ALISA KOBZAR | 11836155

(Name in Blockbuchstaben) | (Matrikelnummer)

# Erklärung

Hiermit bestätige ich, dass mir der *Leitfaden für schriftliche Arbeiten an der KUG* bekannt ist und ich die darin enthaltenen Bestimmungen eingehalten habe. Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, den 23.11.2022

……

Unterschrift der Verfasserin/des Verfassers


Bachelorstudium Computermusik

Studienkennzahl V 033 104

# Sound Object

**WRITTEN PART OF THE ARTISTIC BACHELOR THESIS**

Kunstuniversität Graz

Institut für Elektronische Musik und Akustik

Vorgelegt von: Alisa Kobzar (Matrikelnummer: 11836155)

Betreuer: Univ.Prof. Dr.phil. Gerhard Eckel

# Table of contents

# I. INTRODUCTION

*„Every object of perception is at the same time an OBJECT in so far as it is perceived as a unit locatable in a context, and a STRUCTURE in so far as it is itself composed of several objects.“ (M.Chion [1])*

The starting point of my inquiry was the desire to fabricate the allusion of immaterial objects through dance movements and sounds. Those objects are understood as sound objects in this text (other notions of sound objects will also be referred to).

The affordances of sound objects are meant to be evoked by sound-gestural mappings[2]. As a tool for technology-mediated performance, those mappings are expected to meet both performative intentions and affordances of the controller technology. The use of active gestural control in the further described projects aims to create a strong connection between dance movements and the identification of sound sources or general acoustic characteristics.

The goal of the current text is to share the details of a subjective artistic process and an evaluation of its methods, as well as artistic adaptations of the relevant theoretic statements.

“Floating pointers” is a real-time project for one dancer and one computer music performer (duration: ca. 10-11 min). The word „pointer“ is coming from computer science and means a data object storing a memory address which refers (“points”) to another data object. The main idea of the project is the creative interpretation of the term „sound object“. This project has met two realizations with two different technical setups and modes of interaction.

The conceptual framework of the current text is therefore focused on describing those two realizations of the same piece as an experiment on evaluating the current artistic approaches and interaction models of the artistic practice of the duo rotkäppchen.

---

[1] Chion, “Guide To Sound Objects“, p.58

[2] Altavilla et at.,“Towards Gestural Sonic Affordances“ , p. 61

To describe the creative decision-making process, the needed theoretical bases in the auditory and visual domains are provided.

The auditory and visual terminologies are purposefully divided in this text, reflecting the working method of the rotkäppchen duo from two perspectives: the computer music performer's perspective and the dancer's perspective. This method could be described as a technology-mediated human-to-human interaction while mutual following of the collaborator's output, in particular: the dancer is listening to the sound or watching for the projection surface (or the beamer's rays change) and reacting with her movement sequences. In contrast, the computer music performer/visual artist is constantly watching the choreography and immediately reacting with sound or/and visual art as a response to the movements of a dancer.

In the collaboration process of the rotkäppchen duo, as well as in the current text, the auditory and visual aspects are equally important, and that's why this text contains 2 parts of terminology overview - auditory (part 2) and visual (part 3). Those parts include surveys of the relative literature, key terms, technologies and main approaches.

The part 4 proceeds with a general description of two main modes of the rotkäppchen duo's interaction, - human-to-human and human-computer (involving sensor technologies' mediation).

The following part 5 will be describing a previously mentioned artistic experiment - 2 realisations of the project „Floating pointers" (referred to as (A) and (B)) with 2 approaches - with and without the use of sensor technology. After the description and analysis of both approaches, a comparison and evaluation of them will be presented.

I will conclude with the perspectives of the collaboration and further planned artistic work.

The current text is not aimed at precisely describing and categorizing the sound objects (using some of the existing ontologies) and objects' affordances, which are formed through the movement interaction, as well as not to establish new propositions of classification for interactive sound objects, but to describe and evaluate the potential of fabricating the allusion of immaterial objects through dance movements and their technology-mediated sound

reorganizations, as well as to revise the framework for artistic collaboration in the duo rotkäppchen.

## II. TERMINOLOGY - AUDITORY

This part contains an overview of the most relevant research strands, approaching the different definitions of the central topics of this text. The “sound object”, “listening modes”, and the “soundscape” will be described from points of view relevant to my work.

