of intermedia "contamination" in the collaborative works of Lucia Romualdi and Fausto Sebastiani. As it unfolded by the four artworks, the main features of this contamination are:

1. Intermedial dialogue: a seamless fusion of visual art and music, creating a shared intermedia space where both forms remain distinct yet interconnected.
2. Overlap and Fusion: Sebastiani's music actively interacts with Romualdi's "light scores", forming a unified sensory experience rather than a descriptive one.
3. Creative Evolution: Each work reflects a deepening contamination, with both artists adapting their mediums to enhance interdependence.
4. Experimental and Inclusive Approach: Their work challenges rigid boundaries, embracing fluid artistic expression.

5 Acknowledgements

We thank Fausto Sebastiani and Archivio Lucia Romualdi for the material provided.

References

[Gelli, 2008] Gelli, P., Ed. (2008). *Dizionario dell'Opera 2008*. Baldini Castoldi Dalai.

[Higgins, 1966] Higgins, D. (1966). Intermedia. *Something Else Newsletter*, 1(1).

[Speroni, 2024] Speroni, A. (2024). *Raggiungere l'orizzonte. Significato, valore e funzione della ricerca artistica*. L'Erma di Bretschneider.

[Torelli Landini, 1993] Torelli Landini, E. (1993). Interferenze del sessantotto. *La Repubblica*. 6 Maggio.

FDBK: Real-time Interaction in a Neural-driven Music Ecosystem

Andreas Papaeracleous[1] and Francesco Ardan Dal Rì[2,3]

[1] *MTG - Music Technology Group, Universitat Pompeu Fabra, Spain*
[2] *DISI - Department of Information Engineering and Computer Science, University of Trento, Italy*
[3] *Conservatory "F. A. Bonporti" of Trento and Riva del Garda, Italy*

Abstract

In this paper, we present an interactive performative ecosystem for two musicians, based on neural networks which estimate parameters for a virtual synthesiser mimicking the performer's actions. In line with the conference's keyword, we designed the interaction allowing mutual sonic **contamination** between the musicians. We evaluated the proposed system with four different duets, asking the participants to provide insights in relation to each other and to the system itself.

1 Introduction

The presence of interactive music ecosystems has recently sparked interest towards the interplay between human performers and artifacts, fostering reflections on how musicians and algorithms co-create in a real-time scenario [Waters, 2007]. As such, the research scrutiny has shifted from individual focus to the interactions between musicians and autonomous agents [Rezwana & Maher, 2023]. Within interactive ecologies, performers can explore new forms of creative expression by engaging with machines in a collaborative or co-mediated fashion. Central to this dynamic is the idea of agency being distributed across human and machine participants [Moran, 2017].

We hereby present FDBK, a prototype for a neural-network-driven interactive system for two musicians. The software component extracts audio features from the performers' output in real-time, and exploits neural networks to approximate the parameters of digital synthesisers to mimic the original timbres. The interaction loop is established by cross-exchanging timbral parameters, enabling two musicians to engage in a reciprocal **contamination** process.

The prototype was tested and evaluated with the participation of four duets. The analysis of semi-structured interviews, conducted with the musicians after each experiment, brought out remarks and insights on the system itself and on how it was exploited and perceived by them during a short musical performance.

2 Background

2.1 Agency in Interactive Ecologies

In the last two decades, research on performance ecologies [Gurevich & Treviño, 2007] highlight the intertwined relationships between actors and artifacts. In this context, several frameworks for designing interactions [Rezwana & Maher, 2023] or analysing ecologies [Masu et al., 2019] have been recently proposed.

Interactive ecologies are thriving ground to foster reasoning on the many facets of cognitive processes involved in co-creative music making. Such architectures are characterised by a complex "network of interdependencies among system components" [Di Scipio, 2003], across which musical agency is distributed. As underlined by [Barad, 2007]: "agency is a matter of intra-acting; it is an enactment, not something that someone or something has". Indeed, agency is overall a distrusted force, that does not belong exclusively to human actors. According to [Latour, 2007], "anything that does modify a state of affairs by making a difference is an actor – or, if it has no figuration yet, an actant". These reflections resonate in many pieces of music research. Notable examples include [Melbye, 2021] and [Davis, 2017] practical/theoretical considerations on *shared agency* with their augmented instrument (respectively a self-resonating double bass and an actuated cello), *Ambiguos Devices* by [Stapleton & Davis, 2021], where agency is distributed across the networked human–machine-environment, and the live coding system by [Dal Rì et al., 2023], where musical actions are mutually co-mediated.

2.2 Parameter Estimation

Parameter estimation, or *sound matching*, refers to retrieving the parameter configurations of a synthesiser producing a given sound from the sound itself [Itoyama & Okuno, 2014]. Typically, it involves regression tasks operated by neural networks in a supervised contexts. State-of-the-art methods make use of deep learning networks which model acoustic samples by making the whole synthesiser differentiable ([Engel et al., 2020], [Shan et al., 2022]): these models require specific implementations, and are not regularly available in hardware and software synthesisers. Other networks are oriented towards recreating electronic spectrum, such as DiffMoog [Uzrad et al., 2024], which provides a series of modules allowing to modify the internal architecture, or Universal Synthesiser Control [Esling et al., 2019], which constructs an invertible mapping between auditory and parameter latent spaces. As in sound design, such systems imply a careful choice of synthesis techniques in order to approximate and recreate a variety of timbres. For instance, InverSynth [Barkan et al., 2019] uses subtractive and FM, while [Faronbi & Gilmore, 2021] uses different oscillators, noise, and filters. Despite the fact that some of these systems can be deployed in real-time scenarios, to the best of our knowledge, very little use has been done in performative situations. Usually, in such contexts, smaller networks are deployed, as in the performance FeedbackFeedforward (2022), in which the Feedback Cell duo explored the idea of a “shared instrument” by coupling a Feedback Cello [Eldridge & Kiefer, 2017] with a modular synthesiser. In this performance, the cello signal is analysed and a neural network reconstructs control voltage values for the modular synth.

3 FDBK

In this section, we provide an overall description of the prototype. For further details, please refer to the code: `https://github.com/Darce-One/FDBK`

We defined a simple SynthDef in SuperCollider[1]. It allows for both harmonic and inharmonic timbres by combining several UGens as source signals, namely an 8-component harmonic oscillator, an FM module, and pink noise. The mixed signal is then passed through two filters, high-pass and low-pass, and a fix reverberation is added in parallel before the output. In total, the synthesiser can be controlled by 9 individual parameters.

[1] `https://supercollider.github.io/`

Then, we implemented a feature extractor module in Python using Essentia[2], which analyses the signal coming from auxiliary busses over 8192-samples blocks. We retrieve the fundamental frequency $f0$ using the YIN algorithm [Brossier, 2006] and the amplitude $\mathcal{A}$ with an envelope follower. In addition, we also compute timbral features: 40 MFCCs, 40 Mel-bins, Spectral Contrast, Inharmonicity, Dissonance, Pitch Salience, and Spectral Flatness.

These timbral features are then passed to a Multi-Layer Perceptron (MLP) trained to reconstruct the synthesiser's parameters, with the exception of f0 and amplitude which are passed directly. Dealing with small input/output vectors, the network architecture is quite straight-forward, composed of five Linear Blocks with BatchNorm and Tanh/Leaky Relu/Relu activations, with a total of 31901 trainable parameters. The output vector is sent to the synthesiser via the Open Sound Control (OSC) protocol [Wright, 2005].

3.1 Training Pipeline

We need a dataset containing actual sounds along with their target synthesis parameters for a given synthesiser. We artificially created that by randomly generating 15,000 samples of one second in length, storing each file along with the respective parameters using SuperCollider with FluCoMa[3]. We constrained the random parameter generation process in order to avoid ambiguous situations in which the network would struggle to cope with (e.g. filter frequencies which can conflict with partials amplitude ratios, or oscillator amplitudes near 0 which can overcome the significance of other relative parameters and inject errors).

We wanted our network to be robust to outliers; as such, the objective is to minimize a Huber Loss L_δ in a multivariate regression problem. Given true parameter vectors $\mathbf{y}$ and predicted parameter vectors $\hat{\mathbf{y}}$, the loss function is defined as:

$$L_\delta(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}\ell_\delta(y_{i,j}, \hat{y}_{i,j}),$$

where N is the number of samples, M is the number of parameters, and δ is the threshold between the quadratic and linear behavior in:

$$\ell_\delta(y_{i,j}, \hat{y}_{i,j}) = \begin{cases} \frac{1}{2}(y_{i,j} - \hat{y}_{i,j})^2 & \text{for } |y_{i,j} - \hat{y}_{i,j}| \le \delta, \\ \delta|y_{i,j} - \hat{y}_{i,j}| - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases}$$

The MLP has been initialised with Kaiming uniform distribution [He et al., 2015] and trained on a regular 16GB CPU for 150 epochs using the Adam optimiser [Diederik, 2014], a batch size $bs = 50$, a learning rate

[2] `https://essentia.upf.edu/`
[3] `https://www.flucoma.org/`

$lr = 1e-4$, and a 90%/10% train/test split, achieving top MSE $= 0.04$, MAE $= 0.14$, and $R^2 = 0.46$ on unseen synthetic data.

3.2 Interaction Design

We first created two exact copies of the whole extractor-MLP-synth pipeline, each one dedicated to process the sound produced by a single musician. While $f0_i$ and $\mathcal{A}_i$ are passed directly to the respective synthesiser S_i, we exploited the timbral parameters to establish the interaction between the two musicians M1 and M2 by means of mutual **contamination**: indeed, the timbral parameters P_1 retrieved by the first MLP are passed to the second synth S_2, and vice-versa. Our interaction is designed upon the assumption that $f0$ and $\mathcal{A}$ are clearly recognizable, thus providing a certain degree of determinism for the players, while P is more ambiguous, allowing subtle textural interaction. Finally, we added two high-level controls for the synthesiser's temporal behaviour: a lag value L, which determines the temporal interpolation lag between consecutive parameters, and a delay value D, which injects a delay in the communication.

A visual summary is provided in Figure 1.

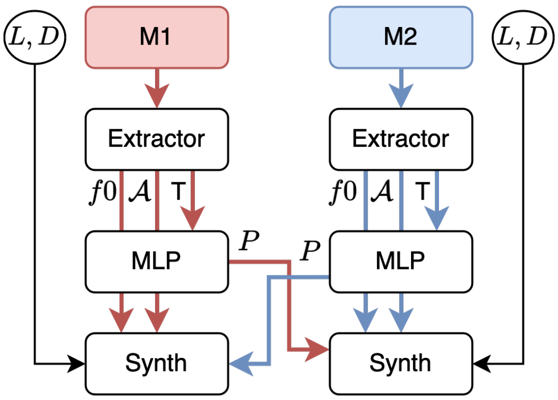


Figure 1: *Block representation of FDBK.*

4 Evaluation with Various Duets

4.1 Setup

We recruited eight volunteer professional musicians (here referred as P1-8) to take part in a preliminary evaluation of the prototype. Despite coming from different musical backgrounds, all of them have many years of experience in improvisational contexts. We arranged the duets covering different scenarios:

- Acoustic instruments (saxophone & double bass - P1 & P2);
- Electronic instruments (Theremin & drum machines - P3 & P4);
- Mixed (viola & synthesiser - P5 & P6);
- Unconventional (electric bass & voice + objects - P7 & P8)

Three of the experiments were conducted in person, while the fourth remotely, providing the software, trained model, and setup instructions. All evaluations were carried out as follows: first, we briefly described the system to the participants and provide some examples on the controls. Then, we allowed each duo to rehearse freely for some time (ca. 10 to 30 minutes) in order to take confidence with the system (Figure 2). During this phase we encouraged spontaneous discussion between the two musicians so that they could brainstorm musical ideas, without providing any insight or specific musical directions. Finally, we asked the duets to perform a short improvisation as if it were a live performance. Speakers were placed beside each performer, each one receiving signals corresponding to their respective musicians.

4.2 The Musicians' Perspective

At the end of each performance, we interviewed the musicians via semi-structured interviews. The recordings of such discussions have been transcribed and analyzed, extracting meaningful topics using thematic analysis [Braun & Clarke, 2012], which we outline here. Direct quotes are reported in quotation marks.

Overall, the participants appreciated playing with the system and found the idea interesting and fun to engage with. However, although the overall concept was easy to grasp, there were some difficulties in how to fruitfully exploit it in a musical way.

Indeed, in terms of **learning curve**, they all had initial concerns in trying to achieve reproducible results: "It's not plug and play" - P1; "It needed attention to grasp all the nuances" - P3; "It caused me some frustration" - P5. Their initial approach was to focus on their individual actions, and only after some time they did start to talk between themselves and collaborate: "Including the other player into the equation was a huge step to overcome" - P3; "The actual struggle was to exploit the interaction between us" - P5. All of them agreed on the necessity to acquire confidence with the ecology, and that additional rehearsing time would have led to more aware musical outcomes. Four of them proposed the idea of creating "specific studies to explore the interaction within a controlled scenario" - P2; "We need to discuss the system together" - P4.

Further reflecting on **interaction**, they perceived their mutual influence as particularly present and invasive at times: "The interaction loop was not irrelevant: it was always there, [...] something we needed to consider" - P1. Although the musicians occasionally struggled to establish a fruitful interaction, they all eventually "embraced this uncertainty" - P4 and "tried to be condescending with the ecosystem, not fight against it" - P6. Overall, the participants primarily focused on interacting amongst themselves, possibly mediated by the coexistence of the software counterpart. However, some of them elaborated more on such interactions and exploited the software to establish a deeper connection with their companion, highlighting the pivotal role of mutual co-creation: "A mental switch is needed to properly engage with such ecology. The risk is to just concentrate on yourself, forgetting about the general economy of the music" - P5; "My interventions in the long run became trivial without the contribution of the other musician" - P6.

This also resulted in different **performative strategies**. All the participants described their performance as mostly exploratory, as they experimented with different ideas such as counterpoint, mutual mimicking, timbral effects, and control mangling. Electronic musicians had more difficulty differentiating their instruments from the software synthesisers: "I was hiding myself beneath the sound produced by the software counterpart" - P4, and often yielded background drones: "I was providing a sort of canvas for my companion to work with" - P6. The majority of participants began to use a more gestural approach aiming at provoking a reaction in the software's behaviour: "I played quick, dilated gestures to simply provoke some weird sweeps in the other player's synth" - P6; "My intervention was intended to cause some glitches and perturbations" - P3.

Figure 2: *Two musicians freely rehearsing with the system before a performative session.*

The musicians deeply discussed **agency**. All of them were initially expecting the machine to be more proactive: "Nothing new was entering the conversation" - P7; "I did not perceive the machine as a third entity" - P1; "It was mostly [...] an extension of what I was doing" - P2. Overall, this tendency of looking for unexpected musical material by the software was much more evident for traditional instrumentalists, while electronic musicians seemed at ease with the system acting "like glue between us, injecting a further meta-level" - P4. As such, performers perceived agency differently in relation to the other musician, in combination with the interactive loop. Some tried to impose their agency and lead the musical direction: "It was me trying to control the improvisation" - P1; "I had the most responsibility for what was happening" - P6, while others adopted a more subordinate role: "I felt like the ecology had agency even without my contribution, due to the actions of the other player" - P4. Indeed, in line with the interaction design based on mutual contamination, all participants highlighted a certain degree of *shared agency* across all ecology components. Nevertheless, a dividing line has again emerged, with traditional musicians loading agency mostly on themselves and electronic ones supporting more distributive approaches. We argue that this is due both to personal backgrounds and to the mapping, direct or indirect, between gesture-sound of their instruments.

Finally, some **limitations** also emerged. The musicians all agreed that the system is more suited to specific musical situations: "It lacks some versatility" - P1. They considered long notes played by both musicians in similar registers as the most problematic scenario, while more musically effective results were obtained with complementary sonorities or gestures. However, some have speculated that this is due to their substantial seek for clarity: "There the result was confusing" - P1; "It was sometimes difficult to understand who was intervening on what" - P3. In such contexts, some also perceived the system as unsupportive: "It wasn't adding too much to the overall outcome, [...] its contribution was not so effective" - P1; "It should add more variations" - P8.

5 Conclusions

In this paper, we presented the design process and preliminary evaluation of FDBK, an interactive musical system prototype based on neural-driven parameter estimation, promoting mutual contamination between two musicians. Beside the prototype itself, the primary contribution of this work lies in providing reflections on how such ecosystems can enhance live performance by blurring the lines of individual agency and fostering novel cognitive and creative processes.

6 Acknowledgments

We would like to thank prof. Ted Moore and prof. Raul Masu for providing insights on this project. We initially developed the system at the Generative Music AI Workshop, hosted by Sound Of AI and MTG in Barcelona in June 2024 alongside with Luca Forcucci and Laura Lopez, to whom we extend our thanks.

References

[Barad, 2007] Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. duke university Press.

[Barkan et al., 2019] Barkan, O., Tsiris, D., Katz, O., & Koenigstein, N. (2019). Inversynth: Deep estimation of synthesizer parameter configurations from audio signals. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(12), 2385–2396.

[Braun & Clarke, 2012] Braun, V. & Clarke, V. (2012). *Thematic analysis.* American Psychological Association.

[Brossier, 2006] Brossier, P. M. (2006). *Automatic annotation of musical audio for interactive applications*. PhD thesis, Queen Mary, University of London.

[Dal Rì et al., 2023] Dal Rì, F. A., Zanghellini, F., & Masu, R. (2023). Sharing the same sound: Reflecting on interactions between a live coder and a violinist. In *Proceedings of the International Conference on New Interfaces for Musical Expression 2023*.

[Davis, 2017] Davis, T. (2017). The feral cello: A philosophically informed approach to an actuated instrument.

[Di Scipio, 2003] Di Scipio, A. (2003). 'sound is the interface': from interactive to ecosystemic signal processing. *Organised Sound*, 8(3), 269–277.

[Diederik, 2014] Diederik, P. K. (2014). Adam: A method for stochastic optimization. *(No Title)*.

[Eldridge & Kiefer, 2017] Eldridge, A. & Kiefer, C. (2017). The self-resonating feedback cello: interfacing gestural and generative processes in improvised performance.

[Engel et al., 2020] Engel, J., Hantrakul, L., Gu, C., & Roberts, A. (2020). Ddsp: Differentiable digital signal processing. *arXiv preprint arXiv:2001.04643*.

[Esling et al., 2019] Esling, P., Masuda, N., Bardet, A., Despres, R., & Chemla-Romeu-Santos, A. (2019). Flow synthesizer: Universal audio synthesizer control with normalizing flows. *Applied Sciences*, 10(1), 302.

[Faronbi & Gilmore, 2021] Faronbi, D. & Gilmore, A. (2021). Synthesizer parameter approximation by deep learning.

[Gurevich & Treviño, 2007] Gurevich, M. & Treviño, J. (2007). Expression and its discontents: toward an ecology of musical creation. In *Proceedings of the 7th international conference on New interfaces for musical expression* (pp. 106–111).

[He et al., 2015] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision* (pp. 1026–1034).

[Itoyama & Okuno, 2014] Itoyama, K. & Okuno, H. G. (2014). Parameter estimation of virtual musical instrument synthesizers. In *ICMC*.

[Latour, 2007] Latour, B. (2007). *Reassembling the social: An introduction to actor-network-theory*. Oup Oxford.

[Masu et al., 2019] Masu, R., Bettega, M., Correia, N. N., Romão, T., & Morreale, F. (2019). Arcaa: A framework to analyse the artefact ecology in computer music performance. In *Proceedings of the 9th International Conference on Digital and Interactive Arts* (pp. 1–9).

[Melbye, 2021] Melbye, A. P. (2021). Resistance, mastery, agency: Improvising with the feedback-actuated augmented bass. *Organised Sound*, 26(1), 19–30.

[Moran, 2017] Moran, N. (2017). Agency in embodied music interaction. In *The Routledge companion to embodied music interaction* (pp. 105–112). Routledge.

[Rezwana & Maher, 2023] Rezwana, J. & Maher, M. L. (2023). Designing creative ai partners with cofi: A framework for modeling interaction in human-ai co-creative systems. *ACM Transactions on Computer-Human Interaction*, 30(5), 1–28.

[Shan et al., 2022] Shan, S., Hantrakul, L., Chen, J., Avent, M., & Trevelyan, D. (2022). Differentiable wavetable synthesis. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4598–4602).: IEEE.

[Stapleton & Davis, 2021] Stapleton, P. & Davis, T. (2021). Ambiguous devices: Improvisation, agency, touch and feedthrough in distributed music performance.

[Uzrad et al., 2024] Uzrad, N., Barkan, O., Elharar, A., Shvartzman, S., Laufer, M., Wolf, L., & Koenigstein, N. (2024). Diffmoog: a differentiable modular synthesizer for sound matching. *arXiv preprint arXiv:2401.12570*.

[Waters, 2007] Waters, S. (2007). Performance ecosystems: Ecological approaches to musical interaction. *EMS: Electroacoustic Music Studies Network*, (pp. 1–20).

[Wright, 2005] Wright, M. (2005). Open sound control: an enabling technology for musical networking. *Organised Sound*, 10(3), 193–200.

Voice Contamination in Contemporary Music: Towards a Posthuman Interpretation of Music Through AI

Paolo Paradiso

University of Pavia, Cremona branch, Italy

Abstract

The technological development of the last decades has opened many new creative and artistic possibilities in the vocal musical field [Garda, 2020]. Among many others, even the concept of "contamination" has been severely impacted by recent digital development. AI, for example, has proven to have a pervasive influence in different aspects of contemporary music: from composition to performance practices and listening habits, both in the commercial and experimental music fields [Guiducci, 2023]. This article aims to provide an insight on the role of machine learning (ML) as a new tool for technological contamination in the field of contemporary vocal music.

The voice has always been a fundamental identity marker for human subjectivity [Connor, 2000]. Recent achievements in the manipulation of NLP and in the practice of cloning voices have contributed to question human identity and subjectivity, and human's role in the making of music [Paradiso, 2024]. The philosophy of Posthumanism provides useful ways to interpret this condition, through a new presentation of society and a redistribution of social roles in it [Braidotti, 2013]. Applying these concepts to the innovative musical asset of today can represent an effective way to deal with the rapid technological progress of recent years and to gain more awareness of its impact on cultural and artistic domains, such as music.

1 Introduction

The notoriety of the topic of AI has undergone a notable surge in recent years and we can therefore speak of a real new "spring" in AI research and development [Quintarelli, 2020]. This situation presents on the one hand a particular fervor of heterogeneous interest on this theme, with a consequent significant production of studies and research on the application of these technologies in the various fields of human knowledge. At the same time, such intense activity brings with it the risk of creating confusion (if not even misinformation), precisely because of the multitude and disparity of voices on these issues, nullifying the efforts of scientific research to clarify and provide tools for a deep understanding of complex subjects. A lack of a proper mindset to confront these topics could potentially do more harm than good [Harari, 2024].

The importance of the investigation into the impact of AI in music increases in scope if we consider that what happens in the musical field can subsequently be reflected in other realities: «therefore, the opportunity to explore the AI challenges only depends on the curiosity and willingness of its actors to examine how a sustainable and equitable music ecosystem responsible for all its human and nonhuman actors is possible and viable» [Clancy, 2023a]. From this we understand how studies on AI can represent: «a rare opportunity for us all to think deeply the meaning of being human» [Huang et al., 2023] and a way to escape from "technological sleepwalking", that is, a complete trust in technological progress (also understood as "technological determinism"[1]) that leaves no room for a deep reflection on its social repercussions [Clancy, 2023b].

[1]Two terms very close to the concept of "Transhumanism" a highly technophilic approach that view technology as a tool to expand human possibilities over its biological limits [Baioni et al., 2021].

2 The Voice of the Machine

2.1 Vocal Identity

Talking about voice is complicated because of its characteristics, which make it both immediate and intangible. It is the daily means of communication, especially – but not only – when crossed by language. One aspect that makes it particularly close to us is its bond with human identity and subjectivity. The voice marks our identity giving more specific information about us than our fingertips: it can indicate gender, age, health status, cultural background, and other aspects of the speaker [Leoni, 2022]. Through the voice the body can express itself in a genuine way (as singers knows very well), thanks to what Roland Barthes defines as "grain of the voice": «the materiality of the body speaking its mother tongue; perhaps the letter, almost certainly signifiance» [Barthes, 1977].

The pre-linguistic sounds of the voice can also represent one of the first signals to determinate the conceptualization of our subjectivity: «our preverbal sounds are a series of increasingly confident vibrational speculations that there is an "I" whom we can hear» [Pettman, 2017]. The voice gives the speaker the ability to recognize himself and being recognized by others [Connor, 2000].

This strong connection between the voice and the subjectivity has been a central point from which the vocal performance art of the 20th century has tried to take distance, through a process of depersonalization and gradual ventriloquism, unachievable without the technological development, especially of the computer and the digital manipulation of sound [Garda, 2020]. Therefore, technology has entered the music scene contaminating the human relationship with the voice, making it more complex and opening new creative and artistic possibilities [Garda, 2023]. Among the last technological innovations, ML represents the richest tool to exasperate this contamination, challenging the traditional definition of human and attracting posthuman ideas to the musical discourse [Di Scipio, 2023].

2.2 The Artificial Voice

Artificial voice recording and reproduction began with the invention of the phonograph by Thomas Edison in the late 19th century [Napolitano, 2022], but the origins of the desire to create an instrument for this purpose and the attraction to the concept of a "talking head" are much older[2]. Today's voice synthesis and cloning technologies (such as systems that exploit deep learning and ML) reveal themselves as the current phase of a long tradition that involves the recognition of different properties in systems for voice reproduction, depending on the case: prophetic[3], authoritarian[4], omnipotent or apparently incapable of telling lies[5].

