#4 EVOLVING MUSICAL EXPECTATIONS: MUTUAL CORRELATION BETWEEN A HUMAN MUSICIAN AND AN AI MUSICAL INSTRUMENT

Koray Tahiroğlu

4.1 Introduction

Artificial intelligence is one of the most active areas of research in the music technology community. There are significant efforts to further our understanding of artificial intelligence's potential applications in music on behalf of many practitioners, artists, musicians, computer scientists, with their own particular approaches, goals and their contributions to the development of AI research in music. One of the main areas of focus has increasingly been on expanding the current use of artificial intelligence to the creative practice of musicians, making new technologies available and accessible to instrument builders, musicians and composers (Eigenfeldt & Kapur, 2008; Tatar & Pasquier, 2019). Currently, there is an ever-growing demand for tools and techniques that allow for the creation of autonomous devices and processes (Tahiroglu, Kastemaa & Koli, 2020). Following the demands of musicians, a considerable amount of research has been devoted to various approaches to manipulating and transforming musical instruments into musical agencies through the autonomous acts of the musical instrument in a collaborative music action between musician and the musical instrument (Tanaka, 2006; Karlsen, 2011). Such musical instruments can incorporate features that allow them to act in mutual co-operation with human musicians in the process of composing and performing music. In this joint activity, the development of algorithms that can contribute to compositional structures as well as composition methods is a promising path towards the integration of more advanced computational technology with the creative practice of musicians.

In this article, I intend to reflect on the potential use of artificial intelligence technologies in instrument building, and subsequently to address the specific challenges and

https:/doi.org/10.34727/2022/isbn.978-3-85448-047-1_4.

opportunities that arise through these technologies. I focus on emerging musical expectations and musical demands in a practical research implementation of our AI-terity autonomous musical instrument. Relating advanced technology to music and performance is not a new phenomenon and has been discussed in more depth (Collins, 2007; McPherson & Tahiroglu, 2020; Magnusson, 2019; Tahiroğlu, 2021), at the same time, not much effort has been spent in examining the following questions: How do artificial intelligence autonomous algorithms and human musicians contribute to new musical expectations? What can be realised through the equal contribution of AI instrument and human musician in a music performance?

Figure 4.1: ***AI-terity instrument (Photo: Koray Tahiroğlu, 2020, CC BY)***

4.2 AI-terity and the composition Uncertainty Etude #2

In our Sound and Physical Interaction (SOPI) research group at the Aalto University School of ARTS we have been building, developing and performing with the AI-terity musical instrument (Figure 4.1), which is a non-rigid instrument that comprises a deep learning model, GANSpaceSynth,[1] for generating audio samples for real-time audio synthesis (Tahiroglu, Miranda & Koli, 2020). Physical deformability becomes the affor-

[1]GANSpaceSynth is a hybrid generative adversarial network (GAN) architecture that we developed in our SOPI research group. It applies the computational features of the GANSpace method (Härkönen et al. 2020)

dances of the instruments for handheld physical actions to be applied. The instrument uses an abstract form of an interface that is responsive to manipulation through bending or twisting, controlling parameter changes in granular synthesis.

We developed the GANSpaceSynth, specifically to provide more control over the spatial exploration of the audio features that are distributed in the latent space. This unique access to the GAN latent space gives musicians the ability to interact with the higher order structures to generate new audio samples. The advantage of this approach is the control of the directions for moving one point to another in spatial dimensions of the latent space. We call this point, the synthesis centre point (SCP).

Following the unique features of the instrument's deep learning model, we developed the instrument's autonomous features further to bring in alternative musical responses in music performance (Tahiroğlu, Miranda & Koli, 2021). In this way we could approach the performance of a music composition as an entity providing an independent variable that could affect the musical context by changing the decision with a non-arbitrary way of generating new sounds. To do that, we built in autonomous features to change the direction of the SCP in the latent space. GANSpaceSynth generates audio samples based on the SCPs, and these points determine the audio characteristics of the samples. The musician can navigate through the latent space by interacting with different parts of the interface, and GANSpaceSynth receives each of the SCPs as input to generate a corresponding audio sample. The idea of the autonomous behaviour is to monitor the musician's state of performing with the instrument and change that confident state of performing to an intermittent state of performing. The autonomous behaviour allows for alternative sound-producing expressions to appear; these are then layered in the changing audio features of the real-time granular synthesis. In this way, the autonomous nature of the instrument can be seen as an autonomous behaviour that aims at keeping the musician in an active and uncertain state of exploration, which allows massive flexibility and instantaneous exploration of an instrument's playability.

