Becoming latency-native

Rebekah Wilson
Amsterdam, The Netherlands
rebekah@loopcontrol.io

ABSTRACT
The potential richness of audio technology on the internet springs from advancements integral to developments driven by the primary concerns of commerce and science, giving rise to operable and affordable bandwidth in newly-accessible geographical areas as well as the growing sophistication of codecs, browser technology and audio frameworks. Yet use of these tools remains unexplored for music performance, with the primary cause of disruption to performance flows being transmission latency. Through the employment of sophisticated tools and processes, musicians may, however, learn to navigate Networked Music systems as a native performance platform.

1. A TOOL AND A HUMAN ACTIVITY
In 1969, 50 years ago as of the writing of this text, UCLA graduate student Charley Kline sent the first digital message to Bill Duvall in Stanford, the other side of California. That communication link was how ARPANET began, leading to today's modern internet. In these last 50 years we have seen digital communications give rise to an expanse of services built on the internet, such the webRTC framework that has prompted the development of an expanse of high-quality services for transmitting real-time audio suitable for performance. This has been made possible thanks to a few key developments: the provision of the OPUS codec with music-quality encoding, the ability to disable echo cancellation, and the ability to process and manipulate transmitted audio using audio APIs. As Heidegger observed, technology is both a means to an end as well as being a human activity [1]. When a technology affords us an expressive outlet, we cannot help but to explore this technology through a creative lens. Armed with this knowledge, any developer may create a platform for long-distance music performance, designed specifically for her purpose.

2. NON-LINEAR PERFORMANCE
What was once limited to complex systems—high-quality, low-latency audio over long distances—now becomes ubiquitous. Networked Music Performance, a real-time synchronous audio system for remotely interacting performances over geographic distances [2], is however subject to limitations specific to its nature: latency, technical uncertainty and the loss of full sensory feedback such as acoustic resonance. Technical uncertainty refers to the oft-complex systems that musicians need to put in place in order to perform together remotely; this particular limitation can be solved by way of improvements in interface design. The lack of full sensory feedback—for example, when we cannot sense a musician's subtle bodily gestures that may otherwise be providing significant information to us in an ensemble performance—can be accommodated for with increasingly sophisticated technology as video streaming improves in speed and quality; VR/AR advances ensure a future rich with immersive remote experiences. Latency however is an endemic property of digital networks and will remain so while data can travel no faster than the speed of light.

Since sound transmission has been possible, extensive research has been conducted on the effects of latency in music. In Networked Music Performance the "overall delay experienced by the players includes multiple contributes due to the different stages of the audio signal transmission" [2], most significantly network packetization and transmission latency, and hardware and software processing latency. Alexander Carôt states that "if the latency exceeds a certain value, a realistic musical interplay becomes impossible" [3] an agreement shared by most literature that musicality is dependent on latency being as low as possible.

And yet, we live in a non-linear world.

In Daniel Chua's poetic text on the operatic echo, he says "music takes time. The echo, by measuring the distance between subject and object, simply stretches the point" [4] In Monteverdi's *Vespers 1609*, an antiphonal echo sounds reverberantly from behind the audience. Echoes are "messy; they don't synchronize or return to their place of origin; they travel unexpected distances to connect unrelated times" (ibid). Similarly, latency in music creates a kind of messiness. As in Monteverdi, the effects can be spectacular.

3. LATENCY AND FLOW
When performing together in composite space—where musicians share both location and time—a great deal of information is exchanged that fosters Csíkszentmihályi's concept of 'flow', a desirable psychological state defined as a "collective state of mind" that occurs when "members develop a feeling of mutual trust and empathy, in which individual intentions harmonise with those of the group" [5]. Flow is experienced when musicians play in time together, and when they resonate with each other in acoustic space. When performing together remotely over a network, not only do musicians lose the experience of acoustic resonance but their interactions, subject to latency, lead to interrupted rhythms and difficulties with synchronisation. Yet, human reaction time is on the order of hundreds of milliseconds and quite variable [6], opening the question that it is not transmission latency that

Web Audio Conference WAC-2019, December 4–6, 2019, Trondheim, Norway.

disrupts musicians as a rule in networked music, but our perception of latency.

Music data is traditionally relational while language is transactional [7]. Transactional data can be packetised, encoded and decoded, while relational data relies on cues and forms of data which may not be able to encoded in an internet transmission, such as the glance of one musician to another to indicate the start of musical section or a change in improvised direction. Latency disrupts relational data; it breaks us into parts and "nobody can tell us whether what we have spilt up can be put together again or not" (ibid). When we lose sufficient information to intuitively interpret data that has been split up, data becomes transactional rather than relational. Any interface that "seeks to engage with our personal act of knowing needs to be able to afford us our relational dimension in balance with our transactional" (ibid); unless we take care in the design of our interfaces to accommodate for this need for balance, we will experience tension when we perform together.

