Henderson, Jordan (2025) *Scores into Bodies: communication between audiovisuals and instrumentalist*. PhD thesis.

https://theses.gla.ac.uk/85595/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

Scores into Bodies

communication between audiovisuals and instrumentalist

University of Glasgow

College of Arts and Humanities

School of Culture and Creative Arts

Scores into Bodies

communication between audiovisuals and instrumentalist

Jordan Henderson

September 2024

Scores into Bodies: communication between audiovisuals and instrumentalist.

Jordan Henderson

Supervisors: Professor Nick Fells and Professor Louise Harris

Submitted in fulfilment of the requirements of the Degree of Doctor of Philosophy

September 2024

Copyright © Jordan Henderson 2024

University of Glasgow

School of Culture and Creative Arts, College of Arts and Humanities, Music

G12 8QQ Glasgow, Scotland

Abstract

Watching people play music together pulls one into the ensemble through their communicative gestures and social interaction. How can we preserve this sense of togetherness when one of the performers is computer generated and perhaps not a human form? This is a portfolio–based music composition research project that explores the design of non-human entities and how they communicate, proposes an initial conceptual understanding for ensemble performance, and creates a technical framework for the composition of multimedia interactive systems that prioritises relationships between sub–systems.

Contents

List of Figures

Acknowledgements

I would like to thank my supervisors for their guidance throughout, all members of the 'crit group' for providing a supportive space to share work and to complain about PhD life, and finally Lucy Fry, whom without, this project would not have been possible — I love you and I like you.

1

Introduction

Watching people play music together in small groups is enthralling. They communicate with each other, perform constant adjustments in response to others, and anticipate the future, all in service of constructing the piece right in front of the audience. Is this experience possible when one of the performers is a computer, and can the use of live generated visuals help tackle this? This research project, comprising of a portfolio of composition, a *SuperCollider* quark, and this written commentary, investigates this issue. Although starting under the premise that what makes small ensemble performances captivating is *equal communication*, this idea is developed and refined into a proposal of audience experience called *ensemblematic experience* (§4.1). Primarily, this project aims to be useful to artists who value the human aspect of musical performance while also pursuing the intersection between art and technology. Secondly, this project might be of interest to those concerned with: animated musical notation (§2), interactive music systems (§4.2.2, §G), audience experience of live small ensemble music (§4.1), and those interested in creating art centred around relationships.

While there are numerous ways in which visuals might be incorporated into a performance, the focus here is on enhancing the appearance of *equal communication* between the human instrumentalist and the computer controlled live audiovisuals. Importantly, whether or not one considers the computer as *actually* communicating with the human (or vice versa) isn't important, only the appearance of such communication is valued in this project, particularly from the audience's perspective — whether or not the performer believes this, is relevant only so far as it impacts their performance. Communication is chosen here due to the overarching aesthetic goal of recreating or imitating the social dynamic found in small ensemble performances, notably small groups of singers where there isn't an apparent hierarchy between performers and everybody appears to contribute equally, creating a palpable sense of teamwork and togetherness. The literature from interactive systems (notably Jordà 2005; Chadabe 1984) that considers interaction as communication will be used to offer some insights regarding this point. One of the main goals for this project is for the audience to perceive the computer generated audiovisuals as being an equal member of the ensemble, which throughout this project's evolution, turned out to be problematic while providing an important aesthetic goal, engendering an interesting way to conceive of and technically implement interaction between humans and computers.

This project, while being completely continuous in practice, is roughly split here in writing into three chapters each containing two compositions. The first (§2) explores using animated notation to provide insight into the relationship and communication between human instrumentalist and computer part, here the assumption is that the performer's relationship with the generated audio is created through the composition but the audience struggles to experience it, and by providing an animated score, the relationship might be observable. The second (§3) is a transitory chapter that starts to explore the possibility of creating audiovisuals which are perceived by the audience as a body that sits on stage with the human performer, along with considering how the performer might interact and understand such an alien body. The third (§4) seeks to refine the idea of audiovisuals as performer, as well as exploring the limitations and issues that arise in doing so.

However, before continuing, it is valuable to identify the perspective from which I am approaching this project. I am, amongst other things, a composer interested in two apparent diverging worlds, computer technology and improvising with others. Whilst these two worlds are quite contradictory, together they cover a broad range of musical possibilities and practices, sonic and otherwise, in which I attempt to embed my own practice. In terms of outcomes, I find technological work capable of provoking larger questions beyond our day–to–day understanding. One of the hallmarks of modern technology is the varying scales it can encompass, from the microscopic world to the macroscopic world. Improvising with others on traditional instruments, however, is always small in scale and revels in the fragile relationships between the performers. It can produce art of and about people and the personal relationships in ensembles. Likewise, there is a disparity between the practice of creating digital technological works and instrumental improvisations: the former is often architectural, requiring a carefully considered engineering approach that is iterative, building upon previous decisions; whereas the latter is fluid, dynamic, easy to alter, occurs in real–time, and thoroughly collaborative. Straddling these two poles is a continued concern throughout this project.

One initial example of how these concerns are addressed can be found in my approach to interaction and communication. Since this project's goal is to create the perception that the human and computer are engaged in the type of interaction one finds in small ensemble works, perhaps one approach might be to choreograph this interaction, carefully dictating each movement and gesture in an attempt to construct the desired outcome. This is not what this project attempts. Instead, the interaction between the human and computer strives to be as close to the fluidity found in small ensemble improvisations. One concern is that by scripting this interaction, the subtleties of the relationships might get lost and require the performer to enact them. Enacting relationships is complicated, requiring great acting skill to do so convincingly, or else the performance runs the risk of becoming 'hyperreal' or even absurd. Of course, authors have made use of these extremes: new complexity requires a form of hyper–instrumentalism/virtuosity and Kagel's music, which does deal with the types of interaction this project seeks, is often absurd, sometimes purposefully, but sometimes because authentically re-enacting Kagel's instructions is challenging. In an attempt to steer away from those outcomes, this project proceeds under the assumption that the best way to create the perception of interaction and communication for audiences is to have the performers be engaged authentically in relationships and communication on–stage, right before their eyes.

1.1 Why Communication?

As mentioned, this project pursues the appearance of communication as found in small ensembles. This is due to my personal experience of attending concerts and being captivated by the way ensembles engage with each other during the performance, such that they appear, not just sonically but across all senses, to be joined together in some way. This is often termed *musical togetherness* and while many authors give differing definitions, some notable ideas can be extracted that suggest togetherness is an experience that creates:

- a communal sense of *we* (Bilalovic Kulset and Halle 2020);
- feelings of social presence, closeness, and joint agency (Bishop 2023);
- feelings of working together and collaboration (Kos 2018).

However, little research has been done on how the audience perceives such togetherness as most of the literature is from music psychology, where the focus is on the performers and their mental state, and pedagogy, where the focus is on the student-as-performer and their social and/or musical enhancement. As an artist-researcher, my primary interest is in the aesthetics of musical togetherness, both as an observed phenomenon from the audience's and performer's perspectives and as a mechanism through which art can be created. For this reason, this project starts with the proposal that observable communication between ensemble members aids in the perception of musical togetherness, as it allows the audience to understand how the ensemble is collaborating and acting as a single entity. Focusing on communication also allows for a mechanism which the computer can attempt to engage in as it has technical connotations.

The question then becomes, what exactly is communication and how is it observed? While communication studies is a relatively young discipline, there are a few definitions available:

- John Durham Peters suggest a colloquial definition might be related to the exchange of information, although acknowledges that phenomenologists would reject this definition (Peters 2001, p. 8);
- one particular intense form of communication, communion, can be defined as 'a sharing of inner experience' (Lowenthal 1967, p. 336);

- John Dewey in the 1920s thought of communication as a ‘pragmatic making-do’ and a ‘participation in the creation of a collective world’ (Peters 2001, pp. 18-19);
- whereas Frank J. Macke proposes that the the status/stasis of the relationship both before and after the exchange is the defining feature of communication (Macke 2014, p. 210).

The main concept of communication used throughout this project is Macke’s proposal that communication is concerned with the dynamics of relationships. This is because it is easily applicable to the audience’s perspective as they can readily observe that a relationship has changed, whereas the others are less obvious to recognise. Following this, three conditions for a perceptible communicative performance from the audience’s perspective arise:

1. the relationship at any given moment between the computer and instrumentalist should be observable;
2. as this project seeks some form of equality between computer and instrumentalist, the relationship must remain on average equal over the duration of the work;
3. the relationship should change with time, implying some evolving dialogue.

Musical relationships, particularly between human and computers is well studied. John Croft outlines five different relationships (called paradigms in their terminology): backdrop, accompanimental, responsorial/proliferating, environmental, and instrumental (Croft 2007, p. 62). While most of these suggest some hierarchy and therefore don’t meet the goal of imitating the relatively equal communication found in small ensemble performances, accompanimental and responsorial/proliferating could be considered as meeting this criteria. Although Croft is considering both 'procedurally live [and] pre-recorded' (ibid., p. 62) electronics, the responsorial/proliferating relationship is often made with the use of delays which require the human instrumentalist take a more prominent role. While this category could theoretically be more equal, in practice it seldom is. Nevertheless, this simple type of relationship between computer and human is too narrow for this project and doesn't cover the myriad of relationships small ensemble music creates. Likewise '[accompanimental] is where the sound from the loudspeaker functions as a kind of accompaniment' (ibid., p. 62), and while this could theoretically be inverted giving the computer a more prominent role, Simon Emmerson questions the use of electronics as an instrumental force as it 'often resulted in apparent "super-instruments", virtuosic and ghostly counterparts to the live performer who interact with an apparently superhuman (and sometimes robotic) force' (Emmerson 2000, p. 207). Again, this implies a clear hierarchy between the *real* human and *ghostly* computer from the audience's perspective.

A further proposal for the relationship between computers and human in art can be found in Agostino Di Scipio's concept of ecosystemic systems (Di Scipio 2003; Di Scipio 2021). Here individual people or agents aren't considered as individuals, but as members of an ecosystem, and the concept of being 'alive' is defined in a microbiological way to mean interacting and participating with and inside an environment. While this has certainly proved an influential idea, it does away with the individual and thoroughly human phenomenon of social communication and participation that this project seeks, replacing it with a top down view, more akin to looking through a microscope and observing from above the complex and fascinating relationships between the agents and their environment. In the final section of this research project a technical approach is finalised that bares similarity to ecosystemic methods, however while this is canonically applied to space, environment and systems, here I am applying it to individuals and their social relations.

For the above noted reasons, neither Croft's nor Di Scipio's concepts are used in this project. Instead this project follows its own path, starting with the assumption that there is a sonic relationship created during composition that simply needs revealing. An example of this is most prominent in contrapuntal vocal music where the sonic abstract relationships contained within the score are enacted by the vocalists as they perform, becoming visible to the audience. For this reason, this project begins by exploring animated notation as a way to reveal this obscured relationship. As noted, this does not produce a satisfactory conclusion for this project but nevertheless forms the initial hypothesis.

1.2 Overview of this Document

This document comprises this introduction, 3 main chapters, and a conclusion. The first main chapter is on animated notation (§2) and focuses on exposing the assumed existing relationships between human and computer to the audience by showing a dynamic score. The two works considered here are: *Together We Tore Through Forested Thickets* (*TWTTFT*) (§A, §2.2), an animated notation for violins and live electronics; and *Mercurial Memory* (§B, §2.3) an animated notation with live electronics in a web browser for flute. *TWTTFT* attempts to create a unified notation which the computer and human performers read from, while *Mercurial Memory* uses the visuals to create an environmental score through which the flautist moves and computer interacts with.

The second main chapter (§3) begins exploring how human and computer might be placed in *relation to* one another. Here the focus is on creating performance directions which set the entities into a changing relationship. The two works are: *Whilst We Speak Through The Wind* (*WWSTTW*) (§C, §3.1) for alto clarinet, bass clarinet and live audiovisuals; and *Glimmer* (§D, §3.2) a work for baroque flute and video score. *WWSTTW* uses a slightly novel notation on a paper score that shows how the performer and audiovisuals should be related to each other, along with exploring ideas around gesture, whereas *Glimmer* focuses specifically on how the human instrumentalist can read and understand the visual as a metaphor for their body in search of a more efficient and concise language between human and computer. The pieces in this chapter are ultimately united by their focus on notation and how the performer reads or otherwise understands the audiovisual computer generated part.

The final main chapter seeks to consider the generated audiovisuals as a performer in their own right, which sit on stage alongside the human instrumentalist. The two works here are: *Breathe* (§E, §4.2) a work focusing on the breath of a vocalist and generated nebulous body; and *Jellyfish* (§F, §4.3) a work for bowed guitar (or cello) and generated 'jellyfish' that seeks to imitate and extend some of the communicative gestures found in small ensemble performances. Additionally, using literature from music psychology and focusing on ensemble performance and dance literature, this chapter proposes a model for understanding audiences' experience of small ensemble music, called *ensemblematic experience* (§4.1).

Alongside this project, several significant contributions have been made to the *SuperCollider* project, which while not directly fitting in with the narrative of this document, nevertheless form an integral part of my practice and are valuable research outputs, and so are included as an appendix (§H).

2

Animated

Animated notation forms the starting point of this project as the initial assumption made was that the structure of the music creates relationships between the human and the computer, and that animated notation might aid in demonstrating these relationships to the audience. Here communication is understood, following **Macke(2014**, to mean changes of relationships, and therefore, as the piece progresses, the roles of the human and computer change, altering their relationship and allowing communication to be observed by the audience. According to Cat Hope (2017) and David Kim-Boyle (2018), animated notation has the ability to show, demonstrate, or otherwise enhance an audience's experience of a musical performance, potentially making it suitable for this task. In the following section, I will discuss the differing approaches authors have towards animated notation, considering both its surrounding literature and practice. As this project progresses, animated notation becomes less relevant as new way of conceiving of the issue are explored, nevertheless the terminology introduced in this topic is foundational for the following chapters.

2.1 Animated Notation

Although the origins of animated notation can be traced back to at least the graphic notation from the 1950s and 1970s, it is in the past 20 years with the rise of computer graphics that animated notation has developed a distinct identity (for a thorough review of this history see Fischer 2015; Smith 2016). However the term is often tied up with many others, muddying its definition. This section of writing will deal with unravelling the following terms often used in animated notation's literature as they relate to this research project: real-time notation, projected scores, environmental scores, action notation, and player scores.

Real-Time Notation

Real-time notation is when some system drives the generation of notation (Freeman 2008). Examples include:

- generating the score from the live sound (Winkler 2010; Essl 1998; Didkovsky 2019);
- a visual gesture detection system creating a feedback loop of control such that as the performer moves the score changes (Dori 2020);
- by the audience as in Jason Freeman's *Glimmer* (2008) or Harris Wulfson's *Livescore* (2006) (see Barrett and Winter 2010).

While the use of computer displays is often employed, it is not required, and many experimental composers from the middle of the 20th century wrote such music. Notable here is: Earle Brown's *Calder Piece* 1963–1966, where a mobile sculpture is both percussively struck, requiring performers to chase after it, and read as a score; and many of Christian Wolff's compositions where the score is derived from reading and responding to other performers — an idea that will reoccur in this project (§3.1).

Freeman raises an important criticism of real-time notation, that of the virtuosic strain put on the performers required to read the new notation instantly and perform it accurately, which is termed 'extreme sight-reading' (Freeman 2008). Nick Didkovsky's *Zero Waste* (see Didkovsky 2019) exemplifies this issue of extreme sigh-reading, using it for its own creative potential. In *Zero Waste*, a pianist is presented with a score to be played, following which, the audio is recorded and translated through machine listening back into a score, which is again shown to the same performer who plays it, thereby repeating the process. In an hypothetical world, the notation would remain the same as the performer correctly realises it and machine correctly identifies it. In reality, notation rapidly begins to drift from the original and small errors are introduced, both by the pianist and the machine listening/transcribing (Whyte et al. 2018).

This research project directly invokes real-time notation in all of the works to a varying degree.

Projected Scores

Projected scores are when the score is displayed such that the audience can see the score the performers are playing from. Although not directly addressed in the literature, having a single score that is read by the performers and audiences, as opposed to each person having their own copy, appears to be pertinent. Two early works that are often mentioned as influential are Earle Brown's *Calder piece* 1963–1966 in which a mobile is visible to the audience and read by the performers, and Mauricio Kagel's *Prima Vista* 1962–1963 in which slide projections are used. When considering the benefits of projected scores, Kim-Boyle suggests that 'the listener is invited alongside the performer/s to playfully engage with a work's structural processes and in turn develop an intimate understanding of the world it explores' (Kim-Boyle 2018, p. 50). Yet there have been concerns raised about whether this could detract from the appreciation of a work. As such, Hope and Vickery have proposed six cases where a projected score can enhance or even be necessary. These are when screen presentation:

> - allows an already existing work to operate more naturally than the media available at the time of composition;
> - conforms to the composer's conceptualization the work as comprising visual and auditory components;
> - adheres to or more closely corresponds with the composer's intentions in regard to permitting conceptual or structural goals to be realized;
> - assists the comprehension of the work by the audience;
> - does not unduly add to the cognitive load of attending the work;
> - does not detract from the dramatic performative aspects of the work.
>
> (Hope and Vickery 2010, p. 9)

However, not all animated notation utilises projected scores, and some composers prefer to keep the animated nature of their scores hidden from the audience, or at least, only displayed in conferences specifically about animated notation. An example of this can be found in many composers who use the *decibel scoreplayer* (Hope, Wyatt, et al. 2015), a tool specifically designed to synchronise scrolling scores across a network. Similarly, there are composers who use virtual reality headsets to display content to their performers, hiding it from the audience (Santini 2020; Kim-Boyle and Carey 2019). Since this research project is only concerned with animated notation in so far as it alters the audience's experience of performances, only projected scores are considered.

Environmental Scores

Although often projected but not necessarily animated, environmental scores are when the performer(s) move through the score in some way, performing the notation as they reach it. Kim-Boyle, while discussing animated notation's ability to develop an intimate understanding of the world a piece explores, points to Jobina Tinnemans's *Imagiro Landmannalaugar* 2017 where a score is wrapped around the walls of a room and the performers walk alongside the walls, reading the notation closest to them. This is similar to an approach employed by Lindsay Vickery and Cat Hope in *Talking Board* 2011 where a large, quickly evolving texture is displayed across the screen that small coloured circles gradually traverse, indicating which areas of the texture/environment each performer should be reading from. Here there is a shift from the physical environment of *Imagiro Landmannalaugar* to the virtual environment of *Talking Board*. Another author who it could be argued uses environmental notation, albeit abstract environments, is Ryan Ross Smith whose radial scores create rotating environments that bring symbolic notation to the performers' virtual position as indicated through some graphic mark on screen. Environmental scores could be relevant to this research project as they create a way for the audience to understand the relationship between each performer as being relative to one another through their position within the world. This approach is employed in both the following compositions *TWTTFT* and *Mercurial Memory*, albeit in different ways.

Action Notation

While an orthogonal concept to animated notation, the notation of physical actions required to produce sounds rather than symbol inscriptions of sound is often employed in animated notation. A historical example of action notation is lute tablature, which records where the fingers should be placed rather than which notes should be heard, while a more contemporary example might be certain works by Helmut Lachenmann (Kojs 2011). By animating the notation, the motion of the symbols in time adds an extra dimension of physicality, allowing for a more visceral and perhaps embodied understanding of the score (Hope 2020b, p. 193). Further, action notation might offer a solution to extreme sight-reading in real-time scores if the intended action can be communicated more succinctly (Shafer 2017). This is perhaps achieved in *Portale* by Giovanni Santini (2019), where augmented reality is used to draw virtual lines or areas that the performer should trace or strike on the tam-tam (Santini 2020).

Action notation, however, isn't a distinct category from symbolic notation as the two frequently co-exist or blend into one another. *Cruel and Unusual* (Hope 2020a) uses an animated scrolling score where lines indicate relative pitch to be performed by a string quartet. As the curved lines move up and down the page, the performer is asked to make the pitch follow the line. Since these are string instruments, the act of increasing the pitch is the same as moving one's hand up the finger board, following the shape of the line. By notating pitch such that it mirrors the instrumental technique, the boundaries between action and symbolic notation are blurred.

In many action–based animated notations, the mapping from the notation to the action is ambiguous, requiring the performer to interpret the visual movement. One work that employs this tactic is Páll Eiðum's work *Saðras* (2013) where a ball rolls off a shelf and hits the floor, and the performer must choose whether to perform the moment the ball hits the floor, or the motion of the ball travelling along the shelf. While the former is more akin to Smith's radial and collision scores (2016) than action notation, the latter draws a parallel to the physical motions involved in instrumental technique, thereby invoking action notation. This approach is directly used in a latter piece in this research project (*Glimmer* §3.2).

Player Scores

Player scores describe an interaction between the score and some live audio(visual) system. Since an animated notation is probably using the same technology to generate or display itself as the live audio(visual) system, communication between the two is possible. There are at least three identified relationships between score and live audio(visual) system. Smith proposes the situation where the animated notation sends information to the live audio(visual) system, effectively 'playing' it (ibid., p. 112), with the score able to be either pre-determined or created in real-time. Although not explicitly animated, Simon Emmerson puts forwards a similar idea of a 'superscore' (Emmerson 2016, p. 127). The third approach is similar to a sequencer, where a third program controls both the score and audio(visual) system. A notable use of this is with web technologies (Fober, Orlarey, Letz, and Michon 2021).

One area that I believe has not been considered is the combination of player scores and projected scores. While the literature of projected scores does focus on the effect the score's presence has on the audience, whether they perceive the performer as reading the score isn't considered. Further, no apparent consideration is given to whether the audience consider the player score to be actually 'playing' anything. The category is a technical one, and has not been applied to the audience's phenomenological experience of the performance.

2.2 *Together We Tore Through Forested Thickets*

The first work of this portfolio, *Together We Tore Through Forested Thickets* is an animated real-time score player with electronics for any number of violins. The notation software is custom-made in *openFrameworks* while *SuperCollider* is used to create the sounds and sequence the score. Standard staff based notation is used but the rhythm is proportionally spaced using the beam of the note to indicate its duration. This approach is similar to *InScore*'s (Fober, Orlarey, and Letz 2012), yet while InScore could have been used, it would require moving the graphics logic into SuperCollider, which although a valid option, would present additional complexities.

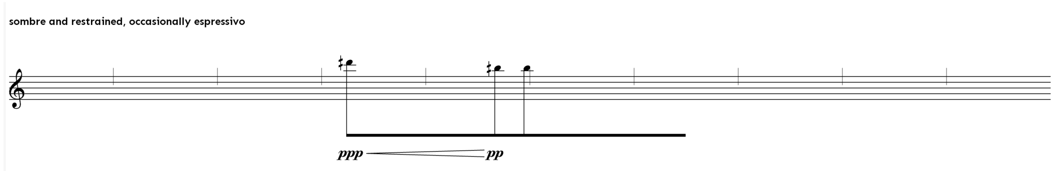


Fig. 2.1. *TWTTFT* example of notation

TWTTFT's score (see figure 2.1 as an example) is displayed on a screen for the performers and mirrored for the audience behind the performers, most likely projected. When reading the score, each violinist is instructed to imagine a playback-head moving across the screen, jumping back to the beginning when the right-hand edge is reached. In this way, the score is similar to the radial scores noted by Smith, but as reading traditional notation upside down is hard, this wrapping approach was a compromise. The speed of this imaginary playback-head is not defined as each performer should progress at a different rate. Therefore, as each violinist reads from a different location on the score at any given moment, the notation forms a changing environment for the performers to traverse. Here, pitches appear and disappear along with dynamics, articulation, and the beam (which is used to indicate duration and articulation when short), thereby creating variety between each performer.

The computer's sounds are implemented by a process that is similar to how the violinists read the score. When a new note appears on screen a synthesis unit (a *SuperCollider* synth) is created, comprised of a modulated ringing filter that operates on the live violin signal. This filter lasts for as long as the note is visible on screen and is modulated by a number of effects controlled by attributes visible in the score, such as the dynamics and articulation. This means that the performers and computer both use the same notational material to make sounds, which is always visible to the audience, creating a similar way for the audience to understand each part. In *Imagiro Landmannalaugar*, there is an interesting moment when the performers physically get close to one another reading from the same location on the score. Here their relationship is at its most clearest. This is emulated in *TWTTFT* by having the filters start ringing when the performers play the same pitch, indicating that they are both at the same location in the score.

While the initial intention was for the score to become an environment through which the audience observe the instrumentalists and the computer moving, upon reflection, it is more likely the audience view the computer as a part of the environment itself. There are at least two reasons for this. First, it isn't possible to observe a concrete location in the score for the computer due to there being as many filters as there are notes on screen. Second, the audio sound is too similar to the violinist and lacks independence. Going forward, this project generally avoids live processing of the instrumentalists' audio. Another potential issue regarding the audience's experience of the instrumentalists is that they are required to understand music notation to begin considering the score as an environment.

For these reasons, *TWTTFT* didn't achieve the goal of this research project, although it has contributed to the technical implementation of how the relationship between computer and human instrumentalists is expressed in the code —as discussed in the following section— and affirmed the importance of keeping the computer's audio distinguishable as a separate sound source from the instrumentalists'.

2.2.1 Technical Approach

Another important way in which this piece contributes to this project is by exploring possible technical approaches to managing the relationship between the instrumentalists, the computer part, and the score. First, though, it is important to understand how the score and computer function.

Score

Made in *openFrameworks*, the score displaying software is similar in design to *InScore*, but has a more restricted and focused set of features, making it easier to use, but coupling it to this specific piece. One example of this coupling is in how time is managed, as it reflects the environmental score approach by placing notes in space on the page rather than in time and existing for some duration. The score accepts 5 types of messages each with individual arguments, sent over OSC (open sound control) (Wright et al. 2003). The important message that reflects how the environmental score influenced the design of this interface is the `/noteOn` message which is structured as follows:

- `/noteOn`: display a note.
  - `Time`: a number between 0 and 1 indicating where on the horizontal axis the note should appear.
  - `Pitch`: a midi number rounded to the nearest 0.5 allowing for quarter tones.

- `Dynamics`: a whole number: 1 through 7 map to ***ppp*** through ***fff***, by adding 8 to the number, hairpins are inserted to the next dynamic, -1 maps to ***sf*** and -2 to ***sfpp***.
  - `Articulation`: a whole number: 0 none, 1 louré (tenuto-staccato), 2 tenuto, 3 wedge accent.
  - `Glissando`: a number: 0, don't start glissando or end previous, 1 start glissando.

- `/noteOff`: remove a note from the display.
  - `Time`: a number between 0 and 1 indicating the point in time where the note to be removed exists.

As can be seen, to create a note the spatial position across the screen is given, and to remove the note the same position must be supplied. This has the added effect that there can only ever be one note in one position. This way of thinking about note placement and time as spatial events across the score is reflected in the sequencing, where shapes are used to control structure.

Audio

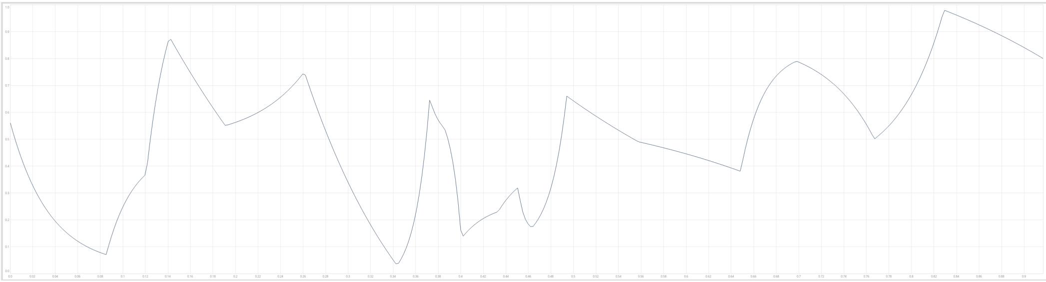
The audio is created using *SuperCollider*. There are a few effect-bus synthesiser units for reverb, global detuning, along with a master out. The main source of sound comes from a series of synthesiser units that are created when each new note appears on the score. These use a ringing filter, some light compression and a wave-shaper, and all multiplied by smooth stereo noise modulating at $0.23Hz$ to create a shifting stereo effect.

Sequencing

The score software, the audio generation, and the human instrumentalists all do nothing by themselves: the score displays an empty stave, the audio remains silent, and the instrumentalists twiddle their thumbs. They all need to be given instructions, which is done through the sequencer. The main advantage of this is in decreasing the amount of *time–mutable state* across the system, making it easier to reason and understand how the piece as a whole changes — which makes composition easier —, and additionally should make jumping to different sections of the piece in rehearsal trivial as these changes don't have to be coordinated across separate systems. This means that each system does not have its own timelines and sequence of events, but is coordinated by this middle piece of software (what I am referring to here as the 'sequencer'). While this isn't entirely possible as audio processing (notably filtering) depends on what happened previously and therefore does store some time–mutable–state, the *goal* is to make the system as time independent as possible, or at least be accepting of the differences caused by jumping to different sections. In *TWTTFT*, the score software and audio software didn't quite achieve this as jumping around in the timeline ran the risk of creating notes then never removing them, however, both are tempo independent, meaning one can easily jump to a section by linearly interpolating through the timeline at an increased pace, both forwards and backwards. This is addressed in the final two works, *Breathe* and *Jellyfish*, where the piece can be performed in any order, and while a different performance will result, nothing technically 'breaks' or otherwise produces results outside the bounds of what constitutes the piece.

The score controller itself consists of a series of sections each controlling certain parameters as they evolve through time. This is done by specifying envelopes which can be assigned to parameters or combined together to produce new envelopes — see figure 2.2. Following this, the parameters undergo some automatic transformations and are routed to either the animated score, the performing electronic part, or both. Examples of musical parameters that are shared between the electronics and the score are: pitch, duration, sustain and position. Since the instrumentalist will ultimately be following the score, this creates a shared interface through which both human and computer can be composed together. An example of this is that the pitches that are set to the 'midinote' parameter are sent to the animated notation and also used to trigger the ringing filters — figure 2.2: line 1. Another example of this is dynamics, which both writes the dynamic on screen and changes the amplitude of the synth. This approach is very similar to both Morton Subotnick's concepts of a ghost score (Whipple 1983) and Emmerson's 'superscore' (Emmerson 2016), however the signal is also being sent to the performer through the score.

After early tests of the work by myself, the interaction between the performer part and the audio appeared incidental and the computer part lacked any appearance of agency, seemingly forming a part of the environment in which the human played the main role. Two possible responses to this are considered. First, keeping the environmental score but making the audio part have some visual effect upon the score. Second, removing the environmental score and directly giving the computer some visual form. The first approach is explored in the following piece, *Mercurial Memory*, and the second in the following chapter with *WWSTTW*.



(a) An envelope.

```
1 \midinote : ~sections[\get].(\03, \falling)
2         .linlin(0, 1, ~note.("b'"), ~note.("g''")),
3
4 \sustain  : Pseed(500, Pwhite(0.0, 20.0, inf))
5         + ~sections[\get].(\05, \smooth)
6               .linlin(0.0, 1.0, 80, 10)
```

(b) *SuperCollider* code for the score controller. Lines 1–3 reads, assign to 'midinote' the following, get an envelope from section 3 called 'falling' and scale it between the pitch B above middle C and the G above that.

Fig. 2.2. *TWTTFT* mapping example.

2.3 *Mercurial Memory*

This piece is for flute with animated notation and fixed audio playback all made inside a web browser — see figure 2.3. The electronics are limited to just sound file playback: they do not use any live signal processing. The animated notation consists of a bundle of black lines that flow across the screen, the bundle can spread out in places, and the lines can change length as they travel across the screen. This gives the impression that there is some form of fluid in which the lines are travelling. The flute reads this by changing their register with the height of the lines, their tone with how spread out the bundles are, while an additional fingering chart is provided below this notation, see 2.3. Conceptually, *Mercurial Memory* plays with a metaphor for how some of the earliest computers' memory worked, e.g., the *Electronic Delay*

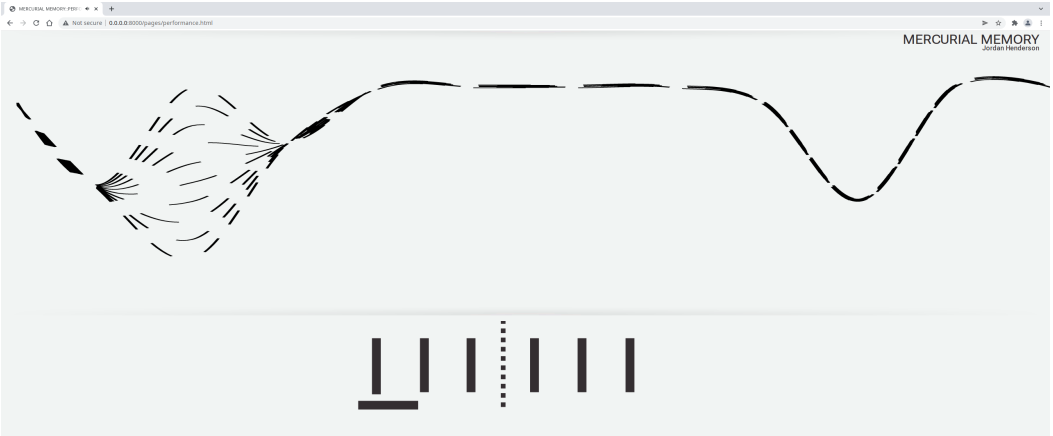


Fig. 2.3. Screenshot from *Mercurial Memory*

Storage Automatic Calculator (EDSAC) (Wilkes and Renwick 1950), in which large tubes of mercury had sound pulses played into one end and recorded from the other. When the recorded sound was amplified and fed back into the tube, the tube of mercury acted like a memory buffer, not too dissimilar from how modern random–access–memory (RAM) works today. This idea of sloshing liquid memory forms an artistic starting point to this piece, where the computer part 'injects' sound into this liquid, causing turbulence which the flautist reads as a score. There are lines suspended in this medium showing these currents — the black horizontal lines visible in figure 2.3. Reading *Mercurial Memory*'s score is done similarly to *TWTTFT* in that the performer imagines a play-head, but here, the play-head's movement is determined by the currents of liquid, which is deduced from the movements of the lines. The vertical displacement of these lines is then read as a pitch contour with a fingering chart provided below.