We wrote a composition for the AI-terity instrument, aiming to idiomatically reflect the autonomous features of the instrument. The work is based on the idea of *uncertainty*, where the instrument moves the SCP across latent space, aiming to find a new target point, but never stays in one particular point long enough to allow the musician to stay in a comfortable and certain state of performing. The composition brings up some confusion and surprise. Figure 4.2 shows the studio recording of the piece.[2] I should clarify precisely what is meant by 'uncertainty' in this case. It doesn't mean that the instrument chooses any 'random' points in the latent space; the jumps in between the SCP that the GANSpaceSynth uses to generate audio samples for the granular synthesis are not random. Instead, the autonomous features are designed to move in the opposite directions on the basis of the latent space centre and gradually introduce new audio samples in a smooth transition that allows the musician to explore the changing timbre of the audio samples during each jump. Through this process the music is composed—and it is unpredictable. You can hear it changing all the time. For a while follows the course of playing in an original set of generated audio samples, but the music is composed for each

on the audio synthesis features of the GANSynth model (Engel et. al. 2019) for organising the latent space using a dimensionality reduction for real-time audio synthesis. The open source code of the GANSpaceSynth is available at https://github.com/SopiMlab/GANSpaceSynth/.

[2]The studio recording is available at https://vimeo.com/514201580

jump with the intention of bringing in a new set of audio samples. It is possible to create a certain combination of sounds, each combination being a result of the SCP. What the performance of the piece attempts to do is to create unpredictable yet original musical expectations for the musician and the audience.

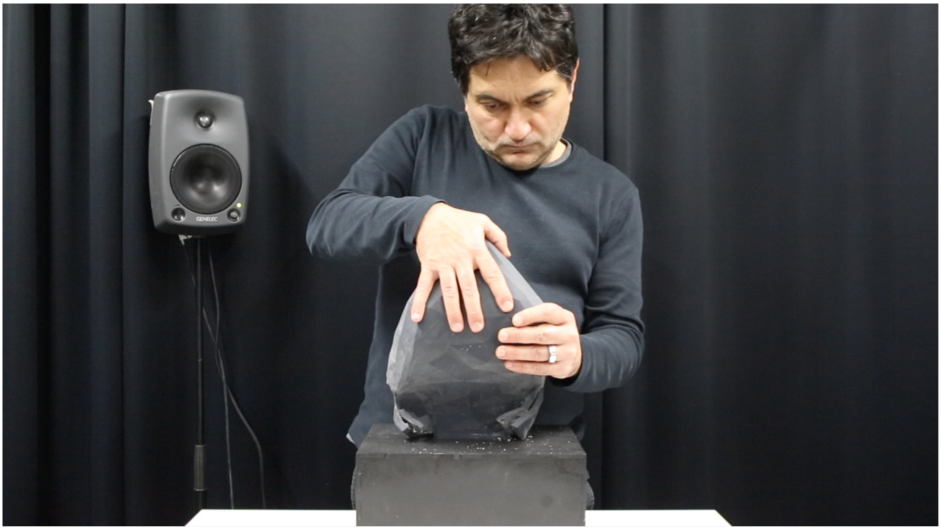

Figure 4.2: ***Studio recording of the composition*** **Uncertainty Etude #2** ***(Photo: Koray Tahiroğlu, 2019, CC BY)***

4.3 Achieving musical expectations

Titon (2009) describes a model for our interaction with music that presents musical performance as the process of creating the active experience of musical works. Following this model, we can see that music performance makes both those musical and aesthetic expectations appear that are already present before the act of the performance, as well as those that are not. In the context of music with an AI powered autonomous instrument, musical expectations appear in relation to the mutual connections between the instrument and the human musician. The performance of the piece *Uncertainty Etude #2* explores how the artificial intelligence instrument can serve as a musical partner, so that the human musician and the instrument can communicate to evolve musical expectations.