As a sort of re-enactment of Eco's myth, technologies fulfill what for a long time has remained in the realm of the imaginary [Garda, 2023]. In the 18th century many talking machines were made, creating a strong attraction for this topic. In his study about the voice, Mladen Dolar describes this specific attraction for these voices as something at times disturbing. The few words and sentences carrying meaning generated by the machine, as described by Dolar, produced the voice as an excess to language and meaning. The attraction for it arose precisely from here: despite the difficulty in understanding the voice produced at low resolution, it creates fascination because it seems to emanate the most human of effects: that of "interiority" [Dolar, 2006]. The delicate role that the voice plays in Dolar's theory is now being questioned by the latest voice synthesis technologies. While the machine voices that Dolar talks about chase a generic conception of voice, technological progress has allowed (especially through ML) the production of customizable artificial voices, less flat, ethereal, robotic and indistinct, but more realistically "human", whose manipulation and technological origin is opaquer [Napolitano, 2022]. Today, in fact, distinguishing a human voice from a non-human, artificial one is an increasingly difficult task. Advances in the fields of AI allow the voice to be manipulated in ways that were unimaginable years ago, overcoming the idea of an impersonal, mechanically produced voice [Napolitano, 2022].

Engaging in conversations with artificial interlocutors is becoming an ever-increasing practice, for example using voice assistants. A new disturbing effect appears today in the accurate reproduction of the human voice through digital instruments, constantly imbued with an acousmatic effect (since its cause does not correspond to its source), that persists even when the source of the sound is visible[6] [Napolitano, 2022]. The tendency to subjectivation of machines is thus taken to the extreme, giving computers a semblance of an individual imprint perceived precisely by listening to their voice and with which we are increasingly led to deal with [Cox, 2017].

[2]This fascination can be traced back to at least the 1st millennium BC with the practice of divinatory rituals with teraphim by the Chaldean people [Pettorino & Giannini, 1999].

[3]For example, in the myth of the oracle of Orpheus [Pettorino & Giannini, 1999].

[4]As it is for the radiophonic voice, equipped with power properties [Garda, 2020].

[5]The popular aura of "safety" around the answers given by ML devices is highly contaminated by the multiple biases that are perpetuated by these probabilistic systems. The algorithms themselves should never be understood as objective, but as "opinions collected in codes" [Napolitano, 2022]

[6]The doubt remains as to whether the interlocutor is human or not. This effect can be described as the acoustic version of the "Uncanny Valley", a term created by Masahiro Mori to explain the disturbing sensation of uneasiness that can occur while watching humanlike robots [Pettman, 2017, Caballar, 2023]

3 Posthuman Voice

Contemporary experimental music is a very active field in the exploration of the impact of new technological developments regarding the artificial voice. Tomomi Adachi[7] and Jennifer Walshe[8] are two composer-performer-technologists that work intensely with ML algorithms in relation to the human voice. *Tomomibot*[9] (2018) by Adachi and *ULTRACHUNK*[10] (2018) by Walshe are two examples of how contemporary music can approach a posthumanist way to perform music in the hyper digital world of today [Paradiso, 2024].

These compositions consist of a live vocal improvisation between a human performer and their digital alter-ego, made using artificial neural networks (ANN), programmed to recreate the voice of the human performer[11] and interact with it during the performance. The result of this approach takes shape (in *ULTRACHUNK* as in *Tomomibot*) with an operational integration between the human and the artificial agencies, which cooperate in a shared space [Di Scipio, 2023].

One of the recurring themes in posthumanist philosophy is the decentralization of the human within the social framework: the humanistic Vitruvian model of worldview is no longer followed, and an attempt is made to approach a non-anthropocentric model, inserting the human within a complex ecosystem including all living beings and at the same time giving greater recognition to non-human agency, such as AI [Braidotti, 2013].

The vision of posthumanism is certainly very articulated, complex and in some ways even provocative, but it aims to reflect deeply on what it means to be human today in the era highly mediated by digital technology [Baioni et al., 2021]. *ULTRACHUNK* and *Tomomibot* are emblematic examples of interdependency between human and non-human agencies, both called to relate to each other and take part in the realization of the performance. The ANN becomes part of the performance as an actor with an active role, within the framework of the concept of distributed creativity, which is the basis of much live electronic music involving the voice, where the singer is often found alongside the electronic interpreter, thus dividing the aesthetic and performative realization between several actors [Sarno, 2023]. This, in turn, becomes more complex by integrating the non-human agency of the ANN into the discourse, whose work is inserted into a network of reciprocal influence with the human interpreter [Di Scipio, 2023]. These experimental compositions represent the posthumanist concepts of interactive network between human and non-human [Braidotti, 2013]. A network (or rhizome) able to create a “unitarian agency”: an agency with an ecosystemic character, with which it is possible to investigate the connection between voice, body and space within a performative device which is itself ecosystemic [Di Scipio, 2023].

In this interactive network, the voice plays the role of connector between human and non-human actors: a voice that is born from a human body and that from this finds a new space and body compared to the original ones; that therefore passes from a biological to a technological origin, contributing to the formation of artists’ artificial alter-egos. In this way, it becomes possible to create an extended body, a hybrid performative device formed by the human performer, the projection of elements of the performer to the outside (thanks to the computer), and the interaction of these in a shared body-environment [Di Scipio, 2023]. The voice remains the element most directly linked to the biological body, but, at the same time, the cloned voice complicates the identity bond I referred to earlier, placing an embodied voice alongside an extracorporeal one [Garda, 2020]. The meeting of these voices in the shared environment forms the posthuman performance, where human identity and artificial agency cooperate in the artistic realization of the performance [Di Scipio, 2023, Paradiso, 2024].

[7] https://www.adachitomomi.com/
[8] https://milker.org/jenniferwalshebiography
[9] https://www.youtube.com/watch?v=0VexyC86F8o
[10] https://www.memo.tv/works/ultrachunk/
[11] The ML approach is used to clone the voice of the performer, using a huge dataset of hours of audio and video material to train the ANN to create a synthetic vocal profile that can generate new vocal sound [Napolitano, 2022, Paradiso, 2024].

4 Conclusion

In this article the concept of contamination has been aligned with some of the most innovative ways to use AI in music. ML has already shown profound ways to contaminate vocal performance, pushing even the musical discourse towards more inclusive representations of society, such as the one postulated by Posthumanism. Most likely, we find ourselves just at the beginning of the potential of AI to contaminate human experience and, specifically, music. We have nothing left but to see what the future (even the most immediate one) holds for us.

References

[Baioni et al., 2021] Baioni, E., Cuadrado Payeras, L. M., & Macelloni, M. (2021). *Abbecedario del Postumanesimo*. Milano-Udine: Mimesis.

[Barthes, 1977] Barthes, R. (1977). *Image Music Text*. New York: Noonday Press.

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

[Caballar, 2023] Caballar, R. D. (2023). What is the uncanny valley? https://spectrum.ieee.

org/what-is-the-uncanny-valley. Accessed: 09/2024.

[Clancy, 2023a] Clancy, M. (2023a). Ethics. whose ethics? approaches to an equitable and sustainable music ecosystem. In M. Clancy (Ed.), *Artificial Intelligence and Music Ecosystem* (pp. 121). Abingdon: Routledge.

[Clancy, 2023b] Clancy, M. (2023b). Philosophy: Amor fati: A theoretical model of the music ecosystem. In M. Clancy (Ed.), *Artificial Intelligence and Music Ecosystem* (pp. 167–168). Routledge.

[Connor, 2000] Connor, S. (2000). *Dumbstruck. A Cultural History of Ventriloquism*. Oxford-New York: Oxford University Press.

[Cox, 2017] Cox, T. (2017). *A ciascuno la sua voce. Come parliamo e ascoltiamo dai Neanderthal all'intelligenza artificiale*. Bari: Dedalo.

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

[Dolar, 2006] Dolar, M. (2006). *A Voice and Nothing More*. Cambridge-London: The MIT Press.

[Garda, 2020] Garda, M. (2020). Arcipelago voce. In L. Cardilli & S. L. Vallauri (Eds.), *L'arte orale. Poesia, musica, performance* (pp. 180–181). Torino: Accademia University Press.

[Garda, 2023] Garda, M. (2023). *La mediazione tecnologica della voce*. Roma: NeoClassica.

[Guiducci, 2023] Guiducci, L. (2023). Musica artificiale. In D. Ventura (Ed.), *Pensiero umano e Intelligenza Artificiale* (pp. 61–94). Roma: L'Asino d'oro.

[Harari, 2024] Harari, Y. N. (2024). *Nexus. Breve storia delle reti di informazione dall'età della pietra all'IA*. Firenze-Milano: Bompiani.

[Huang et al., 2023] Huang, R. S., Holzapfel, A., & Sturm, G. L. T. (2023). Global ethics. from philosophy to practice: A culturally informed ethics of music ai in asia. In M. Clancy (Ed.), *Artificial Intelligence and Music Ecosystem* (pp. 131). Routledge.

[Leoni, 2022] Leoni, F. A. (2022). *Voce. Il corpo del linguaggio*. Roma: Carocci Editore.

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

[Paradiso, 2024] Paradiso, P. (2024). La sperimentazione vocale nell'epoca dell'intelligenza artificiale. In *Memorie Proiettive/Projecting Memories* (pp. 40–47). Venezia-Torino: DADI.

[Pettman, 2017] Pettman, D. (2017). *Sonic Intimacy*. Stanford: Stanford University Press.

[Pettorino & Giannini, 1999] Pettorino, M. & Giannini, A. (1999). *Le teste parlanti*. Palermo: Sallerio.

[Quintarelli, 2020] Quintarelli, S. (2020). *Intelligenza Artificiale*. Torino: Bollati Boringhieri.

[Sarno, 2023] Sarno, G. (2023). Io è/e un altro. In M. Garda (Ed.), *La mediazione tecnologica della voce* (pp. 195). NeoClassica. cit.

Contaminating Music Genres: the Role of Brass Bands

Mario Diani and Silvia Sacchetti
Department of Sociology and Social Research, University of Trento, Italy

Abstract

In this paper we discuss whether inclusive grassroots music associations may reduce the strength of the correlation between social status, cultural capital, and "omnivorousness" in musical taste. Drawing upon data from 810 members of local brass bands in the alpine province of Trentino, we suggest that participation in brass bands and other social and cultural activities reduces the impact of cultural and economic capital over omnivorousness.

1 Introduction

According to the Cambridge Dictionary online, contamination is "the process of making something dirty or poisonous". From the more optimistic and creative perspective adopted in this conference, it means "blending knowledge from different research areas, meeting at the crossroads of hard and soft sciences". Either way, it means combining elements that are usually not expected to be found together. Their combination requires breaking some kind of boundary – or, in cultural terms, canon. In this paper we do not focus on the process of musical contamination itself but rather on its preconditions. We assume, in other words, that a prerequisite of contamination be familiarity with, and openness to, different styles or genres. In the last three decades, reflections on this issue have been largely shaped by Peterson's [Peterson, 1992] well known argument about "omnivorousness". In a nutshell, this thesis originated from a criticism of rigid distinctions between highbrow and popular culture, and particularly their strong association with distinct socio-demographic profiles. In Peterson's words, "within the purview of the elite-to-mass perspective, it is incumbent on members of the cultural elite not only to do the right thing, but as importantly, to absolutely shun all other sorts of cultural practices" (ibid., 245). In contrast with approaches stressing the differences between elites and masses, and the role of cultural tastes in reinforcing social hierarchies, most notably associated with Pierre Bourdieu's Distinction [Bourdieu, 1979], Peterson suggested that people more endowed with economic and cultural capital would be more inclined to engage with a variety of cultural styles and genres, including those normally associated with the lower classes; while the latter would be more inclined to concentrate on a smaller range of styles and cultural forms. In his view, the fundamental divide, while still embedded in socio-demographic inequalities, did not run between high and low culture, but between omnivore and univore styles of cultural consumption. Peterson's argument has inspired innumerable empirical explorations and theoretical discussions that is impossible to discuss here (for a short yet usefull reference: [Hazir, 2021]). We accept his basic claim that higher economic and social capital be conducive to greater omnivorousness, i.e., to greater heterogeneity of cultural (in our case, musical) tastes. Instead, our goal is to explore if membership in voluntary associations active on cultural issues (in our case, local brass bands) manages to reduce the impact of socio-economic status over omnivorousness. Inspired among others by Ostrom's work on the commons [Ostrom, 1990], we regard a broad musical culture, enabling those who share it to engage with a variety of musical styles, as a particular kind of public good; such a good is created and reproduced through intense cooperation between people rather than through market exchange or state education. And of course, social movement analysts have long pointed at the role of grassroots collective action in generating innovation and creating new organizational forms (see e.g. [Rao et al., 2000]). Again, this is no place to discuss these different strands of literature (see [Borzaga et al., 2021]). Suffice to say that recognizing the role of associations in creating public goods prompts our first question: **do inclusive voluntary associations, attracting a socially diverse membership, reduce the impact of socio-economic differences over omnivorousness?**

In terms of musical tastes, omnivorousness can be measured as the number of genres people are interested in. We should also recognize, however, that some genres are closer to each other than others. For example, jazz and blues can be supposed to be closer to each other than classical music and rap, or brass band repertoire and electronic music. Accordingly, in order to measure omnivorousness it may be useful to differentiate between people who engage with different musical tastes belonging in the same broader family, and people that engage with tastes that cut across different musical families. Contamination may indeed be expected to be higher if people follow and practice musical genres that the majority regards as distant from each other, and tends not to combine. Concepts and methods from network analysis have been routinely used recently to explore connections between cultural forms [Pachucki & Breiger, 2010], often with a special attention to music [Vlegels & Lievens, 2017]. In these studies, the tie between two genres is weighted against the number of people practicing both. The people who span genres that are rarely followed jointly – that bridge, in other worlds, "cultural holes" [Lizardo, 2014] - can be regarded as the most likely agents of contamination. They act as "interpreters" [Bauman, 1987], bringing together different musical languages and facilitating their creative combined use. This prompts our second question: **does socio-economic status affect the probability of bridging usually non-communicating musical styles?**

In order to explore our questions we rely on data from 810 members of brass bands, active in different areas of Trentino. It's important to recognize that a proper test should compare members of brass bands with the general public in Trentino. This would enable us to check whether associations actually make a difference. As we lack those data, we'll limit ourselves to check if socio-economic features still contribute to explain some variation in brass band members' degree of omnivorousness.

2 Analysis

Table 1 reports about the popularity of the different music genres among brass band members (figures refer to members who listen regularly and/or practice specific genres). It also reports the corresponding percentages among members holding degrees and belonging either in the professional/managerial class or the educated service class. As the tables shows, there are very few significant differences between members of the higher status social groups and the other members.

Our indicator of omnivorousness is a simple additive measure of the 14 different genres we explored (it averages 4.7 with a s.d. of 2.8). Given its distribution very close to a normal curve we run stepwise OLS regression models, first fitting socio-demographic variables only, and then adding additional variables measuring brass members' involvement in a range of associations and cultural activities in the local community. The first column in Table 2 shows that some sociodemographic traits affect the breadth of musical tastes, but they are not directly linked to status.

Table 1: *Strong interest in musical genres by university degree and class position.*

	Hold a degree	Professional/ manager	Educated service class	Total
Band music	62	56	68*	61
Classica	69*	64	62	58
Pop	49*	59	53	53
Folk	48	51	54*	49
Rock	49	44	53	47
Songwriters	37	46	49	45
Blues	49*	49	45	42
Jazz	43	41	43	40
World	37	39	38	37
Liscio	18	31*	21	20
Electronic	9	10*	6	7
Rap	4*	10*	4*	5
Trap	3	5*	2	3
N	274	61	245	810

* Category differs from total at .05 significance level.

Table 2: *Stepwise OLS regression of band members' traits on omnivorousness (s.d. in brackets; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$).*

	Model 1	**Model 2**
Socio-demographic traits		
Age (ln)	1.27*** (.22)	.87*** (.22)
Secondary school diploma	.51** (.21)	.54** (.20)
Conservatory of music	1.14* (.49)	1.02* (.48)
Artisan/skilled worker	N.S.	.95* (.46)
Associational memberships		
Cultural circles		.88*** (.24)
Other musical ensembles		.59** (.22)
Amateur theater companies		.46* (.20)
Human rights groups		-.84* (.42)
Sport clubs		.59** (.24)
Folklore revival groups		.76* (.33)
Attendance of events		
Cultural festivals (occasional)		.91*** (.23)
Cultural festivals (regular)		1.10** (.46)
Other bands' concerts (occ.)		.74* (.31)
Other bands' concerts (reg.)		.89* (.44)
Constant	.02 (.76)	-3.04*** (.86)
N	748	748
Prob > chi^2	.00	.00
Pseudo R^2	.06	.18

Age is on, as omnivorousness seems to grow with it and the other is education, but not higher education. Apart from holding a conservatory degree (interesting finding in its own right, as conservatories are supposed to be protectors of traditional canons), a high school diploma, not a university degree, seems to predict broader musical tastes. In any case, the explanatory capacity of the model more than triples when we introduce variables that measure involvement in civic and cultural life. Participation in a range of associations in the fields of culture and sport turns out to have a positive effect; so does attendance to fellow bands' performances as well as to cultural festivals in the area.

The positive effect of associational life does not seem to extend, tough, to associations with a more political edge like human rights groups, for which the correlation is actually negative. Interestingly, introducing the new batch of variables results in social status playing a role, but not as expected it is actually a condition of artisan or skilled worker that facilitates broader musical interests.

The next step of our analysis requires identifying subsets of genres that band members tend to perceive as similar. A factor analysis (reported in Table 3) points at four sets. The first brings together the brass band repertoire with classical music on one hand, and different kinds of popular music on the other. The second consists of emerging, mostly youth-based, niche genres from rap to electronic music. The third reproduce the classic association of jazz and blues, and the fourth that between pop, rock, and songwriters' (cantautori) repertoire. We can see these subsets as "musical families", and concentrate on that particular form of omnivorousness that bridges different families.

Table 3: *Factor analysis of musical genres.*

	1	2	3	4
Brass band	.617			
Pop				.809
Rock				.692
Jazz			.890	
Blues			.853	
Folk	.661			
Songwriters				.635
Liscio	.629			
Classica	.583			
World	.509			
Rap		.845		
Trap		.872		
Electronic		.658		

Table 4 below reports the results of two stepwise regressions of the same variables used in Table 2, yet over a different dependent variable, namely, number of bridges across musical families. First of all, socio-demographic variables count for nothing here, barring a very modest, positive contribution of age (model 1). Model 2 performs better, yet again much less than when it was fitted to the simple additive measure of omnivorousness. A more limited number of associational types turns out to be positively correlated with bridging tastes, while there is a slightly more differentiated set of events that play a role, such as local festivals or attendance to museums/exhibitions. Most remarkable however is that, at this level, high cultural capital seems to matter more. Not only conservatory degrees at both level, but also university degrees appear to have a positive effect. This does not extend to class though.

Table 4: *Ordinal logistic regression of band members' traits on number of distant genres bridged (s.d. in brackets; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$).*

	Model 1	Model 2
Socio-demographic traits		
Age (ln)	0.61*** (0.15)	0.87*** (0.22)
Secondary school diploma	N.S.	1.32** (0.49)
University bachelor	N.S.	1.21* (0.51)
University master	N.S.	1.01* (0.51)
Conservatory of music I	N.S.	1.54** (0.59)
Conservatory of music II	N.S.	1.73** (0.66)
Associational memberships		
Cultural circles		0.55** (0.18)
Other musical ensembles		0.59*** (0.16)
Sport clubs		0.69** (0.18)
Attendance of events		
Local festivals (occasional)		0.93* (0.43)
Cultural festivals (occasional)		1.10** (0.46)
Cultural festivals (regular)		0.74* (0.31)
Museums/exhibitions (reg.)		0.75* (0.35)
N	749	748
Prob > chi^2	0.00	0.00
Pseudo R^2	0.01	0.05

3 Conclusions

Regardless of how we measure omnivorousness, measures of socio-economic status play a much more limited role in accounting for it than other variables, reflecting individuals' involvement in civic and cultural life. Why is this so? Our analysis is exploratory and its data do not allow us to draw strong conclusions on this point. What they do suggest, however, is that participation in musical associations goes along with intense involvement in a range of social and cultural activities, and such breadth of engagement also fosters broader musical tastes. Combined with the fact that about 63% of respondents suggested that their band had enhanced their personal creativity, while 78% thought it had expanded their musical culture, this suggests a massive democratizing role for musical associations in levelling social differences. This is particularly evident when we note the positive effect of being an artisan/skilled manual worker when data on participation are brought into the model accounting for omnivorousness. At the same time, we must also note that education, an important measure of social stratification, plays a more significant role when accounting for the bridging more diverse musical genres. This points at the persistence of some role of status differences even within inclusive, cross-class associations like musical brass bands.

4 Acknowledgments

Data for this work come from a survey of members of brass bands in the province of Trento, Italy, conducted in 2021 in collaboration with the Federazione dei Corpi Bandistici del Trentino [Sacchetti & Diani, 2022].

References

[Bauman, 1987] Bauman, Z. (1987). Legislators and interpreters: on modernity, post-modernity.

[Borzaga et al., 2021] Borzaga, C., Sacchetti, S., et al. (2021). Inclusive organizations as agents of democracy. In *Social justice in a global society. Toward new forms of social innovation and economic democracy for a sustainable development* (pp. 227–255). Feltrinelli.

[Bourdieu, 1979] Bourdieu, P. (1979). La distinzione. critica del gusto. *Bologna: Il Mulino*.

[Hazir, 2021] Hazir, I. K. (2021). Cultural omnivorousness. In *Oxford Bibliographies*.

[Lizardo, 2014] Lizardo, O. (2014). Omnivorousness as the bridging of cultural holes: A measurement strategy. *Theory and Society*, 43, 395–419.

[Ostrom, 1990] Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge university press.

[Pachucki & Breiger, 2010] Pachucki, M. A. & Breiger, R. L. (2010). Cultural holes: Beyond relationality in social networks and culture. *Annual review of sociology*, 36(1), 205–224.

[Peterson, 1992] Peterson, R. A. (1992). Understanding audience segmentation: From elite and mass to omnivore and univore. *Poetics*, 21(4), 243–258.

[Rao et al., 2000] Rao, H., Morrill, C., & Zald, M. N. (2000). Power plays: How social movements and collective action create new organizational forms. *Research in organizational behavior*, 22, 237–281.

[Sacchetti & Diani, 2022] Sacchetti, S. & Diani, M. (2022). Il senso di fare banda: Le bande musicali all'interno della comunità trentina (the meaning of playing in a band: Marching bands within the trentino community).

[Vlegels & Lievens, 2017] Vlegels, J. & Lievens, J. (2017). Music classification, genres, and taste patterns: A ground-up network analysis on the clustering of artist preferences. *Poetics*, 60, 76–89.

Contamination in Theatre: from the Original Play to the Cooperative Creative Contributions on Stage

Marta Nicolodi[1] and Silvia Sacchetti[2]
[1] *University of Milan, Italy*
[1] *University of Trento, Italy*

Abstract

Theatre problematizes the distinction between work of art and speech in action: dramatic works attain their ideal expression in performance, while the printed form is often seen as secondary. This scenario makes theatre an interesting field for discussing concepts of contamination in art. The aim of this work is to explore the organizational structure of theatre by identifying the relevant stakeholders involved, in order to highlight the various forms of contamination that can exist within this artistic practice. Additionally, the discussion attributes the reasons for this facilitated interchangeability of ideas to the cooperative environment of theatre organizations, which can help creativity. To better understand the features of this organizational setting, we will also refer to contributions from economic theories that analyse horizontal arrangements in firms. These theoretical insights will aid in examining the nature of contamination in our field and discuss its possible limitations.

1 Introduction to the Idea of Contamination in Drama Theatre

Performing arts and cultural activities are based on collective behaviours that reflect the desire for shared experiences or collective production and consumptions which transcends those of the individual participants involved [Throsby, 2001]. For a work of art to appear in its final form, many preliminary activities must be carried out, involving the presence of multiple stakeholders, including not only the artists but also other individuals referred to as "support personnel" [Becker, 1974]. This complexity becomes particularly evident when examining the staging of a theatrical performance. The necessary activities typically include the presence of a playwright who conceives the idea for the work and writes the script of the play, the casting of suitable actors for the staging of the performance, the creative process on stage related to the creation of the performance and the presence of an audience.

We focus our paper on contaminations in theatre because we believe that this form of art, occurring live on stage and usually involving a collaborative process among various artists, lends itself to being a good area for "contaminations". Even if an actor were alone on stage and able to organize everything independently from supporting staff, there would still be the relationship with the audience, which can also contaminate the artists' work on stage (consider comedies, for example, where the audience laughter creates a positive emotional shift in the actor which can result in a change in performance outcomes).

2 The Organizational Stages of a Theatrical Production Process and the Stakeholders Value Chain

To frame in more technical terms how these contaminations occur within the theatrical sector, we first identify the value chain of theatre, presenting the organizational stages of a drama production and the various stakeholders involved.

The artistic process always starts from a written text (i.e. the script) which cannot be reduced to its literary dimension, as it is conceived with the aim of being performed. The idea of collaboration and artistic contamination in theatre originates precisely from this distinction between the work of art and the speech in action. Drama scripts are, therefore, brought to life on stage by diverse parties – the director, the actor, the producer – other than the playwright. Additionally, in recent times they are often authored via a collaborative process, which decouples the author from the resultant text. There are still cases where the author did not allow their texts to be modified by the actors on stage [McDonagh, 2014, McDonagh, 2021]. However, as theatre is one of the art forms that most embodies humanity, there will inevitably be a personal interpretation of the work reflected in the actors' attitudes on stage. In other terms, simply because the performance is executed by someone other than the original author, elements of contamination to the work are introduced. Even when strictly adhering to the director's or playwright's instructions, actors inevitably provide a personal and emotional interpretation of what was solely conceived as text.

Another relevant stakeholder is the producer, who needs to obtain the playwright's permission to stage their play. This occurs through the acquisition of copyrights [Battelli et al., 2020], where the producer is committed to remunerating the author as well as to representing the work without making changes, in accordance with the author's instructions (Italian Copyright Art, Article 138). The producer and the playwright will need to discuss which parts of the play cannot be modified by the stakeholders involved in the production. To allow the staging of the play, the playwright may need to agree to a compromise, allowing a first "contamination" of their work in this contractual phase. The reasons for these changes can vary and are sometimes purely economic (e.g., the theatre may lack the budget to stage certain aspects of the original work, such as reducing the number of characters or the scenography). In other cases, the playwright is also part of the production, working as director or even as actor, and may be more flexible in adapting their text throughout the performance creation process.