Similar to *TWTTFT*, this piece has not been performed professionally, however myself and a few others have tested the score, and there is a video of the animated notation and electronic playback in the portfolio (§B). One issue with this piece is that while the computer part in *TWTTFT* was positioned as the environment which the instrumentalists traversed, in *Mercurial Memory*, the instrumentalist is still inside the environment but the computer part is now outside of this environment, speaking from above, as if looking down from the heavens and speaking to the Earth below. In a sense this piece has overcompensated as the computer now sits outside rather than within the environment. What was needed was a way of positioning both the human instrumentalist and the computer inside the same environment, side by side, together. This is addressed in the subsequent chapter through *WWSTTW*.

A similar performance dynamic can be found in *Line Crossing* by John Burton (2019) in which a performer turns and faces a projected score that also has live computer generated sound. Technically, in terms of programming and control flow, *Line Crossing* has the score controlling the computer sounds (Burton 2019b), whereas in *Mercurial Memory*, a third program controls both. However reflecting on my experience as an audience member, while my perception aligns with the technical implementation in *Line Crossing*, in *Mercurial Memory*, I perceive the computer audio as affecting the visual score, which in turns affects the instrumentalist. While it would be an exaggeration to state that the control flow inside the program directly affects the audience's understanding of the relationships between the computer and performer, it seems likely that a centralised sequencer offers a more flexible model capable of producing many types of relationships between computer and performer from the audience's perspective. This is primarily because it clearly separates the visual/score production from its sequencing and generation. As this project progresses, this technical architecture is maintained and expanded upon, and begins to mark the movement away from an animated notational approach in this project.

2.3.1 Notation

The notation in *Mercurial Memory* consists of a bundle of black lines that flow across the screen as being pushed by currents. Similar to *TWTTFT* the performer imagines a 'playback–head', however, here, rather than the playback–head moving at a steady pace across the screen, the performer allows it to be moved across the screen pushed by the underlying currents. Then, the black lines at this position form the notation to be read. The height of these lines indicates what register the performer should play in. As the fingerings are specified in a chart below, this is achieved through overblowing. The black lines are in bundles that get displaced vertically - the tighter this bundle, the clearer the tone should be; when it is dispersed, an airy tone should be played. Compared to *TWTTFT* where the audience needed to understand staff notation, this is a much simpler ask of the audience, lessening the required knowledge needed to understand the notation. However, the performer's location within the score is still not obvious, and perhaps is even less obvious than in *TWTTFT*, as the performer moves around with the flow of the graphics.

Yet, the main issue with this notation was the lack of variation in the performance, which was noted in both my testing of the piece and others'. This could have been for two main reasons. First the notation changes too fast, requiring a vast amount of mental energy to read, leaving perhaps insufficient time to either read or perform the nuances of the notation. While this is similar to the concerns raised by Freeman regarding how animated notation can turn into extreme sight reading (2008), it is slightly different as this questions whether the animated nature of the notation is antithetical to the nuance found in static score performances, as might be achieved by the performer repeatedly practising material. Second, and in contradiction to the first point, the notation might be too simple, and in attempting to avoid extreme sight reading, the notation has lost some of its expressive power. This could be resolved through another test, however as this is ancillary to the main research question, this isn't undertaken in this research project.

In the next chapter a different approach is taken to notation. Rather than produce an animated visual score, the computer part itself is given a visual body. *WWSTTW* then uses notation as a way of coordinating how the instrumentalists should respond to the computer's audiovisuals.

2.3.2 Sequencer

The interaction between the performer, audio, and score is implemented using the sequencer approach similar to that in *TWTTFT*, with a slight difference in how the events are managed and conceived of, making it easier to compose and jump to different sections in rehearsal.

Technically, the composition is split into four processes, called 'players'. A player has a timeline and a set of actions. The players are: the graphic's shape; the flow of the graphic, which also determines how the lines are 'stroked', giving the appearance of many small lines; which flute keys are currently depressed; and which sound files should be played. Each of these players is split into two components, a sequence of actions or events and a density distribution.

For the line's shape, some of the available actions include setting the vertical position, applying force to the positions, altering the thickness at certain points, and smoothing the line. The line's flow and stroke style are simply altered with a speed/acceleration and a size parameter. The fingerings can be set or toggled. Finally, the sound system functions by choosing a sound file from a collection, where each is ranked in terms of 'energy', allowing for a more or less energetic file to be played.

Each section of the piece has a fixed time duration and the events are distributed in this section by supplying the number of events and a density function which can be used to shape how the events are distributed across the section's duration. These density distributions can be shared across multiple players. As an example, this could allow a shared increasing distribution (an *accelerando*), but with each component playing a different number of events in the same allotted time creating complex multi-rhythms that accelerate. It is particularly useful for the sequences where players start together and drift apart or start a part and come together. The former happens in the opening section as the audio starting is correlated with the bursts of movement in the score, and gradually drifts apart.

The downside with this sequencer is that it doesn't allow for a concrete description of the relationship between the instrumentalist and the computer audio part. Instead, the actual behaviour of the instrumentalist depends on how they read the score. Ultimately the issue is that what is important in this project, and would therefore be useful to express directly in the score, is the relationships between the elements (computer and human), rather than the elements themselves. From this moment on in the research project, finding a solution to this becomes paramount.

3

Relational

TWTTFT and *Mercurial Memory* had two important issues that this chapter of the research project attempts to address through *Whilst We Speak Through The Wind* (*WWSTTW*) and *Glimmer*, namely:

1. the inequality as perceived by the audience between the generated computer part and the instrumentalist,

2. the failure of the notation to directly encode the desired relationships at a compositional level, meaning the composer and performer have to create these indirectly.

Both of these concerns are addressed in *WWSTTW*, whereas *Glimmer* interrogates the ability of a performer to interpret an abstract video as bodily movement.

Therefore, the initial focus of this chapter is on relationships and how to encode them notationally. For sake of convenience, I will be referring to this as a 'relational' approach to notation, which, while similar to Bourriaud's 'relational aesthetics' (Bourriaud 1998) in that it is concerned with relationships, should be understood as a separate concept, as 'relational art' is often concerned with audience interaction — a topic outside of this research's scope. This is motivated by Macke's emphasis on the importance of relationships in communication (Macke 2014). Before talking about these pieces however, it is worthwhile to quickly explore the terms 'relationship' and 'related'. While definitions are not always the most helpful, the following from Oxford English Dictionary's definitions have proven to be helpful in unpacking these terms (omitting the obsolete and legal terms).

> Relationship:
>
> 1. The state or fact of being related; the way in which two things are connected; a connection, an association.
> 2. A connection formed between two or more people or groups based on social interactions and mutual goals, interests, or feelings.
> 3. An emotional and sexual association or partnership between two people.
>
> (Oxford English Dictionary 2023c)

> Related:
>
> 1. To recount, narrate, give an account of (actions, events, facts, etc.).
> 2. To have reference *to*; to refer to.
> 3. To have some connection with; to stand in relation to.
> 4. With to. To understand or have empathy for; to identify or feel a connection with.
>
> (Oxford English Dictionary 2023b)

Two terms are shared between these and are worthy of consideration themselves, these are 'connection' and 'association'. While a connection implies some affective linkage between entities, association is less obvious.

> Association:
>
> 1. The action of combining together for a common purpose; the condition of such combination; confederation, league.
> 2. A body of persons who have combined to execute a common purpose or advance a common cause; the whole organization which they form to effect their purpose; a society; e.g. the British Association for the Advancement of Science, the National Football Association, the Church Association, the Civil Service Supply Association.
> 3. Union in companionship on terms of social equality; fellowship, intimacy.
> 4. The action of conjoining or uniting one person or thing with another.
> 5. *Ecology*. A group of associated plants within a formation.
>
> (Oxford English Dictionary 2023a)

With these three definitions in mind, a new criticism of the previous two works and environmental scores (as a means of creating relationships) is raised. The third definition of 'related' reads 'to have some connection with; to stand in relation to' which the environmental score achieves by physically (or virtually) standing the performers in space; it is through their location they are connected. However, this type of connection is more analogous to that between oneself and the nearest piece of furniture: while there is a connection, it is not an association as the entities are not 'conjoined', ‘combined’ or in a ‘companionship’ and acting with shared purpose. This explored in *WWSTTW*.

3.1 *Whilst We Speak Through The Wind*

WWSTTW (§C) is a work for bass clarinet, alto flute, and computer generated visuals. The generated visual is a tree that sits between the performers on-stage and sways side-to-side as if each performer’s breath was pushing on the tree. This is done by placing a microphone in front of each performer and measuring the amplitude, which in turn, applies a force to the virtual tree to look as if the player’s breath is blowing the tree. Initially the tree is reactive to the performers, who ignore it, but as the piece progresses, the tree becomes a source of instruction as the performers are tasked with responding to its movement. In the portfolio there is a demonstration of the tree and the electronics as it responds to the sounds from microphones[1]. There is also a score that the performers follow that asks them to respond to each other in specific ways, and at the end of the piece, asks them to respond to the computer generated visuals in a similar way.

This section will consider three things, the notation (§3.1.1), staging (§3.1.2), and finally gesture (§3.1.3). Whereas *TWTTFT* and *Mercurial Memory* sought to create notation that provided extra information demonstrating and explaining how the performers and computer were related for the audience’s benefit, the approach taken in *WWSTTW* is to describe relationships explicitly in the notation such that the performer can enact them. These

[1]This work was set to be workshopped professionally, but due to the pandemic had to be cancelled. In part, this is a reason why the final pieces of this portfolio are performed by myself.

relationships are communicated by describing connections between the ensemble members — at this point in the research project the computer generated audiovisuals start to be considered a performer in their own right. While the staging for the previous works wasn't fully considered, it is with this piece that how the performers (human and computer) are positioned on stage becomes important, both in terms of how this impacts the practicalities of performance, and the audience's perception of the performers and the relationships between them. Gesture is a new concept introduced here and forms the basis for how communication is considered to take place between performers.

3.1.1 Relational Notation

In response to the difficulties encountered in earlier works which attempted to generate dynamic relationships evolving over the course of the piece, *WWSTTW* has turned to using notation as a way to describe relationships between performers through connections. While 'connections' could take on many forms, in this piece causal connections are explored where one player's actions stimulate a response in another. As a simple example of a connection, imagine an ensemble with two performers, A and B, an instruction for performer A might read, 'copy performer B's dynamics', thereby creating a causal connection between A and B. While this is a simple connection, a more complex feedback situation can easily be constructed if B follows A in some way — this approach has many commonalities with *mapping* as is common in the literature of NIME (new instruments of musical expression), a topic that shall be revisited in the subsequent chapter. As one concern of this research project is to make communication observable to the audience it is important that these connections can be observed by the audience, whether this be audibly or visually. In the case of following another performer's dynamics this should be audible to the audience, but the audience might also be able to observe the performers attending to one another by watching their gaze and/or body language, which is explored specifically in a later subsection (§3.1.3). The specific approach to notation here (see figure 3.1) I am calling *relational notation*, however, there is a long history of expressing relationships in scores, some of which are similar to the connection-based approach I have taken.

Perhaps one of the earliest uses of directly expressing relationships in scores came from *For 1, 2 or 3 People* by Christian Wolff 1964, which uses a complex system of graphics to notate cues and other relational information. Yet, this is perhaps more common in text works, with notably many of the works from Pauline Oliveros's *deep listening* employing a similar technique (Oliveros 1971).

Michael Nyman in *Experimental music*, describes this type of musical process as a 'contextual process' which is 'concerned with actions dependent on unpredictable conditions and on variables which arise from within the musical continuity' (Nyman 1999, p. 4). This is contrasted against 'people processes' where each individual acts separately (ibid., p. 5), presumably, without relationship.

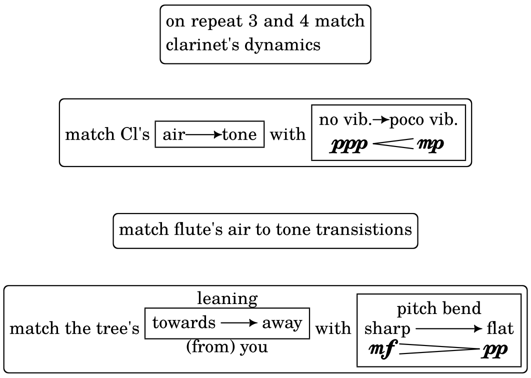
Jennie Gottschalk, while providing a comprehensive review of experimental music, broadly describes similar music as an 'interactive process' (2016, pp. 188 - 217). Here three distinctions are made between improvisation, indeterminacy and interaction, which all share a concern with the unknown (ibid., p. 188). Of note are: improvisation which is a temporal unknown, not knowing what will happen one moment to the next; and interaction which is a social unknown, not knowing because one does not know what the other is doing. Drawing from multiple authors, including George Lewis (2002), Gottschalk ultimately concludes the split between indeterminacy and improvisation is not useful and instead proposes focusing on the 'placement of agency in a musical interaction [which] reveals more about it than any categorization' (Gottschalk 2016, p. 189). This emphasis on agency is quite useful for this research project as it allows one to reconsider how agency is placed within an ensemble of both human and computer members. A similar project focusing on the placement of agency in the more general case of interactive music systems can be found in Murray-Browne (2012). Although not considered in this chapter, agency becomes a more important topic in the following.

James Saunders, drawing from Donelson R. Forsyth (2006), further breaks down relationships in ensembles into several types of processes: formative, influence, performance, conflict, and contextual processes (Saunders 2021). While there is overlap between the processes they can broadly be categorised as follows.

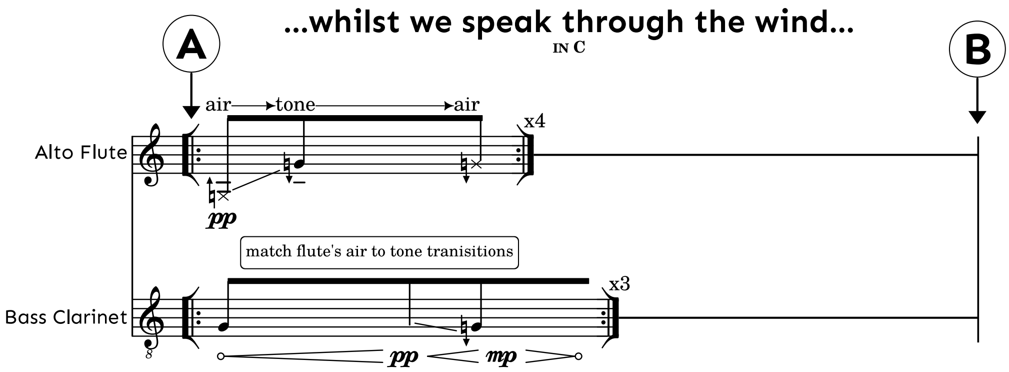
- Formative processes form groups out of individuals into collectives. They have the power to comment on the structure of social space and musical institutions by creating new hierarchies between performers (ibid., p. 16).
- Influence processes deal with how performers exert influence over each other and create power relationships (ibid., p. 17).
- Performance processes consider how a group can or cannot reach their goal, 'the resultant music can be experienced as a sonification of

dynamic interaction between people and their progress towards a goal' (Saunders 2021, p. 18).

- Conflict processes seek to impede a performer's ability to complete a given task by creating disturbing relationships (ibid., p. 19).
- Contextual processes consider social and environment factors of music performance (ibid., p. 20).



(a) Examples of connections



(b) Simplified section A to B

Fig. 3.1. Examples of notation used in *WWSTTW*

WWSTTW attempts to bring some of these ideas together by using text instructions asking the performer to match some sonic parameter of theirs to some observable parameter of another performer. These can be simple, as in 'match A's dynamics', or involve transferring some continuum of sound onto another parameter. See figure 3.1a for examples used within the score.

Using Saunders's five categories, the notation most clearly aligns with the formative process, in that it brings groups of performers together by putting them into relation through connections, but could also be considered a performance process as the resulting sound is a sonification of the dynamic relationship between the performers. However, although making the dynamic relationship observable to the performers is a goal of this project, it alone does not meet the criteria of an association as there is no shared goal that the performers are united to achieve. This concept is addressed shortly through the concept of goal oriented gesture (§3.1.3), however this is revised in the subsequent chapter.

Importantly for working with computers, a notational schema can be created from this model that could be encoded into the computer. This specification is as follows:

1. an observation;
2. a linking process that describes how the observed and action are related;
3. and a resulting action.

To examine this in more detail, I will return to the example of one performer following another's dynamic. Here the *observation* is the other performer's dynamic, the *linking process* is to follow, and the *resulting action* is a change in dynamic. The resulting action is simple for both the performer to understand and for the computer to implement inside the programming language. The linking process can be more problematic in that the computer often has a greater potential to enact more complex linking processes, however, if these are kept simple as in 'follow', or even 'inversely follow' then there exists a middle ground where instructions can be performed by either, thereby creating a compositional technique which is equally applicable to human and computer. However, the observation stage needs unpacking as it would be a mistake to equate the dynamic observed by a computer with either the way a human hears, or the way a human interprets a dynamic marking.

When the instructions in *WWSTTW* say ‘match A’s dynamics’, although the other performer has access to the complete score and could figure out what dynamic marking the other player has, this should be done through the performer’s own perception and understanding. In this way, the phenomena are transduced through the performer’s sensory organs, their perspective within the performance environment, their psychology, and musical expertise. As an example, consider a duet with a trombonist and a harp player. While a score might have both instruments playing ***p***, for the trombonist to observe the harpist playing ***p***, they must have knowledge of the harp’s dynamic range, how that dynamic affects timbre, or use their own instrument as a point of comparison. Further, if the performers can see each other, then how much visible effort each player exerts might contribute to the perceived dynamic, which might depend on the performer’s playing style and body language (e.g., perhaps how much they move their arms in the harpist’s case). This means observations cannot be considered descriptions of absolute objective reality. Instead, each observation should be considered as a reflection of the observer.

Considering the score in this way simplifies the composition of the computer’s part as it turns using machine listening into a creative task where the machine’s sensory organs and perception are created. The microphone then becomes a sensory organ similar to the human’s ears, and the algorithm used to determine dynamic level is a part of its psychology, physiology and musical understanding. While it is common to use human calibrated algorithms for things like loudness —e.g., Loudness Units Full Scale (LUFS)—, by considering the algorithm a representation of how the computer *perceives*, it invites all algorithms to be understood as describing and implementing a type of experiencing 'thing' — perhaps we might call this a body? Although *WWSTTW* didn't fully explore the creative possibilities here, the final two works of this portfolio, *Breathe* and *Jellyfish* have used this idea creatively to construct more elaborate virtual performers.

Although *WWSTTW* was designed so that the relational notation could be given to the computer, the main focus of this work was on finding an easy way to communicate relationships to human performers, and as a result, the computer part has static connection that imitates how a tree moves in the wind. This forms a part of the work's narrative as the two woodwind players ignore the tree for the majority of the piece and only towards the end begin to see how they have been affecting it and to respond to those changes. Theoretically, this could be made dynamic easily and evolve as the piece progresses, however practically doing this in a reasonable way was challenging and is explored in more detail in the following chapter.

Specifically, the computer understands and responds to the performer's dynamics by bending away from the sound as if pushed by the air that comes from the performer's instrument. This is achieved by placing microphones in front of each performer and taking the amplitude. The algorithm here uses a low-pass filtered root-mean-square of the amplitude from the microphones. These two amplitude levels (from each performer) are combined in two separate ways, one for the *sway* of the tree (horizontal motion), and the other for the *stretch* (vertical motion). For the sway an averaged direction is computed. Below is the simple formula for this, where F is the flute, C the clarinet, i the flautist's facing direction, and j the clarinets. As per the score, the two performers sit directly facing each other, meaning the result of the following equation is between negative and positive 1.

$$\mathrm{sway} = |iF + jC|$$

The other metric that is taken from the amplitude is their cross product which is used to control how vertically stretched the tree is. Conceptually, rather than measure the direction of the wind, this measures the overall intensity, such that when both instrumentalists are loud they do not cancel each other out, as they do in sway, but grow together, pulling the tree upwards from its roots.

$$\mathrm{stretch} = iF \times jC$$

While this description of the processing could be simplified to just the sum and the difference of the amplitudes, by describing the direction each performer is facing as three-dimensional vectors, it allows for editing these during sound check, so that the exact physical locations of the performers can be virtually modelled, creating a more realistic blowing effect — the staging implications of this piece are considered in a subsequent subsection.

Both the sway and $\mathrm{stretch}$ are used inside the audio processing as well, which alters pitch shifting and comb filters applied to the instruments' live sound. This can be expressed in the relational notation outlined earlier and used by the instrumentalists:

- sway away from the loudest instrumentalist
- match overall ***pp*** $<$ ***ff*** with shrink $\leftrightarrow$ stretch

While the specifics of the audio processing are neither novel nor important enough on their own, importantly for this project, this connection is possible to express in the notational schema of relational notation, meaning that during composition, it is possible to use exactly the same conceptual tools to manipulate both human and computer. While this wasn't actually used in *WWSTTW*, the intention was for this to be relevant in a later composition. Although the research project drifted away from written musical notation, in the final two pieces this is realised.

In addition to the relational notation explored so far, *WWSTTW* employs a different technique regarding rhythm and synchronisation. As can be seen in figure 3.1b, the score uses proportional notation with lots of repeated melodic fragments. The melodic fragments do not imply any 'vertical' synchronisation, instead the performer need only reach the rehearsal marks at the same time. This means each performer must listen carefully to the other and adjust their tempo so that they finish together. This creates a pattern similar to a large polyrhythm at the fragment level — figure 3.1b is like a 4 against 3 rhythm. To achieve this, the performers are given the full score so that they know what each other is doing, allowing them to speed up and slow down their material so that they both reach the section's end together, thereby forcing the performers to work as a team and form some kind of association united in a common goal. Although this technique is not applicable to the computer, as it would require some form of score following software (such as *Antescofo* IRCAM 2018), a similar process occurs in the final piece *Jellyfish*.

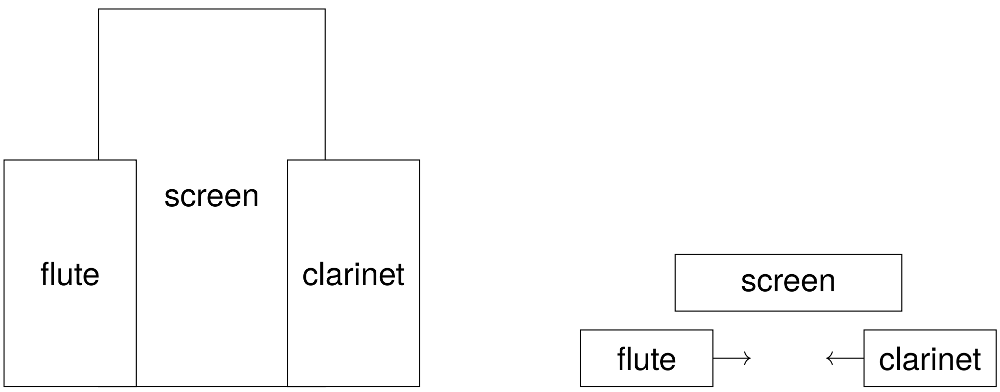
3.1.2 Staging

In *TWTTFT* and *Mercurial Memory* exactly how the computer graphics should be positioned on stage wasn't fully considered. This is in part because they haven't been performed, but also because the computer's role and identity isn't connected to the graphics. *WWSTTW*, however, uses the generated graphics and sounds as a single entity, such that the sound appears as if it comes from where the graphics are placed on stage. The graphics should themselves appear as if they are on stage, thereby asking this composition to use some theatre techniques in stage design. While this is carried on throughout this research project, notably in the final two works, it first becomes a concern here.

There were a few properties the staging needed to achieve: first, the humans must be able to see the computer graphics as required by the score; second, the audience should observe the computer and humans in a similar way, meaning they should be close together on stage and be roughly a similar size; and finally, the computer graphics should be physically positioned inside the ensemble. While there are many existing works for computer and instrumentalist that do consider staging, few directly pursue this goal. Instead they either have the performer turn from the audience to read from the screen, as is common in many animated notation performances, or the visuals are projected behind the performer, which they ignore (Hamel 2015; Corral 2020). Seldom is the space itself - between the visuals and the performer - staged as the audience's main focal point. One work, albeit not involving technology, that has a similar relationship and that proved influential, is Mauricio Kagel's *Match: for 3 players* 1967, in which two cellists sit opposite each other with a percussionist centre stage.

> Both cellists sit on the forward edge of the stage as far apart from one another as possible — on podiums of about 50 cm height — almost in profile to the audience. The percussionist, who plays from a somewhat set-back position at stage center, should be easily visible to the two cellists. The music desks for the cellists must be very low.
>
> (ibid., p. ii, performance notes)

This layout is very similar to that employed in *WWSTTW*, with the exception that the performers sit closer together and do not require podiums and that the tree is positioned in the centre of the stage squarely between the instrumentalists, taking the place of the percussionist. Kagel might have chosen this specific layout to facilitate a certain type of relationship, often described as a musical competition (perhaps tennis?) where the percussionist referees (Griffiths 2010, p. 198; Heile 2006, p. 46). This piece is also influential when considering gesture in the next subsection as the stage design allows for the performers to communicate gesturally, using their bodies to convey information for synchronising their parts, and to navigate the changing relationships as the piece evolves.



(a) Front view

(b) Top view, arrows show facing direction

Fig. 3.2. *WWSTTW* stage layout

WWSTTW is the first piece in this project that attempts to extend the on-stage movements and physical space into the the digital, which is done through the use of the performers' breath blowing on the tree and the encoding of the physical direction in the calculation of the sway and stretch. A similar approach is revisited later in *Jellyfish* where the visuals are designed to appear as if they reach across the stage towards the instrumentalists, thereby creating a connection between the virtual space and the physical.

3.1.3 Gesture and Goal Oriented Communication

One aspect that arose from *WWSTTW* was a focus on gesture and how performers use their bodies to communicate, and whether the computer generated visuals could communicate gesturally. As some form of equality is sought in this project between human and computer performers, whether the computer's movements are experienced by the audience as communicative gestures or incidental movements becomes a concern, as the gestures, eye glances, subtle nods, and other body movements that human instrumentalists naturally make when playing together are visible to the audience and show the performer's participation in the ensemble. If the generated audiovisuals cannot gesture themselves, nor respond to gesture, perhaps they might not be observed as participating equally. This issue of perceived participation could also be a contributing factor to why *TWTTFT* and *Mercurial Memory* didn't achieve the goal as in the first case the computer had no visual presence nor affect, and in the second, they acted from outside of the ensemble space.

This subsection will proceed by looking at Kagel's *Match* and consider two types of gesture, what I am terming, *mutual synchronisation gesture*, and *directed gesture*. Mutual synchronisation gesture is when both performers work together, and directed is instructional without reciprocation or dialogue.

In the opening section of *Match* the cellists 'bounce' pizzicati off each other, alternating in a complex rhythm at a very slow tempo that also includes pauses. To perform this, the players must visibly gesture and communicate with each other to ensure proper synchronisation. Out of the three performances readily available on the internet in video form (Öhman et al. 2016; Unterman et al. 2018; Ensemble Offspring et al. 2010), in all three cases the cellists gesture towards each other throughout this short section indicating where each beat should fall, communicating how long the pauses are, and marking a change of relationship from playing to listening. This is primarily done with head nodding, but also through raising and lowering the shoulders. Empirical studies around simpler pieces of music have shown that similar gestures are used to aid in the synchronisation of rhythmically complex works, where the types of gesture aren't simply instructional but are attempts to build a shared sense of time and to work in synchronicity (Bishop and Goebl 2017). These gestures aren't there just to communicate some information from one performer to another, but may also be means of seeking information about the other performer, as it has been proposed that performers hold mental models of each other in their heads and update these as the performance progresses (Keller 2008). Therefore this type of gesture can be considered a *mutual synchronisation gesture*, as it isn't stating when something should happen, but seeking to construct a moment of shared synchronicity in the future. These types of gestures are not written into the score (Kagel 1967), but are a natural result of the music making process and can therefore be found throughout many different musical practices where performers must construct the piece live. To clarify, these gestures may not exist when working with a metronome, or some other time keeping device. Additionally, while the role of the conductor is complex and extends beyond that of the metronome, it is possible that their presence might inhibit such gestures, as the performer no longer communicates directly, but through a third party. Further, it isn't clear that the gestures given by a conductor are mutual and allow the possibility for the instrumentalist to communicate back, in this case, the conductor's gesture could be considered directed and instructional. In *WWSTTW*, this type of gesture hopefully takes place between the performers by nature of the synchronisation needed to align each section. This was one of the potential functions of this type of notation. However, as this workshop was interrupted by the pandemic, the efficacy of the notation remains questionable. Further, mutual gestures do not occur between the tree and the instrumentalists; this issue will be addressed in the final work *Jellyfish*.

Another important type of gesture that is used in Kagel's *Match* is the directed gesture, which unlike the previous, is a one-directional indication that something is expected to happen and is explicitly notated in the score as solid vertical arrows (Kagel 1967). This type of gesture is often about indicating a specific time that something should happen and does not involve any negotiation. Throughout *WWSTTW* this type of gesture has been avoided but is commonplace in other human–computer interactions. This is also the basis of many different criticisms, including: the distinction between reactive and interactive music systems, Di Scipio's refocusing from 'interactive music composing to composing musical interactions' (Di Scipio 2003), and Alice Eldridge's distinction between pipeline and feedback approaches to interaction (2008). This final point about the difference between pipeline and feedback approaches to interaction almost directly correlates to the directed and mutual synchronous gestures noted here and this is returned to in the next chapter where a more technical implementation is sought.

Gesture raises an issue when applied to the computer directing the human and when there is mutual gesturing. For the audience to perceive a gesture they must understand the movement as communicating something. In the common case of a nod or beating time this movement is common body language and a socially expected norm in many styles of musical performance. However, when the computer generated body isn't that of a human, and perhaps not even an animal, it might be difficult for the audience to perceive these movements as gestures. This concern hence becomes a major research question for this project — how can an audiovisual computer-generated body be designed such that an audience perceives it as communicating gesturally and as an equal partner with the human? The subsequent chapter addresses this issue, although struggles to find a concrete answer (§4).

Turning now to *WWSTTW*, do the movements of the tree and the performer's response to them suggest to the audience that the tree is actively directing the human instrumentalists? Or that there is mutual synchronisation taking place? In both cases, it is doubtful. There ought not to be any mutual synchronisation observed because the tree doesn't respond in turn to the human's gesture. Rather than direct the instrumentalists actively, it is instead passively read by humans: the tree doesn't expect a response and nor does it react to a successful one. This follows on from the previous chapter where the role of the instrumentalist was to read the visuals as a score. In this way, despite placing a visual entity on stage to appear as the embodiment of the computer's agency, the audiovisuals are still viewed as an animate inanimate object — as a score.

To help understand why, some useful insights can be found in the field of animal communication as studied in biology. Drawing from J.L. Austin's *How to Do Things with Words* (1975), Amphaeris et al. (2023) propose a model of communication focusing on what they term 'mutualistic interactions'. These interactions imply the kind of equality this project is seeking, although not applying so readily to the directed gestures discussed before. In summary it is a three step process:

Mutualistic Interactions

1. A signaller A, has a goal G, that motivates the need for a communicative signal S, which is communicated in some modality (e.g., acoustic).
2. The perceiver B, perceives the signal S, and forms an interpretation of the goal $\bar{G}$.
3. For this be to a successful communication, G and $\bar{G}$ must be at least be partially related.

However, this theory doesn't address what a third observer (the audience) experiences, and therefore is of limited use when being used as a framework to create the perception of communication. Here, I shall propose one following a similar model that focuses on when both performers alter what they are playing.