### Sound object

The concept of „sound object“ is coined by P. Schaeffer in his fundamental „Treatise on Musical Objects“. Based on the object definition by Husserl, Schaeffer’s phenomenological approach to the „sound object“ is explored in connection to physics, acoustics, physiology and philosophy in theory and his practice.

He insists, that the sound object does not have a relation to its physical source or material, - it rather departs from a notion of a real or imagined causality, being approached through reduced listening, and merges its abstract and concrete meanings. Schaeffer mentions that the sound object can be studied not only as a totality, but also as a composition of its individual parts or, on the contrary, the sound object could be placed in a larger structural context, which could be perceived as an object in itself. This multidimensionality of approach inspired my artistic work described below.

M. Chion in his „Guide des objects sonores“ provides a commented summary of the „Treatise“. His interpretations of Schaefferian theories are less ambiguous and more distilled than the original ones.

Spectromorphology, as proposed by D. Smalley, redefines the notion of a sound object by Schaeffer. Based on embodied listening, the listener’s identification of the referential human activity, which is rooted in the listener’s gestural experience, is defined as sound source bonding. Here the sound object definition is placed in the gesture-texture parameter space.

R. Godøy also investigates the connection of sound and gestures not deriving them from each other, but rather arguing for a deep connection between morphologies of both gesture and sound.

My perception of the sound object is rooted in Schaefferian-Chion's description (as quoted at the beginning of this part) related to Smalley's embodied listening and incorporating the gestural mediation to fabricate a sound object (in contrast to Godøy). This sound object could be fabricated and represented in many different ways according to the listening intentions applied.

## Listening intentions

„Every object perceived through sound is only so because of our listening intention."[3] - says Schaeffer, intending, that there are multiple meanings, which could be evoked by the sound sensations, and the human intention is the main instrument for analysis of those. Combining the four listening functions (listening, perceiving, hearing, and comprehending), the different aspects of the perceived sound could be evoked.

"*I perceived (ouïr) what you said despite myself, although I did not listen (écouter) at the door, I didn't comprehend (comprendre) what I heard (entendre).*"[4] - says M.Chion, illustrating the four ways of listening.

The sound object, defined by Schaeffer as the correlate of reduced listening, is initially proposed and understood as a foundation of acousmatic reduction based on listening intentions, approached by removing any references beyond a sound per se.

Based on those four ways, Chion proposes its own taxonomy, called modes of listening, comprising causal, semantic and reduced listening.

W. Gaver uses a more generalised ecological approach to the listening modes' taxonomy, defining only two categories for listening to environmental sounds - everyday listening (focused mostly on events) and musical listening (focused on sound characteristics).

Another framework, proposed by D. Huron, explores listening from the emotional aspects and consists of six categories: reflexive (refers to fast

[3] Schaeffer, "Treatise on Musical Objects", p.272

[4] Chion, "Guide To Sound Objects", p.20

responses), denotative (allows to identify the sound source), connotative (allows to identify the source's properties), associative (refers to arbitrary associations), empathetic (allowing to detect emotion from the sound agent), critical (allowing to evaluate the intention of the sound).

K. Tuuri draws on the Hurons' taxonomy and proposes three categories instead: experiential (reflexive and connotative - focused on the sound source), denotative (focused on sound context), and reflective (reduced mode of Chion and critical category by Huron).

It is important to point out that several listening modes by Chion, Huron, Tuuri link the sound to action, which created this sound. Embodied music cognition, explored by Godøy, tends to see the perception of sounds based on actions.

R. Godøy in a series of research suggests, that there is always a gestural component in our minds while listening to the recording of musical sound. According to his theory, there is a constantly ongoing process in our minds of embodied tracing of sound parameters during the listening and imaging of music. That's how the gestural-sonorous images (with kinematic and motor components) are evoked in our minds, constrained by human biomechanics. Furthermore, this process can take an opposite direction - the gestural images could cause certain sonorous images in our minds.

B. Caramiaux et al. draw connections between the level of a sound source's identification and the gestures performed while listening to them, as well as link the modes of listening to the gestural strategies.

Creating the gestures based on the sound identification results and perceiving sounds as results of gestures, or tracing some of the unidentified sound parameters with a gestural component - all these approaches find their place in the „Floating pointers" project.

My approach to the listening intentions in the context of fabricating sound objects is based on an interactively controlled combination of Chion's and Tuuri's approaches (reduced listening with emotional aspect involved), sometimes referencing Gaver (everyday listening), but also including the medium of movements' interaction with the sonorous image. The way the dancer relates to the immaterial (sound) object allows the audience through

empathy to also relate to it, forming an allusion to a sound object, which the dancer perceives and gesturally interacts with.