Then there is the director who is responsible for the scenic transposition of the text, guiding the actors and managing the artistic and creative elements on stage. This stakeholder provides a significant creative input to the final performance, with the possibility to influence and direct the authors according to personal taste – although some restrictions may be imposed by the producer, due, for example, to budget constraints, or by the playwright, as discussed earlier. Given the director's crucial role in determining the outcome of a theatrical performance, the question has also arisen in literature regarding whether, in certain cases, directors should be considered co-authors of the play [McDonagh, 2014, Miccichè, 2006]

Subsequently, actors are selected to form the cast, based on the director's instructions. While the work of the actors is undoubtedly central to the creation and execution of the play, it would not be possible without the support of technical professionals, such as the technical director, stage director, light designer, scenographer, and others. These professionals also contribute to the theatrical production process, offering perhaps less visible but still significant artistic input. Their unique expertise can introduce elements of contamination, enriching the creative process.

Until now, we have referred only to stakeholders involved in the artistic part of theatre production. To provide a complete picture of the organizational complexity behind staging a play, it is worth mentioning that, depending on the size of the theatre, administrative staff also play a role. Although their work is generally separate from the artistic activities, can still "contaminate" production choices, pushing the creative team to make adjustments that could, for example, enhance ticket sales or reduce costs for specific performances.

Overall, what allows the contamination of the original play is the fact that its production is subject to the well-known principle of division of labour, where each stakeholder contributes specialized and complementary skills. The need for interaction in a coordinated way arises precisely because of specialisation. This could not happen without shared expectations about cooperative attitudes. Theatre is, therefore, rooted in collaboration and trust, which enables a continuous and fluid exchange of ideas and contributions.

3 Flexible, Cooperative and Team-based Organizing Structures which Favour the Development of Creativity and Contamination

If specialisation requires coordination and trust among participants in a theatre production, we can now ask how the theatre's working environment facilitates the contamination of the original work of art by other stakeholders. Unlike the division of labour in the factory system, where creativity is traded off for efficiency, theatre is not dominated by strict rules, nor does it typically adhere to a rigid hierarchical structure. This flexible and fluid atmosphere favours creativity to be easily developed. Theatrical productions have evolved into complex organizing structures. Still, they have maintained the governance features, based on teamwork and collaboration, that historically characterized them since their early development in the form of Commedia dell'Arte [Gallina, 2007, Gallina, 2014]. The ease in contamination processes in theatre can be traced to this initial organizational form, where all actors were working on a level playing field, without an underlying script or hierarchical leader, except from the "Capo Comico", a fellow performer among peers, who simply coordinated the group of actors rather than directing them.

To better frame the nature of this organizational context, creative processes in theatre can also be interpreted in the light of the contributions of some management scholars who have mainly discussed the extent to which a certain work environment may favour the development of creativity [Folk, 2023, Fischer & Vassen, 2011]. In their work on establishing the organizational conditions for creativity, Amabile et al. [Amabile et al., 1996] propose five dimensions of the work environment, which contribute to the capacity of individuals and groups to be creative: (i) organizations must encourage creativity, through policies and reward systems, through the role of supervisors as nurturers of the creative spirit, and in the way that work groups function and value creative effort; (ii) the organizational environment needs to provide individuals and work teams with "autonomy and freedom" and "a sense of ownership and control over their own work and their own ideas"; (iii) individuals and teams need to be provided with adequate resources to engage in creativity; (iv) individuals and teams need to be provided with adequate time for such activities; (v) organizations need to consider the nature of internal impediments to individuals and groups being creative in their work (these impediments are said to come in the form of "internal strife, conservatism, and rigid, formal management structures").

However, illimited freedom may also result in dissonant creativity [Dennis & Macaulay, 2003], especially in circumstances where the original input of a single individual is in contrast with the capability of other agents to exert their own creativity. In these cases, some shared rules on how to harmonise or coordinate the different inclinations can be useful. Fillis and McAuley [Fillis & McAuley, 2000] discussed measuring and controlling creativity, and Levitt [Levitt, 2002] emphasized the importance of discipline over the "essential primacy of the creative impulse". We take these considerations into account to highlight the relational element of creativity – that is, the extent to which the interaction between individuals can provide a productive impulse that benefits both the person and others, or, conversely, becomes a limit. Previous evidence on musicians' collaborations indicates that these emerge among people who share a common understanding and are willing to contribute to an artistic project (a music production in this case) [Sacchetti, 2023a]. The idea of creativity, which favours the circulation of ideas and contamination, can be seen with positive connotations. According to Abraham Maslow's psychology, it can lead to individual's self-actualization, which, in his hierarchy of needs, is located at the highest level of psychological development. He defined self-actualization as the desire for "self-fulfilment, namely, the tendency for him [the individual] to become actualized in what he is potentially. This tendency might be phrased as the desire to become more and more what one is, to become everything that one is capable of becoming" [Maslow, 1943]. This idea goes beyond the focus on bettering the self and it is understood as comprising others and selfless needs because, in order to self-actualising, there is always a need to move from self to others [Sacchetti, 2023b, Greene & Burke, 2007].

As modern theatres evolved into organizational forms with complex characteristics similar to those of firms, we can reconstruct the features of their cooperative and horizontal relations and derive the benefits of these types of interactions within an organizational process, referring to Margaret Blair and Lynn Stout's "Team Production Theory of Corporate Law" [Blair & Stout, 2017]. According to these scholars, a public corporation can be viewed as a team of stakeholders, including, in addition to shareholders, also creditors, workers, managers, and communities, who undertake a team production project that requires all to make some form of enterprise-specific investment. This theory recognizes the value of "horizontal interactions among team members", affirming that "if the activities and inputs of those participants are adequately coordinated, their collective output can be qualitatively different and vastly larger than the sum of what each individual could produce separately" (ibid).

4 Limits to Contamination and Conclusive Remarks

The process of perturbating the artwork by external agents can be considered intrinsic to the nature of theatre, where the performance needs to be constructed by a variety of stakeholders which bring a diversity of ideas and experiences to the table. Since merging different perspectives and ideas can inspire innovative interpretations and approaches to a play, contamination acquires positive connotation in this context. This can lead to fresh narratives and interpretations that reflect a broader spectrum of cultural and social perspectives, enriching the theatrical landscape, cultural diversity and circulation of ideas.

At the same time, when multiple stakeholders contribute their ideas, disagreements or conflicts about the artistic direction of the production may arise. Therefore, it is important to set boundaries to contamination, otherwise what initially carries positive connotations may ultimately develop negative implications. Organizational mechanisms for coordinating resources and making decisions, such as market dynamics and hierarchical structures, can help establish these boundaries without excessively curbing the positive aspects of contamination. When applied within certain limits, in fact, these mechanisms can also foster the development of the artistic creativity [Sacchetti, 2023b].

The market mechanism, for instance, can help theatre companies to pay attention to audience preferences and feedback, allowing them to filter out ideas that might fragment the original artistic vision due to the contaminative action of an excessive number of stakeholders. Conversely, the director, who holds a position of authority over the team of actors, can operate as a guide in the creative process, helping to maintain a coherent vision for the production and resolve disagreements among team members.

The issue of contamination in theatre may be even more complex than what has been discussed in this paper and requires further investigation. We are conducting empirical research through interviews with key stakeholders in the theatre sector to gain deeper insight into the research matter. Additionally, exploring interdisciplinary perspectives - drawing from psychology, law, economics, and theatre studies - could be beneficial in providing a more comprehensive understanding of the dynamics at stake.

References

[Amabile et al., 1996] Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the work environment for creativity. *Academy of management journal*, 39(5), 1154–1184.

[Battelli et al., 2020] Battelli, E., Rajani, G., Rossi, V., Dodaro, P., Incutti, E. M., & Ruggeri, F. (2020). *Diritto privato dello Spettacolo*. Giappichelli.

[Becker, 1974] Becker, H. S. (1974). Art as collective action. *American sociological review*, (pp. 767–776).

[Blair & Stout, 2017] Blair, M. M. & Stout, L. A. (2017). A team production theory of corporate law. In *Corporate governance* (pp. 169–250). Gower.

[Dennis & Macaulay, 2003] Dennis, N. & Macaulay, M. (2003). Jazz and marketing planning. *Journal of Strategic Marketing*, 11(3), 177–185.

[Fillis & McAuley, 2000] Fillis, I. & McAuley, A. (2000). Modeling and measuring creativity at the interface. *Journal of Marketing Theory and Practice*, 8(2), 8–17.

[Fischer & Vassen, 2011] Fischer, G. & Vassen, F. (2011). *Collective creativity: Collaborative work in the sciences, literature and the arts*, volume 148. Rodopi.

[Folk, 2023] Folk, L. S. (2023). *The Hidden Barriers and Enablers of Team-Based Ideation*. Springer.

[Gallina, 2007] Gallina, M. (2007). *Organizzare teatro. Produzione, distribuzione, gestione nel sistema italiano*, volume 7. FrancoAngeli.

[Gallina, 2014] Gallina, M. (2014). *Ri-Organizzare teatro. Produzione, distribuzione, gestione*. Milano: FrancoAngeli, 1st edition.

[Greene & Burke, 2007] Greene, L. & Burke, G. (2007). Beyond self-actualisation. *Journal of Health and Human Services Administration*, 30(2), 116–128.

[Levitt, 2002] Levitt, T. (2002). Creativity is not enough. *Harvard Business Review*, 80(8), 137–144.

[Maslow, 1943] Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50, 370–396.

[McDonagh, 2014] McDonagh, L. (2014). Plays, performances and power struggles - examining copyright's 'integrity' in the field of theatre. *Modern Law Review*, 77(4), 533–562.

[McDonagh, 2021] McDonagh, L. (2021). *Performing copyright: law, theatre, and authorship*. London: Hart Publishing, 1st edition.

[Miccichè, 2006] Miccichè, A. (2006). *Legislazione dello spettacolo. Cinema musica teatro*. Roma: Artemide.

[Sacchetti, 2023a] Sacchetti, S. (2023a). *La produzione di cultura musicale che persiste e il sistema delle scuole di musica del Trentino. La vitalità di un sistema di istruzione pubblico-privato*. Rapporto di ricerca, Dipartimento di Sociologia e Ricerca Sociale, Università di Trento.

[Sacchetti, 2023b] Sacchetti, S. (2023b). What can economic coordination do for creativity and well-being? *Creativity and Innovation Management*, 32(3), 378–387.

[Throsby, 2001] Throsby, D. (2001). *Economics and Culture*. Cambridge: Cambridge University Press.

Collaborative Networks in the TV Fiction Industry: Italian Serials and Sitcoms before the Digital Switch Off (1996-2009)

Francesca Odella
University of Trento, Italy

Abstract

In this short paper I propose to analyze the transformations in the Italian TV fiction production industry adopting the sociological perspective of the production of culture approach (Di Maggio, 1977), and specifically focus on the dispute between autonomy and control in the production of popular culture.

First, I will be looking at the relations between the political and regulatory policy of Italian tv broadcasting and the evolution of audiovisual fiction productions in the decade anticipating the so-called 'digital switch- off' (1996-2009) and the proliferation of TV channels. According to Martina and Palmieri [Martina & Palmieri, 2015], during this decade Italian TV fiction 'became the field on which the huge battle, or rather imbalance, between import and export was played out in favor of the latter'; the same battle, however, paved the way for renoval of Italian serial fiction productions before the emergence of digital television offer.

Second, applying the methodology of social network analysis (SNA) to map the space of the Italian television industry, I will show how television broadcasters re-allocated resources in new areas and launched collaborative programs with other national and international broadcasters, developing new organizational forms for managing cultural prominence. Described in organizational studies as constellations, these collaborations under specific institutional and market conditions have proved to incentivize innovation in the diverse cultural sectors.

1 Introduction

In cultural and social sciences the attention of researchers for the historical evolution of the audiovisual television production industry has been constant as there is a unanimous perception that TV fictions and related publics have become important components of the global digital culture. As highlighted by Bourdon [Bourdon, 2011] television is still essentially a medium of national communication and national cases provide an interpretation of the complex evolution in cultural domains and cultural worlds. Tv fictions, specifically, convey and express contents relevant for understanding the structure and reproduction of niche cultures and for generating complex narrative worlds [Zanatta, 2011].

Since the nineties, the European national television systems, in particular, changed their structure and underwent deep transformations in their assets and roles [Franquet et al., 2020]. Privatization of broadcasting services and product segmentation of television fiction products were two most visible effects, while on the background aspects concerning the relationships of the media sector with political and financial systems were becoming evident and subject to public debate. The European Commission in the same years contributed both to innovation and reform of the media sector (the first Creative MEDIA program was in 1991), as well as to merging national television audiences, promoting european circulation of films and joint TV fiction products.

Compared to extensive studies carried on in the American fiction industry - see among them Bielby [Bielby and Bielby, 2009], the Italian case represents a topic rather unexplored; the period situated between the crisis of generalist television (early nineties) and the advent of digital media television contents (the first digital switch-off in Italy was in 2009), in particular, is the most debated and obscure, consisting in the years when Berlusconi dominated both as Italy prime minister and broadcasting media tycoon. The sources of broadcasting data about this period are also scarce because EAO (European Audiovisual Observatory) monitoring started only in 2015.

Among the few sources about the Italian case was the OFI (Osservatorio Italiano sulla Fiction, [Buonanno, 2017]), a private university based foundation that from 1988 to 2000 published yearly guidebooks of Italian audiovisual products broadcasted on the two main TV services (Rai and Mediaset). According to OFI data presented in Figure 1 the quantitative evolution of audiovisual production in Italy (broadcasted hours per season/year) from 1989 to 2008 was marked by a significant decline in production at the end of the eighties, followed by a stable and constant rise during the nineties. In 2009 television series 'made in Italy' reached 40% of broadcasted products in public tv and only in the second half of the 2010's with digital television and pay per service channels the non-domestic fiction became more popular than the Italian one (45% in 2015).

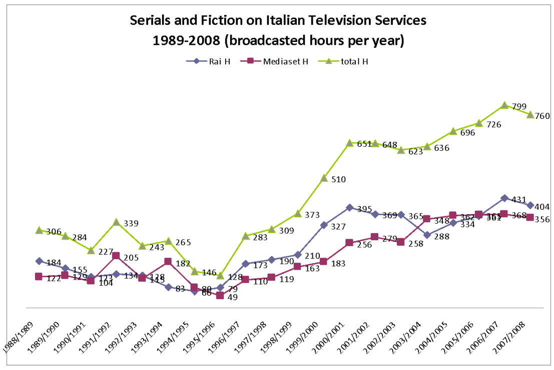


Figure 1: *TV production in Italy (1989-2008).*

However, considering the number of fiction producers for the two main broadcasting services (Rai and Mediaset), a sharp difference is evident between a pre-crisis (up to 1996) and post-crisis period (from 1997 to 2008).

The crisis and the subsequent rise in broadcasted fiction, in particular, had an effect on the industry organization. Specifically, it enlarged the core number of independent and external producers, multiplying the number of companies, of different size and market share and the average number of producers working seasonally with Mediaset and RaiTV channels rose 40-45% from the pre-crisis period. Large and medium Tv production companies that appeared on the market between 1996 to 2000, in particular, were the first ones to embark in the design of original tv fiction products and to develop narrative and organizational standards according to 'modern' tv seriality, adopting new forms of collaboration and partnership with the main TV broadcasters (Rai or Mediaset).

These new organizational forms can be described by means of the concept of constellation. Constellations are organizational alliances with multiple and different partners; among their characteristics are the use of different specialists and services to cope with short term collaborative projects and the employ of multiparty complementary requisites of production. Graphically, constellations can be represented as social networks generated by the practice of relations over time in which relational power is continuously reproduced and generated by dynamic exchanges of relations and resources among actors.

2 Constellations in TV Production

Media industry and in particular the cinematographic industry have been assimilated to constellations and studies carried on early and contemporary cinematographic industry or other innovative sectors such as biotechnology industry put in evidence the pro-innovation aptitude of these collaborative structures [Jones, 2001; Cattani and Ferriani, 2008]. Tv movies are produced under network-based control and the more certain actors are in the 'kernel' of these networks the more powerful, and at the same time dependent, they become on the reproduction and prolongation of already established social (power) relations. They establish themselves as 'critical actor constellations', as described by [Manning & Sydow, 2007] and decide the timing of innovation (or resistance to innovation) until external factors (institutional actors or environmental drivers) modify the context of their relations and the balance of power among the partners. So, how did these mechanisms work in the Italian context? How innovation and control were balanced in the long-term panorama of Italian TV fiction production?

2.1 The Impact of Italian TV Duopoly

First, the duopoly of the Italian television system favored preferential relations with a limited number of 'critical actors', functional to set the seasonal TV fiction panorama according to specific cultural contents and formats. In long-term fiction scheduling the choice of a specific show of any type of fiction (soap, serial, tv film,..) is in fact indirectly associated with its main characteristics (format, subject, etc.) and its (multiple) contributors (screenwriters, directors, etc.). Because the production of fiction and broadcasting on TV are tightly connected, it is rare that a product is rescheduled on a channel different from the one that originally 'commissioned' it and even less probable that the product is 'resold' to competitors. As illustrated in the Figure 2 below, the large/top producers - the green dots - are more interconnected and distribute their activity across multiple years of TV scheduling, while other smaller producers - the red dots - have marginal position and depend on seasonal collaboration with the core players.

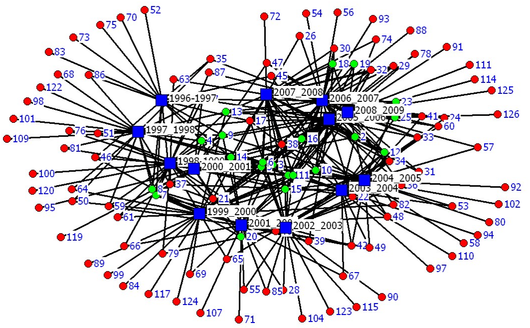


Figure 2: *2-mode network of relations between TV fiction producers (1996-2009).*

The constant confrontation in a institutionalized duopoly - Rai and Mediaset - prevented competition, as well as alliances among middle and small fiction producers, and resulted in a small number of big and stable producers (often financially associated with the two main TV broadcasting companies) involved in the production of consequential tv seasons and in a periphery of small and uncertain firms that operated on a one shot logic.

2.2 The Role of Collaborative Alliances

Second, the innovative component in the system was introduced by collaborative alliances and partnership at the national and international levels. During the so-called period of consolidation of Italian fiction production (1999-2007), in particular, italian companies started to invest in new products and seek partnership and innovative formats [Cardini, 2013]. Of a total of 594 fictions produced by Italian companies in the period starting from tv season 1999/2000 and ending with season 2005/2007, only 51 fictions were developed in partnership among Italian companies, while 119 were produced in collaboration with at least one international media company. Initially (1999-2001) the collaborative products with more than one foreign partner represented 35% of the total; in a second moment (2002 to 2004) the proportion of international projects involving companies based in Europe (above all Germany, France, and Great Britain) or other international context (United States, Canada, and Latin America) rose to 69% of the shows co-produced, then stabilizing to 54%. In the latest period (2008-2009) the collaborative projects diminished and international co-productions stabilized around 40% of the total.

Finally, the proceedings of EU governance in the media sectors accelerated this process aiding the largest media companies, and those already working globally in fulfilling their marketing strategies. The top 6 Italian media companies are in fact involved in 45% of all the collaborative projects and 65% of those involving an international partnership.

3 The Emergence of a New System

To test the aforementioned mechanisms I applied a SNA analysis [Borgatti & Everett, 1997] to the networks of collaborative projects among producers (2-mode network matrix814 nodes/ audiovisual products by 238 nodes/ companies) and according to literature suggestions repeated the analysis on the whole set of data as well as on the sub-set of collaborative productions, looking for specific differences in the network structures.

The longitudinal analysis of the structure of the networks of domestic and international fiction collaborations, showed the high fragmentation of this cultural industry, exemplified by the presence of a distinct core-periphery structure (blockmodelling density matrices with 0,63 and 0,16). Results showed, in particular, that convergence of the two functional mechanisms generated constellations across three phases of growth (1996 to 1999), consolidation (2000 to 2005) and institutionalization (2006 to 2009) of the Italian fiction industry. Each phase represented a step towards a new system of production and distribution of media products, highlighting how the constellations of Italian producers were focused in pursuing individual firms' advantage and innovation when dealing with products designed for attracting new public and the international market.

To exemplify this last statement, I will rely on the results of longitudinal analysis of the OFI dataset, recoded as sequenced relational data to perform Multiple Correspondence Analysis (MCA). The aim of this analysis was to map the progressive change of the TV fiction system in terms of both types of products and organization of production. The bi-dimentionsl cultural space in Figure 3 represents cultural and organizational space of tv fiction in a probabilistic perspective [Roberts, 2020]; each deviation from the two axes representing sets of TV products and sets of TV producers signify a change by direct or indirect consequence of variations in the reciprocal distances between these sets during the five time sequences.

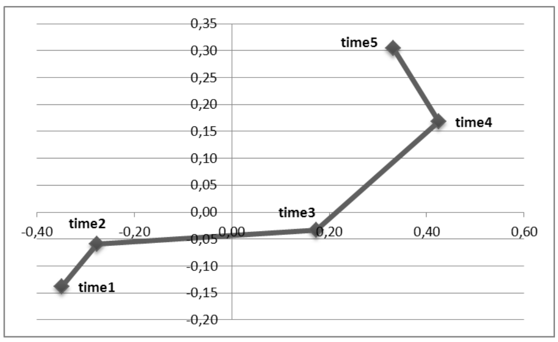


Figure 3: *Plotting of the changes in the cultural space of Tv fiction production (1996-2009).*

While the overall structure of the TV-production industry presents mainly a core-periphery network structure (global tendency), the MCA plot shows that when fragmentation of the TV production sector increased (see the interval between time3 to time5) and the production for foreign markets privileged distinctive genre niches, such as historical melodrama, significant variations in product orientation (genre and type of fiction) and in the forms of collaboration among producers (local effect) emerged. The final variation in the mapping (time4 to time5) finally signifies the presence of new production strategies linking Italian companies and the advent of alliances between the TV production sector and other media companies (internet/multimedia entertainments services).

Concluding, this analysis of network constellations generated by Italian collaborative TV projects shows how re-constructing and documenting national cases of the TV fiction production sector can raise our knowledge about the mechanisms that operate in a specific cultural production domain [Ibert & Muller, 2015]. It also puts in evidence the organizational factors that favor or reduce innovativeness in the media production sector and how the same factors can reduce their impact due to institutional changes (the tv digital switch off, 2009) or unanticipated socio-economic effects (collaboration of media producers and the first European Creative Media program, 1991).

References

[Borgatti & Everett, 1997] Borgatti, S. & Everett, M. (1997). Network analysis of 2-mode data. *Social Networks*, 1, 243–269.

[Bourdon, 2011] Bourdon, J. (2011). *Du service public à la télé-réalité. Une histoire culturelle des télévisions européennes 1950-2010*. Bry-sur-Marne: INA Éditions.

[Buonanno, 2017] Buonanno, M. (2017). *Primo rapporto sulla fiction televisiva in Europa*. Roma: RAI-Eri. Originally published in 1999.

[Cardini, 2013] Cardini, D. (2013). *Serial tv. Come la televisione ha sconfitto il cinema*. Laterza.

[Franquet et al., 2020] Franquet, R., Richieri, G., & Hibberd, M. (2020). *The introduction of commercial broadcasting in Europe*. Whiley.

[Ibert & Muller, 2015] Ibert, O. & Muller, F. (2015). Network dynamics in constellations of cultural differences. *Research Policy*, 44(1), 181–194.

[Manning & Sydow, 2007] Manning, S. & Sydow, J. (2007). Transforming creative potential in project networks: How tv movies are produced under network-based control. *Critical Sociology*, 33(1-2), 19–42.

[Martina & Palmieri, 2015] Martina, M. & Palmieri, A. (2015). Researching television serial narratives in italy. *International Journal of TV Serial Narratives*, 1(2), 89–102.

[Roberts, 2020] Roberts, J. (2020). Correspondence analysis of two mode network data. *Social Networks*, 22(1), 65–72.

[Zanatta, 2011] Zanatta, S. (2011). Famiglie creative. reti professionali nel campo della serialità televisiva. *Etnografia e ricerca qualitativa*, 3(3), 339–365.

Paradoxical Contaminations in Human Development: Immaterial Goods and Cryptocurrency Trusts

Silvia Sacchetti[1] and Andrea Salustri[2]
[1] *University of Trento, Trento, Italy*
[2] *Sapienza University, Rome, Italy*

Abstract

This paper explores and raises questions about the role of diverse forms of economic coordination for human development (which accounts for both material and immaterial human needs). Building on theoretical and empirical findings about the role of relational goods for life satisfaction, we compare the nature and effects of cooperation in the social and solidarity economy with cooperation designed within blockchain technologies. We argue that, despite both economic coordination modalities are argued to root in cooperation, only the first mode of organising production merges material and immaterial needs, focusing on mutual help, deliberative practice, use of creativity and solidarity. Blockchain technology applied to cryptocurrency trust, differently, focuses on the creation of material wealth but crowds out immaterial goods. It follows that cryptocurrency trusts feed wealth on the one hand but subtract from human development on the other. The wealth creation effect can be hence off-balanced by the loss of immaterial goods. We conclude that contaminating Ideas of wellbeing with coordination solutions that mostly ground on material wealth and extractive practices can lead to paradoxical outcomes in our economies.