Observed Mutualistic Interactions

An observer O perceives two entities A and B. O perceives:

1. A making some signal S,
2. B observing S,
3. A and B changing their behaviour together,
4. A and B together reaching some prominent state N,
5. A and B no longer changing their behaviour together,
6. therefore, a causal relationship between A and B.

Working through this list and using it as a mechanism for criticising *WWSTTW* proves useful when considering whether the tree is observed as mutualistically interacting with the instrumentalist. I will consider the moment that the instrumentalists begin responding to the movements of the tree, where the audience is O; the tree, A; one of the human performers, B; and the tree's movements, S.

First, is whether the audience perceives the tree make a signal. Although the tree is definitely observed moving side to side, it has been doing this throughout the performance, which the instrumentalists have been ignoring. Perhaps by ignoring the tree throughout the beginning the performers are signalling to the audience that they should also ignore the tree and consider it as a background or accompanimental feature rather than an equal performer. This invites the hypothesis that the computer-generated body can be given a more equal role from the audience's perspective through observing the performers treat it as one.

The second point, concerning the audience attending to the human instrumentalist responding to the tree's movements, is probably noticeable by the audience, as previously the instrumentalists had been looking at each other and they now turn slightly to look at the tree. Here it is the instrumentalists' attention which has changed and is visible to the audience.

The third point concerns a change of behaviour from both the instrumentalists and the computer. While the tree does respond to the instrumentalists' playing, it is very predicable and similar to the opening section when the tree was being ignored. Therefore it is unlikely the audience notice any substantial change. Although this was intended during the composition — as the continuum between inanimate animate thing and body with agency is a potentially artistically fruitful spectrum to traverse —, with hindsight it does interfere with the project's goals and is not carried forward. The fourth and fifth points about reaching some prominent state and halting their changing behaviour is also unlikely to be observed for the same reason regarding the lack of a definitive behavioural change.

While the final point about noticing some causal relationship between the instrumentalist and the tree could be observed, it is unlikely that is a mutual relationship. Perhaps the tree is likely to continue the accompanimental role that was present in the opening while only a change in the instrumentalists' behaviour is noticed. This echoes a similar concern from *TWTTFT* and *Mercurial Memory* where bringing the computer generated entity forward into the ensemble proves problematic.

Many of these concerns are addressed in the following chapter through *Breathe* and *Jellyfish*, however *Glimmer*, the final piece of this chapter, explores the slightly different issue of how the instrumentalists might move away from 'reading' the computer body, perhaps towards a more physical intuitive response.

3.2 *Glimmer*

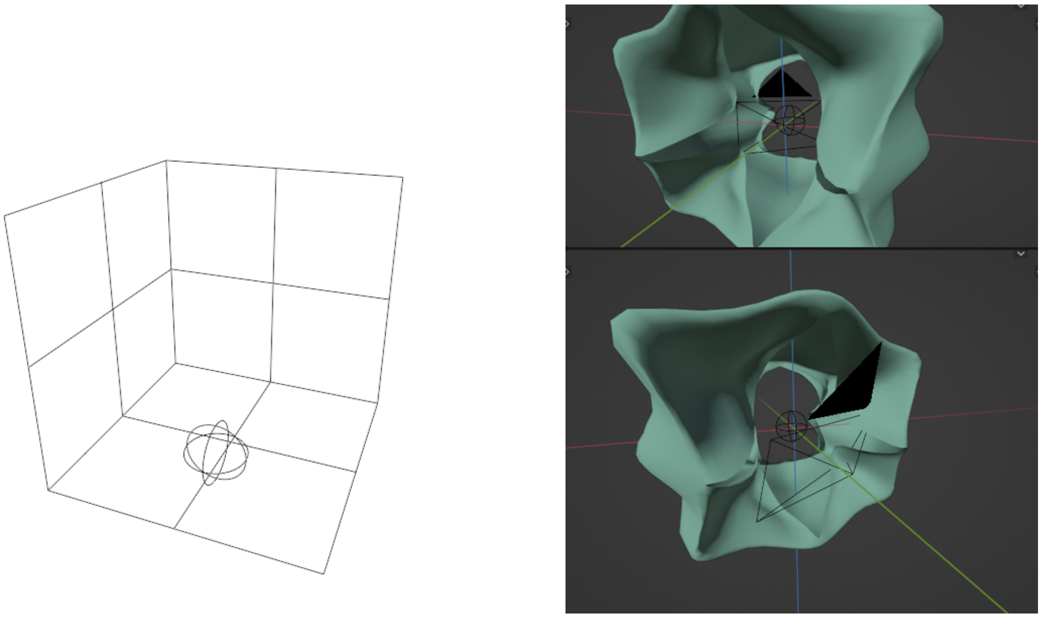
This piece is for baroque flute and projected video score and has been workshopped by Dr Carla Rees (§D). To aid with understanding the visuals, there is a paper score that acts as a possible transcription of a performance. Through this, the performer learns the basic way the visuals should be understood. Midway through the performance, the score fades away, and the performer is asked to carry on the performance in a similar manner. Unlike other pieces, there is no computer sound as *Glimmer* has one particular goal in mind: to explore how a performer might read and understand graphics as physical movement or bodily gesture given some initial instruction. This is based on the idea that action notation might present an easier type of notation to respond to when animated (Shafer 2017). While *WWSTTW* used written notation and was precise, it was also verbose, taking a long time to comprehend. This placed limits on how much information could be put into the score and the nuance of the relationship between the performers. *Glimmer* seeks to find a way of maintaining the nuance while simplifying the reading process by asking the performer to read the graphics in a more intuitive manner given some initial guidance. Initially, it was assumed that this meant sacrificing a large amount of compositional control over the music, however, while the workshop performance did deviate in some unexpected ways, these deviations were only slight and many of the subtleties encoded into the video score during composition were in fact realised.

To aid the real-time reading of the score and to provide the 'fragile' aesthetic this piece sought, a video score was created that was intended to describe the physicality of flute technique. This approach shares similarities with the recent research project undertaken by Scott McLaughlin where composition is approached, not as a process of manipulating sonic objects, but as a process that values the agencies and subtleties of each instrument-player combination (McLaughlin 2022). The main point of departure between *Glimmer* and the pieces published as a part of the project *The Garden of Forking Paths* led by McLaughlin, is that *Glimmer* requires the performer to attend to the video score rather than the instrument.

Following on from *WWSTTW* was a need to explore how a body, or more broadly a sense of physicality, can be communicated to both the instrumentalists, to aid in sight reading, and audience, to help create a perceptible relationship between the two. Thereby, this piece explores whether the instrumentalist can read and understand abstract graphics in a physical manner and respond to them with their own bodily movement, guided by their instrumental technique. This physical interpretation is guided by a paper score which details how the opening section should be realised, thereby demonstrating how the graphics are linked to movement. However, this approach was discovered through a process and a previous attempt was made that sought to describe this in text. This first attempt called, *Memory*, is discussed shortly and followed by *Glimmer*, however, first the complete video score is considered as this was created first.

3.2.1 Visuals and Movement

The visuals are fixed and created in *Blender* (The Blender Foundation n.d.). They appear as a pink swirling nebula, a theme revisited in the following piece *Breathe*. Inside *Blender*, this is a toroidal mesh with an overlaid semi-transparent texture that emits light. A large part of the visuals are achieved



(a) 3–D controls for the shape keys, the round object is moveable, the three grids are orthogonal outlining the cube.

(b) A distortion of the toroidal mesh is shown in green along with the camera which has a black triangle above.

Fig. 3.3. Examples of *Glimmer*'s controls in *Blender*.

through heavy use of a bloom filter and standard post–processing effects. The main controls for the visuals are described in 5 parameters. The first three correspond to the shape of the underlying torus, these interpolate between different shape–keys, which are predefined distortions of the mesh, i.e., each shape–key describes a new set of positions for the vertices. The final two control the texture, which is overlaid across the mesh by altering how much is transparent, and the emission of light. There are further controls for the camera, but these are used infrequently by comparison.

To aid in creating a link between bodily movement, flute technique, and this abstract graphic, the three controls for the mesh were linked to the position of a three–dimensional point inside of *Blender* — see figure 3.3 —, meaning the entire object's shape could be controlled by thinking about how this single point moved in space, thereby allowing a simple but physical interpretation of the otherwise abstract controls. This three–dimensional space was bounded to the limits of a cube, such that the bottom–left corner was a natural resting position where the torus had little distortion, and the top–right corner had the most distortion. To animate the three–dimensional control point, I experimented playing a flute and imagining an elastic force pulling my fingers towards the body of the instrument, meaning pushing my fingers away from the instrument's body was a constant struggle. This struggle was made significantly easier by my unfamiliarity with the instrument. Then these motions were roughly approximated into movement of the three–dimensional control point. This created an metaphorical relationship where movements up towards the top–right corner of the box were a constant struggle and often snapped back to the bottom–left to rest as my fingers tensed and relaxed across the flute's body. Further, the light's intensity and the transparency of the texture is set up so that when the point is in the bottom–left the entire shape is transparent and there is no light, and in the top–right the transparency and how much light is emitted changes rapidly. This spatial relationship combined with the metaphor of struggling against some force which snaps back to a rest position helps form a long crescendo followed by a quick diminuendo structure that runs throughout the video.

3.2.2 *Memory* — version 1

Memory's approach to score design is to break down the interpretation into several axes that should be performed at once. It is a systematic approach where some visual attribute is translated onto some flute technique. As an example, the visual's brightness might be related to how much air is put through the instrument, such that when the visuals are dark, little air is expelled, and when blindingly bright, a large volume of air is put through the instrument.

Memory has three of these dimensions - morphological, illumination, and kinetic - which broadly relate to the corresponding movements of fingering, air & embouchure, and trills. Illumination relates to air and embouchure as follows: when the screen is dim, play dark, airy, bending the pitch flat, pianissimo, sounding in the lower octave; but when the screen is bright, play fortissimo with a bright piercing tone sounding in the highest octaves of the instrument. The kinetic axis relies on the performer's perception of how energetic the visuals are, such that when the visuals are static there should be no trill; when there is slow and gentle movement there should should be a slight wobble of the finger similar to sporadically uncovering a tiny portion of the hole; and when there is rapid movement and the visuals have a lot of energy use a rapid trill. Full information about these two axes can be found in the score (§D.3).

Whilst these two axes are simple enough on their own, they do get more complex when performed together. Yet, it is the morphological dimensions that ended up being the most complex.

When creating the visuals I would play the flute and move my fingers in response to how the visuals appeared to twist and turn through space. The morphological dimensions are an attempt to reverse that process, to take the visual's contortions and have the performer extract similar movements. This was not expected nor intended to be a replicable process, but there should be some similarity in how my fingerings changed in relation to the visual's movement and how the instrumentalist's change. To do this, the performer was asked to imagine the visuals moving around inside a three-dimensional space, the three axes of this space each have a unique fingering placed along them, and as the visuals moved along one of these axes, the performer should interpolate towards the corresponding fingering — see figure 3.4.

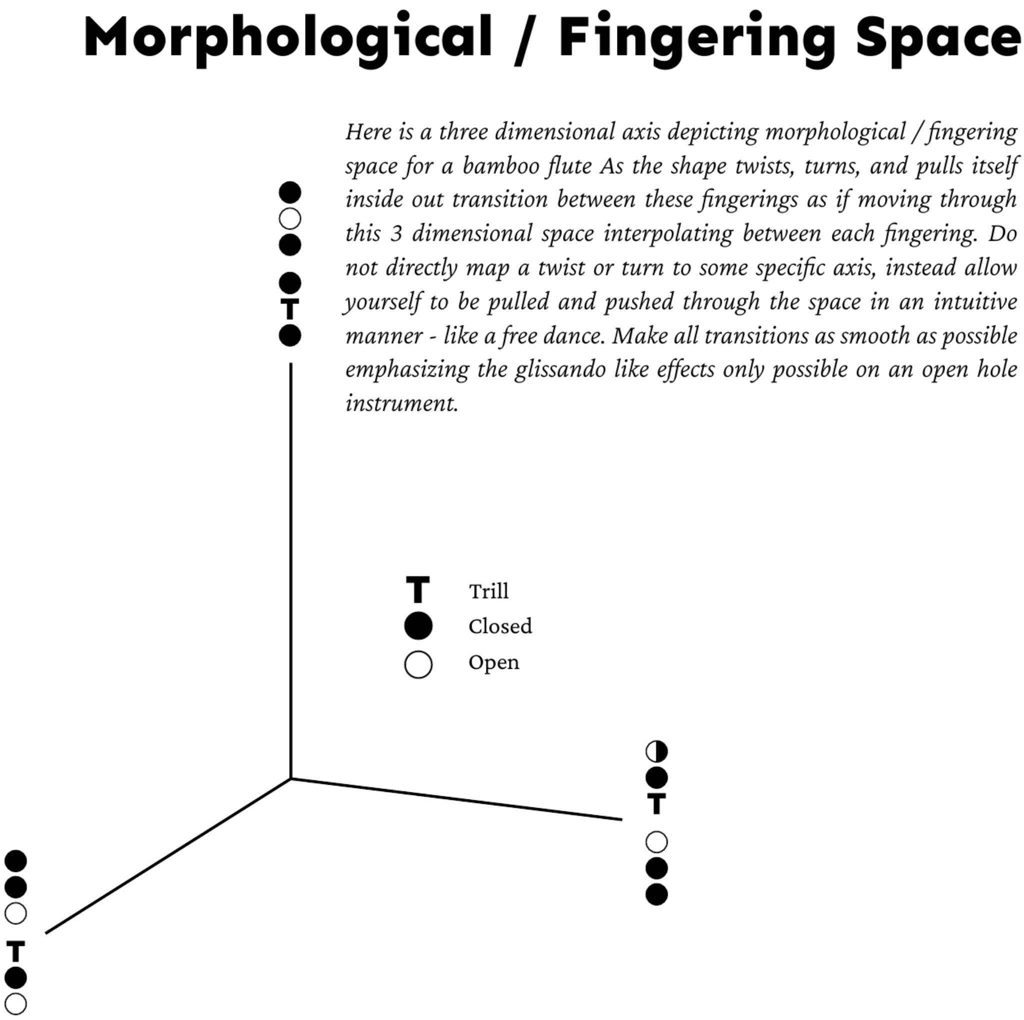


Fig. 3.4. Fingerings dimensions used in *Memory*

One issue was in reading the visuals as a singular entity, in that it isn't obvious that there is a singular visual object as each glowing section comes into life independently of the others, therefore deducing a position which can be tracked isn't possible if no object can be identified. While this wasn't an issue for me, as I knew of the underlying toroidal mesh and its distortions, it isn't obvious from just the video alone. Here the issue lies in how the visuals have been designed, as, although there is a single object in the software, the visual result is a collection of smaller objects that move together like a flock, creating and destroying composite objects through their individual movements and interactions. The knowledge of the underlying mechanism should not have been used as a guide to its interpretation.

This issue aside, the complexity involved in interpolating between fingerings would still have been too difficult as simple linear progressions between fingerings were to be avoided, due to the creation of a sudden shift from one fingering to the other. In figure 3.5, two different fingering states are interpolated between, and three different ways this could be performed are shown. The first version shows a complete linear fade — this behaviour *Memory* is trying to avoid —, the second version shows a better solution, where each finger is interpolated sequentially, however the third version shows the ideal result, as each finger has its own trajectory and the sound produced should neither be a glissando nor a sudden jump. While describing this process is not too difficult, requiring the performer juggle these competing methods of interpreting the video was evidentially going to be cumbersome. For this reason, text based scores were abandoned and a new approach is considered where the desired relations, which in *WWSTTW* were sought through relational notation, are instead demonstrated, being left for the instrumentalist to understand and re-enact.

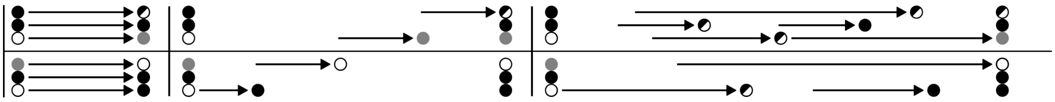


Fig. 3.5. Possible ways fingering interpolation might be performed in *Memory*. Each version is separated by a bar line, they are ordered from least desirable to most.

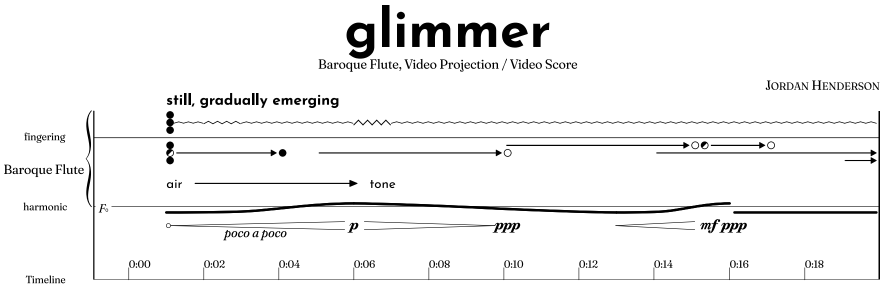
3.2.3 *Glimmer* — version 2

Glimmer approaches this issue by creating a ‘tutorial-zone’ where the performer learns how the visuals should relate to movement and is then tasked with continuing that understanding themselves once the tutorial is over — an idea picked up again in *Breathe* and, to a lesser extent, in *Jellyfish*. This is done by providing an example of how the video should be interpreted, communicated in flute tablature consisting of three staves: one for the fingers and hands, one for the mouth and air, and a timeline (see figure 3.6b). The fingering stave consists of two types of actions. First are the hole/key states, which are made of circles indicating how open or closed a hole is. They come in three fixed states, open, half open, and closed. There are also transition lines between each state, these are the horizontal arrows on the score. The second type of action is a vibrato-like wobble which is achieved by rocking the finger back and forth, gradually opening and closing the hole.

The mouth and air stave is a thick black line indicating pitch. Marks below the lighter stave line indicate a pitch bend flat, and each ledger line (not visible in figure 3.6b) above the stave line indicates a sequential jump in the harmonic series. When the line transitions slowly between harmonics, the overall sound might not change depending on the current fingering. Further, most of the time, the embouchure is bent too flat or sharp to produce the actual pitch with normal clarity.

Whereas in *Memory* the score was created by reflecting on how I made the video and attempting to distil the movement into some rules, here, the movements I made on the flute in response to the video were transcribed into tablature. By having the flautist practice and enact this transcription, the aim was that they would begin to see and understand the visuals as I did. Half way through the work, the notation begins to fade away and the performer is asked to continue this reading by themselves. Importantly, the opening notated section isn't required to be played verbatim during the performance, but is only meant as a guide to the video. While the visuals here are fixed

(a) Two stills



(b) First page

Fig. 3.6. Two stills and the first page of the score from *Glimmer*

media and there is no audio and perhaps no computer-generated performer observable to the audience, the intention was that this approach might be extrapolated and used to facilitate a relationship between the instrumentalist and the computer performer which then must be maintained.

This piece was workshopped with Dr Carla Rees online, from which, two key observations were made. First, unsurprisingly the notation was very hard to perform, but it was possible with enough practice. Second, the score did have a big impact on the improvisation in the second half of the work, and similarities between both Dr Rees' and my understanding of the relationship between visuals and movement were evident in the improvisation.

One section of the piece where our interpretations of the visuals aligned was the ending. At 4:33 the visuals suddenly dim and expand out towards the corners of the screen, here, Dr. Rees slides their fingers across the holes creating a glissando following the movement of the shape, which was exactly the kind of relationship between visuals and movement I was transcribing during the open section. This relationship is repeated at 4:44 where the visuals now bounce between expanding and contracting, mirrored by similar motions of the fingers. This way of understanding the visuals as physical movement was precisely what I wanted to communicate.

While *Glimmer* presents some difficulties in the complexity of the opening notated section, its ability to communicate a physical understanding of the visuals that directly relates to flute technique and movement has been a success as should be evident from the video recording. This raises two questions for this research project: how can this be applied to real-time generated audiovisuals such that they communicate some physical action, and can sensors be used to make the computer understand the movements of the human performer, thereby reversing the communicative process from computer to human and letting the computer take a more equal role in the ensemble? Both these questions are addressed in the following chapter.

4

Bodies Together

So far, this research project has started considering animated notation as a mechanism to expose the relationship between computer and instrumentalist, which progressed in the previous chapter to creating a virtual body that sits on–stage with the rest of the performers. Throughout, several questions remain that shape the final two works, *Breathe* and *Jellyfish*. These are:

- How does the audiovisual design of the graphic shape the performance? Can the performer understand and respond to it as a fellow ensemble member? And does the audience see it as such?
- Can communicative gestures be used to help the audience and performer consider the audiovisual as a separate performer? Can goal-oriented gestures help? Does a more physical understanding of the graphics help?
- Technically, how can such a system be managed and composed with? While there are numerous technologies available, bringing them together into a system that provides a rapid and simple composition interface is challenging.

These three strands of questioning, regarding the design of the body, the types and qualities of the body's gesture and movement, along with the technical challenges, become the main concerns of these final works. However before considering these questions, this research project's original goal of communication needs refocusing into a clearer concept that reflects the previous advances.

4.1 Ensemblematic Experience

This section refocuses the guiding concept of communication, exploring not what communication is nor how it is used, but how it is perceived by an audience of small ensemble music. The specific ensembles in question are those with a few number of players who can all see each other, do not rely on a metronome or other time keeping device, instead choosing to remain in unison by communicating gesturally. As an example of such a performance, consider Profeti Della Quinta's video recording of Josquin des Prez's *Mille regretz* (2021). The purpose of this is to enhance understanding of the visual aspect of human performances, specifically communicative gestures, which can then be imitated during composition when one of the ensemble members is computer-generated.

Musical interactions and relationships are numerous and complex, but because this project is just starting to explore these ideas, attempting to capture, understand, and imitate the totality of ensemble interaction is far too ambitious. Further, since the goal of this project is to implement some model of relationship inside the computer, it is important to consider what is practically possible from the offset when designing such a model; else, the resulting understanding of interaction might be simple enough for a human, but a technical nightmare to implement. To counter this, a single type of musical interaction is chosen that will form the basis of this model and all other types of interaction are considered as unique ways that the model can break. The main reason for doing this was technical. When these ideas are turned into code, having a single interactive model both allows for a simpler implementation and for the system to played with, breaking down the interaction. This serves to gradually separate the human and computer, or if done in reverse, bring them together, standing in opposition to a multi–interactive model in which different types of interaction are interpolated between.

For this reason, and for the aesthetic reason of seeking some type of equality between human and computer, the type of interaction chosen is a 'harmonic' relationship that avoids conflict, tension, or even hierarchical structures. Here, musicians all have their own independence, but work together without conflict or overpowering one another. This type of interaction is viewed as a maximally rich relationship, that is to say, an interactive state that is complex and contains numerous concerns to balance. This reflects my experience of free-improvisation in small ensembles, where exerting power over one another, like taking a solo, is simple, but maintaining a sustained harmonious relationship is like balancing a pendulum upside down, fragile and complex, even more so if the music is changing with time as performers must negotiate these transitions together.

Turning to the literature, refocusing on the visual aspect of music performance is not done in a vacuum. Nicholas Cook (2001; 2014), after promoting a reorientation of musicology that focuses on performance, concludes

> instead of seeing musical works as texts within which social structures are encoded we see them as scripts in response to which social relationships are enacted: the object of analysis is now present and

> self-evident in the interactions between performers, and in the acoustic trace that they leave.
>
> (Cook 2001, p. 9)

This enactment of social relationships correlates with this project's original concept of communication as suggested by Macke as an evolving relationship. However, although Cook's book considers many musical practices, and it does consider small ensembles' social dynamics (Cook 2014, chp. 8), this is primarily done from a musicological perspective, rather than as a phenomenological study from the audience's.

Approaching this topic from a performer's perspective is Jessica Kaiser (2023) who uses Fuchs' (2017) theory of social interaction as a basis for improvisation. Thomas Fuchs' work *Intercorporeality and Interaffectivity*, drawing from Merleau-Ponty, proposes a theory of social interaction based on a concept of *intercorporeality*, which is a

> prereflective, practical knowledge of how to interact with others in face-to-face encounters which is acquired already in early childhood. [. . .] [Intercorporeality] conceives emotions not as inner mental states that have to be deciphered or inferred from external cues, but as expressive, dynamic forces which affect individuals through bodily resonance and connect them with one another in circular interactions.
>
> (ibid.)

This makes group performances a notable place where intercorporeality can be found. Focusing specifically on a pair of dancers, Fuchs suggests that such close resonances between actors may form a ‘dyadic body memory’ (ibid.), which Jessica Kaiser (2023) attempts to create in their improvisational compositions concluding that how the performers ‘make sense of each other’s musical actions is essentially grounded in their bodily relations and interactions’ (ibid., p. 15). However, although Kaiser is using these ideas to create musical works, there is little consideration of the audience’s experience of such intercorporeality, nor how such an experience might interfere or otherwise alter the perception of sound. Further, Fuchs’ theory — without commenting on its plausibility — is of little practical use when designing a computer system. What is sought here in this project needs to be applicable to design interactive computer systems.

Another notable author is Marc Leman who considers in detail gesture and technology in their book *Embodied Music Cognition and Mediation Technology* (2007). However, their work does not directly apply, as although they do consider gesture and movement, it is primarily the movement that is encoded into sound and received through sound by the audience that is of concern (ibid.). Whereas, in this project, the gestures being considered are the practical movements used by performers to enhance or permit playing together, such as head nodding, looking at each other for reassurances or to understand what the other is about to do, or even maintaining a shared pulse through movement. These gestures, including the explicitly ‘goal-directed’, are not considered a part of Leman’s theory of corporeal engagement (ibid., p. 19). Therefore, the work of Leman and their conceptualisation of gesture runs parallel to this endeavour. Further, this research project perhaps points to a criticism in that the audience’s experience of the ensemble and the communicative gestures used there, whether to enable or enhance group performance, have been omitted from a theory of musical interpretation.

Due to the apparent lack of consideration for this topic, here I am going to propose an audience-centric understanding of small ensemble performances based on the literature of both musical gesture as studied in music psychology and dance literature that looks at audience understanding. This proposal for such an experience I shall term 'ensemblematic experience'.

While there is a sizeable body of literature from music psychology studying musical communicative gesture as employed by performers (see Timmers et al. 2021 for an overview), there is little that addresses the audience's experience of such gesture. This literature focuses on: the mechanics of information exchange in ensembles (Keller 2014; Bishop, Cancino-Chacón, et al. 2019), the movements/appearance of solo performers (Davidson 2007; Wapnick et al. 1997; Thompson et al. 2005; Antonietti et al. 2009), and intra-ensemble gestures (Jensenius and Erdem 2021; Bishop, Cancino-Chacón, et al. 2021). However, Bishop (2023) links to the concept of *musical togetherness*, a theory that proposes performers feel a sense of 'we' (Bilalovic Kulset and Halle 2020) which is similar to the concept of *intersubjectivity* as mentioned by Fuchs. This arises through three mechanisms, 'social presence, closeness, and joint agency' (Bishop 2023, p. 2). I am proposing here that this could easily be applied to the relationship between the audience and performers, such that the audience feel a sense of togetherness with the performers, responding to their movements *as if* they were themselves a performer. While closeness is less relevant for the audience as it has to do with their movement in physical space, social presence and joint agency might be applicable. Social presence 'posits that recognizing others' motor intentions allows one to imitate the intended actions' (ibid., p. 3). Joint agency is broadly described as when each person engaged in a collective actively experiences not just a sense of 'you did that' or 'I did that', but of 'we did that' (for an extensive overview of the topic, see Loehr 2022).

Although joint agency as intended by Bishop refers to performers, Reddish et al. (2019) performed a study in which participants were asked to move their arm in time with a recording of someone moving their arm and rate how much agency they felt over the recording. Despite this recorded arm not being physically nor temporally present, participants reported feeling agency over the recording along with the recording having agency over them, implying that the task was perceived 'as a joint one, and hence participants felt that agency was distributed across both individuals' (ibid., p. 406).

Applying this to the audience of a small ensemble performance raises one area of difference. Generally, audiences in such settings do not move with the performers — movement of any kind is often frowned upon in the concert hall. This is important as the act of imitating the movement might be crucial to joint agency. To bridge this gap, a link can be found by drawing on some literature that has extensively considered bodily movement in performance that isn't directly imitated by the audience, namely dance. Edward C. Warburton, drawing from R. Blair's three types of empathy (motor, affective, and cognitive) (2005), describes watching dance as a 'feeling of' dance, where 'even while sitting still, dancers (and others) can feel they are participating in the movements they observe' (Warburton 2011). This argument has also been given a neurological aspect by Foster (2010) and Hagendoorn (2004) who both draw from the work of Gallese (2003) on mirror neurons. An important clarification though is that the observer's past experience and familiarity with such movements can affect how such movements are reciprocated (Calvo-Merino et al. 2004; Rokotnitz 2006), meaning watching someone move in a way that has never been experienced is less likely to be internally reciprocated as a familiar action.

Piecing these lines of reasoning together seems to suggest that an audience member is capable of internally 'feeling' the bodily gestures used by the ensemble for communication 'as–if' they originated from their bodies despite not actively engaging in the movement. Hence, this creates an experience where the audience feel physically drawn into the social dynamics of the ensemble, and respond to the gestural communicative efforts of the performers as if they themselves were involved in the performance. Although Calvo-Merino et al. (2004) raise a criticism that some actions foreign to an observer's body might not be reciprocated, these gestures are likely to be understood since they are commonplace when working in a team performing any physical activity together (e.g. team sports).

Further, this type of physical reciprocation would interact in interesting ways with sound. This is backed up by studies that conclude that perception is multisensory (Stein and Meredith 1993, e.g.), and that 'seemingly extraneous cues from one modality [audio] can substantially alter judgments of stimulus intensity in another modality [visual]' (Stein, London, et al. 1996, p. 501), which entails that sound itself might increase the perceived frequency of gestural communication. Stein, London, et al. (1996) go on to conclude that having audio and visual stimuli occur in synchronisation increases the perceived certainty that some form of event has occurred (ibid., pp. 501–502). Ernst and Bülthoff (2004) reinforce that prior knowledge of the relationship between different sensory mediums (sound and vision) is extensively used by the brain to inform and interpret current experience (ibid.), which implies common gestures like those used in conducting, the ubiquitous nodding gesture, or even obvious moments of eye-contact could reinforce and perhaps even create the experience of a causal link between gesture and sound, even where none exists.

To recap, a number of conclusions can be drawn regarding the impact of gestural communication on an audience's perception of small ensemble performances. Audiences are capable of mirroring the gestures and body language of the performers, by doing so, this may lead to a feeling of participation where some level of agency is felt over and from the movements. This could entail the audience feeling involved in the ongoing communication and social dynamics of the performance. However experience is always multisensory and sound and vision both interact, meaning what is heard can influence what gestures are seen, and what is seen can influence what is heard. Therefore perhaps small ensemble music, where each musician is visible and takes responsibility for synchronising their parts (as opposed to relying on a conductor or metronome) and negotiates this through gesture, can be considered not an art–of–sound but an art–of–togetherness, a multisensory experience with particular emphasis on social activity, notably teamwork and communication which the audience participate in through mirroring physical action. This type of audience experience I have called an 'ensemblematic' experience as it is both an experience *of* an ensemble and of participating *in* an ensemble. Therefore, this research project can then be rephrased, not in terms of communication, but through the question 'how can an art–of–togetherness be created when one of the ensemble members isn't human?'.

The remainder of this chapter, and the following pieces *Breathe* and *Jellyfish*, both take ensemblematic experience as their goals, putting aside communication directly, and viewing it more as a by-product of the actual desired experience.

4.2 *Breathe*

Breathe is a semi–improvised live audiovisual work for the voice. While the computer part moves through predetermined states, these only describe the behaviour in relation to the performer and other parts of the system, allowing the human the freedom to improvise with the computer, which in turn, makes the computer part somewhat unpredictable. The human performer is given some preliminary instructions regarding the opening section, however like *Glimmer*, the performer is asked to continue this relationship throughout the remainder of the piece however they see fit. The visuals appear like a red-hot nebula, which expands as the vocalist inhales and contracts, emitting a growing low-pitched sound when exhaling. Gradually, the cloud begins to develop its own behaviour, sometimes it observes the performer's breathing and sometimes it has autonomy. The performer is invited to follow these movements, reading their expansions and contractions as movements associated with breathing, which they may follow or otherwise use to improvise with. This is done in an attempt towards the audience experiencing some similarity between the computer-generated body and the performer's, which might increase the likelihood that the audience considers the visuals as an independent performer. To explore this, this section starts by briefly considering some of the literature from interactive music systems and attempting to apply it to interactive audiovisual systems, all under the pursuit of an ensemblematic experience. Following this, the technical approach taken in *Breathe* is outlined, which I am calling behavioural mappings, and is developed further in *Jellyfish*, before being finalised in the *SuperCollider* quark *JX* (§G).

4.2.1 Bodies

There is a natural overlap between this project's goals and the wide ranging literature of interactivity. However, this project is pursuing a particular configuration of interactivity — imitating that found in small ensemble performances —, meaning not all of the literature is applicable.