Musical expectations have been discussed in the context of melodic (Margulis, 2005), tonal (Bharucha, 1994), sound and meaning (Clarke & Cook, 2005), biologically learned (Huron and Margulis, 2010), memory and cultural (Curtis and Bharucha, 2009) patterns involved in music that may give rise to expectations and to ‘affective consequences of expectations’ (Huron & Margulis, 2010) for listeners or musicians. Huron and Margulis (2010) mention that ‘familiar experience’ occurs in the nature of musical expectation even between songs devoid of any structural musical relationships. It is an indication that familiarity could play a role in helping listeners to learn how to anticipate future songs, an ability which is often attributed rather to the listener's 'general knowledge' of the music. This expectation, however, is not a direct knowledge, and only shows to the listener the potential that is inherent in the music. By contrast, expectations have a proprietary quality, in that they can be learned from the specific events themselves. Regarding the indications for 'anticipating the future sounds', the question then arises, in what ways can we expect a particular musical expectation to be achieved in a correlation of sounds that even the musician has not heard before performing the music? For the composition *Uncertainty Etude #2*, we trained the GANSpaceSynth model with the audio dataset that has the overall textures of musical sounds, essentially inharmonic and atonal features with electronic ambience patterns. I have provided my own dataset for training the GANSpaceSynth checkpoint. The resulting checkpoint model in the composition tends to generate rather garbled approximations of the original dataset with smeared transients of unfamiliar sounds that do not follow the musical features of tonality or the context of melody as commonly discussed in the forms of musical expectations.

In the performance of the *Uncertainty Etude #2*, it might be argued that the ability to form and achieve expectations about unfamiliar sounds could still be more innate, while at the same time to some extent being dependent upon musical experience. This leads us to discuss whether musical expectation could be formed and developed on the basis of some other behavioural responses, such as intersubjective experience (Fuchs & De Jaegher, 2009), that is, from hearing other listeners. Perhaps it could be learned; indeed, as the structure of the musical language in *Uncertainty Etude #2* changes, some of the features required to form expectations are then learned through the performance of the piece. It might not be too much to argue that achieving musical expectations about the performance of the *Uncertainty Etude #2* will occur through direct acquaintance with the music. The mutual correlation between the human musician and the autonomous AI-terity musical instrument will evolve unfamiliar and surprising musical expectations that the listeners will experience.

Bharucha (1990) questions whether the listeners or musicians give up the element of surprise in musical expectation when they prioritize what they have already known. I am not sure if there is any particular answer to that question, but it might still be worth mentioning that familiarity has the potential to provide an emotional boost to music listeners' enjoyment, which might counteract the effect of surprise in music enjoyment. If this is true, then, in contrast to the music enjoyment of listeners of unfamiliar genres, are the listeners of familiar genres more likely to experience surprise as an unexpected *loss*?

It could be further questioned whether unfamiliar genre-specific audiences have a greater expectation of novelty than listeners of familiar genres. I think this may be more plausible than it would be in the case of other kinds of anticipation and expectation; but even if the question is theoretically open, it would be a tough argument for the listener of a familiar genre to come to the conclusion that any song is more interesting than, say, a song that fits the familiar genre. In the performance of the composition *Uncertainty Etude #2*, there is a unique set of properties for musical expectations, which is the result of the mutual cooperation between the human musician and the autonomous instrument. This set of properties are part of the unfamiliar and surprise nature of the musical performance itself.

4.4 Musical expectations as a framework for composition

The shift of music performance from a process with a master performer to an activity performed by a human musician and an autonomous instrument can also be considered as a particular social expectation that can be used as a framework for a musical composition (Tahiroğlu, 2021). In the performance of the piece *Uncertainty Etude #2* the human musician and the instrument become part of a collective, part of a 'performance ensemble' (Latour, 2005). There is a 'particularity' in this collective in its own way. The particularity, here, can be expressed in its musical and conceptual form, which can be derived from the composition. As part of the performance ensemble, the human musician and the instrument become a 'system of instruments' with their own musical demands, and yet also a 'system of musical compositions'.

Such performance ensemble opens a space for ways of music-making in which unfamiliar and surprise musical expectations become a part of compositional structures. We can consider this process as a compositional framework that offers insight into the creation of compositions that are fully or partly autonomous, or autonomous in an unusual way. We can speak of unfamiliar and surprise musical expectations with an immediacy that the performance would entail. There is a potential for musical exploration that arises from particular musical expectations, in which the musical structure and its material have to do with something that autonomously evolves with unexpected, surprise and unpredictable musical events.