1 Three Dimensions of Human Development

The ways in which the digital transition is still taking place fuels a context in which inequalities and poverties are persistent, both with reference to material and immaterial wellbeing. This is a self-sustaining phenomenon, as the loss of relationships and subjective wellbeing are often the precondition for dynamics that reinforce poverty and exclusion. When we speak of inequality and poverty, we refer not only to income but, more generally, to an inequality of capabilities and opportunities to enact self-determined actions [Sen, 1992], aimed at both private and public happiness or at personal eudemonistic fulfilment. This implies that decisions on how to organise production and work does not aim solely at economic growth but include other dimensions of human development. Following the civil economy paradigm, these dimensions are identified in the pursuit of: a) the common good, b) psychological or spiritual needs related to the need that people have to transcend themselves, c) relational goods [Zamagni, 2017]. The common good is produced if the wellbeing of all those involved and affected by economic decisions is taken into account. It follows that - for this to happen - publics must be included in the identification of economic aims and issues, and be in a position where their contribution and creativity is valued. Interactions are driven by a common expectation around human nature and behaviours, whereby people do not come from Hobbes' wolfy world, but from the "civitas" where people have pro-social attitudes, can trust each other, cooperate, and enjoy proximity with each other, transcending their self. Relational goods come from such proximity and, differently from material goods, cannot be enjoyed alone or exchanged in the market [Gui, 2000, Uhlaner, 1989, Zamagni, 2017]. They reduce isolation, a major source of unhappiness in our growing economies [Bruni & Stanca, 2008]. Because people expect to be able to trust each other, economic coordination solutions are based on cooperation. However, cooperation might act at different Intensities, and we can position modes of economic production on a continuum between two opposites: on the one hand, deep and capacitating cooperation, and on the other hand, what is often referred to as cooperation but which instead can present the characteristics of a superficial and extractive exchange [Thompson, 2015].

2 Trading-off Immaterial Goods and Value Extraction

When economic coordination does not Include a consideration of immaterial needs, the goals and outcomes achieved create negative effects for people and communities, namely social and environmental costs. Sacchetti and Borzaga [Sacchetti & Borzaga, 2021] notice that the level of surplus (π) appropriated by the club that controls the economic decisions is, in these cases, greater than the output produced by the economic activity (Y) by an amount equal to the total social costs (CS'). Moreover, the output of economic activity Y is less than the socially desired levels (W) in terms of the quality or quantity of the service provided by an amount equal to CS". This is an extractive outcome, attributable to the costs generated by the failure of economic coordination to account for wider needs than wealth for the controlling group. The problem of value extraction from people and communities can be formulated, therefore, in this way:

$$\pi \equiv Y + CS' \tag{1}$$

$$Y + CS'' \equiv W \tag{2}$$

The extracted value (for instance the erosion of existing relational goods) thus turns out to be approximated by the total social costs (CS'), while the subtracted value (the failure to generate increasing levels of relational goods, if valued by society) is equal to CS". Through cooperative modes of coordination, following the example of social economy organizations, it is possible to interrupt the extractive dynamic triggered by exclusion, activating, instead, a process of value creation [Sacchetti & Borzaga, 2021]. In contrast, when immaterial needs are also taken into account, the overall output of an economy (Y) includes the production of value BS (relational goods and life satisfaction for example) in excess of what can be appropriated by producers. The output of economic activity is hence closer to the level of wellbeing socially desired.

$$Y \equiv \pi + BS \tag{3}$$

$$W + BS'' \equiv Y \tag{4}$$

It is useful here to make explicit the underlying assumption that, in ordinary situations, extracted value (in terms of wealth for example) is appropriated by small groups that increase consumption of positional goods (goods attached to status and authority), exacerbating cumulative causation (CS" > CS'). In special situations e.g., when people in poverty increase their material wealth using cryptocurrency, in the absence of other opportunities, value extraction (e.g., in terms of environmental costs) creates benefits for marginalized or excluded groups by enabling the consumption of relational goods. The value extracted by the marginalized category might be less than the value subtracted (CS" < CS'). Similarly, when value creation (e.g., in terms of relational goods) is aimed at greater consumption of positional goods, BS" may be lower than BS'.

3 The Social and Solidarity Economy

The three dimensions of human development are central in the social and solidarity economy (SSE) but can be applied more broadly when considering the aims and effects of economic choices, whether these are operated by the public administration, by profit-maximising firms or by organisations of the SSE. The latter is "a path of transformative and systemic change" in which "ordinary people play an active role in defining all dimensions of human life: economic, social, cultural, political and environmental" [RIPESS, 2015]. It emphasises the multidimensional and multistakeholder approach to transformative change based on poverty eradication and reduction of multidimensional inequalities, emphasizing pluralism and grassroots innovation that promotes human flourishing and happiness. Specifically, SSE supports transformative action that "goes beyond a superficial change in which oppressive structures and fundamental issues remain intact" [RIPESS, 2015].

In particular, the SSE includes organisations and enterprises that: a) have explicit economic and social (and often environmental) objectives; b) involve varying degrees and forms of cooperative, associative and solidarity relations between workers, producers and consumers; c) practice workplace democracy and self-management. Furthermore, the United Nations Interagency Task-force on Social and Solidarity Economy (UNTFSSE) emphasises the variety of institutions that take part in the SSE, which includes, among others, "traditional forms of cooperatives and mutualist associations, as well as women's self-help groups, community forestry groups, social supply organisations or "neighbourhood services", fair trade organisations, informal sector workers' associations, social enterprises and community and alternative finance systems". These ways of organising production of services and goods aims at fostering trust, relational goods, immaterial needs [Mosca et al., 2007], and overall it aims at benefiting communities rather than restricted groups.

Through these structures, SSE promotes relationality and deep value-based cooperation [Thompson, 2015, Sacchetti & Borzaga, 2021] grounding in the eudemonistic concept of happiness. The conception of wellbeing that inspires SSE action is capacitating, advocating a radical shift toward a paradigm of solidarity. Underlying SSE is certainly an idea of the common good, with a view to broadening access to the widest possible public, at least with reference to the enjoyment of services of general interest. Finally, SSE have a countercyclical impact on the economy. A vast literature has recognized, for example, the role of cooperatives in promoting employment in times of crisis, precisely with regard to marginalized or excluded groups, as well as to unemployed workers.

However, SSE Is not the prevalent way of understanding economic coordination. While from a normative point of view SSE operates in ways that are decidedly preferable to cryptocurrency trusts, from a positive point of view the effectiveness and efficiency of SSE action, precisely because of the meritorious characteristics of the ends pursued, could be severely limited by the failure to achieve enough scale to offset the initial costs. Moreover, the maturation of social preferences, a necessary condition for appreciating the extended benefits offered by deep cooperation as opposed to shallow cooperation, might be hindered by the individualistic ethics underlying economic choices.

4 Cryptocrurrency Trusts

The rhetoric of trust, cooperation and wellbeing has been recently used also to emphasise the benefits of specific technologies, namely, to understand blockchain technology and cryptocurrency trusts. Bitcoin was introduced by Satoshi Nakamoto in 2008 using blockchain technology, which ensure trust between anonymous counterparties in decentralised systems, without the need for central supervisory authorities in charge of verifying the correctness of records in traceability systems [Tasca & Tessone, 2019]. This understanding is consistent with the current debate on "decentralised autonomous organisations" (DAOs). However, although DAOs are suggested to overturn the hierarchical mode of organising and reduce organisational costs (Wang et al. 2019):

- they focus on growth rather than human development;
- trust is in fact distrust, as it is dependent on technology, not relationality;
- they activate a mode of cooperation that is superficial and does not meet in any way immaterial needs;
- they extract value from other people and society in most cases, rather than contributing to the common good.

The underlying idea of cryptocurrency trusts is that of commodification, that is, a broader distribution of ownership and appropriation rights through the production of greater purchasing power. The emphasis is on the quality of life that can be achieved through greater monetary wealth. By their very nature, cryptocurrency trusts rest on an instrumental conception of wealth, which in fact coincides with personal enrichment and consumption of positional goods and status advantages, at most "tolerant" to the creation of enrichment opportunities for others. Greater consumption of material goods may substitute relational goods. These can be negatively Impacted also by the technological fiat that mediates interactions when using cryptocurrencies. The aims is to reduce transaction costs, functional to wealth creation. It follows that cryptocurrencies promote a superficial type of cooperation, which in fact rests on a hedonistic and exclusive conception of happiness. Moreover, the time spent in "mining" has opportunity costs in terms of relations.

Despite a guise of democracy, the underlying paradigm is one of inequality, as the returns that trust members can earn depend on the initial investment and the level of financial education possessed. Moreover, they implicitly promote self-seeking preferences in adherents, who are in fact the only potential beneficiaries of the value created under network mechanisms. They create unstable and highly precarious employment in the financial sphere. They are de facto labour substitutes, incentivizing resignation processes such as those identified in the post-pandemic period.

Finally, cryptocurrency trusts have an impact on the economy that is pro-cyclical in nature, as the returns they offer benefit from the financial instability and liquidity injections that characterize periods of crisis, generating significant opportunity costs compared to the research of shared solutions. Broadly speaking, they promote a type of financial activity that tends to capture value, even to the point of setting up real scams carried out under a Ponzi scheme.

5 Concluding Remarks: a Risky Contamination

From a human development perspective, the crowding-out effect of coordination solutions that promote financial or technology-driven cooperation leads to reflect on what seem to be quite crucial in current economies. Why are societies collectively adopting economic solutions that do not maximise collective wellbeing?

Some of the considerations that we have advanced indicate, in line with the findings of [Bruni & Stanca, 2008], that people may opt for consuming less relational good if they have unsatisfied material needs and their functionings are threatened. They also point at the dominance of hedonistic preferences and the aspiration to consume positional goods that produce social costs to the detriment of public value. Finally, if relational goods are seen as public goods and, as such, as non-rival and non-excludable [Uhlaner, 1989], we can argue for the role of public policy, civil society and responsible business. These actors can promote forms of economic coordination that do not erode immaterial goods, disincentivise and limit, where excessively risky, the contamination of the idea of "development" with that of wealth or growth creation by means of cryptocurrency trusts on the other. In the medium term, greater adherence of cryptocurrency trusts to responsible standards (such as ESG: Environment, Social, corporate Governance) and a policy of subsidies to SSE to compensate for unaccounted for positive externalities would bring the two worlds closer and might rebalance the structurally favourable gap to a financially driven growth. Finally, in the long run, the public sector has the difficult task of promoting the formation of other-regarding and inclusive social preferences, so that individuals may shift more toward the consumption of relational (rather than positional) goods and may prefer the creation of net public value inherent in cooperation over forms of personal enrichment associated with the production of negative externalities.

References

[Bruni & Stanca, 2008] Bruni, L. & Stanca, L. (2008). Watching alone: Relational goods, television and happiness. *Journal of Economic Behavior & Organization*, 65(3-4), 506–528.

[Gui, 2000] Gui, B. (2000). Beyond transactions: on the interpersonal dimension of economic reality. *Annals of public and cooperative economics*, 71(2), 139–169.

[Mosca et al., 2007] Mosca, M., Musella, M., & Pastore, F. (2007). Relational goods, monitoring and non-pecuniary compensations in the nonprofit sector. *Annals of public and cooperative economics*, 78(1), 57–86.

[RIPESS, 2015] RIPESS (2015). Global vision for a social solidarity economy: Convergences and differences in concepts, definitions and frameworks. https://www.ripess.org. Accessed July 2023.

[Sacchetti & Borzaga, 2021] Sacchetti, S. & Borzaga, C. (2021). The foundations of the "public organisation": Governance failure and the problem of external effects. *Journal of Management and Governance*, 25, 731–758.

[Sen, 1992] Sen, A. (1992). *Inequality reexamined*. Oxford: Oxford University Press.

[Tasca & Tessone, 2019] Tasca, P. & Tessone, C. J. (2019). A taxonomy of blockchain technologies: Principles of identification and classification. *Ledger*, 4, 1–39.

[Thompson, 2015] Thompson, S. (2015). Towards a social theory of the firm: Worker cooperatives reconsidered. *Journal of Co-Operative Organization and Management*, 3(1), 3–13.

[Uhlaner, 1989] Uhlaner, C. J. (1989). "relational goods" and participation: Incorporating sociability into a theory of rational action. *Public choice*, 62(3), 253–285.

[Zamagni, 2017] Zamagni, S. (2017). *Felicità pubblica*. Working Paper 12/2017, AICCON.

Deep: Virtual Reality Installation Influenced by Hemoencephalography Sensor

Consuelo Donati

Independent Researcher

Abstract

Over the past century, cognitive psychology has developed our knowledge of how internal processes occur concerning the external environment through participant studies, computational modeling, neuroimaging, and biofeedback. This relationship between internal processes and external environments, recognized by cognitive psychology, is it feasible in virtual reality environments? The *Deep* virtual reality installation investigates this question through an audiovisual experience, where the chromaticity and intensity of lights are modeled by the user's cognitive and behavioral processes, which are monitored and processed in real-time via a Hemoencephalography do-it-yourself biofeedback sensor. The user is transported to four levels of depth, where s/he can explore the virtual environment by interacting with different fluid forms and sound objects, thus shaping his/her personal perceivable experience.

1 Introduction

Advancements in low-persistence display technology and contemporary spatial tracking methods have made it possible to create widely available Virtual Reality (VR) systems for the general public, ushering in a new era of immersive experiences [Çamcı & Hamilton, 2020]. VR has become an exciting new medium for artists, researchers, and content creators. VR is an immersive experience, and the sense of immersion is partly drawn from the phenomenology of embodiment [Merleau-Ponty et al., 2013]. This theory centered upon the notion of a unified body that anchors the human experience. With advancements in VR technology enhancing resolution, speed, and capacity to simulate environments that closely engage our human perceptual systems, approaching the realism of physical reality, the challenge for creators to effectively represent and interpret perceptions has become increasingly complex. Achieving immersion and presence goes beyond mere technological improvements in pixel density and frame rate [Çamcı & Hamilton, 2020]; it demands new interdisciplinary research efforts spanning cognitive psychology, perception, art, design, science, and philosophy.

Synchronously, the technologies related to brain-computer interfaces (BCIs) have grown dramatically in the last decade, allowing the development of wearable sensors capable of monitoring in real-time different activities of the human body. Brain-computer interfaces (BCIs) are systems that combine hardware and software to create a direct communication link between the human brain and external devices [Angrisani et al., 2023]. This has also supported new forms of artistic expression controlled by these devices, as can be seen in [Nijholt et al., 2018]. Moreover, artists and researchers in VR and BCI started collaborating and creating innovative interactive works, like the one developed by Coogan and He [Coogan & He, 2018]. BCIs can serve as control systems in navigation, communication, and other tasks, and simultaneously, VR generates custom interactive environments that engage multiple senses, stimulate the brain through multisensory inputs, and enhance enjoyment with game-like experiences [Leeb & Pérez-Marcos, 2020]. Combining VR and its possibility of navigating and interacting with virtual objects through handheld controllers and these BCI sensors allows the investigation of possible human-computer interactions at both cognitive and physical levels, incorporating both brain and body in designing immersive VR experiences.

When creating the proposed *Deep* VR installation, my first question was not primarily about the technological and artistic challenges possible with VR and BCIs. However, it was mainly based on a personal reflection from the cognitive psychology research field because, over the last century, cognitive psychology has expanded our understanding of how internal processes interact with external environments through different methods [Barsalou, 2014]. The question that led to the development of the *Deep* VR installation was:

- Can this relationship between internal cognitive processes and external surroundings also be replicated in virtual reality?

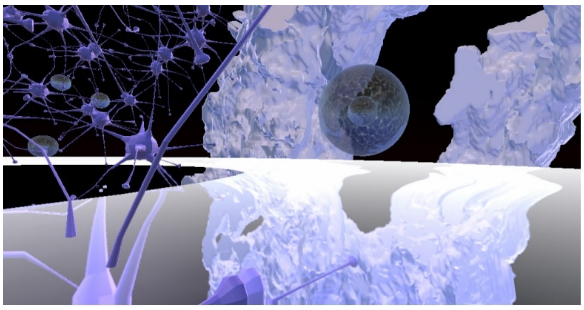


Figure 1: *A screenshot of one of the levels of Deep. In the screenshot, it is possible to see some of the different objects that emit sound and that are graspable by the user.*

2 Brief overview of *Deep*

Deep is a Virtual Reality experience developed in Unity. The Hemoencephalography BCI sensor I used is the "HEGduino"[1], a device that detects regional changes in energy consumption through changes in oxygenation levels in the cerebral bloodstream (particularly the frontal cortex). It is a non-invasive biofeedback method similar to the Electroencephalography (EEG). During the first phase of the project's implementation, I focused on studying the HEG, developed by Joshua Brewster, who is a developer of different open-source HEG devices. The sensor I used to develop this project is worn on the head through a strap. This BCI sensor emits two types of light, red and infrared, which pass through the skull and are absorbed or reflected by the blood and brain tissue. HEG devices detect changes in blood oxygenation in the prefrontal cortex by analyzing the light that is reflected.

The experience is designed as an audiovisual journey within oneself in four levels of depth. This experience, in addition to trying to unite and give artistic expressiveness to the fields of BCIs and VR, carries concepts similar to transcendental meditation in that each user starts on the 1st floor (the highest one). Every 3 minutes, s/he sinks to the lower floor until reaching the lowest one, where the VR experience ends. The visual aesthetic of the entire virtual environment derives from readings by philosophers such as Donna Haraway, Sadie Plant, and the VNS Matrix collective. To compose the music and sound design, I worked on the music composition, realizing it in multiple and different layers. In Deep, there is a musical composition that serves as the background, created through the Cockos Reaper Digital Audio Workstation. At the same time, there is a sound part done of virtual objects represented in the virtual environment with different fluid shapes that can be grasped, and each time they are grasped, they emit different sounds. Each user, wearing the used HEG sensor that has been mapped to model the chrome and lights in the environment according to the monitored values, can interact with these virtual objects and the virtual environment by locomotion (means walking and running) through the use of the handheld controllers.

[1] https://github.com/joshbrew/hegduino.js

References

[Angrisani et al., 2023] Angrisani, L., Arpaia, P., De Benedetto, E., Duraccio, L., Regio, F. L., & Tedesco, A. (2023). Wearable brain–computer interfaces based on steady-state visually evoked potentials and augmented reality: A review. *IEEE Sensors Journal*, 23(15), 16501–16514.

[Barsalou, 2014] Barsalou, L. W. (2014). *Cognitive psychology: An overview for cognitive scientists*. Psychology Press.

[Çamcı & Hamilton, 2020] Çamcı, A. & Hamilton, R. (2020). Audio-first vr: New perspectives on musical experiences in virtual environments. *Journal of New Music Research*, 49(1), 1–7.

[Coogan & He, 2018] Coogan, C. G. & He, B. (2018). Brain-computer interface control in a virtual reality environment and applications for the internet of things. *IEEE Access*, 6, 10840–10849.

[Leeb & Pérez-Marcos, 2020] Leeb, R. & Pérez-Marcos, D. (2020). Chapter 14 - brain-computer interfaces and virtual reality for neurorehabilitation. In N. F. Ramsey & J. del R. Millán (Eds.), *Brain-Computer Interfaces*, volume 168 of *Handbook of Clinical Neurology* (pp. 183–197). Elsevier.

[Merleau-Ponty et al., 2013] Merleau-Ponty, M., Landes, D., Carman, T., & Lefort, C. (2013). *Phenomenology of perception*. Routledge.

[Nijholt et al., 2018] Nijholt, A., Jacob, R. J., Andujar, M., Yuksel, B. F., & Leslie, G. (2018). Brain-computer interfaces for artistic expression. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–7).

[in.tangibile]

Enrico Dorigatti

University of Portsmouth, School of Creative Technologies, Portsmouth, UK

Abstract

This paper presents *[in.tangibile]*, an interactive sound installation commissioned by the Signal To Noise Festival and premiered in January 2024. Following the programme notes introducing the artwork, this paper provides insights into the core technical and philosophical aspects underpinning *[in.tangibile]*, including the relevance of Eco's open work concept.

1 Programme Notes

[in.tangibile] is an interactive sound installation that proposes a reflection on and exploration of the materiality of sound. While often referred to as an object in the sonic arts context, sound nevertheless remains formless and elusive. Yet, it is a physical phenomenon that impacts, albeit usually imperceptibly, the surrounding environment. Is its supposed intangibility, thus, intrinsic to the sound event, or is it related to our perceptual limitations?

By expanding the concept of *opera aperta* (open work) proposed by Umberto Eco, *[in.tangibile]* centralises the role of the audience in the process of artistic creation. According to Eco, an artwork is not an accomplished and unchangeable fact. Rather, it develops from the ecologies of interaction unfolding between the audience and the artefact provided by the artist. Such ecologies range from conceptual (interpretation) to tangible (physical interaction). Through the interactive affordances of *[in.tangibile]*, Eco's notion is expanded, acquiring tangible, physical meaning. Simultaneously, the audience is elevated as co-creators, and their role in shaping the artistic and aesthetic experience is centralised.

2 Overview

[in.tangibile] is a modular interactive sound installation based on custom interfaces leveraging capacitive sensing [Grosse-Puppendahl et al., 2017, Terzic et al., 2012]. Each interface is made of a white rigid canvas covered in tin foil on the back and showcasing an abstract design on the front. The interfaces send data to a Bela microcontroller [McPherson & Zappi, 2015] expanded with a Trill Craft sensor for capacitive sensing.

2.1 Operating Principle

The interfaces are designed to promote touchless interaction. The Trill Craft sensor detects the difference in electrical potential between the conductive backside of each canvas and the body of the person interacting with the installation from around 20 centimetres away. The resulting data streams, one for each interface, are conveyed to the Bela, further processed, and used to control the sound engine.

Although *[in.tangibile]* promotes touchless interaction, it is possible to touch the canvases, exploring them haptically. Either way, the interaction is not dictated by any rule or prescription. However, the abstract designs might be interpreted as graphic musical scores [Cage, 1969].

Abstract art was chosen as the visual scope of reference because, being nonrepresentational, it does not communicate any agreed meaning. Abstract art, therefore, avoids diverting the audience's attention from the artwork, thereby providing greater opportunities for extended subjective interpretation [Aviv, 2014, Leder et al., 2012].

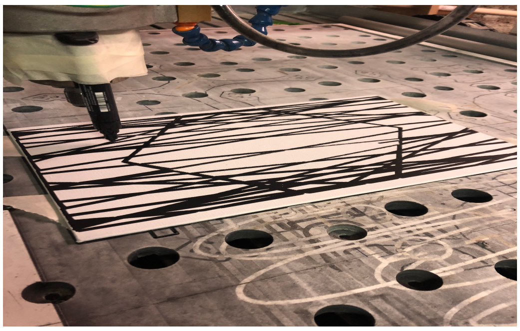


Figure 1: *Abstract designs drawn onto the canvases through a modified CNC machine.*

The abstract designs were realised in TouchDesigner and painted onto the canvases through a modified CNC machine using a black permanent marker in place of the milling head, as in Fig. 1.

2.2 Sound Engine and Mapping

The interaction-driven sonic output of *[in.tangibile]* is produced by a granular sound engine [Roads, 2015] mangling a sound sample realised through databending, a technique used to obtain glitch-based results [Cascone, 2000, Davies, 2004] through unconventional data manipulation [Brunvand, 2019].

A complex mapping between interaction data and sound parameters [Overholt, 2022] was implemented to allow each person to interact with any available interface and fully control the sonic output of *[in.tangibile]*, thus making it possible to experience the artwork intimately or as a collective effort [Blaine & Fels, 2003]. As each interface is a self-contained control surface controlling all the sound parameters, this allows showcasing *[in.tangibile]* with fewer interfaces than the full set shown in Fig. 2.

Figure 2: *The premiere of [in.tangible].*

3 Mixed Artistic Agency

As an interactive artwork, *[in.tangibile]* draws from the concept of open work formulated by Umberto Eco [Eco, 1962]. While Eco deals with the audience's involvement within the artistic experience from a more conceptual, theoretical perspective (understanding, assimilation, and interpretation), digital interactive art brings this concept to a tangible level, making the audience a necessary agent within the creative process. The artist provides an artefact with affordances and constraints [Magnusson, 2010] and, by interacting with it, the audience makes it possible for the artistic experience to unfold.

References

[Aviv, 2014] Aviv, V. (2014). What does the brain tell us about abstract art? *Frontiers in Human Neuroscience*, 8, 85.

[Blaine & Fels, 2003] Blaine, T. & Fels, S. S. (2003). Contexts of Collaborative Musical Experiences. In *Proc. of the Int. Conf. on New Interfaces for Musical Expression* (pp. 129–134). Montreal, Canada.

[Brunvand, 2019] Brunvand, E. (2019). Image Glitching: Manipulating Images with Audio Effects. In *ACM SIGGRAPH 2019 Educators Forum*, SIGGRAPH '19 New York, NY, USA: Association for Computing Machinery.

[Cage, 1969] Cage, J. (1969). *Notations*. New York, USA: Something Else Press.

[Cascone, 2000] Cascone, K. (2000). The aesthetics of failure: "post-digital" tendencies in contemporary computer music. *Computer Music Journal*, 24(4), 12–18.

[Davies, 2004] Davies, H. (2004). Creative Explorations of the Glitch in Music. In *Proceedings of the 2004 IEEE Conference on the History of Electronics*.

[Eco, 1962] Eco, U. (1962). *Opera aperta*. Milano: Bompiani.

[Grosse-Puppendahl et al., 2017] Grosse-Puppendahl, T., Holz, C., Cohn, G., Wimmer, R., Bechtold, O., Hodges, S., Reynolds, M. S., & Smith, J. R. (2017). Finding common ground: A survey of capacitive sensing in human-computer interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17 (pp. 3293–3315). New York, NY, USA: Association for Computing Machinery.

[Leder et al., 2012] Leder, H., Gerger, G., Dressler, S. G., & Schabmann, A. (2012). How art is appreciated. *Psychology of Aesthetics, Creativity, and the Arts*, 6(1), 2–10.

[Magnusson, 2010] Magnusson, T. (2010). Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal*, 34(4), 62–73.

[McPherson & Zappi, 2015] McPherson, A. & Zappi, V. (2015). An environment for submillisecond-latency audio and sensor processing on beaglebone black. In *Audio Engineering Society Convention 138*: Audio Engineering Society.

[Overholt, 2022] Overholt, D. (2022). Designing interactive musical interfaces. In M. Filimowicz (Ed.), *Designing Interactions for Music and Sound* chapter 1. Abingdon, Oxon; New York: Routledge.