An initial valuable insight can be gleaned from considering the difference between a reactive system and an interactive one. In opposition to interactive systems, reactive systems have an absence or lack of cognition (Bongers 2000), and although these systems may be complex, they are 'as predictable as a (very complicated) light switch' (Jordà 2005, p. 86). This creates a difficult tension between the need for the generated body to (at times) move dependently from the performer, and the audience's ability to differentiate the generated body from the performer. In other words, if the generated body is too synchronised with the human performer, it might appear as an extension, mirror, or counterpart to their body rather than a fully formed and distinct ensemble member in their own right. However, there should still be a sense of unity across the ensemble meaning a balance must be struck between an uncoupled relationship, where the performers act individually, and a tightly coupled relationship, where the identities between performers blend into one.

This marks an important difference between all human ensembles and human-computer hybrid ensembles, as when audiences see (not just hear) the former, we recognise that there are individual human actors working together to form an ensemble. Whereas, if we see human and computer actors moving in a similar way, we might stop seeing the computer as a separate entity as its agency is consumed by the human and it is downgraded to a purely reactive role. Tim Murray-Browne suggests that this issue can be reframed by focusing on the audience's perception of agency, its placement and qualities (Murray-Browne 2012). However, what is not considered is how some (perhaps most or all) audience members have a prejudice that humans have agency and computer generated things do not. In Jordà's example of the light switch, perhaps the only reason it is considered reactive is that it isn't considered as innately possessing agency, whereas, if the audience could see another human reacting to the state of the switch there might still be perceived agency, regardless of the qualities of the inter/re-action.

One potential solution to this issue can be found in Donna Leishman's work. Explicitly writing about live audiovisuals, Leishman (2022) drawing from McConachie and Hart (2006), has suggested that live audiovisuals can create the appearance of an entity if the audience is capable of empathising with it, which is defined as 'a mode of cognitive engagement involving mirror neurons in the mind/brain that allow spectators to replicate the emotions of a performer's physical state without experiencing that physical state directly' (ibid., p. 6). Naomi Rokotnitz, writing about theatre performance and drawing from Vittorio Gallese (2001; 2017) places some limits on what can be empathised with: 'in order to interpret the actions or intentions of another, the observer must share the motor schema of the agent — share a bodily knowledge' (Rokotnitz 2006, p. 135). For this reason, *Breathe* uses the sound and motions of breathing which is familiar to all humans.

However, Leishman's proposal might not be so readily applicable in this case as the presence of the human performer on–stage sets up a certain expectation of the role of the computer. After the first performance of this piece I had several comments from audience members, some saw what the intention here was, but others noted that the relationship between myself and the nebula was reminiscent of ventriloquy, and therefore closer to a reactive system without agency of its own. In this way, this piece doesn't quite achieve this project's aims, but does come closer than previous attempts and makes advances in technical implementation. Exactly how to trick or otherwise persuade the audience that the computer generated entity has agency is a sustained and ultimately unanswered question throughout the remainder of this project, although I believe *Jellyfish* improves upon this, which is perhaps due to the less abstract graphics and more easily recognisable form.

4.2.2 Towards Behavioural Maps

While breathing was originally chosen as an attempt to generate the perception of a body, it is also used as an instruction to the performer. Similarly to *Glimmer*, this work employs a tutorial-style opening where the desired motion of the nebula and its relation to the vocalist's breathing is predetermined. This forms a section of the work where the performer can learn how the computer responds and can use to inform the rest of their improvisation. After this, the work progresses through several different relationships between human and vocalist. Managing and creating these differing relationships is a technical challenge and explored here. This approach I am calling *behavioural mapping* and is started here, but is finalised in *Jellyfish*.

Rather than expressing the desired relationship for the performer to pursue as was done in the relational notation put forward in *WWSTTW* (§3.1.1), behavioural mapping creates a set of mappings from input to output which describe a set of possible relationships. These mappings are then interpolated between as the piece progresses, creating the large scale structure of the work. This is achieved by exposing all possible sinks and sources of data to the map, where they are connected together, and therefore completely defines the behaviour of the audiovisual system and its relationship to the performer. However, since the vocalist is instructed to perform with the generated body, this behavioural map has influence over the performer too, meaning, from a compositional perspective all the instruction for both computer and human can be found within the behaviour map. The term 'behaviour' then is used to describe a set of possible time-evolving configurations of the human and computer-generated performers. Additionally, in the system I shall introduce shortly, multiple maps of behaviour can be sequenced together and interpolated between, resulting in the computer's behaviour changing with time. In *Jellyfish*, this is expanded such that behaviours can also change with respect to any other data source and are themselves (to borrow a computer science term) composable, meaning they can be nested within each other and combined together to create new maps. While this feature wasn't desired in *Breathe*, *Jellyfish* makes extensive use of it.

Of course, this is not an entirely new idea and mapping is a well-trodden topic (see Levitin et al. 2002; Miranda and Wanderley 2006), particularly in the literature of NIME[1] where many articles have been published on the subject. This behavioural mapping system I am proposing has two distinct differences from other mapping approaches that I have been able to find, namely, the individual *mappings* are composed together into groups called *maps*, which can themselves be composed into other maps or interpolated between. Further, while the general conception of the authors from NIME is that maps are a layer between the controls and sound source (Hunt et al.

[1] New Instrument of Musical Expression, `https://www.nime.org/`. A crude search of mine through all proceedings from 2018 to 2022 resulted in over 2000 mentions of the term 'mapping'.

2002), *Breathe* is exploring mapping's potential as a technique to create the appearance of another ensemble member, that is a somewhat independent agency from the control source. For this reason a map in *Breathe* is not just a set of relationships between input and output, but also includes the generation of new signals and other complex modulations, and while all these features might appear in the creation of a new musical instrument, it is in their purpose that they differ. Although multiple authors consider unpredictability or some appearance of agency to be useful in the creation of such a 'new musical instrument' (see Eldridge et al. 2021; Magnusson et al. 2022; Magnusson 2019), considering the generated element as an independent agency from the human performer is more readily found in authors interested in *interactive music systems* (see Rowe 1993) and cybernetic systemic or eco-systemic systems (see Di Scipio 2003; Sanfilippo 2020). Alice Eldridge characterises the difference between an interactive music system influenced approach and a cybernetic influenced approach as that between a 'pipeline' based algorithm and feedback based algorithm. A pipeline algorithm takes the input from the human and generates — through multiple processing steps — a response, which ultimately 'reinforces the metaphor of one-way interaction' (Eldridge 2008, p. 43); whereas, a feedback based algorithm does not have 'inputs or outputs in the usual sense, but rather a closed loop of circular causality' (ibid., p. 45), and as Di Scipio infamously noted, marks a shift from interactive composing to compositing interaction (Di Scipio 2003, p. 270).

> This is a substantial move from interactive music composing to composing musical interactions, and perhaps more precisely it should be described as a shift from creating wanted sounds via interactive means, towards creating wanted interactions having audible traces. In the latter case, one designs, implements and maintains a network of connected components whose emergent behaviour in sound one calls music.

(Di Scipio 2003, p. 271)

The above quote hence becomes a guiding compositional principle for both *Breathe* and *Jellyfish* which both use behavioural maps to implement their feedback loops, of which there are two. The first can be traced from the audiovisual system's controls and its output. An example of this is controlling the amplitude of some audio parameter with the visual brightness of the screen, which in turn is set by measuring the initial amplitude. This feedback loop generates difference by exploring the inaccuracies between the setting of data and the measuring of the result, it does not involve sending out into the world but takes place digitally. Poetically, this can be thought of as the difference between the computer's voice and its ability to hear, or its 'body' and ability to see itself. The second loop can be found between the audiovisuals and the human performer. Here the human performer follows the audiovisuals, responding to them as–if they were another human performer, while the audiovisuals themselves are controlled (to a variable degree) by measuring what the human performer is doing. This is a feedback loop from inside the computer to outside and back again, passing through the real-world and potentially, although unknowable to the computer, through the human performers' eyes and ears, into their body/cognition and back out in the form of sound.

Considering the whole system of *Breathe* including the human performer, a diagram can be constructed as seen in figure 4.1, where the circular feedback system becomes apparent. Drawing from Bert Bongers's model (Bongers 2000, p. 42), this also involves a speculative model of human cognition which both the computer and human share. In Bongers's model the flow of data is circular, entering into each system (computer or human), being modified, and exiting as input into the next system, where the two internal processes of ‘senses’ and ‘action’ are identified in the human and computer. This fails to take into account the way our actions and bodies influence our own senses, which, while this aspect of human experience and embodiment might not be relevant for the design of new instruments, is important when attempting to design a system that behaves with as much autonomy as a human.

Important in this figure (4.1) are the double–headed arrows as these are places where feedback occurs. Considering first the top left double–headed arrow between the human’s cognition and breathing, this feedback process arises because we can control our breathing, but our breathing is also felt by our body, relaying information back to our cognition. This is *haptic* feedback, a widely studied type of feedback in instrument design (over 40 papers in NIME mention the term haptic). A different process occurs between cognition and the sensory organs. Here, we receive data from these organs and are capable of selectively filtering data - for example in selective hearing where background noise can be filtered out in order to focus on some sounding object(s).

On the computer side there are several similarities, however the diagram is vertically flipped, as the human’s sensory organs relate to the breath analysis box, and the audio and visual processing correlates to the act of breathing. This is because it is through the breath analysis that the computer experiences the outside world, and it is through the results of audio and visual processing that the computer affects the world external to it. The main difference is that the human cognition has been replaced with the behavioural mapping system, as this is designed to take account of many sources and sinks of data and combine these in a stateful manner, remembering what happened previously just as a human would.

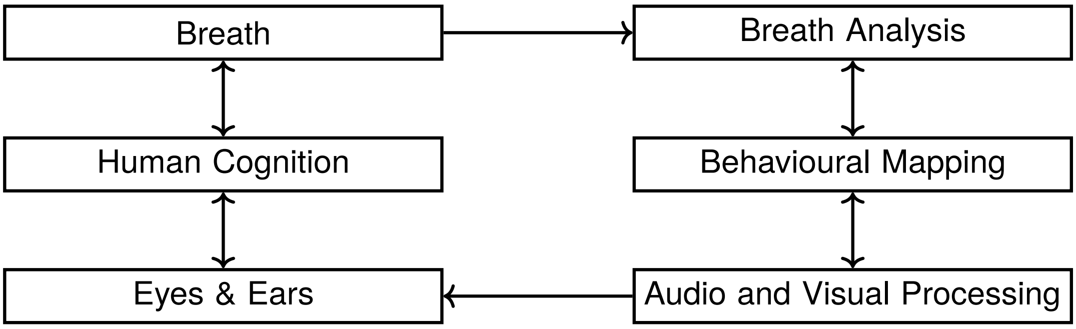


Fig. 4.1. Basic signal flow in *Breathe*. Left hand column represents the human performer, right hand the computer. The arrows imply which direction the data travels in, this may incur latency.

To express behavioural maps in *SuperCollider*, first a list of inputs and outputs are declared. These contain information regarding whether it is an input (source) or output (sink); whether the value should be sent or received across the network or from within *SuperCollider*; whether a single value is expected or an array of values; and any default values. These sources and sinks of data represent all the data in the entire system that can be controlled by the computer. This means the audio and video processing is entirely controlled through behavioural maps as sinks, and information related to the microphone or any other sensor is made available here too as sources. This has the benefit of simplifying the maintenance and development of the system as all of the important states can be found in one central location. Further, there is no need to synchronise the state of the visual system with the audio as they are designed to have no important state of their own.

Once all of this information has been declared, a behavioural map can be created that relates existing or newly generated data sources to some sinks. This map is expressed with a key–value pair where the data sink is the key (left hand side) and the source the value (right hand side), notated in *SuperCollider* in an `Event`, which is delimited with brackets, e.g.:

```
(sink1: source1, sink2: source2)
```

Both sinks and sources of data are accessed via their OSC address but the sources of data need to be accessed through a container named `src`. Meaning the following code,

```
('/speak': src['/meow'] + src['/purr'])
```

creates a mapping that sums the data sources at address `'/meow'` and `'/purr'` and outputs the result to address `'/speak'`. Importantly, all of this happens on *SuperCollider*'s server, meaning standard signal processing techniques can also be used. With this in mind, a more realistic example can be seen in figure 4.2.

```
var voxamp = src['/vocal/amp'].lag(4);
(
'/rumble/freq': voxamp * src['/rumble/amp'],
'/cloud/bright': SinOsc.kr(src['/cloud/src/bright']),
'/vocal/reverb/amp': -15.dbamp,
'/gran/verb/pan': LPF.kr(src['/gran/src/amp'])
)
```

Fig. 4.2. Example of a behaviour map used in *Breathe*. The following are data sinks (places to output data to) and shown on the left–hand side: `'/rumble/freq'`, `'/cloud/bright'`, `'/vocal/reverb/amp'`, and `'/gran/verb/pan'`. Whereas the following are data sources: `'/rumble/amp'`, `'/cloud/src/bright'`, `'/gran/src/amp'`, and `'/vocal/amp'`, which is stored in a variable named `voxamp` for convenience.

This ability to generate signals and to apply arbitrarily complex filtering to inputs and combinations thereof is also found in the prominent *libmapper* project (Libmapper 2023a), which also supports a variety of graphic interfaces, e.g., *webmapper* (Libmapper 2023b). However, as these maps should describe the behaviour of the generated body in *Breathe*, having one static behaviour isn't ideal. By implementing this anew in *SuperCollider* it is possible to blend between different mappings in time. Eldridge would note that this approach to time aligns with a dynamicist understanding of cognition, whereas the 'pipe-line model of information processing allies with a computationalist approach' (Eldridge 2008, p. 47) implying the cognition does not itself change with time. This implementation has some severe limits in *Breathe* as these maps can only be sequenced and interpolated in one order. However, in *Jellyfish* a more complete version is put forth that allows finer interpolation of maps, along with the ability to compose maps from other maps, creating layers of behaviour through nesting which can be modulated by any data source, not just time.

To understand how behavioural maps change in time — along with an additional complexity that occurs when a data sink is omitted from a map —, I shall use the following example consisting of: two maps M_1 and M_2, a set of data sources I, a set of data sinks O, and S representing some sine wave generator. Which together can be notated as follows.

$$M_1 = \left\{ \begin{matrix} O_1 : I_1 \\ O_2 : I_2 \\ O_3 : S_A \\ O_4 : S_B \end{matrix} \right\}$$

$$M_2 = \left\{ \begin{matrix} O_1 : I_2 \\ O_2 : I_1 \\ O_3 : 3 \end{matrix} \right\}$$

Above, the first map M_1 consists of four outputs (sinks) $O_{1..4}$, the first two are mapped to inputs (sources) $I_{1..2}$, and the remaining are controlled by sine wave generators $S_{A..B}$. The second map M_2 only has three outputs $O_{1..3}$, using the inputs in a different order for the first two, and a static value of three for the final.

This raises the question of what should happen when a map has omitted a data sink and the two are interpolated, which I have decided to solve by gradually applying a crude low pass filter to it by interpolating to the previous value, so that the empty sink will cease moving when it has been fully interpolated to. Perhaps this is undesirable in some situations as it means it isn't possible to analytically deduce exactly what the value will be after the interpolation, meaning, in the case of amplitude controls one has to explicitly set the value to zero to achieve no sound. Nevertheless, the aesthetic result appeared to work well in the context of the composition as many of the parameters, such as positional data, lack an obvious 'off' state, meaning that the lack of movement appears more representative of a missing value.

Each data sink can then be understood in the above example to equate to the following equations when interpolated by the time value t.

$$O_1 = lerp(I_1, I_2, t)$$
$$O_2 = lerp(I_2, I_1, t)$$
$$O_3 = lerp(S_A, 3, t)$$
$$O_4 = lerp(S_B, O_4[n-1], t)$$

Notably, the output of this result can be considered a new map $M_3 = lerp(M_1, M_2, t)$ which will be relevant in *Jellyfish*, where it is used to implement complex gestural communication.

4.3 *Jellyfish*

Jellyfish, the final work in the portfolio is for live audiovisuals and bowed guitar (although it could be easily adapted for cello). The visuals are made in *Unity* (2021), sound is made in *Supercollider* (2019), and there is face tracking using a webcam with the *Mediapipe* library (Google 2020) in Python. This is all connected together using some custom *Supercollider* code that has been in development throughout the research project. Artistically, the work is an attempt to bring about some of the features identified in ensemblematic experience (§4.1), notably, the way in with gestures between ensemble members pull the audience into the social dynamics of the ensemble. The overall aim for this work is for the audience to get the sensation that both performers (human and computer) are actively working together. This piece marks the end of the gradual movement away from abstract or non-representational visual to those that are explicitly recognisable as identifiable beings which has been taking place across this research project.

4.3.1 Techniques of Gesture

Throughout this project the concept of body and how the audience experience it has evolved. Initially, the body wasn't displayed to the audience, but implied through music aided with a visual score (§2). This idea was put aside and a display was used to place a computer-generated body on–stage where both realistic and abstract forms of body were experimented with. The previous work of this chapter, *Breathe* (§E), continued the experiment with abstract forms while attempting to create a more complex set of interactions between performer and computer. Throughout this, the audience's ability to recognise, understand, and ultimately empathise with the generated body's movements has grown in importance as it affects their ability to perceive the visuals as an independent body with its own agency. Given the contents of the previous section on ensemblematic experience (§4.1), this work seeks a more easily identifiable body in the form of a jellyfish–like organism. The reason for which isn't to suggest that abstract bodies are impossible for the audience to empathise with, but to limit the variables, stripping the work to more constrained possibilities while focusing on designing gestures and movements for the computer-generated body.

To start this discussion, the main gesture used throughout this piece, that of a 'nod', will be considered. This gesture is commonplace in small ensemble music where one performer nods at the other to indicate some approaching moment in time. At the beginning of the work, nodding is taken literally, but as the performance progresses, the pieces of code created for the nodding mechanism are repurposed. While nodding is a relatively common way to cue another performer in and to give instructions, when I have played music with others — mostly in an improvised setting—, nodding is usually done together, such that one person raises their head waiting for the other person to raise theirs, before together moving the heads downwards, reaching the nadir synchronously, and advancing as one. Nodding then becomes a mechanism for performers to learn about one another, an important aspect of gesture identified by Keller (2008). This means nodding is not a simple directed gesture as considered when discussing goal–oriented communication (§3.1.3), but a complex reciprocal one demanding both agents adapt, thereby mitigating the hierarchical structure between the one who communicates the goal and the one who responds. Further in an improvisational context, the nodding gesture contains little information beyond an attempt at synchronicity: exactly what the other performer should do with a nod might not be defined, only that there is some synchronous moment where the performers act together.

To implement this synchronous nodding gesture in the behavioural mapping, it is necessary to break this down into a series of smaller steps. First, the instrumentalist's movements are tracked in *Jellyfish* using a webcam running Google's *Mediapipe* (Google 2020), where a basic heuristic model is used to guess whether the performer's head is up or down. Changes in this state are referred to as 'nod-up' and 'nod-down' gestures. This model is described in figure 4.3 which tracks the position of each player's head where several events (points in time) are identified. From this there are a few intervals of importance, that is segments of behaviour between events. First, the interval between A and B represents how quickly the computer or human notice that the other is beginning a nod. There are many possibilities here, perhaps one player doesn't notice the other, perhaps they ignore them intentionally, perhaps it takes a while, perhaps it is instantaneous. The second interval is between B and C, this is the amount of time both performers sit there with their heads in the air waiting for each other. While some duration is required, the computer may become stuck here if it fails to recognise the human has put their head down, or incorrectly recognised a head up gesture. The third is between C, D, and E which marks the time between starting the head down gesture and reaching the exact moment of synchronisation. In *Jellyfish*, this final interval is made instantaneous in the computer, meaning it does not wait for the human's head to reach the bottom state. This could be revised in future, however, this is only a minor issue as there is latency between the webcam and receiving the data in *SuperCollider*, which creates the appearance of a delay.

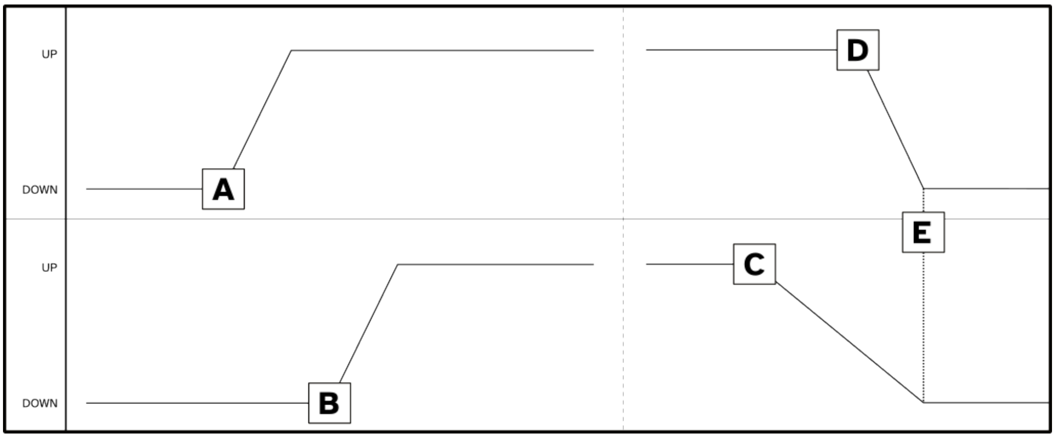


Fig. 4.3. A model of a nodding gesture between two performers. Horizontal lines indicate head position (down and up). There are five events, A is when one player raises their head, B is when the other player notices, C is when one player starts lowering their head, D is when the other player notices it is time to lower theirs, E is the actual moment both players reach the desired point. The vertical dashed line in the middle indicates that the players may switch roles here.

This system, however, needs to be error tolerant as issues can arise. Broadly, these issues fall into two categories: communication and technical. Communication issues would be if one performer attempted to raise their head too soon and the other performer wasn't ready, or perhaps if they were looking in a different direction, or in the middle of a complicated passage and missed a signal. This type of error is actually desired in *Jellyfish* because it reflects how human players engage with each other, and additionally, might require the human performer to act more considerately towards the computer-generated performer, requiring that they fully consider the jellyfish as a fellow performer, rather than a receptacle of instruction. Through playing and practising this piece, this is certainly how I had to approach it, and believe is evident in the video performance provided. The second type of error which is caused by the limits of technology is only acceptable so long as it creates aesthetically relevant results. In this case, that is creating behaviour that matches the design of the jellyfish entity, which is biological as opposed to technological, but not necessarily anthropomorphic. As an example, fast glitch–like behaviour caused by rapidly switching states is to be avoided.

To do this, a state machine was needed that covered the possible pathways, the states are:

- the holding state — before the nod begins,
- the waiting state — after the head has been raised, but before it descends,
- the anxious state — when the head has been up for a while, but the other performer has not yet raised theirs.

In terms of the previous figure (4.3) these roughly correlate to: before A for the holding state, the first half between B and C for the waiting, and the latter half for the anxious. This state diagram can be seen in figure 4.4, which can be summarised as follows: the holding state is entered and held for some duration, ignoring all input from the other performer; the waiting state is entered and a 'nod up' gesture is given, the system is now ready for a nod; the anxious state is entered through one of two paths, either the duration of the waiting state has expired, or the other performer has raised their head; finally either the success or failure outcome is triggered depending on if the system has received a 'nod down' gesture within the duration of the anxious state. Not depicted on the figure is that each state also has a minimum duration which has been done to avoid fast switching that could produce 'glitchy' artefacts.

This is applied in *Jellyfish* by associating each state with a value and using it to interpolate between a set of behaviour maps (§4.2.2). Importantly, this is itself all within a behavioural map, as from *Breathe* the software was expanded to allow maps to be nested inside one another. In the opening section of *Jellyfish*, this can be seen clearest: during the holding state, the jellyfish gradually moves from the top of the screen to the bottom; in the waiting state,

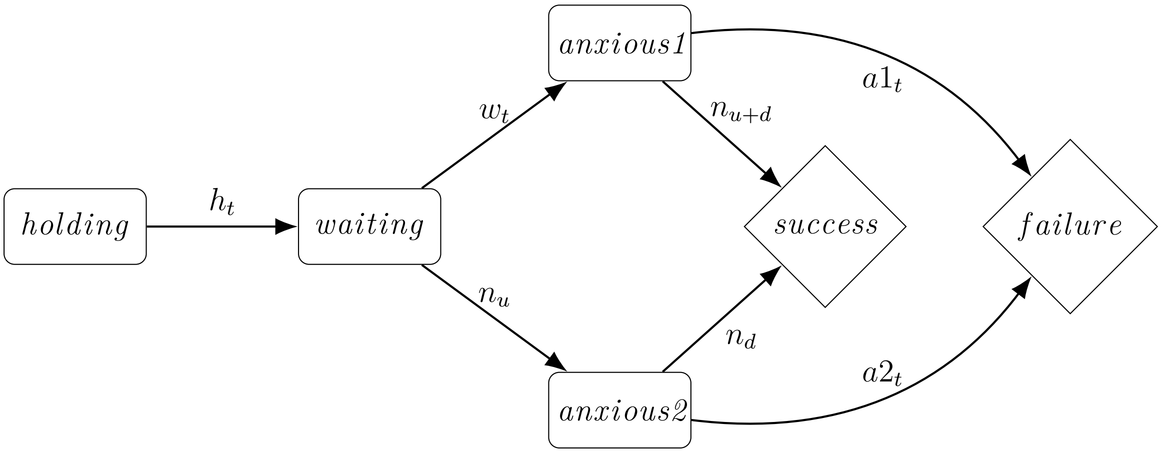


Fig. 4.4. State diagram for the head nodding mechanism. Three states are used, the holding state, the waiting state, and two anxious states. The duration at each state is notated as $state_t$, nodding signals are n_u and n_d for nod–up and nod–down respectively. Outcomes are diamonds. If a state receives a nodding signal that isn't prescribed here, it is ignored.

the jellyfish turns to 'look' at the instrumentalist; in the anxious state, the jellyfish thrashes about; if either a success or failure is reached the jellyfish jumps back to the holding state and the process repeats. Each of these behaviours is a map. In *SuperCollider* this is implemented by defining three maps for each stage, loading them into a class that handles the interpolation called `JHNodder` along with the nodding signal from the webcam and the current state of the jellyfish. The resultant values are accessed by calling 'get' on the nodder object followed by the name of the mapping — see figure 4.5. While there are still some limitations with nesting at this point, these issues are refined in the final version of this software and made into a *SuperCollider* quark called *JX* (§G).

One issue with this approach is that the jellyfish isn't technically capable of responding to a nodding gesture initiated by the instrumentalist, instead, the jellyfish has a strict sequence of events each with a time limit ascribed, meaning it can only respond when in a certain state. While this was done to break situations where the instrumentalist only attempts to control the

```
var holdState = (
'/Jellyfish/position': ...,
'/Jellyfish/width' : ...,
);
var waitState = (
'/Jellyfish/position': ...,
'/Jellyfish/width' : ...,
);
var anxState = (
'/Jellyfish/position': ...,
'/Jellyfish/width' : ...,
);

var nod = JHNodder(webcamInput, currentState,
activeState, waitState, anxState
);

// final map
(
'/Jellyfish/position': nod.get('/Jellyfish/position'),
'/Jellyfish/width': nod.get('/Jellyfish/width'),
'/Jellyfish/rotation': nod.get('/Jellyfish/rotation'),
)
```

Fig. 4.5. *Jellyfish* nested mapping example. A simplified example of how each nodding state is interpolated between in *Supercollider*.

jellyfish, ignoring its agency, it does mean that the algorithm cannot respond to the instrumentalist quite as freely as one might wish. However, from my experience of improvising in ensembles with other people, this is perhaps more realistic, as other people don't behave as passive receptors of gestures, but respond to others on their own terms. To combat this, it may be necessary for performers to have an extended practice session with the jellyfish, learning a little bit about how it behaves, and how to behave with it. Again, this is not uncommon for improvising ensembles. Nevertheless, it might prove fruitful to further develop this sequence, providing different types of gestures and pathways combined with randomness to create a more complex interaction. In the context of figure 4.4, this would involve additional states and connections that trigger under certain conditions. However as this gains in complexity, providing some feedback through the visual so the instrumentalist can understand the state of the computer will be necessary. As this model is simple in *Jellyfish* and the three states broadly align to three distinctive types of behaviour this isn't an issue, but could be as the system becomes increasingly complex.

This three state nodding gesture was used extensively in the opening and ending of *Jellyfish*, but in the middle a different approach was taken, one based on a more dance–like movement of swaying side–to–side. Here the movement and sound of the instrumentalist is tracked and imitated by the jellyfish. This idea of imitation forms a useful method employed in the middle section of the work and is realised through a type of algorithm that I am calling 'imitators', whose purpose is to provide a way to interpolate from directly copying the instrumentalist to gradually developing their own behaviour that is similar, but not identical in nature to the instrumentalist's. Essentially, the goal of an imitator is to provide a way of interpolating between more and less 'freedom'. Two approaches have been used in *Jellyfish*, however more research is needed here, particularly to find a general solution.

The simplest is an *amplitude imitator*. This is designed to produce similar amplitude envelopes as the incoming signal and is a statistical model that functions by measuring the average time between sonic events, their duration, and their attack and decay length, from which a new envelope is generated. In *Jellyfish*, this is used in the middle and end sections of the work, often to control other parameters that have a singular direction associated with them. For example, the intensity of light has a single non-directional axis in which it can move, that being away or towards zero. Like amplitude, it does not make sense for it to be negative. While this has been effective in this piece, it doesn't work for long sustained sounds. From my perspective as the instrumentalist, this requires some practice and awareness to learn when the jellyfish is imitating in this way.

The other imitator used in this piece is a *differential imitator*, which takes the incoming signal, differentiates it, stores the data in a buffer, before playing it back with a slowly changing playback rate followed by multiplying it by very slow noise before finally re-integrating it. This can be understood by imagining the incoming signal whose peaks and troughs are followed, but the length, duration, intensity, and sometimes sign of these are different. This imitator is often applied to signals that behave like positions, as it does not work well with signals that have a rest state or indicate some level of energy. An example of what this can look like is shown in figure 4.6 where both signals have similar, though not exact, behaviour.

In the opening section of *Jellyfish* some of the side-to-side motion of the jellyfish is an imitation of my movement. Here the jellyfish interpolates between its

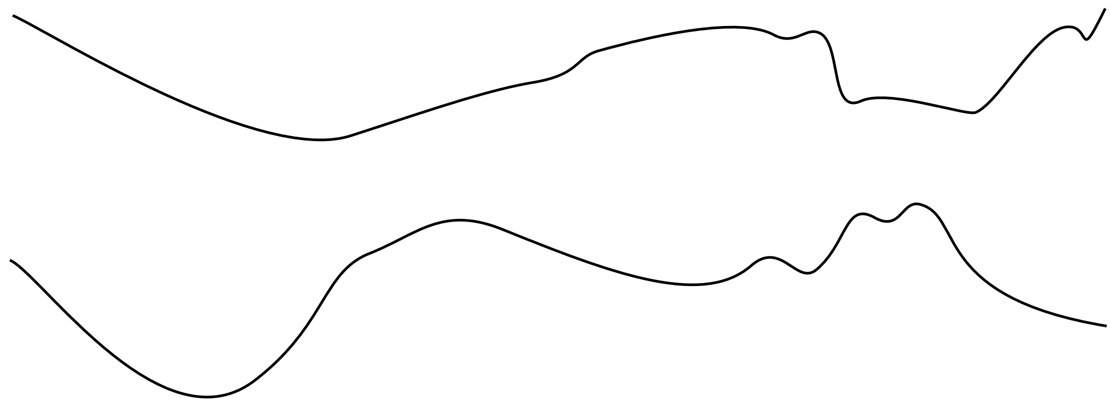


Fig. 4.6. Example of a potential derivative imitator. Bottom path is the input, top path is the output.

own imitation of my movements and exactly copying them. Whereas the nodding mechanics deal with temporally synchronising to some upcoming event, the behaviour created by imitators attempts to allow each entity to drift apart and back together, creating a less structured approach to time. Importantly, as a performer I found that observing the jellyfish imitating my movements before breaking away into their own helped to mitigate the direct sense of control one might have over a more reactive system, particularly as I could not tell at any specific moment if it was following me or just coincidentally moving in the same way.

While the idea behind the imitators has been useful, the actual algorithms proposed here are often unstable and have to be manually tuned for each instance, particularly the amplitude imitator which can either omit very fast or very slow attacks. The derivative imitator is more successful, but also requires tuning to the frequency of the incoming movements otherwise it can become stuck in a particular region of the signal's domain or bounce between the extremities. While this works for *Jellyfish*, as some generalisation about human movement can be made, more research would be needed to produce a more robust solution.