4.5 Conclusion

The intention of this chapter was to present my reflections as an artist and musician performing with an artificial intelligence musical instrument. The question of the unusual musical expectations and further musical demands was also discussed. My main intention has been to introduce the unfamiliar musical expectations that appear through the performance of the composition *Uncertainty Etude #2*. In this composition the integration of an artificial intelligence method with the creative practice of a musician contributes to the autonomous structure of the piece. The performance of such a composition involves further challenges for the musical expectations and musical demands—both for the human musician and the audience. I intended to discuss in what particular way musical expectations and musical demands become present through an autonomous behaviour that was built to enable the appearance of alternative sound-producing expressions, which are layered in the changing audio features of the generated audio samples. I hope this article will contribute to an ongoing discourse about new creative technologies, and especially to the debate around the use of AI technology in music practice and a new way of thinking about composing and performing with musical instruments.

Acknowledgement
This work was supported by the Academy of Finland (project 316549).

REFERENCES

1. Bharucha, J.J. (1994). Tonality and Expectation. In R. Aiello & J.A. Sloboda, J. A. (Eds.) *Musical Perceptions* (pp. 40–60). New York: Oxford University Press.
2. Clarke, E., & Cook, N. (Eds.). (2004). *Empirical Musicology: Aims, Methods, Prospects*. Oxford University Press.
3. Collins, N. M. (2007). *Towards autonomous agents for live computer music: Realtime machine listening and interactive music systems*. [Doctoral dissertation, University of Cambridge].
4. Curtis, M. E., & Bharucha, J. J. (2009). Memory and musical expectation for tones in cultural context. *Music Perception, 26*(4), 365–375.
5. Eigenfeldt, A., & Kapur, A. (2008). An agent-based system for robotic musical performance. *Proceedings of the international conference on new interfaces for musical expression, Genoa, Italy*, 144–149. https://doi.org/10.5281/zenodo.1179527
6. Engel, J., Agrawal, K. K., Chen, S., Gulrajani, I., Donahue, C., & Roberts, A. (2019). GANSynth: Adversarial neural audio synthesis. Paper presented at the International Conference on Learning Representations. Retrieved from https://openreview.net/forum?id=H1xQVn09FX
7. Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences, 8*(4), 465–486.
8. Härkönen, E., Hertzmann, A., Lehtinen, J., & Paris, S. (2020). Ganspace: Discovering interpretable gan controls. *ArXiv Preprint ArXiv:2004.02546*.
9. Huron, D., & Margulis, E. H. (2010). Musical Expectancy and Thrills. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford University Press.
10. Karlsen, S. (2011). Using musical agency as a lens: Researching music education from the angle of experience. *Research Studies in Music Education, 33*(2), 107–121.
11. Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.
12. McPherson, A., & Tahiroglu, K. (2020). Idiomatic Patterns and Aesthetic Influence in Computer Music Languages. *Organised Sound, 25*(1), 53–63. https://doi.org/10.1017/S1355771819000463
13. Magnusson, T. (2019). *Sonic Writing: Technologies of Material, Symbolic, and Signal Inscriptions*. Bloomsbury Academic.
14. Margulis, E. H. (2005). A Model of Melodic Expectation. *Music Perception, 22*(4), 663–714.
15. Tahiroglu, K., Kastemaa, M., & Koli, O. (2020). Al-terity: Non-Rigid Musical Instrument with Artificial Intelligence Applied to Real-Time Audio Synthesis. *Proceedings of the International Conference on New Interfaces for Musical Expression. Birmingham, United Kingdom*.
16. Tahiroğlu, K. (2021) Ever-shifting roles in building, composing and performing with digital musical instruments, *Journal of New Music Research, 50*(2), 155–164. doi: 10.1080/09298215.2021.1900275

17. Tahiroğlu, K., Kastemaa, M., & Koli, O. (2021). AI-terity 2.0: An Autonomous NIME Featuring GANSpaceSynth Deep Learning Model. Paper presented at the International Conference on New Interfaces for Musical Expression. Retrieved from https://nime.pubpub.org/pub/9zu49nu5

18. Tanaka, A. (2006) Interaction, Experience and the Future of Music. In K. O'Hara & B. Brown (Eds.), *Consuming Music Together: Social and Collaborative aspects of music consumption technologies* (pp 267–288). Dordrecht: Springer.

19. Tatar, K., & Pasquier, P. (2019). Musical agents: A typology and state of the art towards musical metacreation. *Journal of New Music Research, 48*(1), 56–105.

20. Titon, J., T. (2009). *Worlds of Music: An Introduction to the Music of the World's Peoples*, first edition Cengage Learning.