[Roads, 2015] Roads, C. (2015). *Composing electronic music: a new aesthetic*. USA: Oxford University Press.

[Terzic et al., 2012] Terzic, E., Terzic, J., Nagarajah, R., & Alamgir, M. (2012). *Capacitive Sensing Technology*, (pp. 11–37). Springer: London.

Navigation in the *Nautilus* Virtual Reality Auditorium as an Approach to Interactive Music Composition

Matteo Tomasetti and Luca Turchet

Department of Information Engineering and Computer Science (DISI), University of Trento, Trento, Italy

Abstract

In the past two decades, several paradigms have been proposed to compose and perform music leveraging the novel possibilities offered by the Virtual Reality medium. In this demo, we propose a new composition approach where the act of navigating within a virtual environment is used to interactively generate sound events and immersive sound spatializations. To implement such an approach, we developed a virtual environment consisting of a large auditorium shaped like a nautilus, where composers move via the use of handheld controllers. During the demo, we will discuss the musical possibilities offered by this system, where basic interaction techniques, such as user navigation and the presence of a network of virtual loudspeakers, shape the composition's overall form and temporal structure.

1 Introduction

The last twenty years have witnessed a considerable expansion in artistic and scientific research in the music scenario using the Extended Reality (XR) medium. This expansion has led to developing a new research area called Musical XR [Turchet et al., 2021], which includes all music processes in virtual and augmented environments. Moreover, recent developments in Virtual Reality (VR) technology have brought back interest in techniques related to spatial audio, such as Ambisonics and binaural audio, which offer significant advantages for immersive audio experiences [Çamcı, 2019]. In general, audio design in VR experiences should be dynamically generated, contain forms of interaction, and be immersive [Atherton & Wang, 2020]. Typically, the 3D interaction techniques that the user in VR can experience are divided into navigation, selection, manipulation, and application control [Berthaut, 2020]. To the best of the authors' knowledge, no works in the literature have considered navigation as an interaction tool for generating and managing music composition in VR. For this reason, we developed a new VR music composition approach, which we detailed in [Tomasetti & Turchet, 2024], that focuses on navigation techniques, mainly walking and running, which means the user can freely and continuously walk or run in any direction in the VE. In this demo, we will show this composition approach, which utilizes the user's navigation to activate in real-time different layers of the composition and manage specific sound parameter attenuations, which involve amplitude, spread, focus, filtering, reverberation zones, and immersive sound spatialization techniques dependent also on the user's head movements. However, in this proposal, we discuss only the part related to managing spatial audio using the virtual Ambisonics approach. To the best of the authors' knowledge, this is the first attempt to exploit the user's navigation in VE to compose interactive and immersive music. This approach investigates the relationship between composition and active user interaction in VEs, exploring the navigation technique's possible musical expressiveness. This demo aims to provide sound designers and composers with a new approach to creating interactive compositions in VR based on the user navigation technique and where spatial audio plays a crucial part. A video of the system in action is available online.

2 The Use of Virtual Ambisonics

We implemented the spatial audio part within the Wwise software and linked it with the Unity game engine to develop the VR system. Spatial audio includes all the technologies, such as Ambisonics and binaural,

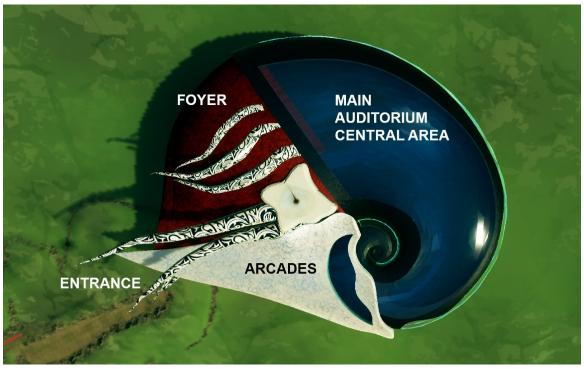


Figure 1: *A screenshot of the 3D model displaying the complete nautilus-shaped auditorium, highlighting various sections such as the entrance tunnel, foyer, main auditorium, and the rear section formed by arcades.*

that can produce the feeling of different sound sources positioned in space, just as we perceive the directionality of sounds in real life. In general, Ambisonics is a method that allows the three-dimensional recording and playback of audio signals, enabling the production of spatial sound content in any playback system of loudspeakers by applying appropriate decoders [Frank et al., 2015]. However, this type of technology needs extensive refined sound system installations. For this reason, the most widely adopted system for reproducing three-dimensional content is binaural audio using head-related transfer functions (HRTFs), which can be appreciated by any listener simply with a pair of headphones [Nicol et al., 2014]. We used the virtual Ambisonics approach to decode Ambisonics signals in virtual loudspeakers. The convolution of these loudspeaker signals with HRTF creates left and right ear signals, allowing binaural reproduction [Noisternig et al., 2003]. We used a virtual Ambisonics approach to encode all the audio signals with third-order Ambisonics. Then, we decoded them for binaural reproduction using the Atmoky Ears binaural software. The VR system utilized each sound we created in Ambisonics by managing the *Positioning* Wwise's property. In that property, we enabled the *Listener Relative Routing* option. Then, we also enabled the *3D Spatialization* option by setting it in the *Positioning plus Orientation*. The latter is the action of panning each sound based on the relative user's position and the different sound emitters' orientations. The Meta Quest 2 head-tracking data sent the relative user's position in real-time. Head-tracking communicates from Unity to Wwise and thus handles the different perceptions related to the spatial position of sounds. This means that whenever the user moves the head, the user hears the sound in the headphones in a three-dimensional surround manner, coming from different positions. Moreover, we designed 3D sound trajectories through automation options for some sound objects of the auditorium's composition. We performed it through the *Position Editor (3D Automation)* option in Wwise. The *Position Editor* allows to write and automate sound trajectories in a classic surround panner interface. Finally, each sound was routed to a binaural bus, where we inserted the Atmoky Ears binaural plugin, which allowed to render all sounds used in binaural. With this setup, every time the user navigates in any direction and moves the head, s/he has the perception of hearing the sounds in a three-dimensional way.

References

[Atherton & Wang, 2020] Atherton, J. & Wang, G. (2020). Doing vs. being: A philosophy of design for artful vr. *Journal of New Music Research*, 49(1), 35–59.

[Berthaut, 2020] Berthaut, F. (2020). 3d interaction techniques for musical expression. *Journal of New Music Research*, 49(1), 60–72.

[Çamcı, 2019] Çamcı, A. (2019). Some considerations on creativity support for vr audio. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 1500–1502).: IEEE.

[Frank et al., 2015] Frank, M., Zotter, F., & Sontacchi, A. (2015). Producing 3d audio in ambisonics. In *Audio Engineering Society Conference: 57th International Conference: The Future of Audio Entertainment Technology–Cinema, Television and the Internet*: Audio Engineering Society.

[Nicol et al., 2014] Nicol, R., Gros, L., Colomes, C., Noisternig, M., Warusfel, O., Bahu, H., Katz, B. F., & Simon, L. S. (2014). *A roadmap for assessing the quality of experience of 3D audio binaural rendering*.

[Noisternig et al., 2003] Noisternig, M., Musil, T., Sontacchi, A., & Holdrich, R. (2003). 3d binaural sound reproduction using a virtual ambisonic approach. In *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS'03. 2003* (pp. 174–178).: IEEE.

[Tomasetti & Turchet, 2024] Tomasetti, M. & Turchet, L. (2024). Handheld controller-based locomotion in virtual reality as an approach to interactive music composition: insights from composers' preferences. *Digital Creativity*, (pp. 1–25).

[Turchet et al., 2021] Turchet, L., Hamilton, R., & Çamci, A. (2021). Music in extended realities. *IEEE Access*, 9, 15810–15832.

SoundDesigner: A Framework for Assisted Sound Composition

Francesco Roberto Dani
Independent Researcher

Abstract

SuperCollider is a powerful platform for audio synthesis and algorithmic composition, widely used by musicians, sound artists, and researchers. Despite its capabilities, SuperCollider presents a steep learning curve, particularly for students of electroacoustic music not involved in computer programming, due to its complex syntax and extensive technical details. To address these challenges, SoundDesigner has been developed, an advanced framework for assisted sound composition designed to fully exploit SuperCollider by providing a complex and structured graphical control framework for managing its synthesis server scsynth.

SoundDesigner integrates multiple paradigms by allowing non-programmers to fully utilize SuperCollider's synthesis server (scsynth) within a more accessible visual environment, fostering contamination between technical programming approaches and intuitive graphical interfaces. This framework aims to simplify the use of SuperCollider, making it more accessible without sacrificing its formalism and unique features, while promoting cross-pollination between users with different skill sets (e.g., people unfamiliar with coding and those comfortable with technical languages), bridging the gap between high-level sound design and low-level audio programming.

1 Introduction

SuperCollider [McCartney, 1996], an open-source platform developed by James McCartney in the mid-1990s, consists of three main components: the programming language (sclang) for creating and manipulating sound processes, the sound synthesis server (scsynth) for real-time audio processing and synthesis, and an integrated development environment (scide) that provides tools for writing and executing code. Together, these components form a robust system used by musicians, sound artists, and researchers worldwide. SuperCollider, despite its robust features, presents a significant learning curve, especially for beginners and students of electroacoustic music.

To mitigate these challenges and foster a more inclusive learning environment, SoundDesigner was developed as a bridge between traditional low-level programming and more intuitive graphical tools. This framework leverages SuperCollider's synthesis server, allowing users to manage and control sound synthesis through a modular and visual interface. By integrating programmable AudioWidgets with scsynth's real-time audio synthesis capabilities and MIDIWidgets to handle events, SoundDesigner promotes contamination between visual sound design and algorithmic sound synthesis, enabling a seamless interaction between graphical and code-based approaches. This integration allows users to harness the full potential of SuperCollider in a more intuitive way, empowering both novice and experienced users to manipulate complex sound structures without being constrained by the intricacies of SuperCollider's syntax.

2 SoundDesigner Architecture

At the heart of SoundDesigner's architecture is the Patch Area, a GUI where users can construct and modify electroacoustic chains inserting a variety of Widgets (Audio, MIDI, or AudioMIDI) and connecting them with cables, storing them as Patches. This environment, which resembles other computer music software like PureData [Puckette et al., 1997] and Max/MSP[Puckette & Zicarelli, 1990] [Zicarelli, 1992], provides a versatile platform for organizing and managing complex audio structures. The PatchArea is paired with a TimeLine, another GUI which allows dynamic modulation of widget parameters over time by means of Parameter Curves to be used for both real-time and fixed media production, offering users detailed control over the temporal aspects of their audio compositions.

The framework is further enhanced by components that integrate timing and synchronization. The Region Line and Region Manager work together to manage playback segments and ensure they are coordinated with the project's timing using the Clock, which operates based on MIDI ticks for precise synchronization. Additionally, Non-Real-Time Rendering allows for extensive editing and high-quality output, making it ideal for complex sound design tasks. Together, these elements form a comprehensive system that supports sophisticated audio synthesis, precise timing management, and effective educational applications.

2.1 Widgets

Widgets are the core of SoundDesigner programming, providing a flexible and powerful means for users to create and manage complex sound synthesis processes. Each type of widget serves a specific function, allowing for a modular approach to sound design.

Audio Widgets represent the main category of Widgets and enable users to integrate custom scsynth-side SuperCollider Synth definitions into graphical objects representing their inputs, outputs, and parameters, thus facilitating the creation of intricate audio synthesis algorithms. These widgets leverage the full capabilities of the scsynth server, allowing users to design unique sound generators and processors tailored to their specific needs. An AudioWidget is composed by a synth definition file (.scsyndef) describing the signal processing chain and a JSON file used to tell SoundDesigner how the synth definition is composed. Every AudioWidget represents a SuperCollider Synth instance (a Node), to which definition it is connected and described graphically; it owns audio inputs and outputs, as well as parameters of types accepted by SynthDefs (like int, float, audio bus, buffer, and trigger). The writing of SynthDef parameters in SuperCollider's side must follow a specific approach with respect to the parameter type inserted: for example, every audio rate argument must be accompained with a selector to tell SoundDesigner whether it has to switch from a fixed number to a Bus, and every audio output must be followed by a control-rate output with the RMS signal of the outputted audio signal.

The following is an example of the SuperCollider-side synth definition of the AudioWidget *SinOsc*:

```
SynthDef(\SinOsc, { |
  out_ch_0=0,
  a_freq=440,
  selector_freq=0,
  a_mul=1,
  selector_mul=0,
  a_add=0,
  selector_add=0 |
  var sig;
  a_freq = Select.ar(selector_freq,
   [Lag.ar(K2A.ar(a_freq), 0.001), a_freq]);
  a_mul = Select.ar(selector_mul,
   [Lag.ar(K2A.ar(a_mul), 0.001), a_mul]);
  a_add = Select.ar(selector_add,
   [Lag.ar(K2A.ar(a_add), 0.001), a_add]);
  sig = SinOsc.ar(a_freq, 0, a_mul, a_add);
  Out.ar(out_ch_0, sig);
  Out.kr(out_ch_0, RMS.kr(sig));
}).writeDefFile;
```

, followed by its SoundDesigner-side JSON definition:

```
"SinOsc": {
  "synth_name": "SinOsc",
  "n_in": 0,
  "n_out": 1,
  "args": {
    "a_freq": {
      "desc": "Frequency (Hz)",
      "type": "audio",
      "min": 20,
      "max": 20000,
      "val": 440,
      "bus": -1
    },
    "a_mul": {
      "desc": "Amplitude (abs)",
      "type": "audio",
      "min": 0.0,
      "max": 1.2,
      "val": 0.75,
      "bus": -1
    },
    "a_add": {
      "desc": "Bias (abs)",
      "type": "audio",
      "min": -20000,
      "max": 20000,
      "val": 0.0,
      "bus": -1
    }
  }
}
```

MIDI Widgets handle MIDI processing, providing interfaces for managing MIDI input and output and processing MIDI messagges. These widgets are essential for users who need to incorporate MIDI controllers and other MIDI-enabled devices into their sound compositions, enabling real-time interaction and control over various sound parameters, as well as loading and playing MIDI scores.

AudioMIDI Widgets offer a hybrid solution, bridging audio and MIDI processing. Written in pure Python, these widgets allow for the integration of custom Python libraries (like PyTorch) used to generate

MIDI messagges or to pilot the instantiation of Synths in the server side. This hybrid approach opens up new possibilities for creative sound design, enabling users to build sophisticated data workflows that are both powerful and intuitive.

2.2 Parameter Curves

Parameter Curves are another fundamental feature in SoundDesigner, enabling precise control over various Widget' parameters over time. These curves can be created and modified in multiple ways to suit the needs of the user, offering a high degree of flexibility and control in sound design. One of the primary methods for creating and editing Parameter Curves is through the TimeLine that supports manual insertion and adjustment of curve points. This intuitive interface allows to achieve the desired modulation effects without needing to write code.

For users who prefer a more programmatic approach, SoundDesigner provides the ability to create Parameter Curves by writing and executing a Python function that iterates through and returns a list of points (math and random libraries are supported, and variables like length, and iteration number are provided – miming the SuperCollider's pvcalc and pvcollect methods). This method leverages the power and flexibility of Python, allowing users to generate complex, algorithmically defined curves that can be easily integrated into their sound design projects. By writing custom Python functions, users can precisely define the behavior and characteristics of the parameter curves, enabling sophisticated modulation techniques that would be difficult to achieve manually.

Additionally, SoundDesigner includes a GUI that supports the loading and modification of numpy arrays into Parameter Curves. The GUI integration with numpy, a powerful library for numerical computing in Python, provides advanced capabilities for processing and analyzing parameter data. Users can modify these numpy arrays within the SoundDesigner interface, providing a seamless workflow for developing and refining parameter curves.

The combination of these methods ensures that SoundDesigner can accommodate a wide range of user preferences and requirements. Whether users prefer the immediacy of manual editing, the precision of Python scripting, or the computational power of numpy arrays, SoundDesigner offers the tools necessary to create and manipulate parameter curves effectively. This versatility makes it an invaluable tool for both beginners and advanced users in the field of electroacoustic music and sound synthesis.

2.3 Region Management

In SoundDesigner, the management of temporal regions of playback is handled through the Region Line and the Region Manager components. The Region Line is a feature that enables the insertion and removal of regions within a project, effectively marking defined temporal segments for playback. This allows users to designate specific sections of the timeline where particular audio elements or sequences should be active, facilitating organized and precise control over playback. The Region Manager, on the other hand, oversees the management of these regions and interacts with the Clock to regulate playback. It ensures that the regions are activated and deactivated according to the timing set by the Clock, coordinating the playback process and maintaining synchronization across the project. Together, the Region Line and Region Manager provide a cohesive system for organizing and controlling audio playback within a project.

3 Methodology for a Validation Test

The primary aim of this article is to present the SoundDesigner software and outline its potential effectiveness in electroacoustic music education, specifically in relation to reducing the learning curve of SuperCollider. While the validation test outlined here has yet to be conducted, it serves as a structured plan for future research, designed to empirically assess SoundDesigner's educational impact.

3.1 Test Participants

The test is planned to involve individuals with a background in sound design, possessing at least a basic understanding of sound design principles. Participants will be selected based on their familiarity with sound design tools, particularly those with prior experience in SuperCollider or other platforms like PureData or Max/MSP.

3.2 Test Design

The proposed test will involve the creation of four patches of increasing difficulty in both SuperCollider and SoundDesigner, as well as the creation of an Audio Widget within SoundDesigner. This design is intended to measure whether and to what extent SoundDesigner can lower the learning curve of SuperCollider in an educational setting.

Participants will autonomously complete the tasks if they have prior knowledge of SuperCollider. Those unfamiliar with SuperCollider but with experience in other sound design software will complete the test with guidance (tutoring). Tutoring will be used as an integral part of the test to assess how much intervention is needed to explain procedures, rather than being seen as a biasing factor [Chi et al., 2001] [Koedinger & Aleven, 2007].

3.3 Research Questions

The test is structured around the following key Research Questions (RQs):

1. Can SoundDesigner be effectively used as a complement to SuperCollider in electroacoustic music education, thereby reducing its learning curve?
2. What are the most critical points in learning SuperCollider that can be simplified through the use of SoundDesigner? These points may include SynthDef creation, execution order management, Routine and automation creation, and MIDI handling.

To answer these questions, participants will be asked to compare their experiences working in both SuperCollider and SoundDesigner. After completing each patch in both environments, participants will quantify the effort required and annotate the most critical points encountered. For the final task, involving the creation of a new Audio Widget in SoundDesigner, participants will compare the difficulty of its creation with that of building a SynthDef in SuperCollider.

4 Conclusion

In educational contexts, SoundDesigner represents a resource for teaching audio synthesis concepts. Students benefit from hands-on experience with practical sound design tools and techniques. Notably, SoundDesigner integrates with SuperCollider, allowing users to create Audio Widgets directly within SuperCollider's SynthDefs and import them into SoundDesigner as JSON definitions. This functionality bridges the gap between theoretical knowledge and practical application, offering learners the opportunity to experiment with and refine their sound design skills in a real-world context. By working with features like the TimeLine for temporal modulation and the Region Manager for organized playback, students gain a deeper understanding of audio synthesis principles and their application.

The future integration of a Patch Sequencer - a graphical interface dedicated to the composition in the time domain of different Patches at the same time within the same project - will allow users to go beyond the traditional patch-like composition which characterizes current audio programming software, allowing to enable/disable single patches at certain times, as well as sharing inputs and outputs.

5 Acknowledgements

I would like to thank Sonia Cenceschi for her continuous support and Matteo Marco Markidis for the time spent in conversations about computer music, without whom this work would not have been possible.

References

[Chi et al., 2001] Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive science*, 25(4), 471–533.

[Koedinger & Aleven, 2007] Koedinger, K. R. & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19, 239–264.

[McCartney, 1996] McCartney, J. (1996). Supercollider, a new real time synthesis language. In *Proceedings of the 1996 International Computer Music Conference* (pp. 257–258).: The International Computer Music Association.

[Puckette & Zicarelli, 1990] Puckette, M. & Zicarelli, D. (1990). Max-an interactive graphic programming environment. *Opcode Systems, Menlo Park, CA*.

[Puckette et al., 1997] Puckette, M. S. et al. (1997). Pure data. In *ICMC*.

[Zicarelli, 1992] Zicarelli, D. D. (1992). An extensible real-time signal processing environment for max. In *Proc. Int. Computer Music Conf. 98* (pp. 463–466).

Gravitational Music: an Installation for the Intuitive Perception of Gravitational Waves

Alessandro Bile[1] and Angelica Speroni[2]

[1] *Art of Images and Contemporary Art, Doctoral School of Aesthetics, Sciences and Technology of the Arts, Paris 8, Paris, France.*
[2] *Department of Fundamental and Applied Sciences for Engineering, Sapienza University, Rome, Italy*

Abstract

With the relentless advancement of scientific progress, knowledge is becoming increasingly confined to an audience with a strong mathematical background. This process seems to further distance the worlds of Science and Art, which were once intertwined. It is up to modern technology to find a way to make scientific knowledge universally accessible once again, by leveraging the language of Art. In this article, we present Gravitational Music, an installation capable of translating the physical phenomenon of gravitational wave propagation into sound.

1 Introduction

Physics and music are generally considered to be very distant forms of knowledge, yet they are actually closely connected [Gordon, 2015]. Every physical phenomenon can be musically represented, and every sound can be described by mathematical laws [Barrass & Kramer, 1999]. In fact, everything around us is energy, and energy is also music. Isn't music itself an organization of energetic events moving toward a new configuration?

The rapid progress of science has necessitated an ever-increasing specialization of knowledge, seemingly distancing the artistic-musical world from the scientific-technological one. With Gravitational Music, we propose an alternative way of investigating complex physical phenomena by leveraging the intuitiveness of sound. The language of sound is universal and accessible to all, as we experience it every day. Among the most fascinating physical phenomena in recent years are gravitational waves (GW). Gravitational waves, theoretically predicted by Einstein [Einstein, 1916] and first measured by [LIGO and Virgo collaboration, 2016], are ripples in spacetime caused by the motion of massive asymmetric bodies in rapid acceleration, which can generate perturbations in the fabric of spacetime. Since they are wave phenomena, it is possible to translate their physics into music. The physical principle that describes auditory perception is, in fact, similar to wave phenomena.

Many sonification algorithms have been proposed based on this similarity [Zanella et al., 2022]. Gravitational Music works differently. The signal is not "transformed" but recreated through the movements of the people experiencing it. Their movements are recorded and converted into sounds, all in real time. Thus, people immerse themselves in a universe crafted from artificial sound, becoming an integral part of it. The sounds produced depend on the specific mass of the individuals and their position in space (this is why they are called TSS, Temporal Space Sound). Physical characteristics of the individuals, such as height and weight, are converted to model stellar objects capable of emitting gravitational waves (GW). Each person is therefore associated with a star with unique characteristics and dimensions. TSS allows the complexity of gravitational wave propagation to be understood through sound. The movement of stellar masses, linked to individuals, mimics the changes in the spacetime fabric in which they are located, producing a signal similar to that of gravitational waves. The parameters of this signal are then converted into sound.

The spacetime fabric is represented through the spatialization of a background sound model. The TSS alters this background sound, allowing their propagation to be recognized through variations in the sound. Thus, Gravitational Music transforms listeners into sources capable of creating perturbations in the sound-based spacetime, much like the presence and movement of stellar objects generate perturbations in spacetime.

The sounds produced depend on the specific interactions between individuals and the environment in which they live.

2 Theory of Gravitational Waves

Gravitational waves were predicted in the Theory of General Relativity [Maggiore, 2007]. GWs are produced by the asymmetric motion of stellar masses. In this article, we consider the phenomenon of coalescence between two stars. The GW signals produced in this case provide us with a reference waveform, allowing us to compare the sound emitted during a properly scaled GW emission process with the sound produced by human movement [Bile et al., 2024].

Let us consider two compact bodies, such as black holes or neutron stars, moving in an approximately circular orbit [Rosado, 2011]. These form a system that emits GWs, and the waveform is characterized by a "chirp" (so named due to its resemblance to bird chirping). Over time, the system loses energy due to the emission of GWs, leading to the two masses drawing closer together and a gradual increase in the amplitude and frequency of the emitted signal. If the masses are distant, i.e., within the limit where stable circular orbits are allowed, the GW signal will be:

$$h(t) = A(t)\cos\left[\varphi(t) + \varphi_0\right] \tag{1}$$

Equation 1 represents the mathematical behaviour of a general time-dependent wave signal with an amplitude $A(t)$ and a phase $\varphi(t)$ which must be specialised for GW. The expression of the amplitude $A(t)$ of the signal can be written taking into account the fact that it depends on the combination of the masses (the so-called chirp mass M, which in Gravitational Music is related to the masses of the performers), the instantaneous frequency of the chirp, the luminous distance D of the observer from the source, and finally a geometric factor G, which is a factor related to the position and orientation of the cosmic object with respect to the detector. M can be calculated by combining the total mass $M = m_1 + m_2$ and the reduced mass $m = \frac{m_1 m_2}{M}$ as $M = m^{3/5} M^{2/5}$.

$$A(t) = G \cdot \frac{4\,\mathcal{M}^{5/3}\,\pi^{2/3}\,F(t)^{2/3}}{D} \tag{2}$$

In order to have detectable effects with our technology, catastrophic events are required, such as those that occur in the Universe, typically far from Earth, involving massive objects, like the death of massive stars, collisions of compact objects, or emissions from pulsars or magnetars. Therefore, the physical characteristics of individuals must be scaled in terms of stellar quantities.