Overall, the two approaches of synchronous nodding and imitation offered useful mechanisms for thinking about the interaction between the instrumentalist and the computer, as the computer managed to retain a level of unpredictability that required myself as performer to continuously pay attention to its movements, thereby allowing the audience to see that I am turning my head or redirecting my eyes out towards the jellyfish, and to see that the jellyfish in turn responds to my movements. This is evidenced in the accompanying video. The ultimate aim here was to encourage the possibility of the audience perceiving the performance *ensemblematically*. However, this raises an interesting question which relates back to the previous consideration of the jellyfish's body: how does the jellyfish look back, or rather, how can the audience experience the jellyfish looking out towards the other performers in the ensemble?

4.3.2 Ensemblematic Experiences as the Audience

One important element of ensemblematic experience is the audience's ability to notice when a performer is interacting with another performer. Perhaps the most evident way this occurs is when two performers look at each other. In *Breathe* the visuals were an amorphous nebula and discrete parts of the body couldn't be identified and its orientation was unknown. While it might be possible for an audience to experience such a form actively *looking*, the occurrence of a head or some limbs might help. Rather than create a literal head, in *Jellyfish* the physics simulation creates a series of strands (arranged in a cylinder at the beginning) where the top point of each strand is pinned in location leaving the rest free to move around according to the simulation. This creates a perceptible ring at one end of the jellyfish which can be expanded/contracted and translated. In the opening section of the work this resizeable ring is used analogously to an eye, where in the holding state the ring is closed, reflecting how it ignores input from the user; and, in the waiting and anxious states the ring is open and positioned in such a way that it points across the stage towards the instrumentalist, creating the perception that it is looking or attending to the performer. This relates back to the continuing issue of stage design, first considered in *WWSTTW*. Yet it is being addressed here by using gestures and movements as a means of creating connections across the stage in an attempt to release the jellyfish from its virtual aquarium into the physical on-stage world of the ensemble. However, this gesture is only used in the beginning of the piece — to create enough interest and structure across the piece's duration, I opted to change the form of the jellyfish into something more abstract, thereby raising an unanswered artistic question regarding balancing computer generated audiovisual permanency and structural development, which is one area for further research. Further, this eye–like ring is used in quite a limiting way. This perhaps could have been used with more nuance to create greater expression in the communication.

Audience reception of both *Breathe* and *Jellyfish*, live performances and recordings, has been mixed regarding experiences of the audiovisual entity's autonomy. Although no concrete study was performed and these opinions are only based on myself speaking to audiences, experiences broadly fell into two camps: those that were drawn into the interplay of communication between the performer and the audiovisuals, and those that experienced the audiovisuals more like a puppet that was reacting and being controlled by the performer. The term 'puppetry' or 'ventriloquism' is particularly important as it was introduced during a group conversation amongst peers after the performance of *Breathe* and represents an experience this project has been actively seeking to avoid. Further, there were a few individuals who stated outright that they couldn't consider the entity to be acting with any kind of agency as they knew it was digitally generated. The first opinion is perhaps a critique of the audiovisual design, whereas the second appears to point to something insurmountable where audiences are unable or unwilling to consider computer-generated things as equal ensemble members. In this latter case it might be worth attempting to blur the distinctions between the human and computer-generated performers, perhaps through more stage design, or augmenting or altering the human's appearance. On the other hand, at a similar meeting of peers, the videoed performance of *Jellyfish* was shown and several people, unprompted, used terms like 'conversational' and 'teamwork' to describe the performance.

Due to these mixed results, it isn't possible to definitively state that *Jellyfish* has achieved the creation of ensemblematic experience. However, there are a few hypothetical counter-cases that could be considered:

- what happens if the entity is (unknowingly to the audience) being controlled by a human?
- how do people experience online video meetings, is simply knowing that there is a human somewhere on the other end enough?
- what if the computer generated entity looks like a human? — this is particularly salient given the increasing powers of artificially generated video
- what effect does the social situation of the music concert have?
- is there something inherent in the accepted formats of musical performance that suppresses the audience's ability to empathise with the digitally created entity, whereby a different setting/sociocultural expectations might engender different results?

Clearly these edge cases point to some blurring of the lines between the definitively human and the definitively virtual. As such, further study is needed to explore these topics, both from an artistic standpoint considering their implications, and from an empirical one considering their impact on audience experience. An obvious example of an aesthetic that engages with these edge cases is that of the cyborg. Nevertheless, given that some audience members did find performances of *Jellyfish* to evoke a sense of teamwork reflective of ensemblematic experiences, it seems likely that the nodding gesture and imitation techniques employed here offer some way to bring the human and computer closer.

4.3.3 Recognisable Bodies

Over the course of this research project how the visuals are considered has gradually changed from abstract forms towards something more concrete and purposefully recognisable as an entity or body. This process started in *WWSTTW* but is developed here in *Jellyfish* by attempting to turn the visuals into an form identifiable as a jellyfish–like entity. Although this isn't maintained throughout the entire work, this is how the audience first encounters the creature and how it appears for the majority of the performance.

The are three major reasons for this change. First, there is a technical urge to simplify the project, as the abstract form of the visuals so far could be considered a barrier limiting the audience's ability to empathise with and consider the visuals thoroughly a part of the ensemble. Second, the issue of stage layout and the limits of what projection technology was accessible played an important role. As the goal of this project is some type of equality between human and computer performer, they should appear similar in size from the audience's perspective so that one does not dominate the other. By adopting a body–like form that sits on stage, the possibility for drastic size changes is reduced as adherence to physical laws becomes important to convincingly create the illusion of an entity. Although these topics were first considered in *WWSTTW*, they become imperative in this final work. Third, although this project started as a somewhat technical question, it has begun to consider the wider social (or even political) possible meanings within such a harmonious equal performance. Paramount to this is the design of the visuals, do they look and behave robotically? naturally? do they use human gestures? or require the human adapt to an alien body language? This subsection in the writing will expand upon these ideas and relate these questions to the existing literature primarily from posthumanist studies.

The understanding of the term 'post–humanism' as employed here mirrors that of Karen Barad, who uses the term to critique 'human exceptionalism' (Barad 2007, p. 136) and to not take for granted the distinction between human and nonhuman (ibid., p. 32). Applying this understanding to the works in this project so far, the first two works have a clear distinction between human performers and notation. The humans moves through the score which is considered as an environment. The middle two works begin to dissolve the boundaries; first *WWSTTW* places a clear form on stage that the performers must watch and respond to, even if that form ignores them. *Glimmer* goes in an opposite direction and explores concepts of understanding abstract visuals as movement and physicality. At the time, *Glimmer* was an exploration of communication through notation, but retrospectively, it can be understood as the first steps towards the computer possessing a body of its own. *Breathe*, the former of the final two works tentatively explores this idea, but uses an abstract form that isn't easily recognisable as a single entity. *Jellyfish* is the first work that commits to this idea and attempts to create the appearance that the generated visuals are a distinct performer. In this way, the human's dominance on stage is, while not overcome nor directly challenged, mitigated through the need to share the space collaboratively with the non–human entity.

Key to this collaboration is the human and non–human's ability to communicate through some gestural language. One issue that arose when considering gesture (§3.1.3) was that of anthropocentrism, which is also a concern of post–humanist writing as it is one feature of human exceptionalism. In that previous section, the limits of what the audience might perceive as a communicative gesture were explored and one concern that arose was whether they might recognise a gesture as communicating something or discard it as some ancillary movement. One of the proposals that was suggested was that the gestures should resemble familiar human gestures, such as nodding or pointing, thus raising the concern of whether this design decision turns the visual entity into a human–like avatar, rather than complete non–human entities in their own right. By nodding, the jellyfish in this final piece does employ this design strategy; in fact, designing some anthropomorphic gestures into the entity was a conscious decision. There were two reasons for this. As this gesture appears clearly at the beginning, it might readily establish the audience's expectation for the remainder of the piece, allowing less anthropomorphic forms of gesture to be experimented with, which if presented first, might not be considered communicative. This is also a concern for the performer who is tasked with forming their own understanding of the entity's gesture so that they might work together in the performance. By initially establishing some shared gesture as the main method of communication the performer is able to continue this understanding as the visual entity takes on abstract forms, reconfiguring its body, and with that, gradually morphing into different types of gestural communication. In this way, *Jellyfish* can be considered as a transition away from human forms of gesture, a gradual dissolution of that human exceptionalism, in which the audience are invited to partake through following and engaging with ensemblematic experience.

Another key aspect of post–humanism, is technology, complexity, and systems. When dealing with any system, one's ability to reason about the outcome of one's actions quickly dissipates as the system grows in complexity and interconnectedness. Computer science attempts to deal with this issue by reducing the complexity of any system. To help with this, a metric was proposed call 'cyclomatic complexity', which (to simplify) is the measure of how interconnected the program is (McCabe 1976). This is a measure of how many independent paths there are in a program, for example, if there is a branch (an `if` statement) where the program's execution may go in one of two directions, then this has a complexity of 2. Later authors have expanded upon this to recommend acceptable cyclomatic complexity values: less than 10 is a simple and safe procedure, 10 to 20 is moderate, 20 to 50 is a complex and unsafe procedure, and more than 50 should be considered unstable and unsafe (McCabe Jr. 2008). What is attempting to be limited in a cyclomatic complexity analysis is the number of different paths throughout the program that must be individually tested to ensure the program works as expected. In real-time interactive art, specifically that related to music, three techniques are often used that compound creating massive complexity, cyclomatic and otherwise. The first is processes that use states from the past. These include common audio filters, but encompass any type of processing where the past influences the future, for example smoothing a signal in time. This creates a 'cycle', a dependency between what happened previously and what will happen next. The second type of complexity arises through signal combination: often explicit branches are not used in audio processing to avoid incorrect branch predictions and instead, interpolation between signals is common place. This means that often interpolation or combinations of signals can be considered branches in the program. In *Jellyfish* the number of unique signals is about 30 and, although not all used at the same time, they are often combined in complicated ways. To conceptualise the scale of this complexity, imagine taking two normalised signals and considering the space of possibilities they create. This is a two–dimensional unit plane. With 30 signals, this creates a 30–dimensional space of possibilities. When the signals are combined, the space of possibilities is distorted, curving it, making sections unreachable, and potentially creating pockets from which one cannot escape. This type of complexity is the continuous equivalent to the discrete cyclomatic complexity. The third type of complexity comes from two different feedback loops (cycles), one internal to the computer in the form of traditionally feedback signal processes, and the other, between the computer's internals, out into the external world, through the human performer, and back into the computer — see figure 4.1. These types of feedback loops are infamously complicated to reason about, forming the bases of an entire field of study called *cybernetics*. Together, these three instances of complexity drastically diminish one's ability to reason about the program whilst composing.

Returning to the idea of anthropomorphic gestures, this complexity implies that even if a human gesture was intended to be programmed into such a complex system, getting a simplified version of this as an output would be difficult, never mind the complications of designing a procedure that sufficiently mimics the nuances of human gesture. This turns *Jellyfish*'s system from a traditional program whose purpose is to be easily comprehensible into a cybernetic system, a system based on feedback with the environment which is infamously hard to predict and reason about. This instability and unpredictable quality of the jellyfish is used compositionally as it provides and important source of resistance for the human performer's improvisation, as sometimes the jellyfish will behave predictably, and sometimes it appears to have a mind of its own, completely ignoring any input. Together all of this complexity and instability reach such a degree that in systems like *Jellyfish* it is often practical to speak of the system as possessing agency of its own, a view shared by many artists in this field (Magnusson et al. 2022). Post–human authors such as Robert Pepperell speak directly to this loss of control:

> 'Humans have imagined for a long time that the ability to develop and control technology was one of the defining characteristics of our condition, something that assured us of our superiority over other animals and our unique status in the world. Ironically, this sense of superiority and uniqueness is being challenged by the very technologies we are now seeking to create, and it seems the balance of dominance between human and machine is slowly shifting.'
>
> (Pepperell 2003, p. 2)

The other famous contribution from post–humanism that should be addressed is that of the *cyborg* — is *Jellyfish* a cyborg? Donna Haraway describes a cyborg as a 'cybernetic organism', a 'hybrid of machine and organism' (Haraway 2016, p. 1). While it is certainly true that the jellyfish is made via machine means, there are a few discrepancies that can be considered. The jellyfish is created and powered by a computer (by a machine) but doesn't contain any organic mechanism. When a human player performs with it, an organic element is introduced, and the jellyfish does imitate and copy some of the behaviour and movements made, but it is technically possible for the human to be replaced by another machine and the system in *Jellyfish* will continue working regardless; further, a lot of the behaviour it exhibits comes from itself and does not require external stimuli. There are also specific points where the jellyfish ignores the human's movements and gestures, this is not only *without* organic input, but purposely *in spite of it*. It is possible to resolve the complications with the term cyborg so that it can be applied to *Jellyfish*, for example, the appearance of the jellyfish could be considered to constitute the organic element, but this is a different type of relationship between machine and organism from what Haraway appears to be suggesting. In Haraway's cyborg, these two elements form an integral part of the being's ontology, in *Jellyfish*, it is more like the machine is striving towards organism–hood through its own machine means, rather than forming a hybrid with an existing organism. In terms of the jellyfish's appearance, it doesn't draw from the aesthetics of the cyborg, whom proudly wear their machinery as we proudly wear arms and legs, instead, the jellyfish is designed to be closer in appearance to something organic, soft, flowing, suspended in the ocean, suspended in the ensemble. In fact, through what means the jellyfish is made matters not to the final performance, as a computer is used only because it is the most practical and economical to bring to life non–human things. Additionally, Haraway in 1985 considered the cyborg in the context of the 'space race, the cold war, and imperialist fantasies of technohumanism' (Haraway 2012, p. 2): the term was used as an analytical device to critique the state of politics from a feminist perspective. The jellyfish is a critique, but not of our relationship to technology. Instead, it comments on our relationship with everything else that isn't human.

Writing in the 21st century, Haraway introduces a more general term that seems applicable to *Jellyfish*, that of the 'companion species' (Haraway 2012;

Haraway 2010). Companion species are considered not as surrogates for theory, but as beings to be lived with, as '[p]artners in the crime of human evolution' (Haraway 2012, p. 5). The point being that companion species are beings we evolve *with*, we change, and so do they through our agency and being together. Despite the application of cybernetic techniques, this term suits the jellyfish far more readily than 'cyborg'.

Throughout *Jellyfish*'s composition, the workflow has followed a predictable cycle of: deciding to make an alteration/addition, making the change, reviewing the change through performance with the jellyfish, noticing what I intended to achieve wasn't realised, and then reviewing whether what the jellyfish actually did created a desirable or interesting relationship. This meant I (as composer) was never truly in control over the outcome. Reviewing the changes had to be done as a performer, meaning I had to commit to and accept what the jellyfish was doing so that the ensemblematic relationship sought might arise. If I was criticising too much then such a relationship wouldn't be able to form, as I would no longer be a committed member of the ensemble. In this way, it is hard to claim sole authorship of *Jellyfish*, as the jellyfish itself played the role of a collaborator, influencing the composition, expressing agency over its own being, and in doing so, shaping the human performer, both their role and their way of interacting and engaging with the world on stage.

This to and fro between myself and jellyfish is what Haraway, drawing from Thompson 2007, calls an *ontological choreography* (Haraway 2010, p. 65). '[O]ntological choreography is both what makes us who/what we are and also what we must engage. We must engage —must dance— ontological choreography if we are to live and die well with each other in the troubles' (Haraway and Wolfe 2016, p. 224). Importantly though, the jellyfish and the performer/composer are both the dancers and the danced, the created beings that arise from dancing. This is slightly different from Andrew Pickering's 'dance of agency' (1995) in which human and materials (often technology) take turns dancing, which brings into being a third artefact, as the jellyfish and I (as performer/composer) dance ourselves into new forms. Dance is both an action and the mechanism through which we become. Composition is ontological dancing.

5

Conclusion

This research project has explored how including computer-generated visuals in mixed ensemble performances can enhance the perception of communication for the purposes of creating a more equal relationship between the computer-generated element and the human instrumentalist. Initially, communication was considered in terms of sonic relationships that are created during composition and an animated score was created to increase the visibility of this relationship to the audience. While this approach did produce some interesting art, after the first two pieces it became clear that this understanding of communication did not fully capture the breadth of the concept sought. Although animated notation might be capable of informing the audience about relationships between performers, this would be a different experience from directly experiencing the relationship as one does when attending to small ensemble performances.

In the middle two works, this branched out into creating a computer-generated performer that sat on stage with the instrumentalists — hence the title of this project *Scores into Bodies*. Here, the focus was on how can such a new performer be generated and how would the instrumentalist read, understand, and interact with the visuals. As a part of this approach, a way of creating and expressing relationships in scores was developed (§3.1.1). Here the goal was to create connections between the human and computer, requiring them to respond to each other, which from the audience's perspective would be visible and show the changing relationship between them. This approach has only been tested at a small scale as the workshop of the piece was cancelled due to the 2019 pandemic. Nevertheless, enough information was gathered to conclude this approach was too verbose and made creating intricate relationships difficult. *Glimmer* explored a solution to this issue by giving partial instructions which demonstrated a physical interpretation of abstract visuals. This work was workshopped and proved highly successful in validating this method as the performer was able to continue the initial section and interpret the visuals as indicating physical movement.

Following this, the final two works of this project have slightly moved away from the concept of communication, and instead focused on imitating the experience audiences have when watching certain types of small ensemble performances, where the performers are actively gesturally communicating with each other. Whilst there is a sizeable quantity of research considering gestures from one performer to another, and a vast amount of literature considering an audience's experience of purely sonic phenomena, there is less considering their intersection — that is, what impact performers' communicative gestures have on audiences. Therefore a proposal of small ensemble music experience was suggested, which I have termed *ensemblematic experience*, which emphasises the audience's perception of active involvement when they can see the bodily movements of ensembles. While this is only a proposal, the hope is that with further research, ensemblematic experience could be turned into a more robust theory of audience engagement with performed music, which perhaps helps restate the value of in–person performances, the unique affordances offered by existing and traditional music practices, and that music when performed by people is primarily a social experience.

Breathe and *Jellyfish* have explored how this type of experience could be recreated when one of the performers is computer-generated. The main issue here was how to make the computer-generated entity be perceived as an equal to the human inside the ensemble. While creating a more realistic visual could have been one path forward, this project has instead pursued making the behaviour of the entity more human and imitating the types of gestures found in ensemblematic experience. *Breathe* explored using the motion of breathing, the rise and fall of the shoulders and chest. The goal was to allow the audience to physically understand the audiovisuals by seeing the similarities with the human's movements. Here, the issue of ventriloquism arose as the computer body is too easily seen by the audience as mirroring or being controlled by the performer, which is in stark contrast to my experience as a performer where the exact opposite is felt. After, *Jellyfish* attempts to imitate some of the gestures used by small ensembles, namely the nodding gesture and a swaying motion. Although some audiences have reported similar concerns to *Breathe*, the numbers were fewer and several people, unprompted, used words like 'teamwork' to describe performance, this at the very least suggests the type of experience sought in this project is possible, if perhaps uncommon. While providing a definitive method to create this type of audience experience is not within the scope of this project, as its artistic utility is questionable, future empirical research might wish to probe this claim, focusing on the backgrounds of audience members and their reaction to such computer-generated bodies.

Nevertheless, since ensemblematic experience is about empathising with the performers, feeling their communicative gestures, along with a sense of agency in the performance, this might provide a useful method for those involved in any interactive system where it is desirable for the audience to experience some agency and physical involvement. One example of this could be extended reality, which often lacks some haptic response, making the user's actions feel arbitrary. By creating one or more virtual interacting bodies engaged in some social–like activity for the user to empathise with, they might be able to feel some physical response through it, which could help create sensations otherwise impossible in extended reality, such as weight or resistance.

Throughout this whole project there has been an intense focus on the use of technology, how these works are constructed, and how they evolve to create a more sustainable practice. This culminated in the approach I have called behavioural mapping, which deals with creating groupings of mappings that can themselves be composed, interpolated between, and controlled using familiar signal processing techniques. A complete implementation of this approach is shown in the *SuperCollider* quark *JX* (§G). Ultimately, this approach is about grouping mappings between outputs and inputs into objects (behavioural maps) which are themselves composed, interpolated between, combined, and nested inside one another. This has two main benefits from a technical perspective: the consolidation of state, and the lack of nested timelines. As all the subsystems, audio, visual, etc., are controlled through this system, it is quick and easy to grasp how it is behaving at any one moment. Further, since time is represented as a data source which one creates mappings with, it allows this data to be replaced by something else, meaning one generally avoids creating nested timelines which need synchronising. This makes jumping to different sections of the piece trivial, a necessary consideration for fast rehearsal and composition.

To address the research question head on, this project concludes that one way forward to enhance the perceived equality of communication between computer generated entity and instrumentalist is by creating an *ensemblematic experience* through imitating existing behaviour and relationships found in small ensembles.

To achieve this, the main method employed in this project is the use of communicative gestures, of which there are three parallel issues: gesture recognition (human to computer); getting the computer to create similar gestures back to the human (computer to human); and finally, convincing the human performer to respond to these gestures in a meaningful and respectful manner.

While the first has existing literature and an ever developing set of software implementations and techniques, the second (computer to human) does not. Although there is literature regarding computer–human interaction generally, it is almost entirely concerned with relaying some information, as opposed enabling teamwork. In this project, this has been approached by considering the design of the entity's body, its capacity for gesture, and their similarities to existing gestures as employed in ensembles. This can be seen in how *Breathe* used the movements of breathing and an inflatable red cloud, whereas *Jellyfish*, which focused on nodding, used an entity with a head that could be positioned such that it points across the stage. However, these are only initial experiments with such an approach. One possible way forward might be considering the evolution and environment in which these computer-generated bodies arose, where real-world evolutionary processes might be emulated. One example could be found by considering dolphins and how their physicality reflects the way they communicate through whistles, clicks and buzzes, all of which is directly reflected in their bodies, i.e., their large heads allow them to whistle in order to communicate, yet without expelling vital air, a pronounced advantage for aquatic beings. Such an approach might give more realism to the bodies, potentially helping the audience believe in their existence.

Whilst there is active research in how ensembles function and the nuances of gestural communication, these studies are often considered from a psychology perspective and not designed for turning into a guide for recreating, imitating or engendering such behaviour. Further, doing so in a way that is readily applicable on a computer adds a further level of complexity. This project has only just started to put this into practice and has considered only the simple case of a nod in detail, and even then, only one model of a nodding interaction of which there could be many others. Additionally, when humans communicate they use many types of gestures concurrently and this project has not considered having multiple types of gestures at once, although using the nesting of behavioural maps would make this a simple task.

The final point regarding the human's response has in the final two pieces been mostly avoided as I have been the performer in question, and therefore knew what the intended relationship should be. However, each stage of the jellyfish's response has certain timing constraints in an effort to help the performer consider it as a separate entity (§4.3.1) rather than merely a reactive piece of technology. This approach is essentially about balancing the relationship between the computer generated entity and the human performer so that they are similar enough in behaviour that they appear together, but not too similar that they appear as a single entity. As the performer, I experienced this as the entity ignored me, but there is a risk that another performer could experience this as a flaw in the system. It might be worth making this more apparent in the physicality of the body, for example, it could turn away from the performer when it is ignoring them implying that this is done on purpose, and adding more importance when the entity directly looks at the performer.

This project has explored communication and experiences of social togetherness in ensemble performances as applied to computer generated performers, in which initial proposals for a conceptual method, compositional approach, and technical framework have been considered. By considering ensemblematic experience, the social value created in small ensemble music is championed and applying this to computer-generated performers enables music performance to be taken beyond the human, probing into the nature of communication and social experience.

Bibliography

Amphaeris, Jenny, Daniel T. Blumstein, Graeme Shannon, Thora Tenbrink, and Arik Kershenbaum (2023). “A multifaceted framework to establish the presence of meaning in non-human communication”. In: *Biological Reviews*. DOI: 10.1111/brv.12989.

Antonietti, Alessandro, Daniela Cocomazzi, and Paola Iannello (2009). “Looking at the Audience Improves Music Appreciation”. In: *Journal of Nonverbal Behavior* 33.2, pp. 89–106. DOI: 10.1007/s10919-008-0062-x.

Barad, Karen (2007). *Meeting the Universe Halfway*. Duke University Press. 542 pp. ISBN: 082233917X.

Barrett, G. Douglas and Michael Winter (2010). “LiveScore: Real-Time Notation in the Music of Harris Wulfson”. In: *Contemporary Music Review* 29.1, pp. 55–62. DOI: 10.1080/07494467.2010.509594.

Bilalovic Kulset, Nora and Kirsten Halle (2020). “Togetherness!: adult companionship the key to music making in kindergarten”. In: *Music Education Research* 22.3, pp. 304–314. ISSN: 1469-9893. DOI: 10.1080/14613808.2020.1765155.

Bishop, Laura (2023). “Focus of attention affects togetherness experiences and body interactivity in piano duos.” In: *Psychology of Aesthetics, Creativity, and the Arts*. ISSN: 1931-3896. DOI: 10.1037/aca0000555.

Bishop, Laura, Carlos Cancino-Chacón, and Werner Goebl (2019). “Eye gaze as a means of giving and seeking information during musical interaction”. In: *Consciousness and Cognition* 68, pp. 73–96. ISSN: 1053-8100. DOI: 10.1016/j.concog.2019.01.002.

– (2021). “Beyond synchronization”. In: *Together in Music*. Oxford University Press, pp. 182–188. DOI: 10.1093/oso/9780198860761.003.0023.

Bishop, Laura and Werner Goebl (2017). “Beating time: How ensemble musicians’ cueing gestures communicate beat position and tempo”. In: *Psychology of Music* 46.1, pp. 84–106. DOI: 10.1177/0305735617702971.

Blair, R. (2005). “Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations”. In: *Consciousness and Cognition* 14.4, pp. 698–718. DOI: 10.1016/j.concog.2005.06.004.

Bongers, Bert (2000). “Trends in Gestural Control of Music ,” in: ed. by M.M. Wanderley and M. Battier. IRCAM - Centre Pompidou. Chap. Physical Interfaces in the Electronic Arts: Interaction Theory and Interfacing Techniquesfor Real-time Performance, pp. 41–70.

Bourriaud, Nicolas (1998). *Relational Aesthetics*. Les Presse Du Reel,France, p. 125. ISBN: 9782840660606.

Burton, John (2019a). *Line Crossing*. Fluid Corpus Manipulation. URL: https://www.youtube.com/watch?v=hYaZq1JCTQo&t=273s.

– (2019b). *Traces of Fluid Decomposition: John Burton (aka Leafcutter John), 'Line Crossing'*. CeReNeM | Centre for Research in New Music. URL: https://www.youtube.com/watch?v=r1LsSFEe4yM&t=980s.

Calvo-Merino, B., D.E. Glaser, J. Grèzes, R.E. Passingham, and P. Haggard (2004). “Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers”. In: *Cerebral Cortex* 15.8, pp. 1243–1249. DOI: 10.1093/cercor/bhi007.

Chadabe, Joel (1984). “Interactive Composing: An Overview”. In: *Computer Music Journal* 8.1, p. 22. DOI: 10.2307/3679894.

Cook, Nicholas (2001). “Between Process and Product: Music and/as Performance”. English. In: *Music theory online* 7.

– (2014). *Beyond the Score: Music as Performance. Music as Performance*. Oxford University Press, p. 458. ISBN: 9780199357406.

Corral, Alba G. (2020). *A Taste of Nature*. URL: https://blog.albagcorral.com/2020/10/07/a-taste-of-nature-casa-seat/.

Croft, John (2007). “Theses on liveness”. In: *Organised Sound* 12.1, pp. 59–66. DOI: 10.1017/s1355771807001604.

Davidson, Jane W. (2007). “Qualitative insights into the use of expressive body movement in solo piano performance: a case study approach”. In: *Psychology of Music* 35.3, pp. 381–401. DOI: 10.1177/0305735607072652.

Di Scipio, Agostino (2003). “Sound is the interface: from interactive to ecosystemic signal processing”. In: *Organised Sound* 8.3, pp. 269–277. DOI: 10.1017/s1355771803000244.

– (2021). “Thinking Liveness in Performance with Live Electronics”. In: *Sound Work*. Universitaire Pers Leuven, pp. 171–194. DOI: 10.2307/j.ctv1ccbg96.12.

Didkovsky, Nick (2019). *Zero Waste*. Composition. Written for Kathleen Supové. URL: http://www.punosmusic.com/pages/zerowaste/index.html.

Dori, Gil (2020).
"Using Gesture Data to Generate Real-Time Graphic Notation: a Case Study". In: *Proceedings of the International Conference on Technologies for Music Notation and Representation*. Ed. by Rama Gottfried, Georg Hajdu, Jacob Sello, Alessandro Anatrini, and John MacCallum.
Hamburg, Germany: Hamburg University for Music and Theater, pp. 68–74.
ISBN: 978-3-00-066930-9.

Eiðum, Páll (2013). "Sáðrás". In:
URL: https://www.youtube.com/watch?v=7A4z-iXVoiI.

Eldridge, Alice (2008). "Collaborating with the Behaving Machine:Simple Adaptive Dynamical Systems forGenerative and Interactive Music".
PhD thesis. University of Sussex.

Eldridge, Alice, Chris Kiefer, Dan Overholt, and Halldor Ulfarsson (2021). "Self-resonating Vibrotactile Feedback Instruments: Making, Playing, Conceptualising". In: PubPub. DOI: 10.21428/92fbeb44.1f29a09e.

Emmerson, Simon (2000). *Music, electronic media, and culture*. English. Aldershot;Burlington, USA; Ashgate.
– (2016). *Music, Electronic Media and Culture*. Taylor and Francis Group. ISBN: 9781315596877.

Ensemble Offspring, Geoffrey Gartner, John Addison, and Claire Edwardes (2010). *Ensemble Offspring: Mauricio Kagel - Match*. Performance and Online Recording. Performance. URL: https://www.youtube.com/watch?v=uXZb2OKGCtM.

Ernst, Marc O. and Heinrich H. Bülthoff (2004).
"Merging the senses into a robust percept".
In: *Trends in Cognitive Sciences* 8.4, pp. 162–169. ISSN: 1364-6613.
DOI: 10.1016/j.tics.2004.02.002.

Essl, Karlheinz (1998). *Champ d'Action*.
URL: https://www.essl.at/works/champ.html.

Fischer, Christian (2015). "Understanding Animated Notation".
In: *Proceedings of the First International Conference on Technologies for Music Notation and Representation*. Ed. by Marc Battier, Jean Bresson, Pierre Couprie, Cécile Davy-Rigaux, Dominique Fober, Yann Geslin, Hugues Genevois, François Picard, and Alice Tacaille. Paris, France, pp. 31–38.
ISBN: 978-2-9552905-0-7.

Fober, Dominique, Yann Orlarey, and Stéphane Letz (2012).
"INScore - An Environment for the Design of Live Music Scores".
In: *Linux Audio Conference*. Stanford, United States, pp. 47–54.
URL: https://hal.archives-ouvertes.fr/hal-02158817.

Fober, Dominique, Yann Orlarey, Stéphane Letz, and Romain Michon (2021). "A Web-Based Environment Embedding Signal Processing in Musical Scores". In: *Proceedings of the International Conference on Technologies for Music Notation and Representation*. Ed. by Rama Gottfried, Georg Hajdu, Jacob Sello, Alessandro Anatrini, and John MacCallum. Hamburg, Germany: Hamburg University for Music and Theater, pp. 221–226. ISBN: 978-3-00-066930-9.

Forsyth, Donelson R. (2006). *Group dynamics*. Thomson/Wadsworth, p. 682. ISBN: 0534368220.

Foster, Susan (2010). *Choreographing Empathy*. Routledge. DOI: 10.4324/9780203840702.

Freeman, Jason (2008). "Extreme Sight-Reading, Mediated Expression, and Audience Participation: Real-Time Music Notation in Live Performance". In: *Computer Music Journal* 32.3, pp. 25–41. URL: https://muse.jhu.edu/article/245518#info_wrap.

Fuchs, Thomas (2017). *Intercorporeality and Interaffectivity*. Oxford University Press. DOI: 10.1093/acprof:oso/9780190210465.003.0001.

Gallese, Vittorio (2001). "The 'Shared Manifold' HypothesisFrom Mirror Neurons To Empathy". In: *Journal of Consciousness Studies*.

– (2003). "The Roots of Empathy: The Shared Manifold Hypothesis and the Neural Basis of Intersubjectivity". In: *Psychopathology* 36.4, pp. 171–180. DOI: 10.1159/000072786.

– (2017). "Visions of the body. Embodied simulation and aesthetic experience". en. In: *Aisthesis. Pratiche* linguaggi e saperi dellestetico, Vol. 10 No. 1 (2017): Ways of imitation. DOI: 10.13128/AISTHESIS-20902.

Google (2020). *Mediapipe*. Online. URL: https://mediapipe.dev/.

Gottschalk, Jennie (2016). "Information, Language, and Interaction". In: *Experimental Music Since 1970*. Bloomsbury Academic. DOI: 10.5040/9781501396328.ch-005.

Griffiths, Paul (2010). *Modern music and after*. Oxford University Press. ISBN: 9780199740505.