The „Floating pointers“ involves a process of listening not only as a way to interact with the sound objects but also as a performative act - an intentional choreographed process. The goal of this process is to evaluate the impact of the previously performed movements on the sound.

## Soundscape

As previously mentioned, a unit and a structure are understood by Schaeffer-Chion as two points of view on the sound object. Therefore, structure, consisting of one or several sound objects, could be „unfolded“ into separate sound objects, forming the result into a zoomed-in micro-space. The resulting space consists of sound-„atoms“, perceived as objects, sound entities „stretched“ in space and a noise-based space in-between. This approach of zoomed-in analysis is considered to be central for the second part of the „Floating pointers“ project.

Speaking about the larger structural context, which could be perceived as an object, it is useful to refer to the theory of soundscape, introduced by R. Murray Schafer. Based on an ecological approach, an acoustic environment (real or imagined), filtered through human perception, is defined as a soundscape. Barry Truax in his "Soundscape Studies“[5] mentions the perceptual and cognitive primitives, which together with the physical parameters form the basis of the soundscape. The purpose of the soundscape composition genre, which grew out of the World Soundscape Project, is to evoke memories and associations in the listener, related to the combination of real and virtual soundscapes, which can be placed in the continuum between the „found sound“ and „abstracted“ approaches.

N. Barrett speaks about two categories of the reductionist approach - intrinsic and extrinsic, distinguishing the inner structure of the sound (spectrum, morphology, structural organisation - spectral evolution) and its spectral relations to something else (such as objects, symbols, gestural or spatial implications).

---

[5] Truax, "Soundscape Studies“, p. 37

T. Wishart, speaking of extra-musical landscapes and sound objects, separates its three main components:

(1) the nature of the perceived acoustic space;

(2) the disposition of sound objects within the space;

(3) the recognition of individual sound objects.

Transforming the recognition of the sound source or sound object because of the interplay of recognisability and lack of it - was one of the compositional methods, suggested by Wishart and artistically interpreted in my own work.

The purpose to evoke memories and associations was an initial idea for my own creative work - a motivation to use the environmental sound recordings as sources, oscillating the sound result between being recognisable by the recipients (more closely treated perceptually as a soundscape component) and not recognisable (leaning towards the sound object definition by P. Schaeffer).

## III. TERMINOLOGY - VISUAL

In the following chapter, I would like to describe the basis of the visual side of our performances - human-to-human and human-computer interaction with the related modalities and instruments.

During our performances, embodied listening is triggering choreographic interpretations, and the choreography changes the auditory part of them. The performer identifies the sound objects or creates an immaterial environment for their identification or further interaction. The latter could be related to the term „choreographic object“, with which W. Forsythe identifies the places, which are intended to evoke the choreographic movements. In „Floating pointers“ the inner structure of „unfolded“ sound objects could be treated as imaginative choreographic objects.

But how are those objects and environments created?

### Human-to-human interaction

Humans perceive the world through their senses (touch, sight, hearing, smell and taste; additionally one can mention spatial awareness and movement balance). Those senses are forming the base for sensory modalities. Sensory

modality is understood here as a way information from the world (light, sound, temperature, taste, pressure, smell, etc.) is transduced.

Any natural communication between humans is multimodal. That means, the different modalities are integrated into a mutual „flow“, while potentially influencing each other.

For example, while interacting with the person socially we don’t only listen to the words, but also watch the gestures and facial emotions, his/her body position in the space and its changes, we listen to the tone of the voice in order to understand his/her feelings.

To create a response to sensory perception, we use the body, hands, face and voice as actuators in the context of the environment.

Human-action modalities could be relatively easily interpreted in their combination by another human for interfacing further interactions.
Human-to-human interaction is understood here as a type of audiovisual face-to-face social interaction between two humans without mediating external technologies.

The benefits of this way of intercommunication are fast adaptation and fast reaction of the agents and variability of multiple combinations of sub-modalities.

The duo rotkäppchen uses the vision and hearing sensory modalities in the process of creating, evaluating and performing.

It is worth mentioning, that even with a wide range of accessible solutions for mediated interaction, human-to-human or face-to-face interactions are described as the most efficient and informational ones because they engage more human senses than mediated interaction.