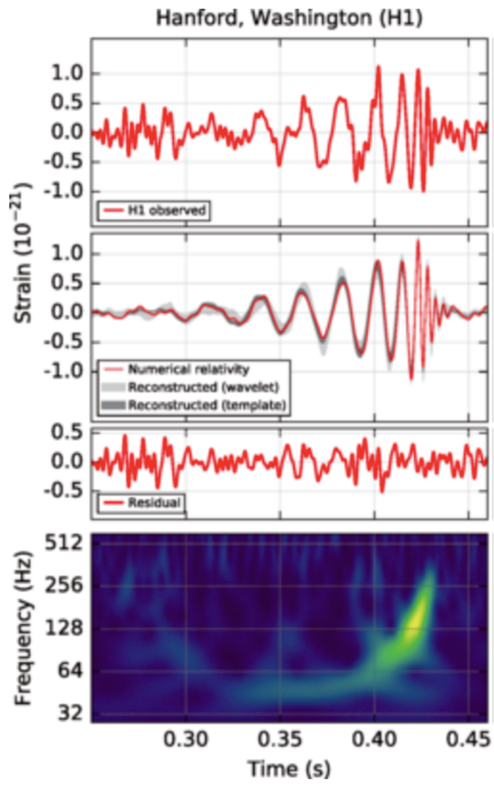


Figure 1: *The GW event GW150914 observed by the LIGO Hanford. During the merging phase the GW frequency increases from ∼30 Hz to ∼500 Hz.*

3 Theory and Code of Gravitational Music

The basis of Gravitational Music is people and their behaviour. The structure of the algorithm on which the installation is based must therefore start precisely from movements. Two different methods have been implemented to detect positions. In this article, we present the first based on the accelerometers already present on everyone's smartphone. Positions are handled through the dialogue between Matlab apps and Matlab code. Currently, two different instances are used, one for handling each person. The main instance is also responsible for sound generation. Smartphones communicate time-dependent acceleration values to the instances.

These can run simultaneously on the same computer, but receive signals from two different devices. Instance 1 has a simple code that collects data from executor 1 and sends it to instance 2. The second instance is in charge of receiving the data from instance 1, sampling the data from the second executor and processing it. The parameters to be entered before starting the installation are the mass in kg and the height in cm. The height determines the radius of the performer's gravitational field, which is realised computationally by simulating a sphere. Before sampling the data, there is a time interval required for instrument calibration. This step is essential to eliminate the instruments' inherent noise. To clean the measured data, a low-pass filter was implemented to eliminate all high frequencies associated with noise.

Thus, by exploiting Einstein's equations for the generation of GWs, the main instance uses the input data (people's positions and physical parameters) and parameterises the propagation of a 'sound gravitational wave'.

The output data from Matlab is serially sent to the Pure Data software. A patch (code name PureData) was developed in this environment. As input, it receives the relative positions of the centre of mass, the amplitude and frequency of the gravitational waves emitted by the two persons and uses them to manage a system of 4 comb filters (2 feedword, 2 backword), the pointer speed for scrolling through a buffer of sounds to obtain the physical-acoustic phenomenon of granulation [Xenakis, 1992], and various types of sound beats [Aygün & Aydın-Güç, 2019].

The comb filter consists of adding to the signal at time t, an older version of it (t-n), at an earlier time of n time steps. The filter is called a comb filter because its frequency response is characterised by equally spaced pulses like the teeth of a comb. A comb feedforward filter is governed by Equation 3, while the schematic operation of the comb filter with feedback is described by Equation 4.

$$y[n] = x[n] + \alpha x[n - R] \quad (3)$$

$$y[n] = x[n] + \alpha y[n - R] \quad (4)$$

Granular synthesis involves constructing complex sounds from the creation of fundamental sounds called microsounds (grains). Grains are sound entities that last between 1 and 50 milliseconds. They can be combined and played together by varying their speed, phase, and amplitude. The result is a sound phenomenon characterized not by a single tone but by a new array of sounds, whose parameters can be managed just like those of a single sound. Beating is a physical phenomenon that characterizes any wave phenomenon and stems from the principle of superposition. Two waves can overlap in three different ways: in phase (constructive interference), out of phase (destructive interference), or somewhere in between.

The two main questions we asked ourselves while writing the code were: what perceptual objective do we want to achieve? How should the sound interact with the user?

In Gravitational Music, the audience finds themselves immersed in a pre-existing spacetime fabric, which is perturbed by the gravitational waves emitted by people in motion, causing various perceptual changes from the initial situation. The signal recorded by the sensors increases monotonically, analogous to the chirp associated with celestial bodies: the frequency of the signal increases progressively as the orbits gradually shrink. The chirp is thus perceived and recognized on an auditory level. The frequency values are responsible for varying the reading speed of the memory buffer. The resulting sound is a fragmentation of the audio file into small amounts: the sound grains (scrolling through memory buffers at different speeds is an effective strategy for achieving granulation).

4 Applications

Gravitational Music is an installation born from the interaction between art and science. It is capable of conveying the complexity of a physical phenomenon, such as the propagation of gravitational waves, by leveraging the intuitiveness of musical language. To enhance the focus on the sound action, the installation operates in a low-light environment. The only light sources are from the smartphones, and in this way, the light trajectory reinforces the auditory perception.

Thus, the understanding of physics is transposed from the mathematical plane to one of sensory perception, which is universal and free of socio-cultural barriers. For this reason, the installation has also been proposed as an educational and scientific dissemination tool.

Recent research on educational methods has demonstrated that technology can facilitate the learning process [Bile, 2022]. In particular, even the learning of scientific subjects, such as mathematics and physics, has benefited from alternative teaching methods that make use of technology [Bile, 2022].

From a didactic perspective, it is important to highlight that Gravitational Music does not require any supervision but can guide people learning autonomously. Once a type of stellar event is selected, the algorithm maps the room where the individuals are located, records their physical parameters, and assigns them stellar characteristics. If the individuals make movements that "do not correspond" to the selected physics, the generated sound event will have a strong dissonant component. On the other hand, if their trajectories resemble those typical of the physical phenomenon, a musical sequence will be generated—still entirely dependent on their movements but self-consistent. In this way, the installation is able to guide participants toward learning the actual physical phenomenon while still granting artistic freedom to those who experience it.

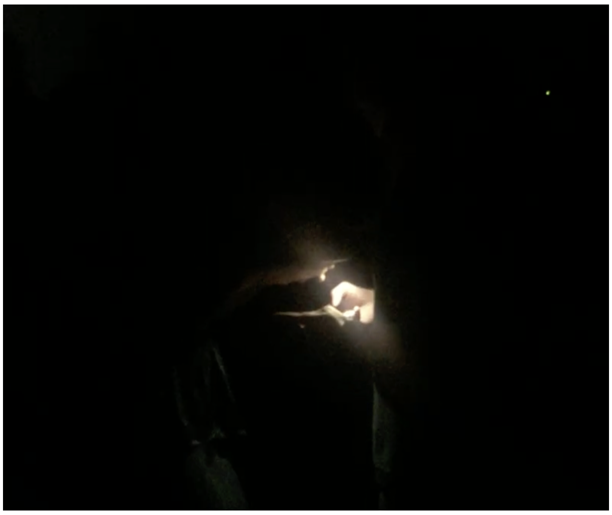


Figure 2: *The moment of the coalescence between the two performers.*

5 Conclusions

Gravitational Music was conceived as an intermedia research experiment. In this case, mathematics, physics, computer science, music, and psychoacoustics work together to create a dynamic sound structure that captures the essence of a complex physical process. Regarding gravitational waves, many attempts have been made to associate them with musical language, but all have stopped at their sonification. Gravitational Music introduces a poietic component, transferring creative power directly into the hands of those experiencing it. For this reason, its communicative strength increases exponentially. Through the installation, we ourselves create a language through which to narrate the life of gravitational waves.

References

[Aygün & Aydın-Güç, 2019] Aygün, M. & Aydın-Güç, F. (2019). Superposition of sound waves: Beats. *Physics Education*, 54, 043007.

[Barrass & Kramer, 1999] Barrass, S. & Kramer, G. (1999). Using sonification. *Multimedia Systems*, 7, 23–31.

[Bile, 2022] Bile, A. (2022). Development of intellectual and scientific abilities through game-programming in minecraft. *Education and Information Technologies*, 27, 7241–7256.

[Bile et al., 2024] Bile, A., Santoboni, R., Frasca, S., & Astone, P. (2024). Gravitational music: a mathematical-model for the popularization of gravitational waves. *Physics Education*, 59(6).

[Einstein, 1916] Einstein, A. (1916). *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, 1, 688.

[Gordon, 2015] Gordon, P. R. (2015). Teaching physics with music. *The Physics Teacher*, 53(7), 415–418.

[LIGO and Virgo collaboration, 2016] LIGO and Virgo collaboration (2016). Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, 116, 061102.

[Maggiore, 2007] Maggiore, M. (2007). *Gravitational Waves*. Oxford University Press.

[Rosado, 2011] Rosado, P. (2011). Gravitational wave background from binary systems. *Physical Review D*, 84, 084004.

[Xenakis, 1992] Xenakis, I. (1992). *Formalized Music: Thought and Mathematics in Musical Composition*. Stuyvesant, NY: Pendragon Press.

[Zanella et al., 2022] Zanella, A., Harrison, C., Lenzi, S., et al. (2022). Sonification and sound design for astronomy research, education and public engagement. *Nature Astronomy*, 6, 1241–1248.

It Takes Two – Three Shared and Collaborative Virtual Instruments for the Musical Metaverse

Alberto Boem[1], Damian Dziwis[2], Sascha Etezazi[2], Matteo Tomasetti[1], and Luca Turchet[1]
[1]*Department of Information Engineering and Computer Science (DISI), University of Trento, Trento, Italy*
[2]*KreativInstitut.OWL Detmold University of Music, Detmold, Germany*

Abstract

It Takes Two is a set of three prototypes exploring the concept and design of shared and collaborative Virtual Reality Musical Instruments (VRMIs) within the Musical Metaverse (MM). These instruments are specifically developed for dual-player interaction, leveraging the unique affordances of multi-user virtual environments such as spatial audio, data sonification, and the use of embodied three-dimensional avatars.

1 Introduction

The Musical Metaverse (MM) [Turchet, 2023] offers a new platform for collaborative music-making, where multiple performers can play together in the same virtual space. Shared and collaborative musical instruments refer to a specific category of instruments that enable simultaneous performance by multiple individuals. These include not only facilitating interactions between performers and the instrument but also enabling interaction among the performers themselves [Jordà, 2005]. To explore the possibilities, properties, and implications of shared, collaborative virtual instruments in the Musical Metaverse environments, we developed *It Takes Two* [Boem et al., 2024]. This project comprises three prototypes of VRMIs designed for two players to be used in the MM, because these instruments can function only if more than one player is using them (see Figure 1). Each instrument we are going to demonstrate embodies a concept that leverages unique characteristics of the MM: the first makes use of shared spaces and sound spatialization; the second exploits the interactions between users and virtual environments through an approach based on real-time sonification; the third examines embodied social interactions by using avatars as musical interfaces.

2 Design and Implementation

The three instruments have been designed as web-based VRMIs, and can be used with WebXR compatible browsers, including the ones available in consumer Head-mounted Displays (HMDs), such as the Quest Browser and Wolvic. In all of virtual instruments that composes *It Takes Two* the players are represented as three dimensional avatars placed in a minimalistic environment. Players need an HMD to access and use the instruments, and hand-held controllers to move and interact with the environment. However, the three instruments differ in how they deal with the concept of shared control and collaboration within virtual environments.

1. Spatial Instrument: This instrument was created to explore the possibilities of sound spatialization in the context of shared control. A sound synthesizer is generating a high-range (C4-B4) and a low-range (C2-B2) ambient sound. The pitches can be selected through an interface designed as two-part keyboard placed on the two opposite sides of the virtual space. With this instrument, one player is responsible for selecting the tones. On the contrary, the other handles the spatial composition. The sound generated by the first player is emitted as a spatialized sound source that corresponds at the avatar of the second player. Therefore, the position in space of the player determines the movement/position of the sound source.

2. Sonification Instrument: In this instrument, a sonification approach is used. Here, the sound synthesis is controlled by parameters derived from the spatial relationships of players' avatars, such as their distance. This allows players to generate music through their exploratory movements without relying on visual interfaces. The sound of the instrument is based on a saw-tooth oscillator. The distance between the avatars is interpreted as the wavelength of such oscillator. Other parameters such as the distance between players and the walls, the vertical position of

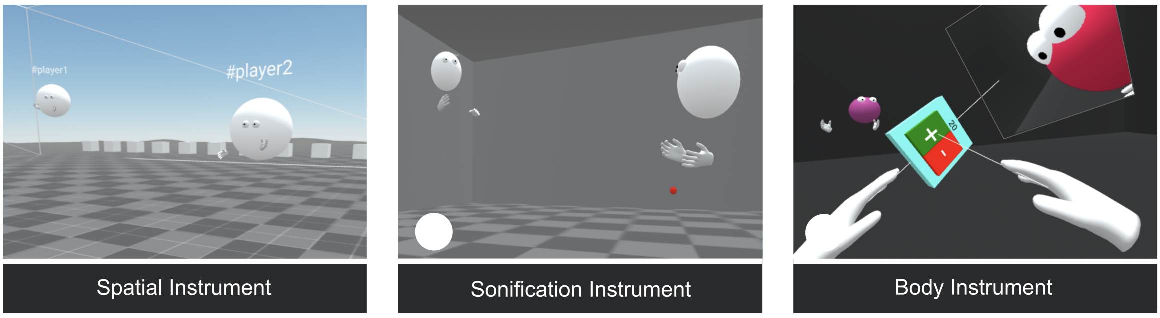


Figure 1: *The three instruments of* It Takes Two.

the avatars and their movements in space are used to control various parameters of a synthesizer, such as a low-pass filter cutoff frequency, frequency modulation, modulation depth, and volume.

3. Body Instrument: This instrument focuses on a more intimate form of interaction: it uses the virtual avatar as the primary interface for collaborative music creation. By touching the avatar's head of a player, a percussive-like sound is produced. However, each user can change the size of their avatar. The bigger the avatar, the lower the sound pitch. This approach aims at blurring the lines between the performer's actions and the musical results.

All instruments were implemented using PdXR [Dziwis et al., 2023], a framework that leverages the Pure Data (Pd) and Networked-AFrame (NAF) [1] for creating interactive, multi-user virtual musical environments in the browser. For audio spatialization, PdXR integrates the Resonance Audio Spatializer for A-Frame [2]. We developed additional JavaScript components to meet the specific requirements of our prototypes, such as controlling spatialized sound sources, processing sonification data, and extending avatar interactions.

3 Conclusion

It Takes Two show three ways for musicians to engage with sound in a collaborative and shared fashion, as well as provide a platform for exploring the boundaries of composition and live performances through VRMIs. Future developments will focus on increasing the complexity of interactions, and expanding the instruments' capabilities to support larger ensembles for richer collaborative experiences.

The source code of the project can be accessed here: `https://github.com/CIMIL/It-Takes-Two`. The instruments were tested successfully on Quest Browser and Wolvic on Meta Quest 2, Meta Quest 3, and Meta Quest Pro. A video of the three instruments can be seen here: `https://youtu.be/d3W0ZxfDjMY`.

[1] `https://github.com/networked-aframe/networked-aframe`

[2] `https://github.com/mkungla/aframe-resonance-audio-component`

4 Acknowledgment

The authors Alberto Boem, Matteo Tomasetti, and Luca Turchet received support from the MUR PNRR PRIN 2022 grant, prot. no. 2022CZWWKP, funded by Next Generation EU. The work by Damian Dziwis and Sascha Etezazi was realized as a part of KreativInstitut.OWL, funded by the Ministry of Economic Affairs, Industry, Climate Action and Energy of the State of North Rhine-Westphalia, Germany.

References

[Boem et al., 2024] Boem, A., Dziwis, D., Tomasetti, M., Etezazi, S., & Turchet, L. (2024). "It Takes Two" - Shared and Collaborative Virtual Musical Instruments in the Musical Metaverse. In *Proceedings of the 5th IEEE International Symposium on the Internet of Sounds, [TO APPEAR]*.

[Dziwis et al., 2023] Dziwis, D., Von Coler, H., & Pörschmann, C. (2023). Orchestra: a Toolbox for Live Music Performances in a web-based Metaverse. *Journal of the Audio Engineering Society*, 71(11), 802–812.

[Jordà, 2005] Jordà, S. (2005). Multi-user instruments: models, examples and promises. In *Proceedings of the 2005 conference on New interfaces for musical expression* (pp. 23–26).

[Turchet, 2023] Turchet, L. (2023). Musical Metaverse: vision, opportunities, and challenges. *Personal and Ubiquitous Computing*.

Demo of *Esteso*: an AI Music Duet Based on Extended Double Bass Techniques

Domenico Stefani[1], Matteo Tomasetti[1], Filippo Angeloni[2], and Luca Turchet[1]
[1] *Department of Information Engineering and Computer Science, University of Trento, Italy*
[2] *Independent - Composer, Musician, Italy*

Abstract

Esteso is an interactive improvisational system for double-bass based on player-idiosyncratic extended techniques. This system was created in collaboration with the contemporary double-bass player and composer Filippo Angeloni and tailored for his personal vocabulary of extended techniques. In *Esteso*, AI agent and the player engage in a duet, taking turns in the performance. The system replies with a manipulation of the real double-bass, achieved live through a timbre-transfer neural network, granular synthesis, and reverb. The timbre-transfer network was trained on a public double-bass dataset, resulting in a peculiar *hybrid* sound. Machine listening is integrated through a classifier of extended techniques played on the double-bass, whose output controls sound processing to affect various techniques differently. We present a demonstration of a performance where the double-bass player interacts with Esteso, creating a back-and-forth interplay between the acoustic and virtual elements.

Figure 1: *Esteso during a live performance. Picture by Alberto Boem.*

1 Project Description

Esteso was presented at the Interational conference on New Interfaces for Musical Expression, Utrecht, The Netherlands ([Stefani et al., 2024]). This demo will however allow attendants to see the inner workings of the system in detail.

The system demoed here is related to the "player" paradigm outlined by Rowe [Rowe, 1992], which defines an "artificial" player able to interact with human players.

Videos of the system during a test session and a live performance can be found online[1,2]. The software can be found on Github[3]. This demo fits the theme of this year's DIMMI as *contamination* can be seen in the multifaceted nature of the sonic output of the system, which steams from the musician's signal, his way of playing, and the sound that the timbre-transfer model learned from generic data. In addition, this represents a multidisciplinary collaboration between a musician and researchers where different interests merge into one output.

The system was implemented as a Cycling'74 Max patch. It is organized into three parts: (1) a technique recognition section, (2) a mechanism to manage the duet-like interplay, and (2) a sound processing stage. The purpose of the recognition section is to detect the use of different extended techniques from the musician and subsequently affect the system's response. Secondly, the duet-mechanism is responsible for the action-reaction nature of the duet, enforcing simple rules that start and stop the system's response. Finally, the sound processing stage is responsible for the sonic nature of the system's response. Figure 3 depicts the architecture.

The technique recognition system is composed of a

[1] www.youtube.com/watch?v=HEhJXAgFiXM
[2] www.youtube.com/live/Vrywo3fpALw at 1:31:00s
[3] www.github.com/domenicostefani/Esteso

Figure 2: *Player interacting with Esteso.*

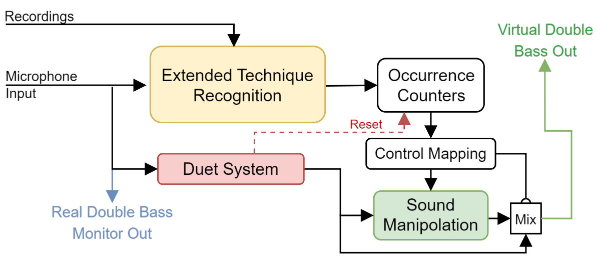


Figure 3: *Architecture of* Esteso.

feature extractor, an onset detector, and a machine-learning classifier. The encoder of a RAVE model was used as a feature extractor, while a simple peakamp object served as an onset detector. Repeated samples from the latent space of RAVE are collected upon onset detection and fed to the classifier. The classifier used was a K-nearest neighbor method. We focused on detecting the three extended techniques (i.e., *"Brushed" Jeté*, *Sfregato con legno*, and *Percussive*) and used the classification result to select different parameters for the sound processing stage.

During the performance, the duet-mechanism enforces a basic rule: *Esteso* starts listening as soon as sound is detected by the microphone; audio is recorded to a short buffer (ten seconds); when the signal from the double-bass is quieter than a set silence threshold for one second, the system feeds the recorded buffer to the sound processing stage. The processed buffer constitutes the response of the agent to the human player. Buffer length and silence thresholds were found through experimentation with the double-bass player.

Finally, the response of the AI agent is obtained through the alteration of the short recordings produced by the duet mechanism. The processing pipeline is composed by a granular synthesizer[4], a timbre-transfer model (RAVE [Caillon & Esling, 2021]), and a reverb effect[5]. The granular synthesizer was chosen as it can morph the temporal structure of input recordings, providing the musician with novel responses. Secondly, the timbre transfer model, trained on generic double-bass recordings, generated unusual sounds that only somewhat resembled an actual double-bass. We chose the model as we felt its sound was interesting and conceptually made it similar to self-sabotaged instruments [Dannemann et al., 2023]. We then used reverb to grant *Esteso* the feel of different acoustic spaces, in contrast with the dry double-bass sound. Selected parameters of the sound processing pipeline are affected by the results of the extended technique recognition classifier. This was mapped through a sound design process, where dry double-bass recordings were fed through the effect pipeline to parameterize it in different ways that would highlight each technique. Therefore, sets of parameter values were mapped to each possible output of the classifier.

Apart from the performance proposed here, the system was formally evaluated through three performance sessions aimed at understanding the musician's reaction and testing different sound processing mappings. New Interfaces for Musical Expression (nime) practices and techniques were adopted in the design of the experiments, the formal evaluation, and the extraction of themes from the musician's comments.

[4]The Max/MSP package "Petra": https://github.com/CircuitMusicLabs/petra

[5]"Reverb-2" from the BEAP module package, by Matthew Davidson: https://github.com/stretta/BEAP

References

[Caillon & Esling, 2021] Caillon, A. & Esling, P. (2021). RAVE: A variational autoencoder for fast and high-quality neural audio synthesis. *CoRR*, abs/2111.05011.

[Dannemann et al., 2023] Dannemann, T., Bryan-Kinns, N., & McPherson, A. (2023). Self-sabotage workshop: a starting point to unravel sabotaging of instruments as a design practice. In *Proceedings of the International Conference on New Interfaces for Musical Expression* Mexico City, Mexico.

[Rowe, 1992] Rowe, R. (1992). *Interactive music systems: machine listening and composing*. MIT press.

[Stefani et al., 2024] Stefani, D., Tomasetti, M., Angeloni, F., & Turchet, L. (2024). Esteso: Interactive ai music duet based on player-idiosyncratic extended double bass techniques. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME'24)*.

Extended Reality and Movie Theater Experiences: the Sound Liveliness Problem

Stefania Colonnese[1a], Paolo Simonetti[2], Andrea Baiocchi[1a] and Paola Panarese[1b]

[1] *Sapienza University of Rome, Italy, [a]DIET Dept., [b]CoRiS Dept.*
[2] *St. Louiss College of Music, Rome, Italy*

Abstract

Among different multimedia services, cinema provides unique atmosphere experiences, inviting viewers to engage deeply in a darkened environment while also allowing for a Brechtian sense of distancing. New technologies, particularly streaming on Extended Reality (XR) devices, dramatically transform the vibrancy of the movie theater experience. Although extensive research has explored theater and musical experiences in XR, the literature often neglects the potential and limitations of watching films within an XR context. This paper focuses on a crucial aspect of adapting films for XR devices: the role of sound and music. After reviewing both the literature and the relevant technological frameworks, we explore the potential of stereo sound spatialization through a case study featuring a small virtual orchestra, created using a Digital Audio Workstation, and comparing different audio formats of the same musical piece. Time-frequency analysis of multiple virtual recordings shows that stereo speakers used with XR devices can deliver richer spatialized sound and improved isolation from ambient noise. XR devices introduce an unmatched level of liveness for digital cinema. Future research will aim to design a campaign for evaluating the quality of experience, comparing XR-based viewing to the distinctive cinema experience.

1 Introduction

Cinema exerts a timeless allure, with some of its features summarized in [Barthes, 1975], where the author observes that the darkness, the relaxation of the body, which ultimately synthesizes a "cinema situation", is pre-hypnotic, and the blackness of the room induces a "twilight reverie". Therein, the author also observes that this pre-hypnotic impact is accompanied by other "awakening"aspects: the grain of the sound, the room, the blackness, the dark mass of other bodies, the rays of light, the entry, the exit. In this sense, cinema offers the possibility of the so called Brechtian distancing.also called the "Verfremdungseffekt" or "alienation effect". This is a theatrical technique developed by German playwright Bertolt Brecht, to prevent the audience from becoming emotionally absorbed in the narrative, and to foster critical awareness in the audience, highlighting theater's potential as a tool for social change.

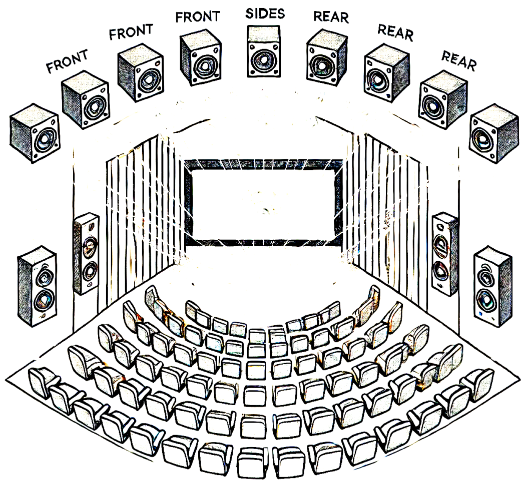


Figure 1: *Movie theater spatial sound.*

It is relevant to ask whether new technologies like Extended Reality (XR) allow for modes of consumption that preserves the liveliness of the movie theater representation, let alone the critical distance. While extensive literary studies have been dedicated to theater and music experience, and even to the critical distance, the literature does not explore the potential and limitations of XR in replacing or complementing movie theater performances. Among the variety of as-

Table 1: *Comparison of Audio-Visual Devices.*

Device	Audio Speakers	Spatial Sound	Resolution/eye	Frame Rate	Bitrate
Movie Theater	Up to 16	Multi-channel	∼4K×4K	50–60 Hz	∼200–500 Mbps
Oculus Quest 2	2	Stereo	1832×1920	60–90 Hz	∼10–100 Mbps
HoloLens 2	2	Spatializer SW	1440×936	>60 Hz	∼10–100 Mbps
Apple Vision Pro	6	Built-in	∼4K×4K	90 Hz	∼20–100 Mbps

pects that characterize movies adaptation for fruition on XR devices, this paper addresses the impact of sound and music.