Hagendoorn, Ivan (2004). "Some Speculative Hypotheses about the Nature and Perception of Dance and Choreography". English. In: *Journal of consciousness studies* 11.3, pp. 79–110.

Hamel, Keith (2015). *Corona*. Performed by Megumi Masaki. URL: https://www.youtube.com/watch?v=3LFn3U7nY5w.

Haraway, Donna (2010). *When species meet*. [Nachdr.] Posthumanities 3. Hier auch später erschienene, unveränderte Nachdrucke. Minneapolis. Minn. [u.a.]: Univ. of Minnesota Press. 423 pp. ISBN: 0816650454.

– (2012).
The companion species manifesto. Dogs, people, and significant otherness. 6th print. Paradigm 8. Hier auch: 6th print 3/2015. Chicago, Ill.: Prickly Paradigm Press. 100 pp. ISBN: 0971757585.
– (2016). “A Cyborg Manifesto”. In: *Manifestly Haraway*. University of Minnesota Press, pp. 3–90. DOI: 10.5749/minnesota/9780816650477.003.0001.

Haraway, Donna J. and Cary Wolfe (2016). *Manifestly Haraway*. University of Minnesota Press. ISBN: 9780816650484. URL: http://www.jstor.org/stable/10.5749/j.ctt1b7x5f6 (visited on 06/25/2025).

Heile, Björn (2006). *The music of Mauricio Kagel*. Aldershot, England Burlington, VT: Ashgate. ISBN: 9780754635239.

Hope, Cat (2017). “Electronic Scores for Music: The Possibilities of Animated Notation”. In: *Computer Music Journal* 41.3, pp. 21–35. DOI: 10.1162/comj_a_00427.
– (2020a). *Cruel and Unusual*. URL: https://www.youtube.com/watch?v=UoQl352fl3A.
– (2020b). “The Future is Graphic: Animated notation for contemporary practice”. In: *Organised Sound* 25.2, pp. 187–197. DOI: 10.1017/s1355771820000096.

Hope, Cat and Lindsay Vickery (2010). “The Aesthetics of the Screen-Score”. In.

Hope, Cat, Aaron Wyatt, and Lindsay Vickery (2015). “The decibel scoreplayer: New developments and improved functionality”. English. In: *41st International Computer Music Conference, ICMC 2015*. Ed. by Richard Dudas. 41st International Computer Music Conference, ICMC 2015: Looking Back, Looking Forward - Proceedings. International Computer Music Association, pp. 314–317.

Hunt, Andy D., Marcelo M. Wanderley, and Matthew Paradis (2002). “The Importance Of Parameter Mapping In Electronic Instrument Design”. In: DOI: 10.5281/ZENODO.1176424.

IRCAM (2018). *Antescofo 1.0*. URL: https://antescofo-doc.ircam.fr/.

J.L, Austin (1975). *How to Do Things with Words*. Clarendon Press, Oxford.

Jensenius, Alexander Refsum and Çar Erdem (2021). “Gestures in ensemble performance”. In: *Together in Music*. Oxford University Press, pp. 109–118. DOI: 10.1093/oso/9780198860761.003.0014.

Jordà, Sergi (2005). “Digital Lutherie: Crafting musical computers for new musics performance and improvisation”. PhD thesis. Departament de Tecnologiade la Universitat Pompeu Fabra.

Kagel, Mauricio (1967). *Match: for 3 players*. Universal Edition.

Kaiser, Jessica (2023). “Exploring Musical Togetherness: An embodied approach to relational interpretation in Maurizio Pisatis Sette Duo”. In: *Music and Practice*.

Keller, Peter (2014). “Ensemble Performance”. In: *Expressiveness in music performance*. Oxford University Press, pp. 260–282. DOI: 10.1093/acprof:oso/9780199659647.003.0015.

Keller, Peter E. (2008). “Enacting intersubjectivity: A cognitive and social perspective to the study of interactions”. In: ed. by Peter E Keller, F Morganti, A Carassa, and G Riva. IOS Press. Chap. Joint action in music performance, pp. 205–221.

Kim-Boyle, David (2018). “Reframing the listening experience through the projected score”. In: *Tempo* 72.284, pp. 37–50. DOI: 10.1017/s0040298217001243.

Kim-Boyle, David and Benedict Carey (2019). “Immersive Scores on the HoloLens”. In: *Proceedings of the International Conference on Technologies for Music Notation and Representation*. Ed. by Cat Hope, Lindsay Vickery, and Nat Grant. Melbourne, Australia: Monash University, pp. 1–6. ISBN: 978-0-6481592-5-4.

Kojs, Juraj (2011). “Notating Action-Based Music”. In: *Leonardo Music Journal,* 21, pp. 65–72.

Kos, Ronald P. (2018). “Becoming music teachers: preservice music teachers early beliefs about music teaching and learning”. In: *Music Education Research* 20.5, pp. 560–572. ISSN: 1469-9893. DOI: 10.1080/14613808.2018.1484436.

Leishman, Donna (2022). “Presence and Live Visual Performance”. In: *Live Visuals*. Routledge. Chap. 11, pp. 267–286. ISBN: 9781003282396. DOI: 10.4324/9781003282396-14.

Leman, Marc (2007). *Embodied Music Cognition and Mediation Technology*. DOI: 10.7551/mitpress/7476.001.0001.

Levitin, Daniel J., Stephen McAdams, and Robert L. Adams (2002). “Control parameters for musical instruments: a foundation for new mappings of gesture to sound”. In: *Organised Sound* 7.2, pp. 171–189. DOI: 10.1017/s135577180200208x.

Lewis, George (2002). “Improvised Music after 1950: Afrological and Eurological Perspectives”. In: *Black Music Research Journal*, pp. 215–246.

Libmapper (2023a). *libmapper v2.4.4*. Online. Software. URL: https://github.com/libmapper/libmapper.

– (2023b). *webmapper v2.4.4*. Online. URL: https://github.com/libmapper/webmapper.

Loehr, Janeen D. (2022).
"The sense of agency in joint action: An integrative review".
In: *Psychonomic Bulletin and Review* 29.4, pp. 1089–1117. ISSN: 1531-5320.
DOI: 10.3758/s13423-021-02051-3.

Lowenthal, Leo (1967). "The Human Dialogue: Perspectives on Communication".
In: ed. by Floyd W. Matson and Ashley Montagu. Free Press.
Chap. Communication and Humanitas, p. 336.

Macke, Frank J. (2014).
Experience of Human Communication Body, Flesh, and Relationship.
Fairleigh Dickinson University Press. ISBN: 9781611475494.

Magnusson, Thor (2019). *Sonic Writing*. Bloomsbury Publishing Plc. 304 pp.
ISBN: 150131386X. URL: https://www.ebook.de/de/product/31797217/professor_thor_professor_in_future_music_university_of_sussex_uk_magnusson_sonic_writing.html.

Magnusson, Thor, Chris Kiefer, and Halldor Ulfarsson (2022).
"Reflexions upon Feedback". In: *NIME*. PubPub.
DOI: 10.21428/92fbeb44.aa7de712.

McCabe, T.J. (1976). "A Complexity Measure".
In: *IEEE Transactions on Software Engineering* SE-2.4, pp. 308–320.
ISSN: 2326-3881. DOI: 10.1109/tse.1976.233837.

McCabe Jr., Thomas (2008). "Software Quality Metrics to Identify Risk".
In: *Deparetment of Homeland Security: Software Assurance Working Group*.

McConachie, Bruce and Faith Elizabeth Hart, eds. (2006).
Performance and cognition. Theatre studies and the cognitive turn.
London: Routledge. 236 pp. ISBN: 9780415763844.

McLaughlin, Scott (2022). "The Material Clarinet".
In: *Rethinking the Musical Instrument*. Cambridge Scholars Publishing.

Merleau-Ponty, Maurice (2004). *The world of perception*.
London New York: Routledge. ISBN: 0203491823.

Miranda, Eduardo Reck and Marcelo M Wanderley (2006).
New digital musical instruments : control and interaction beyond the keyboard.
A-R Editions.

Murray-Browne, Tim (2012).
"Interactive music: Balancing creative freedom with musical development".
PhD thesis. Queen Mary University of London.

Nyman, Michael (1999). *Experimental music. Cage and beyond*.
Cambridge University Press, p. 196. ISBN: 0521652979.

Öhman, Karolina, Erika Öhman, and Thomas Grossenbacher (2016). *Match für drei Spieler by Kagel*. Performed Basel 2016, Online recording. Performance. URL: https://www.youtube.com/watch?v=9Ep31Jk3Vjo.

Oliveros, Pauline (1971). *Sonic Meditations*. Smith Publications.

Oxford English Dictionary (2023a). "association, n." In: *Oxford English Dictionary*. Oxford University Press. DOI: 10.1093/oed/5963744713.

– (2023b). "relate, v." In: *Oxford English Dictionary*. Oxford University Press. DOI: 10.1093/oed/1027224046.

– (2023c). "relationship, n." In: *Oxford English Dictionary*. Oxford University Press. DOI: 10.1093/oed/1697777834.

Pepperell, Robert, ed. (2003). *The posthuman condition. Consciousness beyond the brain*. [New ed.] Includes bibliographical references (p. 189-193) and index. - Description based on print version record. Bristol: Intellect. 199 pp. ISBN: 9781841508832.

Peters, John Durham (2001). *Speaking into the Air. A History of the Idea of Communication*. University Of Chicago Press, p. 304. ISBN: 9780226662770.

Pickering, Andrew (1995). *The Mangle of Practice*. The University of Chicago Press. 296 pp. ISBN: 0226668037. URL: https://www.ebook.de/de/product/3627919/andrew_pickering_the_mangle_of_practice.html.

Profeti Della Quinta, Josquin des Prez, Giovanna Baviera, Doron Schleifer, Jacob Lawrence, and Elam Rotem (2021). *Josquin des Prez / Mille regretz*. [accessed 23/9/24]. URL: https://www.youtube.com/watch?v=dkfVzCZ68_Q.

Reddish, Paul, Eddie M. W. Tong, Jonathan Jong, and Harvey Whitehouse (2019). "Interpersonal synchrony affects performers sense of agency". In: *Self and Identity* 19.4, pp. 389–411. ISSN: 1529-8876. DOI: 10.1080/15298868.2019.1604427.

Rokotnitz, Naomi (2006). "Performance and cognition". In: ed. by McConachie, Bruce. and Hart, Faith Elizabeth. Routledge. Chap. It is required/You do awake your faith: Learning to trust the body through performing The Winters Tale, pp. 122–146.

Rowe, Robert (1993). *Interactive Music Systems: Machine Listening and Composing*. MIT Press. URL: https://wp.nyu.edu/robert_rowe/text/interactive-music-systems-1993/.

Sanfilippo, Dario (2020). *Complex musical behaviours via time-variant audio feedback networks and distributed adaptation: a study of autopoietic infrastructures for real-time performance systems*. DOI: 10.7488/ERA/491.

Santini, Giovanni (2020).
"Action scores and gesture-based notation in Augmented Reality".
In: *Proceedings of the International Conference on Technologies for Music Notation and Representation*. Ed. by Rama Gottfried, Georg Hajdu, Jacob Sello, Alessandro Anatrini, and John MacCallum.
Hamburg, Germany: Hamburg University for Music and Theater, pp. 84–90.
ISBN: 978-3-00-066930-9.

Saunders, James (2021). "Group behaviors as music". In: *Together in Music*. Oxford University Press, pp. 13–23.
DOI: 10.1093/oso/9780198860761.003.0002.

Shafer, Seth (2017). "Performer action modeling in real-time notation".
In: *Proceedings of the International Conference on Technologies for Music Notation and Representation*.
Ed. by Helena Lopez Palma, Mike Solomon, Emiliana Tucci, and Carmen Lage.
Universidade da Coruna, pp. 117–123. ISBN: 978-84-9749-666-7.

Smith, Ryan Ross (2016). "A Practical and Theoretical Framework for Understanding Contemporary Animated Scoring Practices".
PhD thesis. Rensselaer Polytechnic Institute. URL: http://ryanrosssmith.com/papers/RyanRossSmith_PhD_Dissertation_2016_Animated_Notation.pdf.

Stein, Barry E., Nancy London, Lee K. Wilkinson, and Donald D. Price (1996).
"Enhancement of Perceived Visual Intensity by Auditory Stimuli: A Psychophysical Analysis".
In: *Journal of Cognitive Neuroscience* 8.6, pp. 497–506. ISSN: 1530-8898.
DOI: 10.1162/jocn.1996.8.6.497.

Stein, Barry E. and M. Alex Meredith (1993). *The merging of the senses*.
A Bradford book. Bibliography: p. [175]-201. Cambridge, Mass: MIT Press.
211 pp. ISBN: 0262193310.

Supercollider (2019). *Supercollider 3.12.2*.
URL: https://supercollider.github.io/.

The Blender Foundation (n.d.). *Blender*. URL: https://www.blender.org/.

Thompson, Charis (2007).
Making parents. The ontological choreography of reproductive technologies.
Inside technology. Cambridge, Mass. [u.a.]: MIT Press. 360 pp.
ISBN: 9780262701198.

Thompson, William Forde, Phil Graham, and Frank A. Russo (2005).
"Seeing music performance: Visual influences on perception and experience".
In: *Semiotica* 2005.156. DOI: 10.1515/semi.2005.2005.156.203.

Timmers, Renee, Freya Bailes, and Helena Daffern, eds. (2021). *Together in Music*.
Oxford University Press. DOI: 10.1093/oso/9780198860761.001.0001.

Tinnemans, Jobina (2017). “Imagiro Landmannalaugar”. In: URL: https://www.youtube.com/watch?v=OMxgUJXg-JU.

Unity Technologies (2021). *Unity*. URL: https://unity.com/.

Unterman, Michael, Jan Mueller-Szeraws, and I-Jen Fang (2018). *Match, for two cellos and percussion Mauricio Kagel*. Performance and Online Recording. Performance. URL: https://www.youtube.com/watch?v=yYW8SWzEOwE.

Vickery, Lindsay and Cat Hope (2011). *Talking Board*. DOI: 10.26180/5ceb41e023365.

Wapnick, Joel, Alice Ann Darrow, Jolan Kovacs, and Lucinda Dalrymple (1997). “Effects of Physical Attractiveness on Evaluation of Vocal Performance”. In: *Journal of Research in Music Education* 45.3, pp. 470–479. DOI: 10.2307/3345540.

Warburton, Edward C. (2011). “Of Meanings and Movements: Re-Languaging Embodiment in Dance Phenomenology and Cognition”. In: *Dance Research Journal* 43.2, pp. 65–84. DOI: 10.1017/s0149767711000064.

Whipple, Harold W. (1983). “Beasts and Butterflies: Morton Subotnick’s Ghost Scores”. In: *The Musical Quarterly* 69.3, pp. 425–441. ISSN: 00274631, 17418399. URL: http://www.jstor.org/stable/742180 (visited on 03/27/2023).

Whyte, David, Nick Didkovsky, and Stefan Hutzler (2018). “Zero Waste: Mapping the Evolution of the Iterative Sight-Reading of a Piano Score”. In: *Music Theory Spectrum* 40.2, pp. 302–313. DOI: 10.1093/mts/mty019.

Wilkes, M. V. and W. Renwick (1950). “The EDSAC (Electronic delay storage automatic calculator)”. In: *Mathematics of Computation* 4.30, pp. 61–65. DOI: 10.1090/s0025-5718-1950-0037589-7.

Winkler, Gerhard E. (2010). “The Real-Time-Score: Nucleus and Fluid Opus”. In: *Contemporary Music Review* 29.1, pp. 89–100. DOI: 10.1080/07494467.2010.509598.

Wolff, Christian (1964). *For 1, 2 or 3 People*. C.F. Peters Corporation.

Wright, Matthew, Adrian Freed, and Ali Momeni (2003). “OpenSound Control: State of the Art 2003”. In: *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03)*.

Appendices

A

Together We Tore Through Forested Thickets

Title:	*Together We Tore Through Forested Thickets*
Forces:	Animated Notation, Live electronic, small ensemble of violins
Date Finished:	2019
URL for Video:	`https://www.youtube.com/watch?v=FZaFBrDnp5w`

A.1 Screenshots of Instructions

Together We Tore Through Forested Thickets

For 3 - 6 Violins, and Live-Electronics (Stereo)

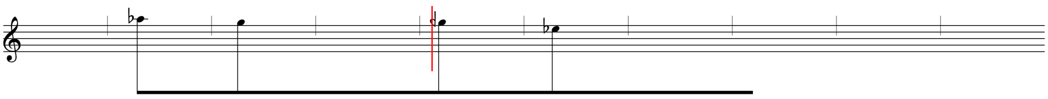
Start SuperCollider

Start Score

Violin Notes
Electronic Notes

▲◀▶ navigation

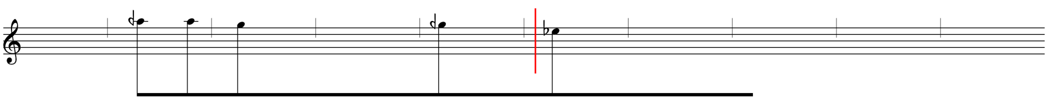
Rhythm



Each player imagines a playback-head (illustrated in red) moving from left to right at a unique tempo
that takes between 10 and 30 seconds to traverse the page.
The "barlines" break up the page and are only there to ease the imagining of the playback-head.
The notes under the playback-head are played and sustained for the length of the beam.
The absence of a beam indicates a rest; if you are playing a note and the beam disappears,
end the note and rest.
Combined, the violins create a varied cannon like texture: the material changes by the time
another performer reaches it.

▲◀▶ navigation

Pitch

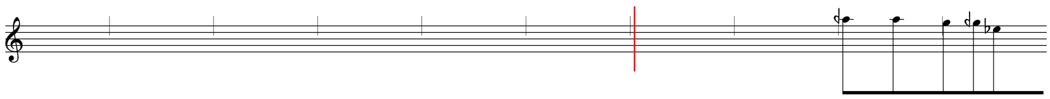


Pitch is in quarter tones.
Accidentals only apply to the immediate following note, they are not persistent.
The second note of this musical fragment is an A natural.

▲◀▶ navigation

Articulation

a gentle vibrato, subtle bow changes.



All notes should be articulated, even repeated pitches.
This should be determined by the rhythm, for example notes with short beams may be staccato.
Each performer should develop their own individual approach to articulation.
Bow changes should be subtle in the first section, but later, when indicated, more obvious.

▲◀▶ navigation

Undefined Behaviour

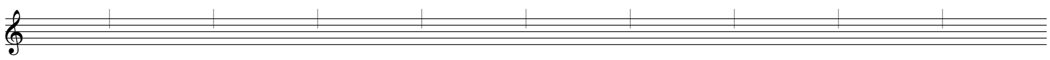
a gentle vibrato, subtle bow changes.

Sometimes there might be conflicting information resulting from the changing score.
What happens if a hairpin appears whilst you are mid way through a note, or a glissando, for example?
There is no right answer, instead each performer is encouraged to develop their own approach.

▲◀▶ navigation

A Final Word

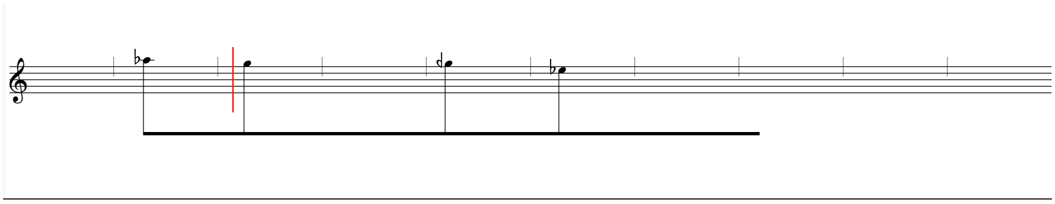
a gentle vibrato, subtle bow changes.

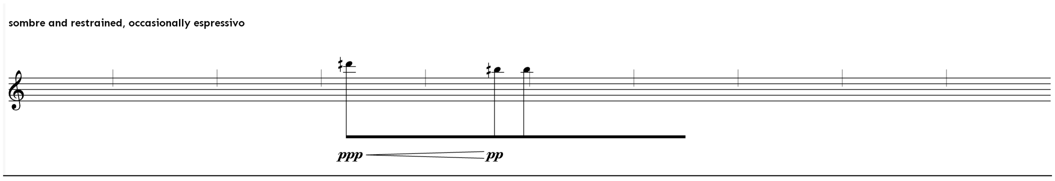


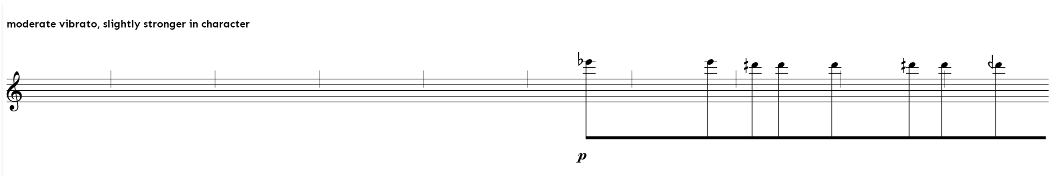
This work is not to be approached as an extreme sight-reading exercise, although there is certainly
an element of difficultly.
When the material changes too close to where your playback-head is and you do not have time to read it,
just choose the simpler option that allows you to continue.
There is no hurry to play all the notes as somebody in the ensemble will be in a different position
to you and have adequate time to play it.

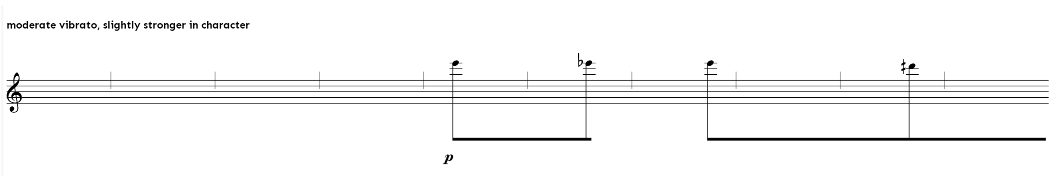
Each performer may start at a different place on the page.
If you notice you are playing in unison for a sustained length of time you may insert a fermata,
if this is inadequate you may jump to a different location on the page.

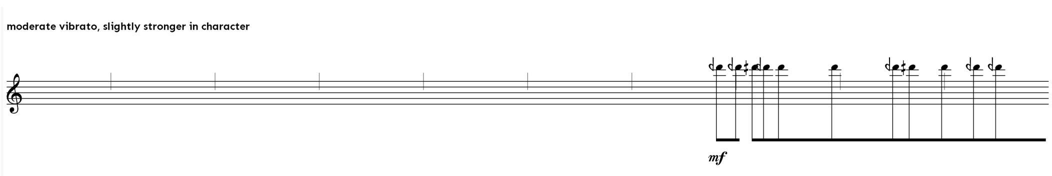
A.2 Further Examples









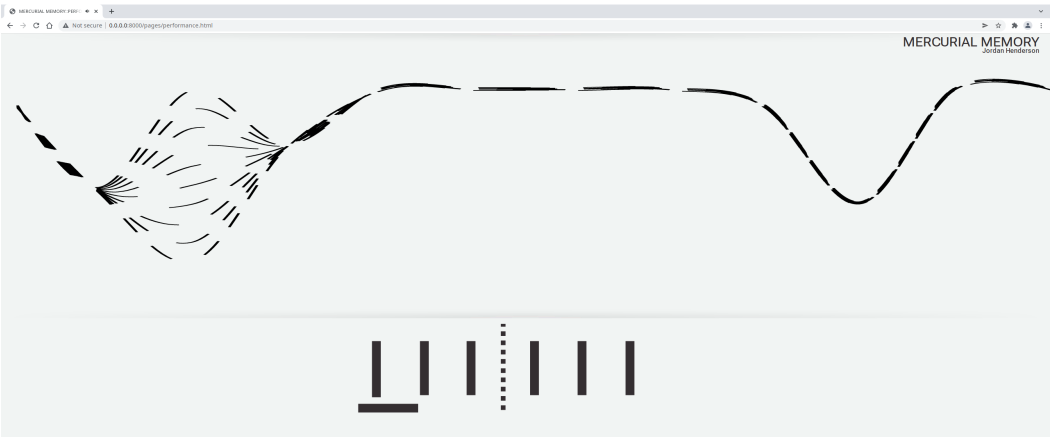


B

Mercurial Memory

Title:	*Mercurial Memory*
Forces:	Animated Notation, Live electronic, solo flute
Date Finished:	2020
URL for Video:	`https://youtu.be/vNfej9VTF6g`

B.1 Screenshots of Notation




MERCURIAL MEMORY
Jordan Henderson
Not secure
0.0.0.0:8000/pages/performance.html

C

Whilst We Speak Through The Wind

Title:	*Whilst We Speak Through The Wind*
Forces:	Live audiovisuals, Bass Clarinet, Alto Flute
Date Finished:	2020
URL for Video Demo:	`https://youtu.be/1y7-mspuM7c`

C.1 Score

...whilst we speak through the wind...

for Alto Flute in G, Bass Clarinet in B♭, Video Projection, and Stereo Electronics

Jordan Henderson

2019 - 2020

Technical Requirements

Equipment List

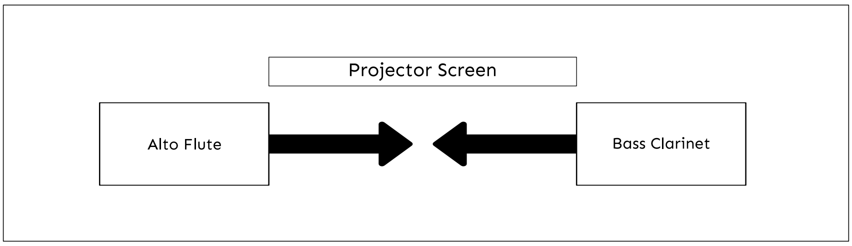
laptop	1
2 in 2 out audio interface	1
microphones plus cables	2
loudspeakers	2
projector	1
projector screen	1

On the Required Software

The laptop must be running Linux or macOS with Supercollider installed. The executable program 'whilst_we_speak' is required and can be obtained from the composer. It controls supercollider and displays the visuals.

On the Projection

Each player must sit to either side of the projection whilst facing each other, so it appears as if they are blowing the tree around as they play. Exactly how the screen is position is left open, however the performers must be able to see it without turning their bodies too much.



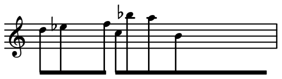
Each player must face each other as if in conversation

Performance Notes

On Time and Rhythm

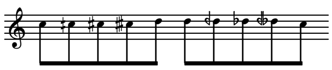
The score is in proportional notation, the duration of each note is shown by the length of the beam across the page. The exact duration of each beam is less important than their relative durations against one another. Emphasise the small changes in duration to help avoid meter.

The whole work should take between 12 and 15 minutes. To achieve this, it is recommend to practice with a timer.

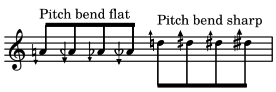


On Pitch

There are two types of accidental, defined quarter tones, and undefined pitch bends. Quarter tones use symbols whereas the pitch bends use arrows. When choosing quater tone fingering, choose one with the darkest colour.



Pitch bends are a timbral effect. Pitch bends should be as extreme as possible to change the sound quality. The resulting pitch is not notated.



On the flute when bending flat the instrument should be rolled towards the player with the bottom lip covering most of the embouchure hole creating a dark and muted sound. Likewise pitch bends sharp should have the instrument rolled away from the player creating an airy sound. On the clarinet both bending sharp and flat is done with vowels, and should be as extreme as possible. This will involve adopting a 'bad' embouchure position.

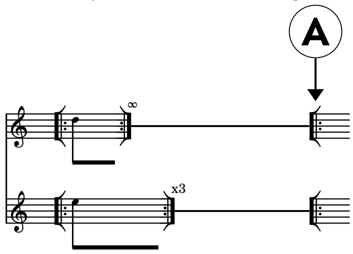
On Repetition and Synchronisation

Repeats come in two forms, counted and uncounted. Counted have the number of repetitions above their final barline, whereas uncounted do not.

A thick black line extends from the repeat, this line does not indicate time proportionally, it serves to aid the proportional spacing.

In both cases the repeat should finish at a synchronisation mark (a circled letter) which both players reach simultaneously. This will require altering the tempo in the case of counted, and altering both the tempo and number of repeats in the case of uncounted.

Both players should adjust their tempo by listening and observing each other's body language during the performance to reach the synchronisation mark together.



In this example, the top stave repeats as many times as needed whilst the bottom staves repeats exactly 3 times. Both performers must reach the end of their repeated material together, so the tempo must be stretched to accommodate.

It is vital that neither player take the role of conductor nor soloist. This may result in small errors, which are both players' responsibility to fix by altering the tempo, or in extreme cases inserting a brief pause. These are not mistakes but features of the piece.

This is all employed to expose the struggles and inherent dynamics of ensemble playing in a way the audience can observe.

No extra action beyond what is necessary to communicate efficiently should be used: this process should not be dramatised.

Performance Notes

On Conditional Instructions

In rounded boxes there are additional performance instructions that are conditional on the other ensemble members, human and electronic. These instructions last until the next synchronisation mark.

match clarinet's dynamics

There are two important key words, 'match' and 'inverse'.

Match requires listening to a transition and replacing it with a different transition. For example, the below instruction should be read as such: as the flute transitions between air and tone the clarinet should transition between mezzo forte with no vibrato and fortississimo with poco vibrato.

match flute's tone → air with no vib. → poco vib. *mf* < *fff*

Inverse is a type of 'match' but when the process is reversed: as the clarinet moves to air, the flute should move towards tone, and vice versa.