4. LATENCY, US AND MACHINES

Between continents, latency is too high even on dedicated networks for traditional performance. Consider efforts to colonise space—even to the moon the latency is around 1.25 seconds. So, what kind of music will we make together when we are on Mars where, depending on the position of the planets, we will experience a latency of 3-20 minutes? Fortunately, when working within the realm of the technology system itself we may make use of the means of transmission and employ machines to work on our behalf not only as transmitter but as interpreter of data.

Nevejan says that where 'latency breaks us, it makes us stronger' [Nevejan, personal communication]; consequently, technology gives us opportunities to create new realities. However, to make new realities "you have to be bilingual" (ibid). Becoming bilingual with networked music means becoming latency-native and adapting ourselves as performers and adapting our idea of what is considered musical. Contemporary musicians are well-suited to this experience; as pianist Reinier van Houdt notes, musicians have been "disciplined enough to be free and expect the unexpected" [van Houdt, personal communication]) and are already challenged by dangerous body movements, fatigue, masks, cramps, extreme temperatures, live processing; the "experimental mind is not worried about the conditions but sees them as a challenge, a medium for research for new experiences." (ibid).

A number of Networked Music explorations have experimented with strategies to explore performative issues specific to the platform, such as Ethan Cayko's toporhythmic patterns [8], Jacktrip's wavetable mode [9], Andrew MacMillan's system for remote parameter control [10], and contextual analysis through machine learning [11]. Machine learning tools shows promise in detecting musical information and is a growing source of great innovation. In a Networked Music Performance context, achieving flow may be reached though the awareness and interaction with latency's effect on time. Taking a cue from Nevejan's concept of "finding the first beat" [12], where moving together in time occurs when recognisable rhythmic patterns begin with clear starting point, we can take advantage of the machine's ability to detect and analyse the patterns of how we move together as musicians, aiding us to strengthen and reinforce rhythmic cohesion while encouraging new aesthetic perspectives and sonic realities. By more closely interfacing our performances and our rhythms with the Networked Music machine-system itself, we become closer to becoming bilingual and latency-native.

Yet we must remember that the machine's purpose is determined by the engineer at the time of execution, while meaning is created at the time of observation by a human. Machines are only ideal for data mining and processing—not for creating meaningful observations; any meaning that is created is tangential to the machine's purpose. A machine has no tacit knowledge from which to create meaning—it can only expose explicit information to be acted on. Can these machines and algorithms be designed to act in ways humans cannot? What can we encode, construct, transform and transmit with these machines? How will these machines co-exist with multiple authors? Will it give rise to something beautiful, when we experience the technological illusion of being present in a place other than our physical selves and actions?

5. REFERENCES

[1] Heidegger, M., 1977. *The Question Concerning Technology and Other Essays*. Harper & Row, New York, NY

[2] Rottondi, C, Chafe, C., Allocchio, C. and Sarti, A. 2016. "An Overview on Networked Music Performance Technologies," *IEEE Access*, vol. 4, pp. 8823–8843

[3] Carôt, A., 2009. "Towards a comprehensive cognitive analysis of delay-influenced rhythmical interaction". *Proceedings of the 2009 International Computer Music Conference*, ICMC

[4] Chua, D. 2005., "Untimely Reflections on Operatic Echoes: How Sound Travels in Monteverdi's L'Orfeo", *The Opera Quarterly*, Volume 21, Issue 4, Pages 573–596

[5] Sawyer, R., 2003. *Group Creativity: Music, Theater, Collaboration.* Taylor & Francis.

[6] Chafe, C., 2011. "Living with net lag," in *Audio Engineering Society Conference*: 43rd International Conference: Audio for Wirelessly Networked Personal Devices

[7] Gill, S., 2015. *Tacit Engagement: Betwixt and Inbetween* Springer International Publishing

[8] Cayko, E., 2016. *Rhythmic Topologies and the Manifold Nature of Network Music Performance*. Dissertation, University of Calgary.

[9] Cáceres. J.-P. and Chafe, C. 2010. "JackTrip: Under the Hood of an Engine for Network Audio," *Journal of New Music Research*, vol. 39, no. 3, pp. 183–187.

[10] MacMillan, A. and Wilson, R., 2019. "Being less un-together". *JONMA,* vol 1.

[11] Wilson, R., 2019. "Towards Responsive Scoring Techniques in Networked Music Performance". *Proceedings of the 2019 TENOR Conference*, Melbourne.

[12] Nevejan C., Sefkatli P., Cunningham S., 2018. *City Rhythm*. Delft University of Technology