After a review of both the literature and technological framework, we analyze and compare the main sound features in movie theater and on XR devices. We present examples where we produce spatialized sound and we analyze the impact of its adaptation to stereo reproduction. Based on this, we examine the main methodological aspects for a systematic analysis of user experience, discussing the potential of an XR device in sound reproduction.

2 Bridging Movie Theater to XR

Movie theater representations thrive on rich acquisition, delivering and displaying technology. A technical overview of the main features is reported in the Digital Cinema System Specification [dcs, 2024] defined in the framework of the Digital Cinema Initiatives [Digital Cinema Initiative Consortium, LLC, 2002]. The literature lacks an in depth analysis of the technical and experiential changes that arise when the content is viewed either in a movie theater or on an XR device. Table I reports the main technological features of different supports, including resolution, bitrate, audio channels. The audio is one of the most critical issues. In the next section we focus the analysis on the audio component.

XR performances raise many challenges, let alone the technical ones [Chakareski, 2020]. In [Brilli et al., 2023], the authors examine how digital technologies affect theatre's presence and audience engagement, with focus on online theatre initiatives and their drawbacks on presence performances. Liveness in music, where the mediatization started with recording technology, enforces authenticity in musical performance in an increasingly wider audience [Auslander, 2022] and the music industry develops new platforms while conventional media experience an unrelenting diffusion growth [Geo]. The work in [Bay, 2023] investigates the impact of widespread adoption of XR on democracy. Companies that produce content for social media are now addressing XR technologies with personalized content, and this may affect the creation of memories without clear boundaries between real and virtual experiences. The work in [Baía Reis & Ashmore, 2022] discusses in a dialogic form a variety of social experiences, including theater performances, in the metaverse and more generally in online frameworks. A special role is played by music in XR [Weinel, 2020], and the wide variety of technologies and performances nowadays available are reviewed in [Turchet et al., 2021].

3 The Spatial Audio Technology

Digital audio signals, sampled at 44.1 KHz or 96 KHz and represented with 12-16 bits per sample, are the basis for broadcasting and streaming over fixed and mobile networks. Since the late 1980s, encoding techniques have aimed to reduce bit rates to around one bit per sample while maintaining perceived quality by minimizing perceptual redundancy. Audio signals are efficiently encoded by dividing them into frequency sub-bands, reflecting the human ear's sensitivity to different frequencies. Masking, where stronger tones hide weaker ones or louder sounds obscure quieter ones, allows for discarding inaudible components. Since the late 1990s, multichannel audio standards have evolved, supporting both natural and synthetic audio across various bit rates and complex multimedia applications. In multichannel audio encoding, channels are mixed before encoding, allowing backward compatibility with stereo decoders while preserving additional channels for advanced decoders. Masking effects in mixed channels can disappear after decoding, causing distortion in individual channels, and care must be taken during mixing to account for potential unmasking. Multichannel technology boosted the interest in the spatialization of sound, continually sought by composers and engineers since the inception of analog and digital music technology.

Early recordings were produced in mono, where the audio was captured from a single source, typically with one microphone, resulting in no sense of spatial directionality. Mono recordings dominated the early 20th century and were the standard for both phonograph and radio until well into the 1950s. These recordings conveyed a clear and focused sound but lacked the ability to provide listeners with any sense of depth or location in the sound field. The shift from mono to stereo began in earnest with technological advances in the 1930s and 1940s. The first experiments in stereo sound were conducted by British engineer Alan Blumlein in 1931. Blumlein, working for EMI, not only developed stereo recording techniques but also created the concept of stereo reproduction, where separate left and right channels give the listener a perception of directionality. His work laid the foundation for what would later become a revolution in sound recording and reproduction. By the late 1950s, stereo sound became commercially viable and began to replace mono as the standard for music recording and playback. Stereo recording divides the sound field into two channels—left and right—creating the impression of spatial width and giving the listener a more immersive experience. The Beatles' early albums, for example, were released in both mono and stereo, illustrating the transitional period in popular music. With the advent of fast tools for digital processing, the development of spatial audio has moved beyond simple stereo. Today, algorithms allow us to simulate the positioning of instruments within complex environments, such as concert halls or cathedrals, replicating the way sound interacts with the space around it. These technologies model reflections, reverberations, and acoustic properties, providing a highly realistic auditory experience that conveys both spatial depth and detail. Even in XR sound generation, spatialization has been addressed to enhance immersivity, using detailed models of the user's head anatomy [Cho et al., 2024, Berger et al., 2018]. In the following, we present a proof of concept relating to the cinematographic audio experience and its possible adaptation to XR devices.

4 A Stereo Sound Design Case Study

In this study, we explore the potential of stereo sound spatialization by comparing different audio representations of the same musical composition by a small virtual orchestra produced using a modern Digital Audio Workstation (© Parallax Audio). The orchestra consists of a collection of multi-sampled instruments. The spatial arrangement, illustrated in Fig. 1, is as follows: on the left, the red points represent four violins, and the two orange points are the violas; in the center, the yellow points represent one cello and one upright bass; on the right, the green points represent a set of woodwinds—two horns, one bassoon, one flute; finally, a piccolo shown in white. The central gray circle and the two side circles represent the directional responses of the mono and stereo microphones, respectively.

Firstly, the traces of the single instruments were produced. Then, three registration environments were simulated, considering the cases of a) a single microphone, b) a pair of microphones in an open, anechoic ambience, and c) a pair of microphones in an echoing environment, as it would be heard in the

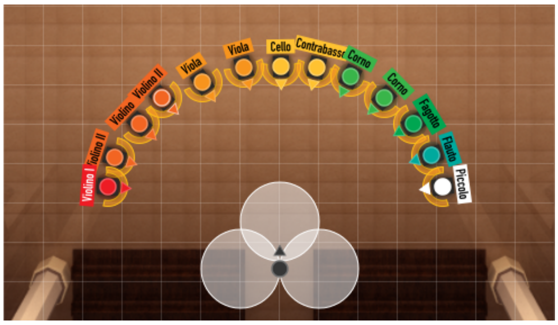
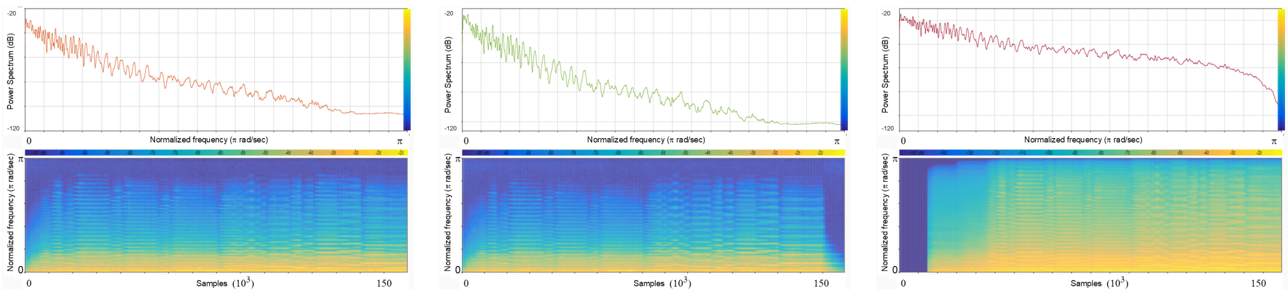


Figure 2: *Geometry of the sound generation. The strings appear as red, orange and yellow points; the woodwings appear as green to white points.*

acoustical space of a church. For case c), we used an algorithmic reverb, an approach that uses mathematical models to simulate how sound behaves in a physical space, based on user inputs such as room size, relative instruments and microphones position and surface material. We modeled a stage of 14mx19m in a 24x29 space with 15m height; the reverb time was set between 3 to 4s; the pre-delay (time to first sound reflection on the wall) was set to 60-80ms. The three recordings were then saved in stereo format, since this is the format adopted in most XR devices.

In Fig. 2 (a), we see the time-frequency representation of the signal measured on a single microphone placed centrally in relation to the virtual ensemble: the signal occupies wide bands across the entire audible spectrum. In Fig. 2 (b), we refer to the stereo version of the signal, recorded in a virtual anechoic environment. The graph shows the differences between the right and left channels, more noticeable at low frequencies. Finally, Fig. 2 (c), shows the left and right differences in the presence of virtual echoes generated according to the model described above: the difference spectrum significantly expands in bandwidth, indicating that the stereo representation, like human binaural representation, captures the richness of space sound.

To sum up, using stereo speakers associated with XR devices offers the potential for enriched spatialized sound listening. Additionally, compared to television playback, listening is enhanced by the potential isolation of the speakers from background noise. This, combined with highly immersive visuals, provides the potential for elevated "liveness". XR devices could offer an unprecedented level of liveness for digital cinema, which is known to have a unique subjective impact on viewers. This opens the door to designing campaigns for subjective quality of experience assessments, comparing fruition on XR devices to the unique cinema experience.



(a) **(b)** **(c)**

Figure 3: *Spectrogram for the mono channel (a) and for the L/R stereo channels differences in open ambience (b) and in a virtually recreated echoing environment (c).*

5 Conclusions

Cinema has a timeless appeal, but adapting content to new technologies like streaming on Extended Reality (XR) devices significantly alters the liveliness of the traditional movie theater experience. While substantial research has been conducted on theater and musical performances in XR, the literature has yet to fully explore the possibilities and limitations of viewing films in an XR environment. In this study, we examine the potential of stereo sound spatialization through a case study involving a small virtual orchestra, produced using a Digital Audio Workstation, and comparing different audio representations of the same musical composition. Time-frequency analysis of various virtual recordings demonstrates that stereo speakers paired with XR devices offer enhanced spatialized sound and better isolation from background noise. XR devices provide an unprecedented level of liveness for digital cinema. Future work will focus on designing a campaign to assess the quality of experience, comparing XR device viewing to the unique cinema experience.

6 Acknowledgments

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on “Telecommunications of the Future” (PE00000001 - program “RESTART”).

References

[dcs, 2024] (2024). *The Digital Cinema System Specification (DCSS)*. Technical report, DCI Spec.

[Auslander, 2022] Auslander, P. (2022). Liveness and discourses of authenticity in music.

[Barthes, 1975] Barthes, R. (1975). *En sortant du cinéma*, volume 23.

[Bay, 2023] Bay, M. (2023). Arendt in the metaverse: Four properties of extended reality that imperil factual truth and democracy. *Convergence*, 29(6).

[Baía Reis & Ashmore, 2022] Baía Reis, A. & Ashmore, M. (2022). From video streaming to virtual reality worlds: an academic, reflective, and creative study on live theatre and performance in the metaverse. *International Journal of Performance Arts and Digital Media*, 18(1).

[Berger et al., 2018] Berger, C., Gonzalez-Franco, M., Tajadura-Jiménez, A., & Florencio, D. (2018). Generic hrtfs may be good enough in virtual reality: Improving source localization through cross-modal plasticity. *Frontiers in Neuroscience*, 12, 21.

[Brilli et al., 2023] Brilli, S. et al. (2023). Theatre without theatres: Investigating access barriers to mediatized theatre and digital liveness during the covid-19 pandemic. *Poetics*, 97, 101750.

[Chakareski, 2020] Chakareski, J. (2020). Viewport-adaptive scalable multi-user virtual reality mobile-edge streaming. *IEEE Transactions on Image Processing*, 29, 6330–6342.

[Cho et al., 2024] Cho, H., Wang, A., Kartik, D., Xie, E., Yan, Y., & Lindlbauer, D. (2024). Auptimize: Optimal placement of spatial audio cues for extended reality. *arXiv preprint*, arXiv:2408.09320.

[Digital Cinema Initiative Consortium, LLC, 2002] Digital Cinema Initiative Consortium, LLC (2002). Digital cinema initiative. https://www.dcimovies.com.

[Turchet et al., 2021] Turchet, L., Hamilton, R., & Çamci, A. (2021). Music in extended realities. *IEEE Access*, 9.

[Weinel, 2020] Weinel, J. (2020). Cyberdreams: Visualizing music in extended reality. In *Technology, Design and the Arts - Opportunities and Challenges*.

Retro Sounds, Modern Aggression: Unpacking Nintendocore

Andrea Spontoni[1] and Enrico Dorigatti[2]
[1]*Independent Researcher*
[2]*University of Portsmouth, School of Creative Technologies, Portsmouth, UK*

Abstract

Early game consoles such as the Game Boy have become symbols of the recent past, defining an aesthetic scope grounded in low-fidelity (lo-fi). While early video games music was shaped by technical limitations, with composers creatively employing the limited affordances of sound chips, today, artists deliberately use these retro sounds, blending them with contemporary music styles and evoking a sense of nostalgia. This paper, acknowledging the limited scholarship on the topic presented and drawing from the extensive experience of the first author as a practitioner in the field under the moniker *Cowboy Bitpop*, aims to introduce and frame the scope of interest while outlining its breadth. Nintendocore is specifically investigated as a prominent musical thread resulting from the contamination between modern instrumental music genres and chip-derived sonorities.

1 Introduction

Early game consoles were characterized by the presence of sound chips equipped with a set number of monophonic channels. Composers creating music to be played on such consoles would employ the channels available as instruments, typically reserving one or more pulse-wave channels for the melody, a triangle wave channel for the bass, and a noise channel for percussive sounds [Schartmann, 2015].

Currently, the Game Boy is among the most renowned consoles for music-making, and most of its versions are equipped with sound chips providing two pulse channels, one noise channel, and one wave channel for building custom waveforms and loading samples. Other consoles, such as the Sega Genesis equipped with a Yamaha YM2612 sound chip, allow for more complex sound synthesis techniques, such as frequency modulation (FM) synthesis [Chowning, 1973]. While the distinctive timbre of early video games music was due to technical limitations rather than artistic choices [Schartmann, 2015], nowadays, artists deliberately explore the sonic features of early consoles [Carlsson, 2017, Yalvaç, 2019].

1.1 Chiptune

The music realised by means of consoles or by following the aesthetic of early video games soundtracks is primarily known as chiptune [O'Leary, 2019, Tau, 2022]. This music genre has a strong connection with retro gaming culture, internet culture, and lo-fi aesthetic [Monacelli, 2023]. As it evokes a sense of nostalgia in the listeners, sometimes it is also used for commercial purposes [Suominen, 2008].

1.2 Bitpop

Bitpop is used within the chiptune context as an umbrella term for music releases that combine sound chips sonorities with other instruments through modern music production techniques [Sundiam, 2019].

2 Metalcore

Metalcore comprises many subgenres, each distinguished by specific stylistic traits [Smialek, 2023]. Common features that metalcore bands typically exhibit are the drop tuning of guitars, frequent tempo changes, and breakdowns [Gamble, 2019], namely a dramatic slowing of the tempo.

3 Nintendocore

Among the styles within the bitpop scope, this paper focuses on Nintendocore. Frequently, Nintendocore blends musical elements from metalcore and post-hardcore [Tau, 2022] with chiptune sounds, sometimes sampled from retro video games soundtracks. Nintendocore as a music genre-defining term comes from an interview where the singer of *Horse The Band* used it to describe the music of the ensemble [Márquez, 2012]. Some influential Nintendocore artists are *As the World Fades*, *I Am Error*, *Dungeon Elite*, *A Parade of Bleeding Bullets*, *Fagasaurus Sex*, and *Ditto's Blood*. Given that Nintendocore is an underground trend, artists often have limited recognition and are primarily known within the local music scene or specialised online communities.

The styles of Nintendocore artists are nuanced, usually but not exclusively drawing upon some of the most prominent metalcore features as outlined in 2. Beyond metalcore, such features are often common to other heavy music genres, such as metal, another scope from which Nintendocore artists draw to create unique stylistic and aesthetic contaminations. For instance, *Math The Band* and *Horse The Band* employ sounds and elements closer to post-hardcore; *The Minibosses*, *The Megas*, *Powerglove*, and *Vomitron* reinterpret retro video games soundtracks in different metal styles; finally, *I Am Error*, *I Shot The Duck Hunt Dog*, and *100 DEAD* mix chiptune sonorities with grindcore. It is also worth mentioning that contaminations involving chiptune are not limited to metal(-core), but may embed elements of other genres. For instance, *Bondage Fairies* and *FUCKING WEREWOLF ASSO* incorporate features close to synth or pop-punk.

Due to the limited scholarship on the matter and the extreme versatility and openness of Nintendocore in terms of novel explorations and contaminations, nuances characterising subgenres within Nintendocore may be difficult to spot [Marquez, 2014]. In addition, a framework may fail the objective of establishing an effective and comprehensive Nintendocore taxonomy. Specialised online communities play a substantial role in this as they forge and diffuse new and arbitrary labels with which to identify musical experiments fitting the broader Nintendocore scope. It is also worth mentioning that effectively recognising different Nintendocore subgenres may be difficult as it requires a broad musical and listening experience [Phillipov, 2012] in relation to the musical contexts intersected by Nintendocore artists, such as metal and metalcore music.

Sometimes, Nintendocore features fragments of early soundtracks, commercials, or viral internet videos. Examples can be found in music pieces by *As the World Fades* and *Nekonomicon*. For instance, the former uses the death music theme from the Super Mario franchise *One More Game, Please Mom!*[1], while the latter makes use of sound samples from commercials in *Those Creatures from Ganon are Pretty Bad*[2].

[1] *One More Game, Please Mom!* by *As The World Fades*: https://www.youtube.com/watch?v=YbTKwghJEps (last accessed: October 22, 2024).

[2] *Those Creatures from Ganon are Pretty Bad* by *Nekonomicon*: https://www.youtube.com/watch?v=GHHYj3Wv2bY (last accessed: October 22, 2024).

3.1 Nintendocore Bands and Fakebit

Nintendocore is usually played by ensembles inspired by the lineup of traditional rock or metal bands: electric guitar, electric bass, drums, and vocals, often distorted through scream or growl singing [Tau, 2022]. These bands are accompanied by chiptune sonorities emulating the sounds of early consoles, often used for melodic lines or solos. Using original hardware, however, is not strictly necessary, as chiptune sounds can be obtained through software synthesizers [McAlpine, 2018], a strategy known as fakebit [Polymeropoulou, 2014]. Recreating retro console sonorities through a fakebit approach unlocks the potential of modern software synthesizers, detaching music production from the limited affordances of early hardware and allowing for a broader range of sound synthesis and manipulation techniques. Given the versatility of modern sound production tools, and considering that a substantial contribution to defining the iconic aesthetic and character of chiptune resulted from the limitations of early sound chips [Magnusson, 2010], it is the prerogative of the fakebit artist to determine how faithful to the sonorities of a specific game console the result will sound.

Although original or recreated chiptune sonorities are usually used with melodic or soloistic roles, they are also layered upon other acoustic or electric instruments in some circumstances. Often, this involves percussive sounds obtained through a genuine or recreated noise channel, which are then stacked upon recorded acoustic drums during production. This approach is frequently explored by artists who blend chiptune with music genres less heavy than metal and metalcore, as in these genres, the sound of a real drumkit is usually preferred. *Bondage Fairies*, for instance, have made extensive use of layering.

4 Conclusions

Nintendocore taps into the nostalgia of retro gaming culture while incorporating features of other musical genres, most notably the aggressive energy and heavy sonorities of metalcore, post-hardcore, and other subgenres of rock and metal. Sounds once defined by hardware constraints are now purposefully explored, recreated, and merged with features of different music genres, giving rise to a diverse range of musical fusions. Chiptune artists deliberately engage in the creation and experimentation with innovative sonic palettes and mixtures based on sonic elements characterising chiptune music. Nintendocore stands as a testament to how digital culture, artistic experimentation and innovation, and retro influence can converge, creating new, dynamic, and to some extents elusive forms of artistic (sonic) expression that transcend traditional and established boundaries and definitions.

References

[Carlsson, 2017] Carlsson, A. (2017). Chip music: low-tech data music sharing. In *From Pac-Man to Pop Music* (pp. 153–162). Routledge.

[Chowning, 1973] Chowning, J. M. (1973). The synthesis of complex audio spectra by means of frequency modulation. *Journal of the audio engineering society*, 21(7), 526–534.

[Gamble, 2019] Gamble, S. (2019). Breaking down the breakdown in twenty-first-century metal. *Metal Music Studies*, 5(3), 337–354.

[Magnusson, 2010] Magnusson, T. (2010). Designing constraints: Composing and performing with digital musical systems. *Computer music journal*, 34(4), 62–73.

[Marquez, 2014] Marquez, I. (2014). Playing new music with old games: The chiptune subculture. *G| A| M| E Games as Art, Media, Entertainment*, 1(3).

[Márquez, 2012] Márquez, I. V. (2012). Nostalgia videolúdica: un acercamiento al movimiento chiptune. *Trans. Revista Transcultural de Música*, (16), 1–21.

[McAlpine, 2018] McAlpine, K. B. (2018). *Bits and pieces: A history of chiptunes*. USA: Oxford University Press.

[Monacelli, 2023] Monacelli, E. (2023). *The Great Psychic Outdoors. Lo-Fi Music and Escaping Capitalism*. London, UK: Repeater Books.

[O'Leary, 2019] O'Leary, J. (2019). A corpus-assisted discourse analysis of chiptune-related practices discussed within chipmusic. org. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (pp. 1294–1294).

[Phillipov, 2012] Phillipov, M. (2012). *Death metal and music criticism: Analysis at the limits*. Lexington Books.

[Polymeropoulou, 2014] Polymeropoulou, M. (2014). Chipmusic, fakebit and the discourse of authenticity in the chipscene. *WiderScreen*, 1.

[Schartmann, 2015] Schartmann, A. (2015). *Koji Kondo's Super Mario Bros. Soundtrack*. Bloomsbury Publishing USA.

[Smialek, 2023] Smialek, E. (2023). Contempt-of-core: A reception history of metalcore subgenres as abject genres. In *The Cambridge Companion to Metal Music* (pp. 281–298). Cambridge University Press.

[Sundiam, 2019] Sundiam, J. (2019). *Video Game Music and Legitimacy* [Bachelor's thesis, California State University].

[Suominen, 2008] Suominen, J. (2008). The past as the future? nostalgia and retrogaming in digital culture. In *Proceedings of perthDAC2007. The 7th International Digital Arts and Cultures Conference. The Future of Digital Media Culture*, volume 15 (pp.1̃8).

[Tau, 2022] Tau, M. (2022). *Extreme Music: Silence to Noise and Everything In Between*. Feral House.

[Yalvaç, 2019] Yalvaç, H. E. (2019). *Creating Music "Inside The Box": Do Chiptune's Limitations Have Aesthetic And Creative Value?* [Master's thesis, Istanbul Technical University, Institute of Social Sciences].

Prototyping a Larsen Glass-harmonica: Concept and Setup of a Relatively Stable Instrument

Maurilio Cacciatore

Conservatory "A. Steffani" of Castelfranco Veneto, Italy

Abstract

The exploration of acoustic feedback has been a key element in my compositions for many years. This research establishes a protocol for generating and controlling feedback using a system of sound exciters applied to the piano soundboard. A set of glasses has been used to create a tuned glass harmonica. By incorporating both traditional and advanced playing techniques, the prototype bridges the gap between electroacoustic instruments and their integration into instrumental music composition with electronics. The prototyping of the *Larsen Glass-harmonica* focuses on creating a stable, functional instrument. This involves refining the design to ensure reliable performance while allowing for a wide range of tuning possibilities and interactions with other instruments or digital processes.

The setup integrates glass containers of varying sizes, a brass frame, and strategically placed sound exciters, all interacting with a piano soundboard or other resonating bodies. Microphone placement and handling techniques are key to capturing and manipulating the acoustic feedback, while live electronics and signal processing can further expand the instrument's sonic palette.

The goal is to strike a balance between flexibility for experimental applications and the stability needed for practical performance use.

1 Introduction

Since *Lost in Feedback*[1] (2014), a piece for electric vibraphone, stage performer, and live electronics, I have been working on controlling acoustic feedback as a key element in my music. Over time, this exploration has expanded into the development of instrumental prototypes, software controls, custom diffusion setups, and the creation of new music fonts.

This research was supported by the Conservatory Agostino Steffani of Castelfranco Veneto during the 2022/23 academic year. It builds on the outcomes of a 2020 residency with the Syntax Ensemble[2] at the Festival Milano Musica[3], where I worked on optimizing software for sound exciters applied to acoustic instruments.

Initially, my goal was to identify the most effective positions on various grand pianos for intentionally generating acoustic feedback using vibrating loudspeakers (also called sound exciters or contact loudspeakers) placed against the soundboard[4]. This investigation has since evolved into controlling the pitch of feedback, leading to the design of a new prototype for an acoustic musical instrument.

2 Diffusion Setup

The diffusion setup I used for this project consists of two contact loudspeakers of 40W each and a bass shaker. It was developed through extensive testing on the piano soundboard. After evaluating various configurations by recording white noise, the most effective setup was determined as illustrated in Fig. 1.

This placement was chosen for its ability to produce the richest spectrum. However, the exact position of the exciters may vary slightly depending on the piano model due to differences in the layout of

[1] Commission of the ensemble Hanatsu miroir, Strasbourg. Edizioni Suvini Zerboni, Milan, ESZ 147788.

[2] https://syntaxensemble.com

[3] Sponsored by Regione Lombardia. https://milanomusica.org

[4] https://www.daytonaudio.com/category/180/exciters-tactile-transducers

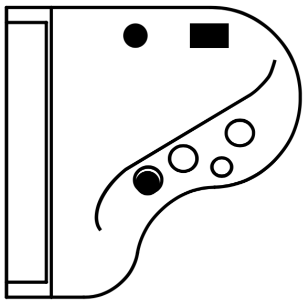


Figure 1: *Sound exciters placement.*

stress bars and resonance holes. Nonetheless, a similar placement close to the optimal setup can generally be achieved across different models.