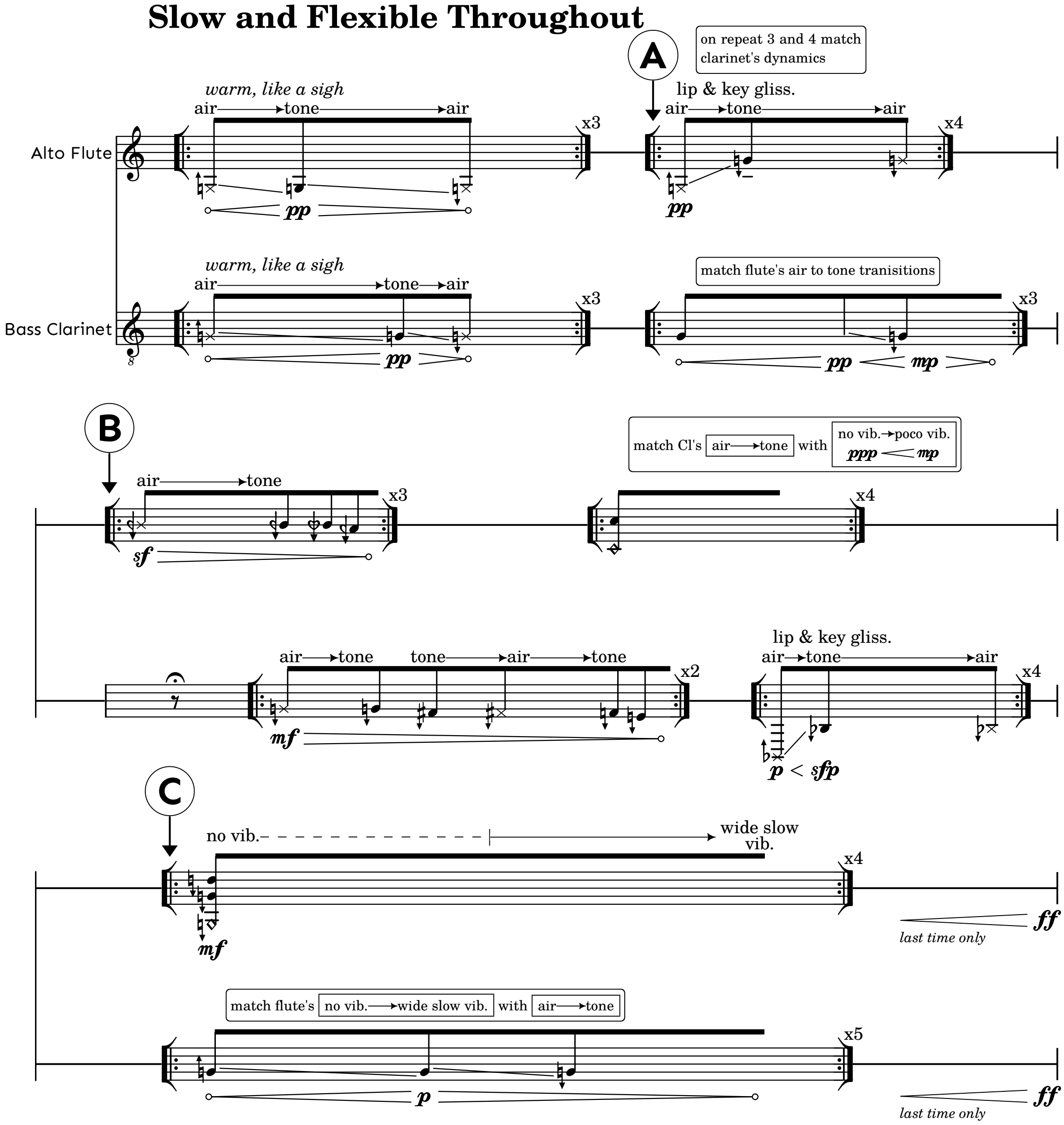
inverse clarinet's air → tone

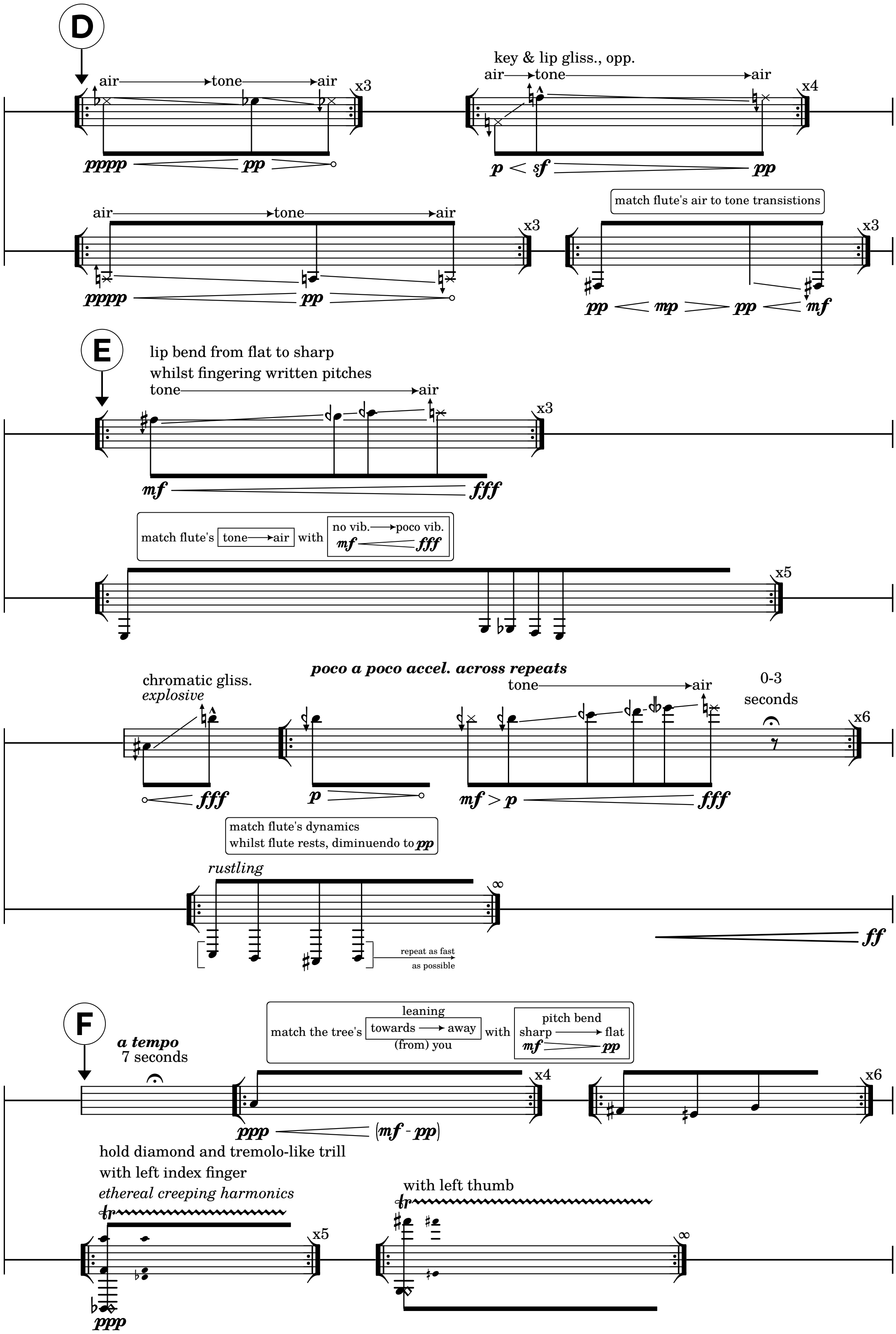
...whilst we speak through the wind...

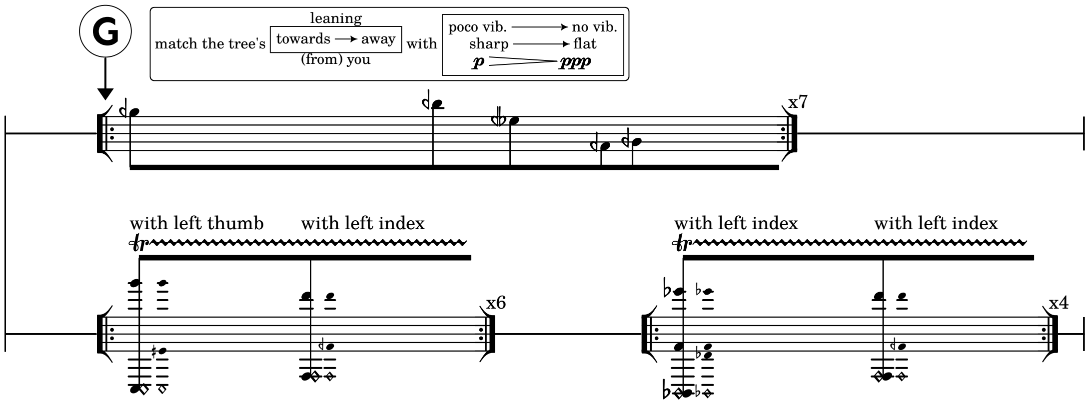
Full Score in C

Music by:
Jordan Henderson

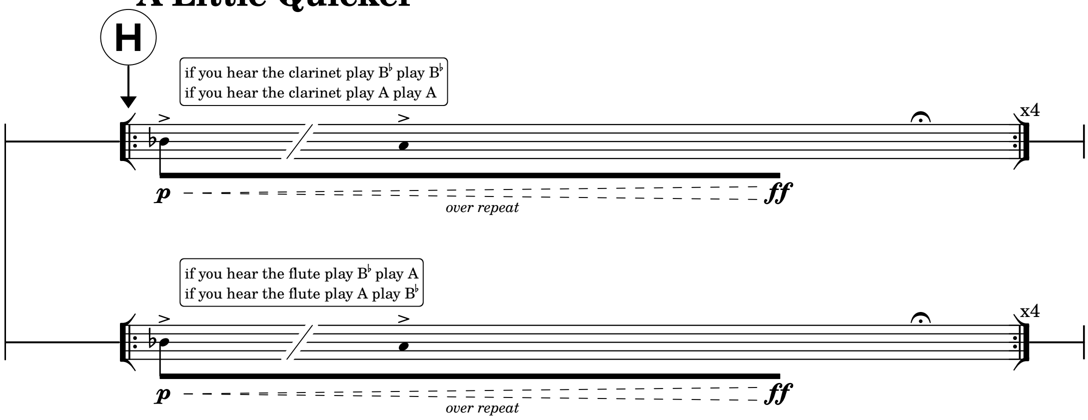
Slow and Flexible Throughout

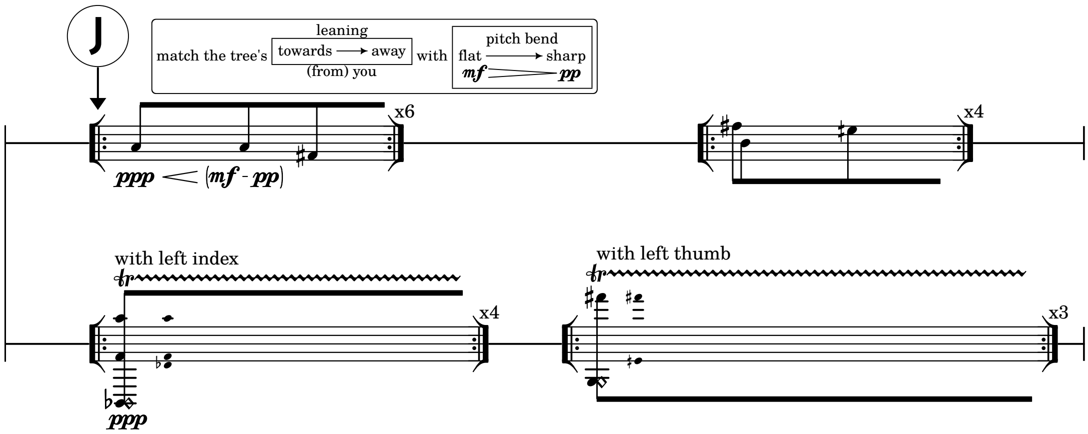


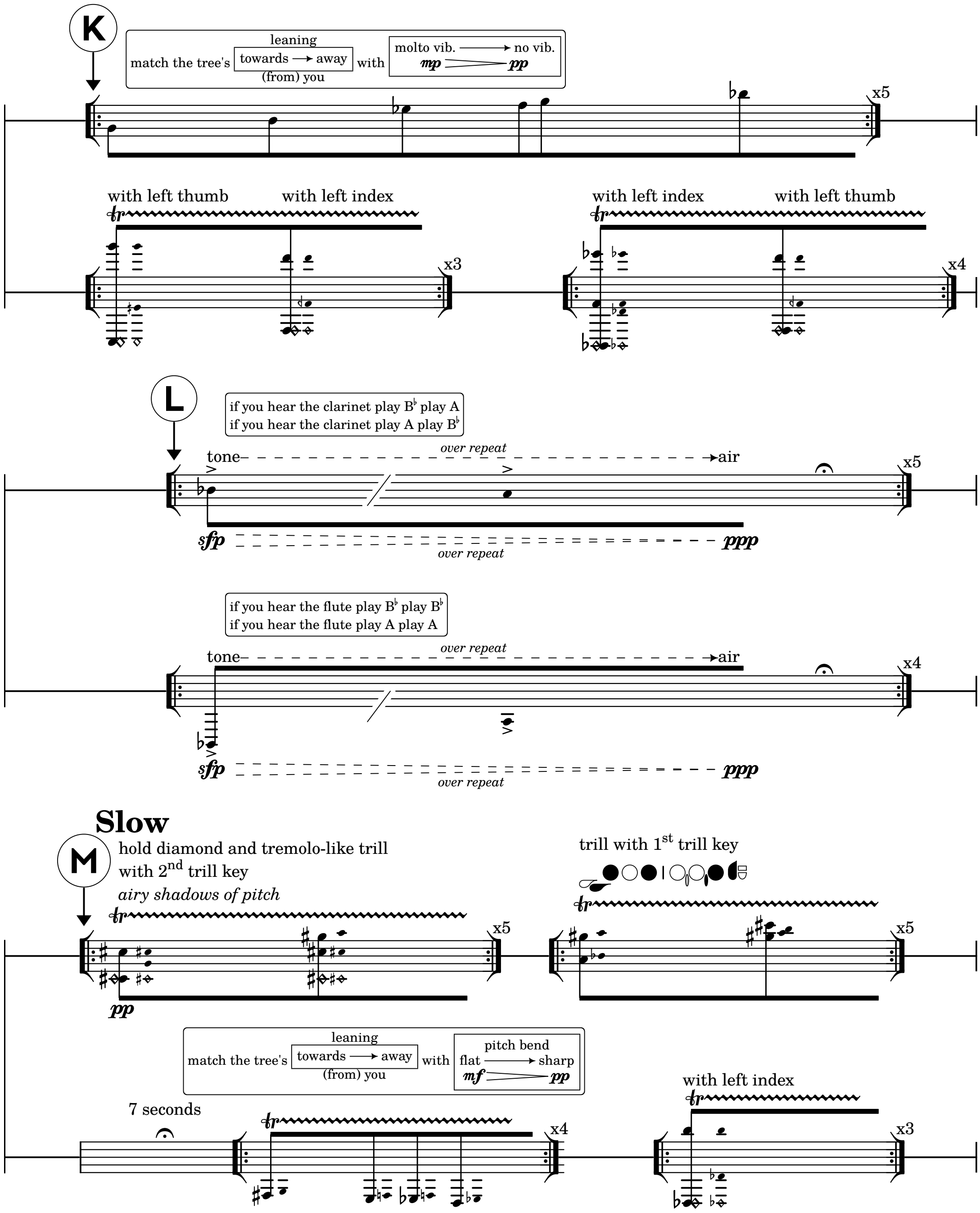

D
air tone air
x3
pppp pp
key & lip gliss., opp.
air tone air
x4
p sf pp
air tone air
x3
pppp pp
match flute's air to tone transitions
x3
pp mp pp mf
E
lip bend from flat to sharp
whilst fingering written pitches
tone air
x3
mf fff
match flute's tone air with no vib. poco vib. mf fff
x5
chromatic gliss.
explosive
fff
poco a poco accel. across repeats
tone air
0-3
seconds
x6
p
mf > p fff
match flute's dynamics
whilst flute rests, diminuendo to pp
rustling
∞
repeat as fast
as possible
ff
F
a tempo
7 seconds
match the tree's leaning towards away (from) you with pitch bend sharp flat mf pp
x4
ppp (mf - pp)
x6
hold diamond and tremolo-like trill
with left index finger
ethereal creeping harmonics
tr
x5
ppp
with left thumb
tr
∞

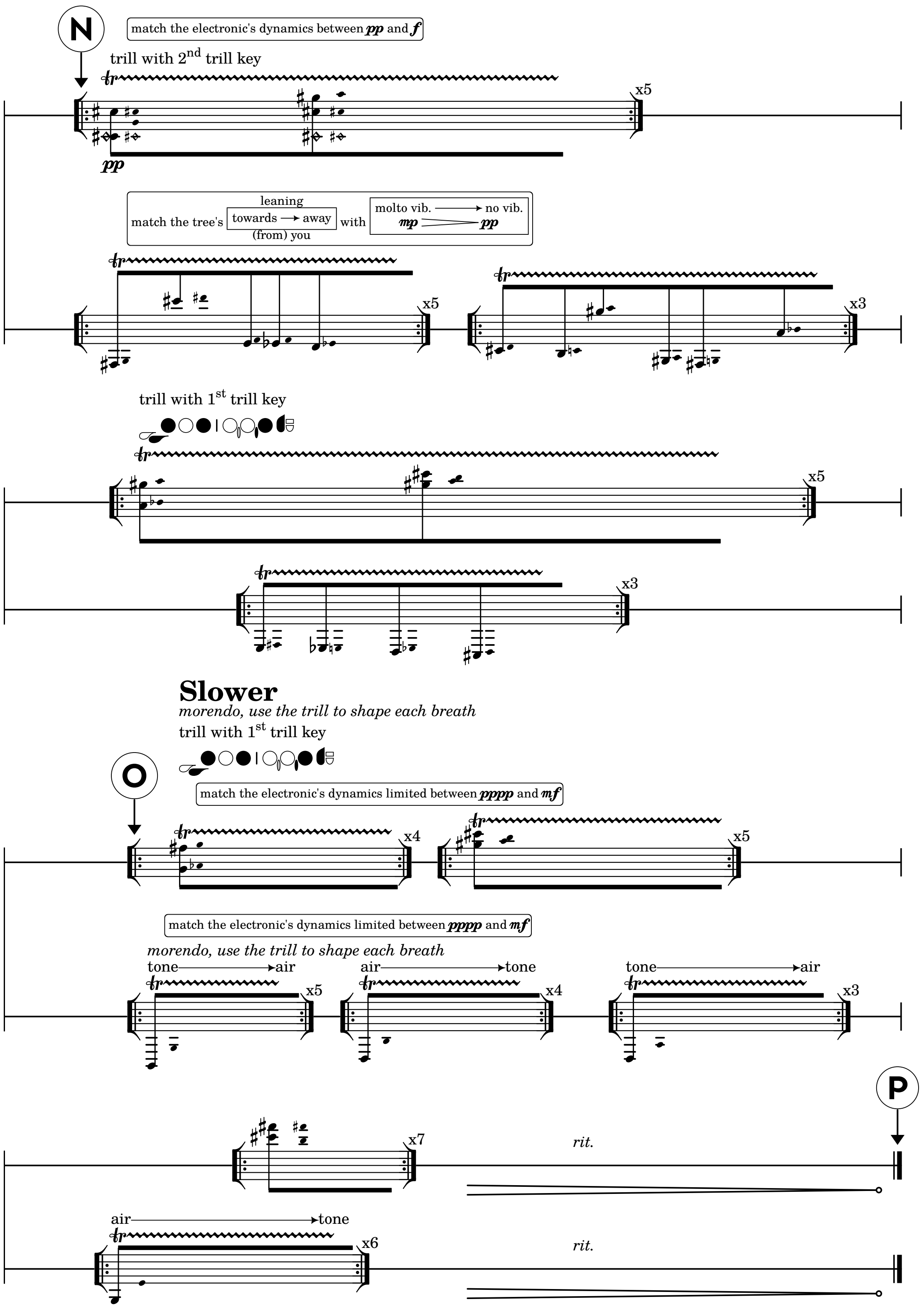


A Little Quicker





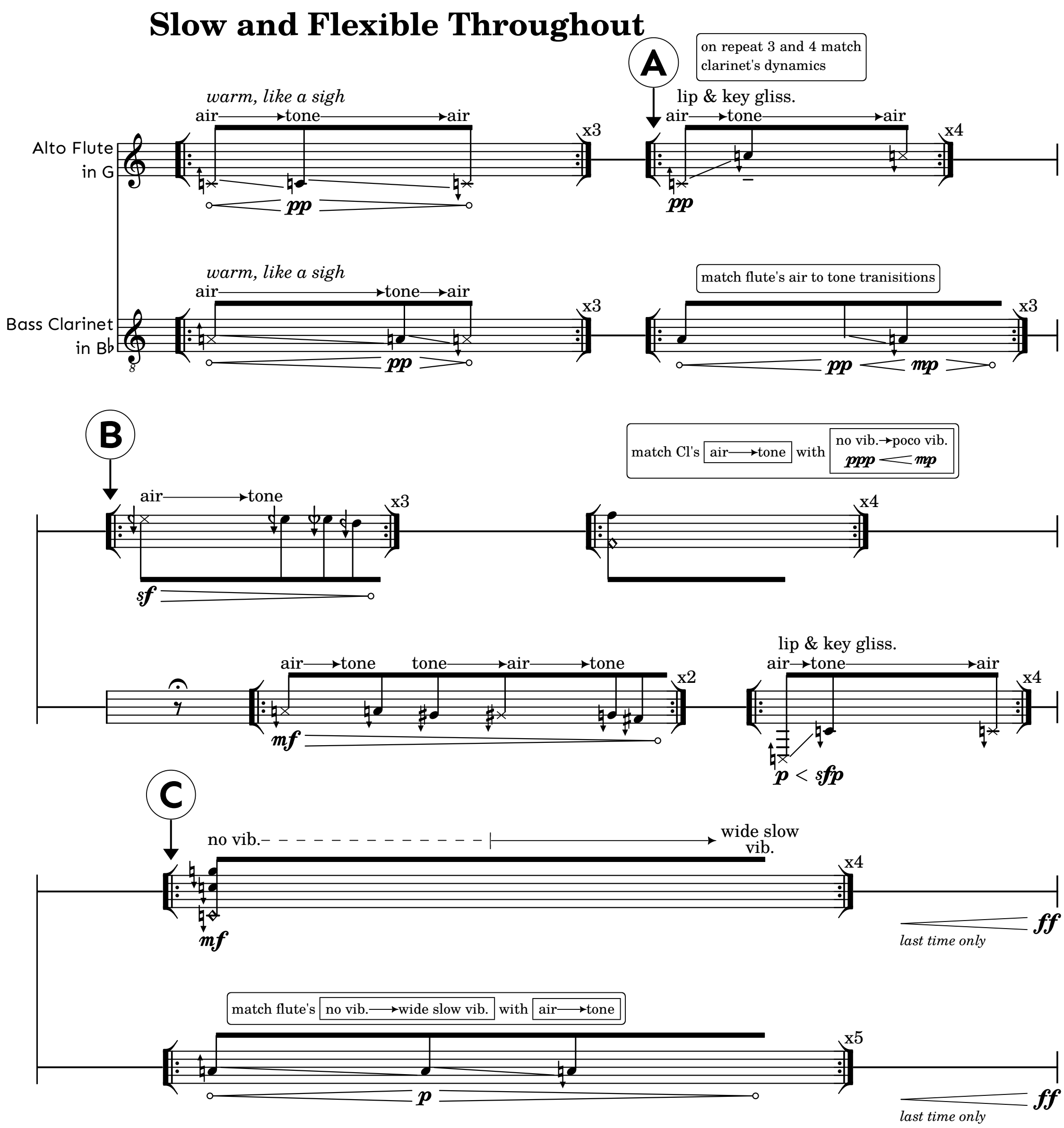

K
leaning
match the tree's towards → away with
(from) you
molto vib. → no vib.
mp pp
x5
with left thumb
with left index
tr
x3
with left index
with left thumb
tr
x4
L
if you hear the clarinet play B♭ play A
if you hear the clarinet play A play B♭
over repeat
tone– → air
x5
sfp
ppp
over repeat
if you hear the flute play B♭ play B♭
if you hear the flute play A play A
over repeat
tone– → air
x4
sfp
ppp
over repeat
Slow
M
hold diamond and tremolo-like trill
with 2nd trill key
airy shadows of pitch
tr
x5
pp
trill with 1st trill key
tr
x5
leaning
match the tree's towards → away with
(from) you
pitch bend
flat → sharp
mf pp
7 seconds
tr
x4
with left index
tr
x3

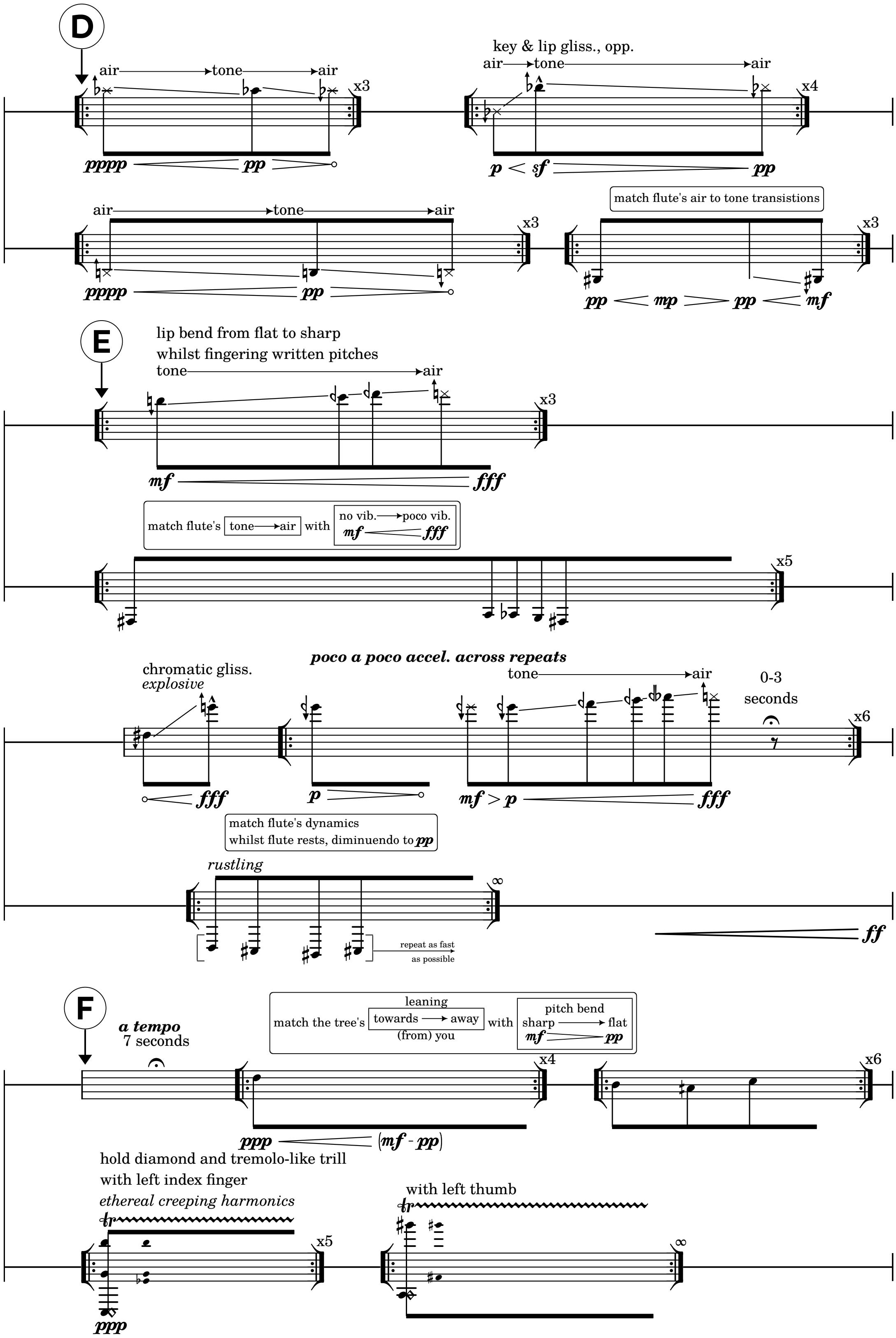

N
match the electronic's dynamics between pp and f
trill with 2nd trill key
tr
x5
pp
leaning
match the tree's towards → away (from) you with molto vib. → no vib. mp pp
tr
x5
tr
x3
trill with 1st trill key
tr
x5
tr
x3
Slower
morendo, use the trill to shape each breath
trill with 1st trill key
O
match the electronic's dynamics limited between pppp and mf
tr
x4
tr
x5
match the electronic's dynamics limited between pppp and mf
morendo, use the trill to shape each breath
tone → air
tr
x5
air → tone
tr
x4
tone → air
tr
x3
P
x7
rit.
air → tone
tr
x6
rit.

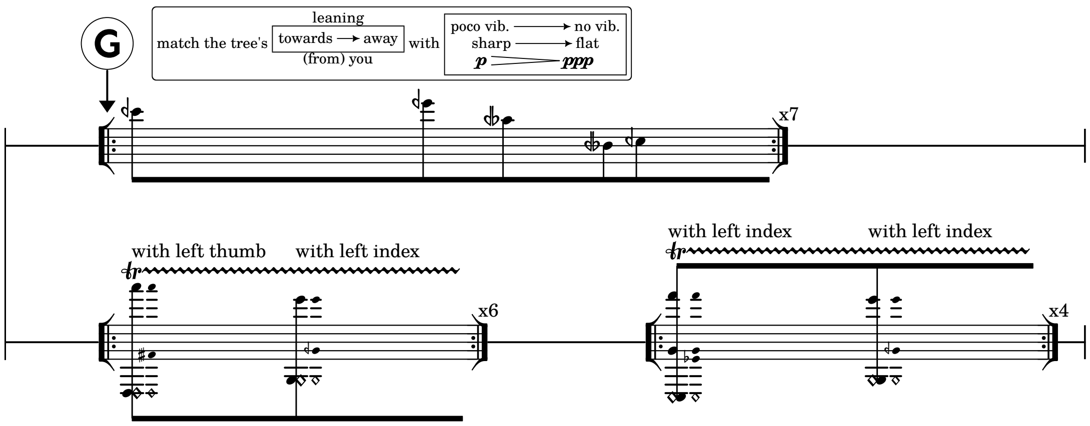
...whilst we speak through the wind...

Fully Transposed Score

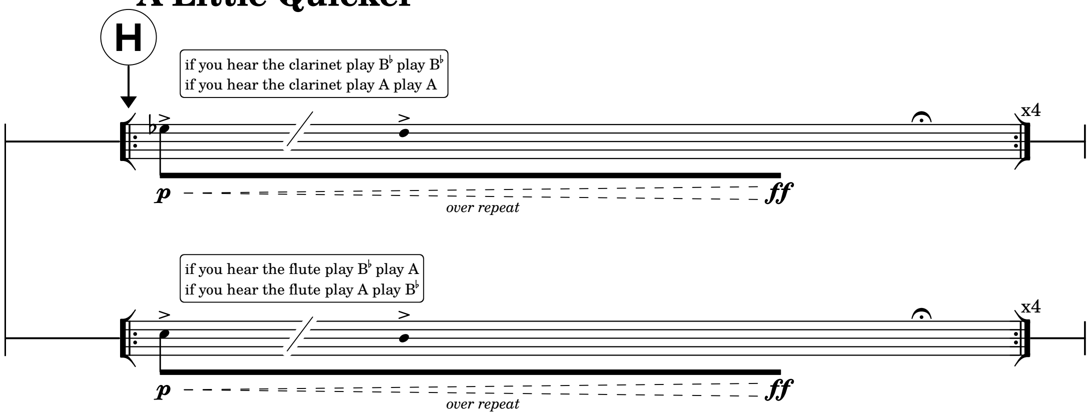
Music by:
Jordan Henderson

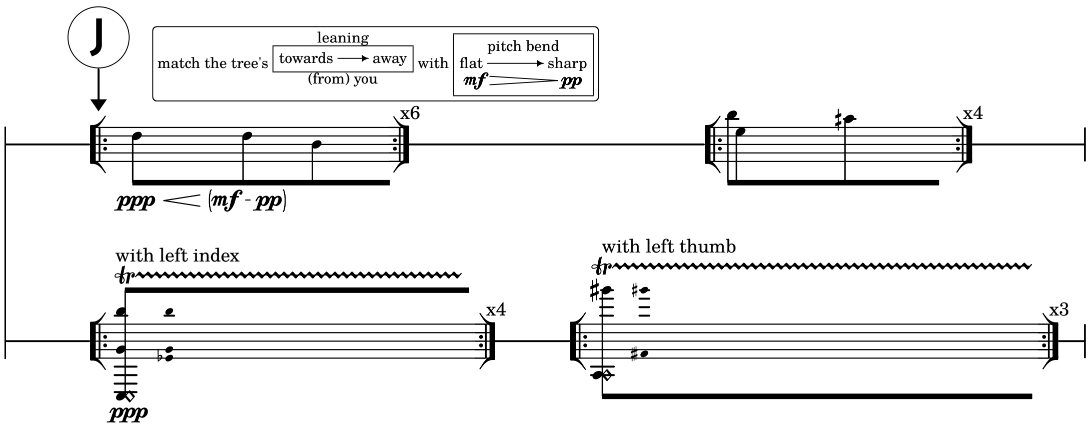


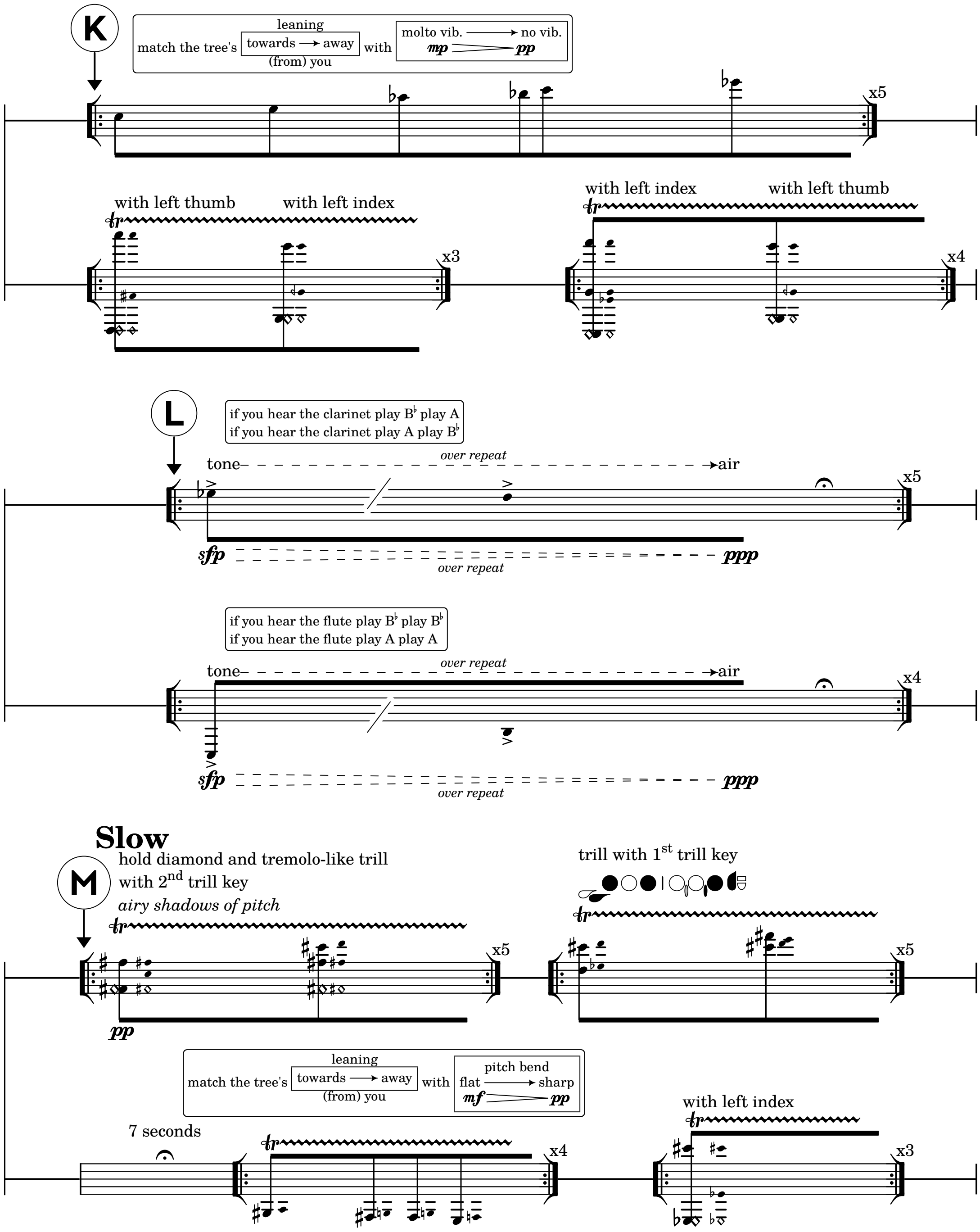

D
air
tone
air
x3
pppp
pp
key & lip gliss., opp.
air
tone
air
x4
p < sf
pp
air
tone
air
x3
pppp
pp
match flute's air to tone transistions
x3
pp
mp
pp
mf
E
lip bend from flat to sharp
whilst fingering written pitches
tone
air
x3
mf
fff
match flute's tone air with
no vib. poco vib.
mf fff
x5
chromatic gliss.
explosive
poco a poco accel. across repeats
tone
air
0-3
seconds
x6
fff
p
mf > p
fff
match flute's dynamics
whilst flute rests, diminuendo to pp
rustling
∞
repeat as fast
as possible
ff
F
a tempo
7 seconds
match the tree's
leaning
towards away
(from) you
with
pitch bend
sharp flat
mf pp
x4
x6
ppp
(mf - pp)
hold diamond and tremolo-like trill
with left index finger
ethereal creeping harmonics
tr
x5
ppp
with left thumb
tr
∞