## Human-computer interaction

The computer is not confined to the human senses. Although there are existing robotic sensors as representations of the basic human senses, computers are widely used for sensing other parameters, which humans can not, - for example, EEG.

Since human-computer interaction (HCI) involves humans as agents, the preferred natural multimodal interaction should be considered. In comparison to

human-to-human interaction (HHI), while interacting with a computer, in most cases one is limited to the modalities implemented in the current device.
The perception of computers is called machine perception. They use attached hardware and software to imitate human sensory systems and beyond them. Computers also can, for example, accurately estimate the position of the human hand in space and measure accurately the subtle changes in the electric activity in the human brain or muscles. The computer sensors are understood as interfaces to receive input from the real world and convert it into something understandable for the computer data flow.

The human-action modalities, used for HCI, are in most cases based on human movements of hands and body. The motion force, acceleration, orientation, and angular velocity could also be used for the input flow.

Wide-spread computer-sensing modalities for HCI (according to R. Sharma[6]) could be divided into the following categories:

1) position and motion sensing,
2) audio sensing,
3) visual sensing,
4) tactile and force sensing,
5) neural sensing.

rotkäppchen in the „Floating pointers“ (A) uses position and motion, visual and neural sensing via the selected gestural controllers - Leap Motion (position and motion - visual sensing) and Myo Armband (position and motion - neural sensing).

Human-computer interaction is understood here as an interface for open-ended communication between humans and computer technologies. The notion of dialogue likens HCI to HHI. The flow of information in HCI is referred to as the loop of interaction.

Wanderley and Depalle[7] mention, that for digital music instruments (in comparison to acoustic ones) there is no direct coupling between gesture energy and acoustic energy. That concept could also be extended beyond these instruments, and the relation between input and output data streams could be

---

[6] Sharma, „Toward Multimodal Human-Computer Interface“, p. 857

[7] Caramiaux et al, „Mapping through listening“, p. 34

set arbitrarily. In order to redefine this disassociated link, the gestural electronic music performance as a part of a wider domain of wearable interaction tries to bring the body back into electronic music.

While performing with gestural sensors, as well as with wearables, concepts of embodiment come into the light, recontextualizing motion patterns, which become conscious and adapted in order to accommodate the action space of the sensors, as well as the sound processes.

## Gestural controllers

The ability to track human movements and gestures could be achieved with different tools or controllers, such as:

1) Visual Based (single, stereo, depth cameras, VR-headset, etc.)
2) Glove or suit-based (wearables - wired, wireless, with IMU - inertial motion units)
3) Biosignal-based (EMG - electromyography).

These sensors use various algorithms, which allow tracking and interpreting the motion of any scale - neuronal, muscular, finger, facial, - up to the full-body motion.

rotkäppchen in their performances uses different controllers for obtaining the positional and dynamic data, including the Optitrack system (visual-based), the Leap Motion controller (visual-based), and the Myo controller (glove and biosignal based). The last two are used in the „Floating pointers“ (A), and that's why I would like to describe them more in detail.

Leap Motion [8] (now - Ultraleap) is a contactless sensor device for tracking hand and finger motions with a 0.7mm precision. It should be placed on a physical desktop or HMD for VR. Using IR cameras and LEDs, it generates 200 FPS of reflected data. That data is sent via USB to a computer, where the algorithm of the accompanying software synthesizes 3D position data by comparing 2D frames, captured by cameras. The small (less than 1m) hemispheric observation area is compensated by the high resolution of the device, which differentiates it from the Kinect depth camera. Kinect is more suitable for full-body tracking in a space sized as a living room.

[8] https://www.ultraleap.com/ . Accessed November 1, 2022.

Myo Armband [9] ( by Thalmic Labs, produced until 2019) is a wireless Bluetooth device aimed to track hand movements using EMG signals from muscles' tension by wearing a band. Initially conceptualised for the bio-medical industry and prosthetics, it contains 8 sensors and a 9-degrees IMU (3-axis accelerometer, gyroscope and magnetometer), and recognises gestures and arm movements. An accelerometer measures the energy of movements and spatial orientation, a gyroscope measures rotational speed, and a magnetic sensor is measuring the absolute orientation in relation to the external frame of reference.

## IV. INTERACTION MODELS OF THE DUO ROTKÄPPCHEN

In 2019 together with the Austrian dancer Lisa Mc Guire, I founded the duo rotkäppchen, which became an environment for developing multimedia compositions. In this collaboration, I program visual and sound components for different kinds of interactions while Lisa Mc Guire is creating choreography and performance. The duo rotkäppchen regularly works in close collaboration to develop pieces together where the auditory and visual media interact organically with the dance. Since 2019, a series of projects were realised, mainly involving real-time electronic music, dance and visual art.