The 40W exciters positioned on the left side of the piano soundboard can be doubled with an additional device if needed, depending on the size of the concert hall. For this project, I used Dayton Audio exciters, secured with Schertler putty.

These passive exciters require external pre- amplification to broadcast the output from an audio interface. The connection schematic is shown in Fig. 2. The output volume should be adjusted according to the acoustics of the performance space. Sound exciters have a custom frequency response and phase behavior depending on impedance levels.

A specialized frequency compensation is required before the sound output. A Gen -based Max package[5] for optimizing sound exciters was developed by Simone Conforti and Alberto Gatti during the Syntax Ensemble's residency, supported by the Festival Milano Musica.

3 Sound Capturing

A wide range of microphones was tested to generate acoustic feedback, with the most promising results achieved using the DPA 4060, a miniature microphone with an omnidirectional pattern. This choice not only meets the need for a small, easy-to-handle device but also facilitates the generation of short fade-in and fade-out acoustic feedbacks, allowing the performer greater control.

However, despite its features, controlling the feedback was initially challenging without additional precautions. The microphone capsule's sensitivity made it difficult to generate feedback at a sustainable loudness, both for the performer and the audience. The use of a radio system like the Sennheiser XSW 1 ME2 significantly improved performance, thanks to its limited frequency range and the internal DSP of the wireless system.

[5] You can download the Syntax package at: https://github.com/SICO-SonicIrrationalComputedOrigami/SyntaxMaxPackage

By sending the signal from the radio transmitter at an amplitude of -30 dB and applying software compression with a threshold of -11 dB and a ratio of 50, the resulting acoustic feedback achieves a loudness comparable to that of a chamber music instrument. According to the product specifications, the radio receiver's audio frequency response experiences a dynamic loss of 3 dB. As a result, the audio interface receives the signal with a -33 dB attenuation. The device setup for capturing and diffusing sound follows this schema:

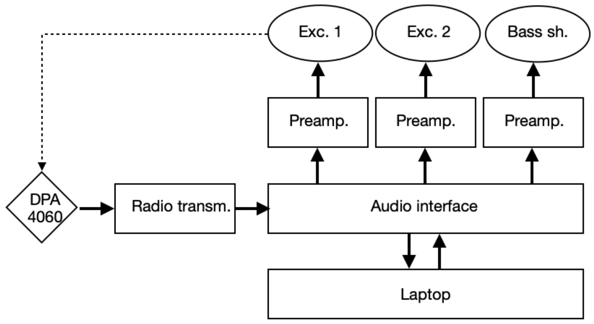


Figure 2: *Complete device setup.*

4 The Larsen Glass-harmonica

Regardless of the piano model, feedback with a clear pitch can be generated from various angles of the soundboard. The feedback range in all cases extends from approximately E3 to G7. These tests were conducted on various models of Yamaha, Kawai, and Steinway & Sons pianos, including Quarter, Baby Grand, and Grand Pianos. For each model, two or three sessions were completed. When repeating tests on the same piano model, the tuning remained consistent. However, while the setup confirmed consistent tuning for each piano model, every model seemed to have its own unique feedback characteristics. This is likely due to variations in the shape of the brass frame and the structure of the stress bars, which greatly limit the use of acoustic feedback in chamber music where predictable and deterministic outcomes are essential for score writing.

While, in theory, tuning could be calculated by measuring the distance between the microphone and the sound exciters, this approach proves impractical for live chamber music performances. Achieving precise tuning by positioning the microphone near the stress bars or other parts of the piano body would require a complex series of measurements that are difficult to apply in a standard concert setting. The diverse shapes of piano soundboards on the market, combined with the personalized body positioning of each performer, make consistent control nearly impossible.

The unpredictability of tuning when interacting with the piano's brass frame through audio devices led me to develop an additional setup to address this issue. I utilized Bormioli jar glasses to construct a series of Helmholtz resonators. By applying the same setup used for generating and controlling acoustic feedback on these glasses, I was able to stabilize the pitches.

The glasses were positioned on a music stand near the stress bars, close to the performer's body. I used three sizes of glasses: 100 cl, 50 cl, and 25 cl. The pitch of each glass is determined by the amount of water inside. According to Helmholtz resonator theory, for cylindrical and rectangular necks like Bormioli jar glasses, the cross-sectional area of the neck follows the formula:

$$f = c \cdot 2\pi AV \cdot L \quad f = \frac{c}{2\pi}\sqrt{\frac{A}{V \cdot L}} \quad f = 2\pi cV \cdot LA \tag{1}$$

where A is the cross-sectional area, V is the volume of the neck, and L is the effective neck length with end correction. Taking into account the mass density and the speed of sound in a gas, the resonance frequency in a closed tube is:

$$f = c4L \quad f = \frac{c}{4L} \quad f = 4Lc \tag{2}$$

I discovered that the pitch calculated from this formula applies when the microphone is positioned at the center of the glass. Another stable pitch can be produced when the microphone is aligned with the neck hole. Additionally, with the 100 cl glasses, a third pitch is achievable by placing the microphone as far inside the cavity as possible. For the 100 cl glasses, varying water levels allow for multiple tunings. For smaller glasses, I focused on the most significant water levels: 1⁄4, 1⁄2, and 3⁄4 full. From these experiments, I compiled the following chart:

Figure 3: *Jars pitches chart.*

This chart was developed for Baby Grand Pianos. The tuning described may vary by approximately a half tone higher or lower depending on the piano model: half a tone lower for Grand Coda Pianos, and half a tone higher for Quarter Coda Pianos. The resulting pitches are influenced not only by Helmholtz resonator theory but also by the distances between the microphone, glasses, and sound exciters. The positioning of the glass series on the stress bars is crucial for achieving optimal sound. I tested this setup outside of a piano, but the sound quality was poor due to the absence of the resonances provided by the piano's soundboard. The best results occur when the piano's resonance pedal is pressed continuously, allowing the soundboard to resonate freely. According to the chart of Fig. 4, several combination of glasses are possible. The closest combination to a chromatic scale can be:

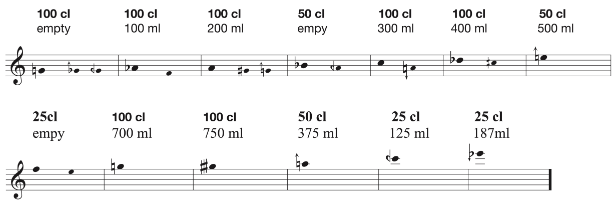


Figure 4: *Jars arranged by tuning.*

Although, theoretically, a small increase or decrease in the water level should result in a proportional change in the tuning, in practice, this is not the case. Instead, the tunings seem to be drawn to specific pitches, almost as if a magnetic-like force is pulling them into place. I've also observed that the tuning of the sound produced by a glass can be influenced by playing one of the corresponding notes on a piano keyboard. This effect is particularly noticeable with 100 cl glasses. The arrangement of the glasses on the stress bar can resemble the setup shown in the image of Fig. 5:

Figure 5: *Exemple of jars arrangement.*

Based on observations made while testing the tuning of the acoustic feedback, placing the microphone near the corners of the stress bars revealed that the positioning of the glasses on other (stable) parts of the brass frame can yield different tunings. As I previously mentioned, the resulting pitches depend on the distance from the sound exciters, which, in turn, are influenced by the size of the piano and the shape of the brass frame. The prototype also allows for some flexibility in experimenting with the piano model used for performance, which can help expand the range of tunings based on the parameters described earlier. This testing margin could further enrich the palette of available tunings.

5 Microphone Handling

This prototype offers a range of techniques for producing standard sounds or other desired effects. The handling of the DPA 4060 microphone can be applied in three main positions:

1. DPA 4060 aligned with the neck of the glass;
2. DPA 4060 at the middle of the glass;
3. DPA 4060 as far inside the glass as possible.

Each note can be produced individually. The attack and release times of the sound envelope are approximately 100 ms, with no significant difference between sounds in the low or high register. It is possible to transition smoothly (legato) between two notes produced by the same glass. The time it takes to transition between two pitches generally ranges from 200 ms to 500 ms. During these transitions, a middle sound with a timbre resembling ring modulation may occur.

Another technique involves holding the DPA 4060 cable between the middle and ring fingers With the palm facing upward. When the microphone is positioned close to the sound exciter (with the player's arm near the body, towards the piano soundboard), a higher pitch emerges, depending on the proximity of the player's body to the sound exciters. Raising the fingers at a 90° angle from the palm dampens the sound until it stops.

A highly effective wah-wah effect can be achieved by quickly opening and closing the glass. The resulting pitch corresponds to the one produced when the microphone is positioned without fully closing the neck opening. The pitch can vary depending on the position of the microphone, as illustrated in the chart in Fig. 3. All of these techniques work significantly better when the piano's sustain pedal is pressed.

The future direction of this project will follow its compositional applications. If the use of the Larsen Glass-harmonica continues to be integrated with the piano soundboard to capture the instrument's resonance, more precise calculations will be necessary for each major piano brand.

Exploring the application of the Larsen Glass- harmonica on other instruments, such as timpani or bass drum, could yield interesting results. Additionally, the development of a stand-alone instrument with a custom resonating body and a resonance pedal is another promising direction for further exploration.

Benefits of Interdisciplinary Contamination in Complex Encoding Systems Research

Giulio Velardo[1,2]
[1]*Conservatory "F. A. Bonporti" of Trento and Riva del Garda, Italy*
[2]*Department of Psychology and Cognitive Science, University of Trento, Italy*

Abstract

"Encoding systems" can be defined as ways to translate data or information in symbolic or abstract form, which can be decoded only by having access to the encoding rationale. According to said definition, it is possible to consider any form of structured and abstract communication as an encoding system. However, not every encoding system is capable of communicating the same level of information. In order to better understand the differences and similarities between encoding systems, I am set to argue for the need to classify encoding systems as "complex" (Complex Encoding Systems, or "CES") or "simple" (Simple Encoding Systems, or "SES"), based on the variety of information they are capable of encode, namely semantic, syntactic and prosodic (or pragmatic, on a case to case basis). Such classification is made possible by numerous pieces of evidence suggesting that semantics and syntax processing might not be supported by domain-specific mechanisms, while also considering the interaction between prosodic/pragmatic aspects of communication and said syntactic and semantic processing. In the current article, I will be focusing on CESs: by understanding the domain-general cognitive processes underlying CES decoding, we could hope to gain new insight on how languages are learned and understood, while also opening new doors to speculate on how languages might have come to be. I argue that the prime candidates to conduct cross-disciplinary research on are verbal and musical languages, two CESs that present functional overlapping in many aspects of semantic, syntactic and prosodic processing. I also present the argument that we must allow psycholinguistic research to be more thoughtfully contaminated by a musically proficient insight, in order to explore the topic to its full potential.

1 Introduction

Musical and verbal languages share numerous commonalities, while also differing from other encoding systems. Initially, neuropsychological research considered the two as distinct, but more recent evidence suggests otherwise [Li et al., 2023, Steinbeis & Koelsch, 2007, Koelsch et al., 2007, Slevc & Okada, 2014, Koelsch, 2011, Atherton et al., 2018]. "Encoding systems" can be defined as tools that translate data or information in symbolic or abstract form, that can be decoded only by having access to the encoding rationale. In order to better understand the similarities and differences among encoding systems, I argue for a categorisation based on the system's (in)capability to convey syntactic, semantic and prosodic information. I define "Complex Encoding Systems" (or "CES") as encoding systems capable of communicating all three types of information described above. On the other hand, I define "Simple Encoding Systems" (or "SES") as systems lacking at least one between semantic, syntactic or prosodic information. For example, programming syntax lacks prosodic information, as the code lines only express semantic and syntactic information: whether one might read programming syntax employing different intonations, the meaning of the code is not altered. These characteristics lead programming syntax to be classified as SES. On the other hand, the semantic information contained in musical and verbal languages interacts with the player's interpretation and the speaker 's prosody respectively, as the meaning of a musical phrase can differ based on the articulation, dynamics and tempo chosen by a musician and the phrase "John, come here!" can have different meaning based on the speaker's voice inflection (e.g., whether the speaker is angry, pleading or scared). The point of understanding encoding systems according to the type of information conveyed is to trace more clearly which modes of communication can be more directly compared, serving two purposes: 1. It allows researchers to treat as similar different languages in cross-discipline studies; 2. It constitutes an attempt to reveal domain-general processes of information extraction. So far, the scientific community has mainly focused on the syntactic and semantic aspects of encoding systems. While the prosodic aspect of communication is key to the proposed classification model, the article will thus mainly focus on semantic and syntactic processing mechanisms, interpreting the presence (or absence) of prosodic information encoding as a tool to better discriminate between CES and SES.

2 The SSIRH Hypothesis

The theoretical and empirical justification to classify musical and verbal languages as comparable encoding systems come from Patel's Shared Syntactic Integration Resource Hypothesis (or "SSIRH") [Patel, 2003]. By formulating the SSIRH, Patel set out to explain the divergence between the neurofunctional and behavioural aspects of verbal and musical language processing. While previous neuroimaging and neuropathological evidence had to that point supported the hypothesis that musical and linguistic processing were encoding system-specific, Patel's EEG studies and behavioural trials suggested that semantics and syntax processing in verbal and musical languages might operate at the same level while being supported by the same cognitive mechanisms, independently from the encoding system presented [Patel, 2003, Li et al., 2023, Koelsch et al., 2005].

On the other hand, evidence has emerged that, while syntactic processing resources are shared between these two encoding systems, [Li et al., 2023], semantic integration appears to pertain exclusively to verbal languages [Steinbeis & Koelsch, 2007]. The fact that semantic integration seems specific to verbal languages appears to discredit the SSIRH, but I argue for a different explanation: we have reason to consider musical syntactic integration as equivalent to musical semantic integration processes. In verbal languages, syntax and semantics are partially separated, as semantic information does not strictly convey syntactic one and vice versa. On the other hand, the musical language does not encode syntax and semantics in a way that allows the listener to distinguish between the two. In other words, musical semantics are intrinsic to the syntactic structures presented. To cite a musical element frequently studied in the current literature [Patel, 2003, Li et al., 2023, Steinbeis & Koelsch, 2007, Slevc & Okada, 2014, Atherton et al., 2018], a cadence marks the end of a musical phrase, thus making it syntactic information. However, listeners also understand cadences as semantic objects (according to their specific functions), thus allowing for a categorisation of syntactic units and rules as semantic elements. It would be disingenuous to not address extra-musical semantic information induced by music: it is understood that music is capable of conveying semantic information by cultural or emotive associations [Barraza et al., 2016], but these modalities are not specific to music and are too subjective to be classified at the same empirical level or "true" musically semantic information. I will then not consider this phenomenon in the current discussion. I argue that the main difference between musical and verbal encoding systems resides in how semantics and syntax interact intra-code. Taking the western tonal system as an example, a single musical element, such as a note or a chord, is incapable of communicating semantic information by itself: semantic information is achieved only through an appropriate syntactic context. This suggest that, while comparing chords to single words might seem intuitive, such comparison is inadequate: a word is linked rather strictly to an object, or a category of objects, which do not change based on the syntactic context they are inserted in (a "cat" is still a "cat" whether the phrase heard or read is "The dog is chasing the cat" or "The dog is chased by the cat"), while the same is not true for chords. For instance, the chord F major does not mean anything if played by itself, but acquires semantic value when inscribed in a cadence: for example, F major carries out the function of a "subdominant" chord if played in the progression "F - G - C" and the function of a "tonic" chord in "Gm - C - F". In other words, terms like "subdominant" or "tonic" refer to the function played by a chord and has syntactic value, while also serving as a label to classify a vast number of musical events, thus making it a semantic unit.

To summarise, we could consider verbal languages' semantics as "descriptive", as they refer to a entities or events present or relative to the physical world, while musical semantics can be defined "non-descriptive", as they convey information that acquires meaning only when considered inside its own encoding system.

3 Neurological Evidences Supporting the SSIRH

Neurological correlates between the shared resources employed in both musical and verbal language semantic and syntactic processing mainly come from EEG studies. By employing trials involving semantic and syntactic violations paired with musical syntax violation trials it has been demonstrated that syntax and semantic processing mechanisms operate as domain general networks [Steinbeis & Koelsch, 2007, Koelsch et al., 2005]. Musical semantics violation elicit the N500 and Early Right Anterior Negativity (ERAN) responses and a general P600 response. On the other hand, language syntax violations elicit Left Anterior Negativity (LAN), while semantic violations elicit a N400 response, and both types of responses are followed by increased P600 [Steinbeis & Koelsch, 2007, Koelsch et al., 2005, Koelsch, 2011]. Evidence about the interaction of such responses come from a variety of studies. Koelsch et al. [Koelsch et al., 2005] found that the ERAN interacts with LAN (namely, LAN was found to be clearly reduced when words were presented simultaneously with music-syntactically irregular chord functions), while Maess et al. [Maess et al., 2001] found that ERAN associated with harmonic processing originates from Broca's area, suggesting functional overlap with syntactic processing. Moreover, P600s was found to be indistinguishable in musicians when listening to syntactic incongruent harmonic sequences and sentences [Patel, 2003]. Steinbeis and Koelsch set out to investigate the interaction between simultaneous presentations of musical and linguistic semantic and syntactic violations [Steinbeis & Koelsch, 2007]. While linguistic and musical syntactic violations were already understood to interact (as shown in the works of Patel [Patel, 2003, Koelsch et al., 2005]), Steinbeis and Koelsch's results suggest that the N500 response is associated specifically to tension-resolution patterns violation, while also interacting with linguistic semantic violations. Crucially, the N500 response did not interact with linguistic syntactic violation, suggesting that music is able to convey strictly semantic information, although tension-resolution patterns are still to be considered syntactic information in the musical context. This leads to the conclusion that musical syntax and semantics are not encoded differently and functionally overlap during the decoding process, corroborating the SSIRH.

4 How Music Can Change the Way We Understand Language

Musical languages lend themselves to be interpreted as more abstract than their verbal counterpart. The fact that syntax and semantics are not distinguishable in musical languages implies that music can be employed to study encoding and decoding processes as regularity-extraction mechanisms, making it a useful research tool. I argue that researching music would help us understand which cognitive mechanisms are involved in language processing syntactic and semantic integration at earlier stages than the ones usually considered by modern psycholinguistics' paradigms. Generalising the empirical and theoretical paradigms of psycholinguistics to musical encoding systems (by borrowing trials, models and laws), we could revisit how we interpret language, leading to a wider interpretation of how languages work, are implicitly learned and how they might have come to be as a psychological entity. By exploring this possibility, it is possible that music might eventually force us to interpret language learning as a specific and peculiarly refined product of the same regularity extraction mechanisms that govern the rest of our epistemological experience, ultimately leading to the reduction of psycholinguistics to object perception and processing models.

As of now, verbal languages are still widely considered to be linked to specific cognitive abilities [Chomsky, 1975, Laurence & Margolis, 2001, Goldberg, 2008], while other researchers interpret language as the results of the combination of various cognitive abilities [Bruner, 1985, Behrens, 2021]. The position taken into the current article is an appeal to consider more favourably the constructivist stance on language development, as identifying the domain-general mechanisms underlying regularity extraction in CESs means understanding languages as a result of simpler cognitive mechanisms concurring to "make sense" of extremely articulated encoding systems. The added value provided by revisiting psycholinguistics in such a way would be to broaden our understanding of how humans interact with one another and how they interpret the external world.

5 Conclusion and Suggestions for Future Research

In this article, I set out to present the argument that psycholinguistics would benefit from integrating different language modalities in their research paradigms. I argue that psycholinguistics should encompass languages that share similar characteristics and complexity (i.e., Complex Encoding Systems). The prime candidates to play such a role are musical languages. In this article, only the western tonal system was discussed, both for the sake of brevity and to not deviate from my personal area of expertise. I would be wise, however, to extend the discussion in order to include different musical systems also. This article's proposition relies first and foremost on the involvement of musically proficient researchers inside psycholinguistic paradigms: to expand the SSIRH and its models, we need to take into account numerous musical variables that have not been considered as of yet, but that are also capable of conveying syntactic (and thus, semantic) information. Notably, aspects such as of voice leading, phrasing, melodic profile, articulation and dynamics have not been investigated empirically as of yet in relation to the syntactic information they are capable of communicating. Excluding harmony, the only syntactic aspect of the western tonal musical language that has been investigated is timbre, which has been proven to indeed carry syntactic information [Fiveash et al., 2018]. As the aforementioned musical variables are arguably more nuanced and subtle than harmonic progressions, I argue that psycholinguistics ought to be contaminated by a more nurtured dialogue with musically proficient researchers, in order to avoid incongruences and faux pas.

Speaking of harmonic progressions, an aspect that has not been researched upon to this day is long term integration, a phenomenon that is also relied upon in verbal languages' semantic processing. As it pertains to verbal languages, "long term integration" refers to the fact that the early understanding of a sentence can be influenced by subsequent information. I argue that a similar phenomenon could present itself in musical encoding systems as harmonic modulation, a process that shifts the semantic interpretation of an harmonic context to a new one.

6 Limits

I have thus far discussed musical and verbal languages under the assumption that they constitute the only examples of complex encoding systems. This assumption is limiting, mainly due to the fact that both music and language mainly rely on auditory processing in order to convey information. Due to this, even if further evidence might prove encouraging, there would be no way of knowing whether the identified regularity extraction mechanisms are modal or amodal. Sign languages would offer an interesting counterpart to the encoding systems discussed here, as they rely on visual elaboration in order to convey syntactic, semantic and prosodic information. Sign languages also present a few peculiarities that functionally overlap with other specific CESs, making them particularly suited to be studied alongside musical and verbal languages. Much like verbal languages, sign languages distinguish between syntax and semantics. On the other hand, communicating through sign languages requires complex motor planification, similarly to playing musical languages. I argue that by also focusing on exploring this double dissociation, we could hope to understand CESs even further and more thoughtfully than by only considering musical and verbal languages. Moreover, it is understood that sign languages are also capable of conveying pragmatic information [Bellugi & Fischer, 1972, Newman et al., 2010], justifying further their value in CES research. As stated previously, the current article has refrained from discussing prosodic encoding in CES in detail, due to a lack of cross-discipline studies in the current literature. While our knowledge on CES-specific prosodic encoding is vast, I suggest it is time to integrate said data and compare the underlying cognitive processes.

References

[Atherton et al., 2018] Atherton, R., Chrobak, Q., Rauscher, F., Karst, A., Hanson, M., Steinert, S., & Bowe, K. (2018). Shared processing of language and music. *Experimental Psychology*, 65(1), 40–48.

[Barraza et al., 2016] Barraza, P., Chavez, M., & Rodríguez, E. (2016). Ways of making-sense: Local gamma synchronization reveals differences between semantic processing induced by music and language. *Brain and Language*, 152, 44–49.

[Behrens, 2021] Behrens, H. (2021). Constructivist approaches to first language acquisition. *Journal of Child Language*, 48(5), 959–983.

[Bellugi & Fischer, 1972] Bellugi, U. & Fischer, S. (1972). A comparison of sign language and spoken language. *Cognition*, 1(2–3), 173–200.

[Bruner, 1985] Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, 1(1), 111–114.

[Chomsky, 1975] Chomsky, N. (1975). *Reflections on Language*. Pantheon Books. URL: https://ci.nii.ac.jp/ncid/BA20658933.

[Fiveash et al., 2018] Fiveash, A., Thompson, W., Badcock, N., & McArthur, G. (2018). Syntactic processing in music and language: Effects of interrupting auditory streams with alternating timbres. *International Journal of Psychophysiology*, 129, 31–40.

[Goldberg, 2008] Goldberg, A. (2008). Universal grammar? or prerequisites for natural language? *Behavioral and Brain Sciences*, 31(5), 522–523.

[Koelsch, 2011] Koelsch, S. (2011). Toward a neural basis of music perception – a review and updated model. *Frontiers in Psychology*, 2, 110.

[Koelsch et al., 2005] Koelsch, S., Gunter, T., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An erp study. *Journal of Cognitive Neuroscience*, 17(10), 1565–1577.

[Koelsch et al., 2007] Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An erp study of music perception. *Psychophysiology*, 44(3), 476–490.

[Laurence & Margolis, 2001] Laurence, S. & Margolis, E. (2001). The poverty of the stimulus argument. *The British Journal for the Philosophy of Science*, 52(2), 217–276.

[Li et al., 2023] Li, D., Wang, X., Li, Y., Song, D., & Ma, W. (2023). Resource sharedness between language and music processing: An erp study. *Journal of Neurolinguistics*, 67, 101136.

[Maess et al., 2001] Maess, B., Koelsch, S., Gunter, T., & Friederici, A. (2001). Musical syntax is processed in broca's area: an meg study. *Nature Neuroscience*, 4(5), 540–545.

[Newman et al., 2010] Newman, A., Supalla, T., Hauser, P., Newport, E., & Bavelier, D. (2010). Prosodic and narrative processing in american sign language: An fmri study. *NeuroImage*, 52(2), 669–676.

[Patel, 2003] Patel, A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, 6(7), 674–681.

[Slevc & Okada, 2014] Slevc, L. & Okada, B. (2014). Processing structure in language and music: a case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, 22(3), 637–652.

[Steinbeis & Koelsch, 2007] Steinbeis, N. & Koelsch, S. (2007). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, 18(5), 1169–1178.

Chairs

Fabio Cifariello Ciardi
Conservatorio “F.A. Bonporti” di Trento e Riva del Garda
Nicola Conci
Department of Information Engineering and Computer Science, University of Trento
Lara Corbacchini
Conservatorio “F.A. Bonporti” di Trento e Riva del Garda
Sara Favargiotti
Department of Civil, Environmental and Mechanical Engineering, University of Trento
Silvia Sacchetti
Department of Sociology and Social Research, University of Trento

Co-organizer

CONSERVATORIO DI MUSICA F.A. BONPORTI
Trento - Riva del Garda | Italia

Supported by

UC UNICITTÀ

FONDAZIONE CARITRO
CASSA DI RISPARMIO DI TRENTO E ROVERETO

Patronage

iprase
apprendimento ricerca innovazione

AIMI

siem
Società Italiana per l'Educazione Musicale
Rappresentante italiana dell'ISME International Society for Music Education