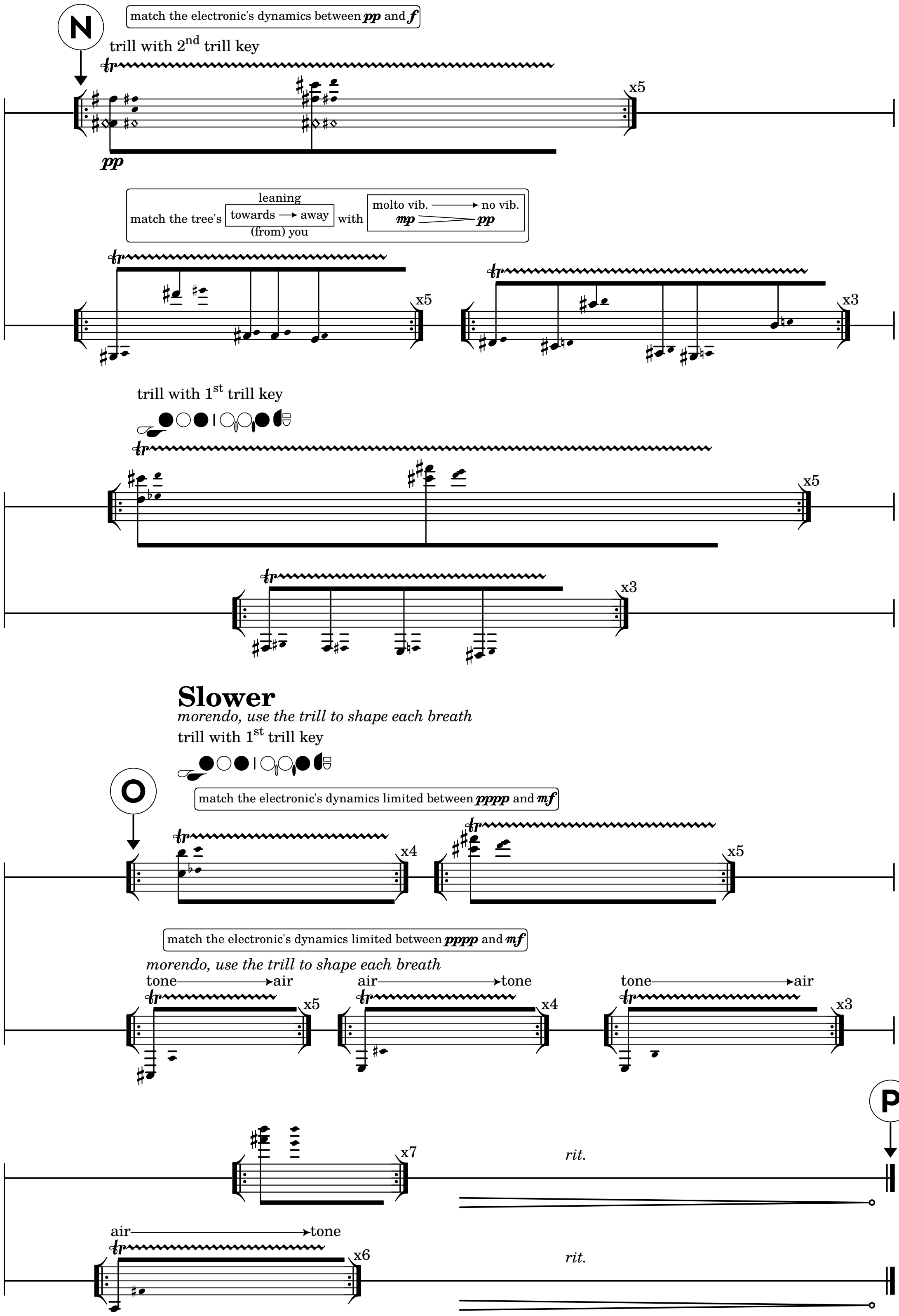


A Little Quicker






K
leaning
match the tree's towards → away with
(from) you
molto vib. → no vib.
mp pp
x5
with left thumb
with left index
tr
x3
with left index
with left thumb
tr
x4
L
if you hear the clarinet play B♭ play A
if you hear the clarinet play A play B♭
tone– over repeat →air
x5
sfp over repeat ppp
if you hear the flute play B♭ play B♭
if you hear the flute play A play A
tone– over repeat →air
x4
sfp over repeat ppp
Slow
M
hold diamond and tremolo-like trill
with 2nd trill key
airy shadows of pitch
tr
x5
pp
trill with 1st trill key
tr
x5
leaning
match the tree's towards → away with
(from) you
pitch bend
flat → sharp
mf pp
7 seconds
tr
x4
with left index
tr
x3

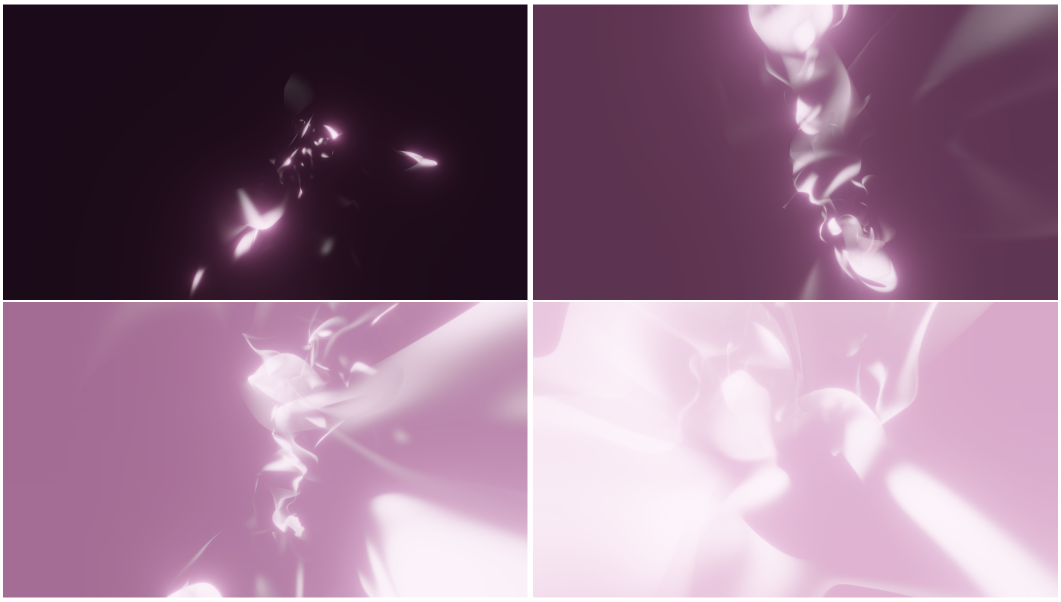

N
match the electronic's dynamics between pp and f
trill with 2nd trill key
x5
pp
leaning
match the tree's towards → away (from) you with molto vib. → no vib. mp > pp
x5
x3
trill with 1st trill key
x5
x3
Slower
morendo, use the trill to shape each breath
trill with 1st trill key
O
match the electronic's dynamics limited between pppp and mf
x4
x5
match the electronic's dynamics limited between pppp and mf
morendo, use the trill to shape each breath
tone → air
air → tone
tone → air
x5
x4
x3
P
x7
rit.
air → tone
x6
rit.

D

Glimmer

Title:	*Glimmer*
Forces:	Baroque flute, Video Projection
Date Finished:	2021
URL for Workshop Recording:	https://youtu.be/Kau_BH0O0hE
URL for Lilypond Code:	https://github.com/JordanHendersonMusic/glimmer
URL for Video:	https://www.youtube.com/watch?v=MUc52c5S_Eo

D.1 Video Screenshots



D.2 Paper Score

glimmer

Baroque Flute
Video Projection/Score

JORDAN HENDERSON

2021

push with the warmth of a palm guiding first steps
pulled through vibrant magnetism tracing interwoven paths
pull supple limbs over contorted bodies of light
pushed beyond, outside of, enlightened from...

feel each twist and turn over xenomorphised muscle and skeletal inscriptions
fractured movements between unfurling bodies
focus out inwards towards resonant cavities filled with
flecks of glimmering sympathies

Performance Instructions

Fingering Stave

covered hole ●
half hole ◐
open hole ○
key *k*
transition ⟶

Although transitions are all marked with a linear line, the rate of transition should depend on the movement on screen. It has been left out for clarity.

shake slow
shake normal 
shake fast 

Slow shakes are slight wobbles of the finger over the hole, the motion should be similar to a string player's vibrato, or a finger vibrato, depending on whether the hole is open or closed. Normal shakes are similar to slow shakes, but they reveal more of the hole. Fast shakes are rapid trills.

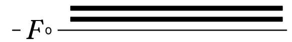
Harmonic Stave

Blowing is indicated with a thick black line.

If the thick line goes through the thin stave line, attempt to play the fundamental harmonic, F0, of the currently held fingering.

Each ledger line above the thin line represents a higher harmonic. If the fingering was a low D, this would be the octave above.

... and this would be the multiphonic of the octave and the fifth above that.



Because the fingerings are, almost always, unconventional, the harmonic series will not be followed. This means it can be hard to tell which harmonic is being played, if there is any doubt, aim for a harmonic roughly in the register that would be produced if the fingering was a low D.

When the thick line dips below the thin line, pitch bend flat.

Articulation, which is left to the performer's discretion, should be used throughout, but not constantly. It is recommended to use a variety of articulation to both punctuate changes in the visuals, and to aid in the realisation of the dynamics through a 'stippling' effect.

The notation that follows is only half the score, the rest is the video.

The video should be projected, so the audience can see it, and again (perhaps on a monitor) so the flautist can see it. The flautist must face the audience.

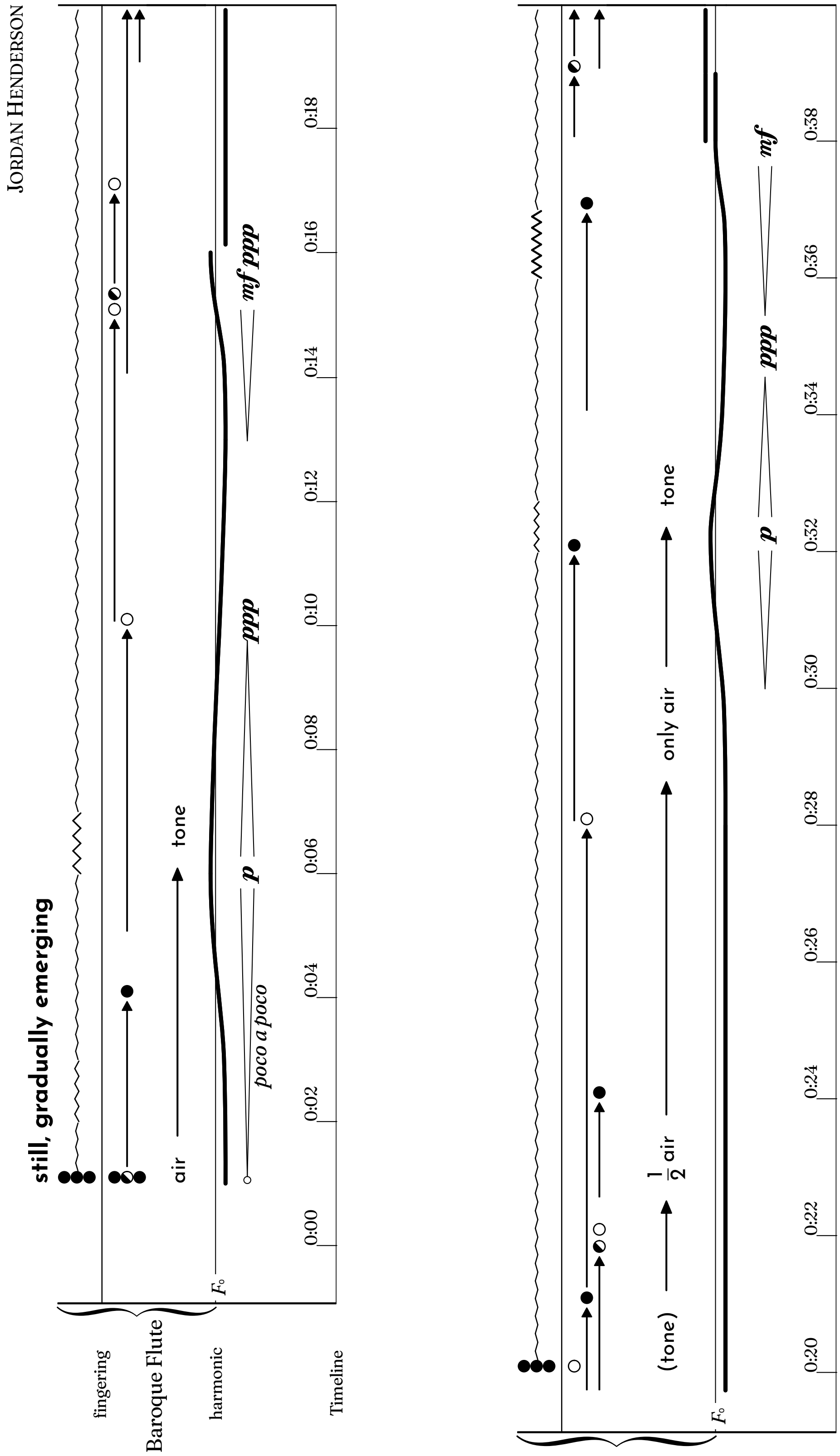
Each mark in this document corresponds to motion or change in the video. It illustrates a way of using the motions of flute technique to dance with and, more importantly, 'as' the object in the video. Therefore, a form of understanding is encoded in this notation.

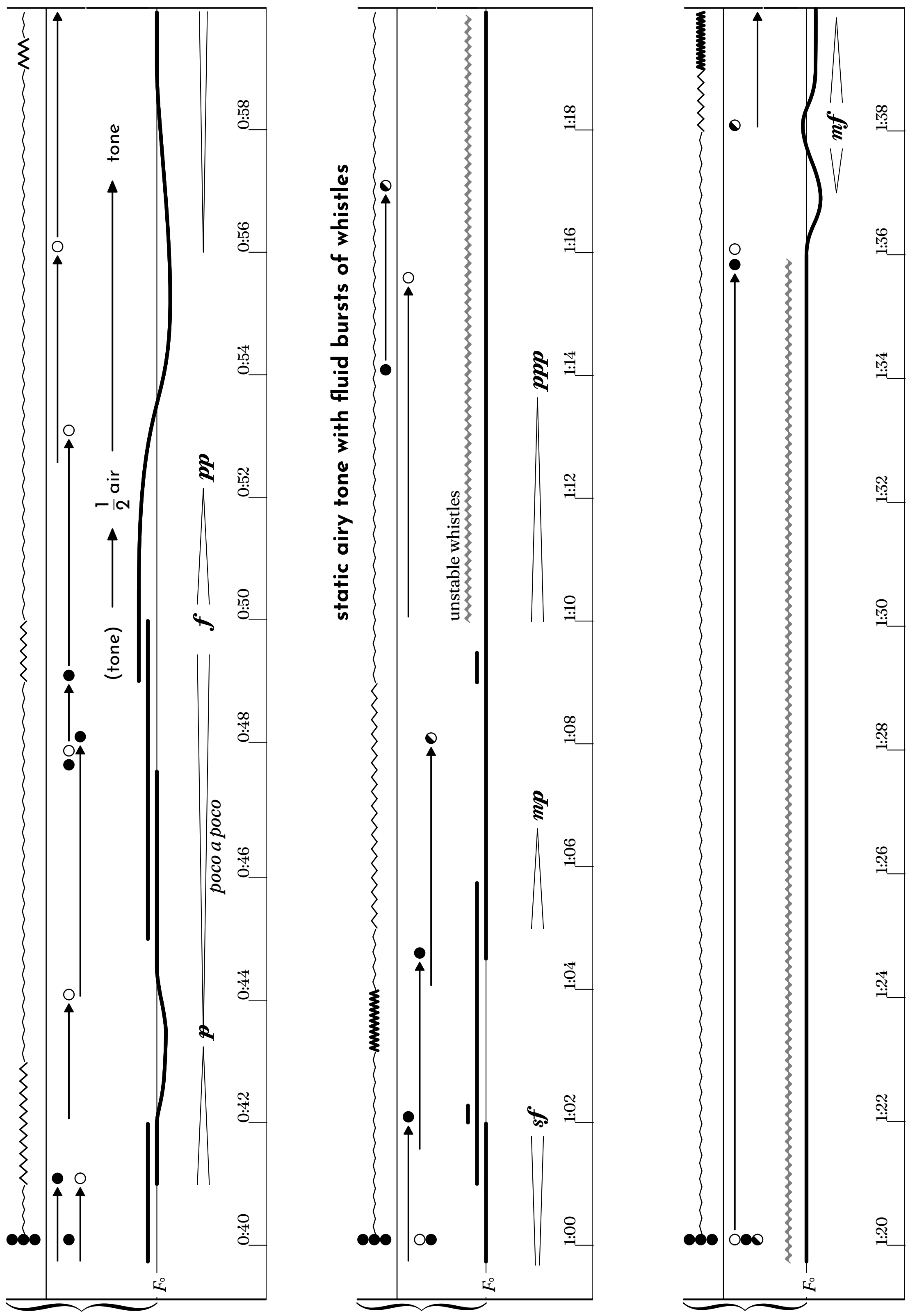
After about 2 minutes, the notation gradually fades ...

... continue this understanding.

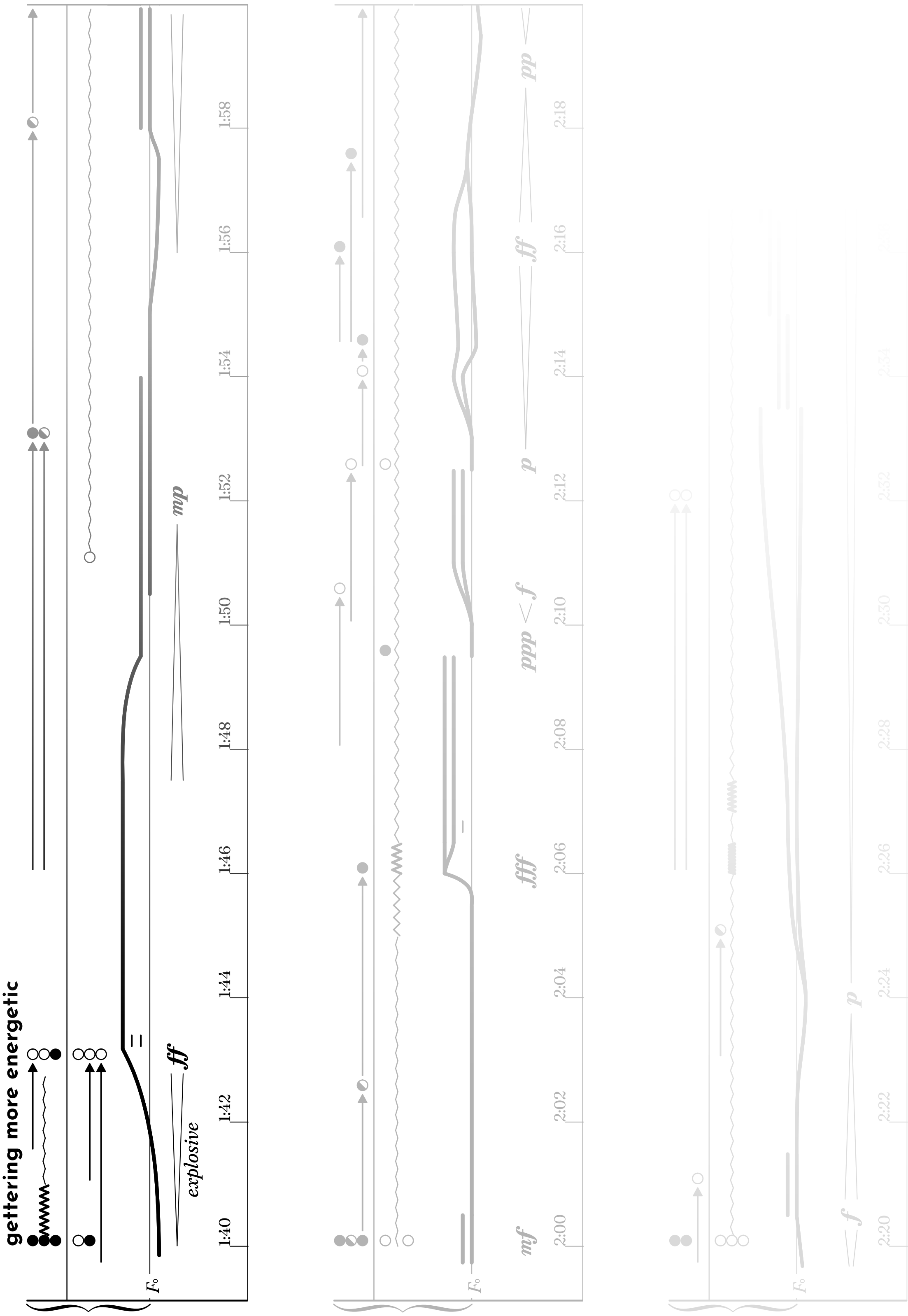
glimmer

Baroque Flute, Video Projection / Video Score





static airy tone with fluid bursts of whistles
(tone)
1/2 air
tone
poco a poco
p
f
pp
sf
mp
ppp
mf
unstable whistles
F_0
0:40
0:42
0:44
0:46
0:48
0:50
0:52
0:54
0:56
0:58
1:00
1:02
1:04
1:06
1:08
1:10
1:12
1:14
1:16
1:18
1:20
1:22
1:24
1:26
1:28
1:30
1:32
1:34
1:36
1:38


gettering more energetic
explosive
ff
mp
1:40
1:42
1:44
1:46
1:48
1:50
1:52
1:54
1:56
1:58
mf
fff
ppp < f
p
ff
pp
2:00
2:02
2:04
2:06
2:08
2:10
2:12
2:14
2:16
2:18
f
p
2:20
2:22
2:24
2:26
2:28
2:30
2:32
2:34

D.3 *Memory* Full Score

The following pages are a reproduction of the score for *Memory*, a precursor to *Glimmer*.

memory

Jordan Henderson

for projected video and any open hole flute

2020

Requirements

Video
Flute

Score Instructions

Light
Fingering
Kinetic

Requirements

Video

The video file is both the score and a visual object for the audience to see

It should be projected behind or to the side (but not on top) of the flautist

It should also be mirrored facing the flautist, this may be done using a second computer monitor and mirroring the main projected display, or a literal mirror may be placed on the back of the hall

If using a second computer monitor it is preferable to use a large display placed behind the audience or otherwise disguised so that the audience can not see it, or at least not obstructing their view of the flautist

The room should be dark, with a soft spot light on the performer

Flute

This work is for any open holed flute

Whether that be one of the varieties of bamboo flute (Dizi, Bansuri...), baroque, or western

If an option is available choose a lower flute but not a 'bass' flute

3 fingerings are given on the following page for a 6 hole instrument (such as the Dizi)

You will need to devise your own fingerings for your instrument

Each fingering must involve changing at least one open hole, although preferably more

The trilled key must always be an open hole

This is all so your fingers can be 'pulled' and 'pushed' over the holes in a linear progression

Score Instructions

The video contains a shape that moves and glimmers which is the score

There should be little to no time delay between the shape moving and the performers reaction this will require practice

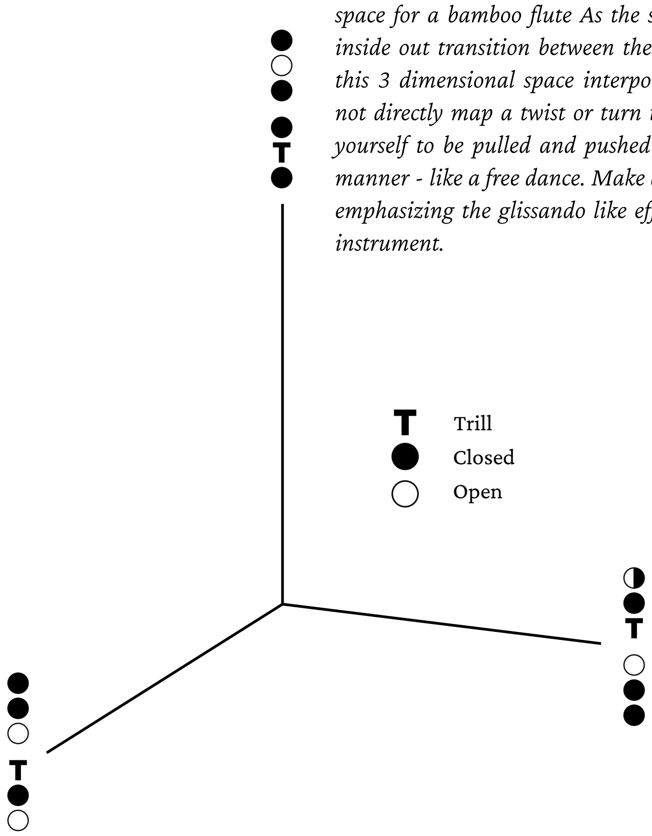
The shape's behavior must be mapped into three separate 'spaces'

In order of importance, these are:

1. *Morphological / Fingering space*
2. *Illumination / Character space*
3. *Kinetic / Trill space*

Morphological / Fingering Space

Here is a three dimensional axis depicting morphological / fingering space for a bamboo flute As the shape twists, turns, and pulls itself inside out transition between these fingerings as if moving through this 3 dimensional space interpolating between each fingering. Do not directly map a twist or turn to some specific axis, instead allow yourself to be pulled and pushed through the space in an intuitive manner - like a free dance. Make all transitions as smooth as possible emphasizing the glissando like effects only possible on an open hole instrument.



Illumination / Character Space

As the shape changes illumination, from pitch black, to partially lit with shadows, to finally blinding bright, gradually transition between the following states as gradually as possible.

pitch black		shadows		blinding bright
dark and airy in tone	→	*fragile & delicate*	→	*piercing*
lowest pitch, pitch bent as flat as possible	*harmonic sweep gliss.*	mid harmonic with a low shadowing pitch	*harmonic sweep gliss.*	high harmonic or multiphonic optional flutter tongue

ø — *poco a poco* — ***fff***

Kinetic / Trill Space

As the shape changes kinetic energy, gradually transition between these three trill speeds. This may be the least accurate of the three spaces.

static		slow		rapid
no change	→	slide fingers over trill hole gradually, like a glissando	→	rapid trill

E

Breathe

Title:	*Breathe*
Forces:	vocalist and live audiovisual
Date Finished:	2022
Performed:	GLEAM (Glasgow Electronic and Audiovisual Media Festival) 2022
URL for Performance:	`https://youtu.be/rcpPwGceL4I`

Jellyfish

Title:	*Jellyfish*
Forces:	Bowed guitar (or cello) and live audiovisuals
Date Finished:	2023
Performed:	GLEAM (Glasgow Electronic and Audiovisual Media Festival) 2023
URL for performance:	`https://youtu.be/tRFPDGTPHSk`

G

JX

Title:	*JX*
URL:	`https://github.com/JordanHendersonMusic/JX-supercollider`

Throughout this PhD there has been a continued focus on how relationships and communication is expressed in code or otherwise encoded into the computer, which is due to my compositional method being primarily experimental and improvised as opposed to planned. If the work was planned it would be possible to view coding as an implementation of the plan, but instead each step is incremental, where the current state of the code influences what could happen next. While there are some planned moments or ideas, these are either general, as in deciding on the nodding gesture in *Jellyfish*, or are quickly discarded once composition begins. Due to this, it is important in my practice to consider the code as an instrument of composition which has certain musical affordances. However, code is extremely fragile and can quickly become unsustainable if not carefully managed, particularly when pieces develop quickly. This problem is confounded when there are multiple systems involved, such as a sound producing system, a visual system, and perhaps a gesture recognition system, as the independent state of each system must be managed and made to interact with the others synchronously. Each successive piece in this portfolio has worked towards creating a better coding framework that addresses both the issue of continued experimental development and distributed state, which has culminated in a *SuperCollider* quark called *JX*.

JX is split into two halves, the behavioural mapping and a way of managing resources on the server, such that each synth has its own OSC address reflecting its place in the server graph and each resource is addressable within that. An example of this can be seen in the following code snippet where OSC sources and sinks are defined within the synth and are addressable at `/src/amp` and `/sink/amp`, which can be directly used within the behaviour mapping system. Further, if these synths are within a group, the address of the group is prepended, thereby causing the server's node structure to be mirrored in the OSC namespace.

```
JXSynthDef('/src', {
  var sig = ....;
  JXOscSrc.kr('/amp', Amplitude.kr(sig, 0.3, 0.1) )
});
JXSynthDef('/sink', {
  var amp = JXOscSink.kr('/amp', 1);
  ...
});
```

The behavioural mapping system is also refined from *Jellyfish* as more helper methods are defined to make different types of interpolation possible and to further ease the nesting of behavioural maps. A small example of this can be seen in the following code snippet where bilinear interpolation is used (X-Y plane) and maps are combined using operators.

```
JXOscMapperSynth({|src|
  var a = mapAmk.makeMap(src);
  var b = mapBmk.makeMap(src);
  var c = mapCmk.makeMap(src);
  var d = mapDmk.makeMap(src);

  // X-Y plane interpolation
  var xy = [MouseX.kr, MouseY.kr];
  var x_y_interp = JXOscMapBiLinearX(xy, a, b, c, d);

  // calculate the mean of (a + b) and (c + d).
  var sum_mean = (a + b) <+/> (c + d);
});
```

H

SuperCollider Source Code Contributions

This section includes significant contributions made to the *SuperCollider* project while working on this research, documentation changes and minor bug fixes have been omitted.

Title: ***SCLang, ClassLib: Add *kwargs***

Status: Merged

Github ID: #6339

Allow extra keywords to be caught by function calls, this means object prototyping and other function wrapper now work with keywords. A significant rewrite of the VM's calling mechanism and first substantial change to the syntax in over a decade. Fixes several long standing bugs with keywords.

Examples of change.

```
A {
  someMethod { |...args, kwargs| [args, kwargs] }
}
A().someMethod(1, 2, a: 10, b: 11)
// returns: [[1, 2], [a, 10, b, 11]]
```

```
a = (
   \foo: { |self, a, b... args, kwargs|
        [a, b, args, kwargs]
  }
);
a.foo(1, b: 2, c: 3)
// returns: [1, 2, [], [c, 3]]
```

Title: ***Scsynth & Plugin: Remove c cast from CALCSCOPE and fix GVerbs roomsize***
Status: Merged
Github ID: #6262

Fix 15+ year old issue in one of the reverb plug–ins.

Title: ***SCLang: Make 64-bit slot use nan boxing***
Status: In Review
Github ID: #6365

Use ‘nan-boxing’ to reduce memory usage on 64-bit systems by half.

Title: *ClassLib: fix and refactor SynthDef compiler*

Status: In Review

Github ID: #6405

A complete rewrite of the `SynthDef` compiler, fixing numerous longstanding bugs with a core *SuperCollider* system and providing significant performance benefits in certain situations, at the cost of longer compile times.