Interactive systems, built for each of our projects, are allowing both improvisation and independence, while through the indirect performer's control, they are becoming interdependent. A sectional structure of the performances allows variable approaches to be applied to the control of interactive systems - they can be controlled directly (either by me or by Lisa) or indirectly (I control the system according to Lisa's cues or vice versa).

Interaction involves mutual influence, - which means both performers should be able to have equally distributed functions. As Ch. Dobrian says, both „must have the capability to respond to input that is not previously known to it . . . and must be capable of producing results that are not fully predictable“[10]. The input could be human-induced behaviour, performative cues, or a result of

---

[9] https://kinemic.com/de/band/kinemic-band-im-vergleich/thalmic-labs-myo-kinemic-band/ . Accessed November 1, 2022.

[10] Dobrian, “Aesthetic Considerations in the Use of ‘Virtual’ Music Instruments.” p. 30.

computer-controlled algorithms. Furthermore, as input could be treated the data, derived from the gestural controllers, which are interfacing with the performer's movements.

Working mostly within an HHI model, the duo also experimented with HCI mediation, which means allowing the data, derived from controllers, to influence certain aspects of the multimedia result. A general taxonomy of nonverbal cues for informational exchange between performers within the HHI or HCI model includes in most cases bodily activity and its expressive quality, subtle changes of the visual and auditory layers (which could be freely treated as cues), position within the performance space, light, etc.

rotkäppchen's interaction is working in several directions at the same time: with the help of mutual following, the dancer controls the sounds and the visuals by adapting to what she is hearing and partly seeing on the projection, and I, as the electronic music and visuals performer, am adapting to what I see from the side of movements and their combination with the projected visual layers. The projection itself is always directed at the dancer and behind her, making the real shadow and real body become other dimensions of the performance.

We habitually record our rehearsals to be able to analyse them afterwards from an outside point of view. Our performances are multilayered, loaded with several layers of meaning and open to many interpretations. We performed in different venues and always the audience claimed to be completely immersed in the multimedia experiences we created.

It has always been interesting for the audience to guess our interaction model: which media is now controlling and what is being controlled.

## Action - reaction in the context of co-following

rotkäppchen duo, as mentioned above, is using nonverbal cues for informational exchange during live performances. Those cues could be treated as triggers for certain human or computer reactions.

Since the duo uses mutual following during its performances (referred to as co-following in this text), I would like to describe the logical order of interaction, called action-reaction.

As an asynchronous series, action-reaction means that one cue is causing or triggering another. This sequential logic could be applied within the same medium (for example, only in dance) or cross-media (a cue in dance - the dance action is triggering an auditory reaction or vice versa). The time distance between the action (cause) and reaction (result) could vary from milliseconds to under 10 seconds. The longer the distance between action and reaction, the more the reaction tries to imitate the action's structure or/and expressiveness.

As a logic of simultaneity, action-reaction in our collaboration could be compared to a negative vector of the force pair (as referring to Newton's third law). This pair forms a single interaction, both forces are simultaneous and neither force exists without the other (like the book and table are pushing towards each other). As a result of co-following, the action-reaction simultaneous logic in the audio-visual domain is being always applied across media. This creates the needed connections, which the audience can follow. The moments of connections, as well as their frequency, are always improvised, after a vocabulary of non-verbal cues is formed.

In the context of HHI, I use my vision and hearing modalities to react musically or visually to the dance and check the cross-media integration. Lisa at the same time uses her hearing, vision and body movements to react to the output I produce (either while she sees the projection on her body or the projection surface, or while she watches towards the beamer, and sees the rays' changes).

Apart from several experiments (like „Floating pointers" (A)), where HCI is considered the central framework, we tend to limit the influence of computers on the creative output from our real-time performances.

Not allowing HCI to conquer the main space of artistic decisions, I understand the role of HCI mediation in most of our collaborations as technological means, that assists, shapes, and influences the processes of creation in the duo's collaboration.

This mediation during the performance happens in 3 ways: activating or deactivating the gestural control; changing or scaling the mapping presets; indirectly treating a gestural data output as an impulse for auditory reactions without assigning direct control to them.

## Technological incorporation

In order to establish interaction and obtain positional and dynamic data, rotkäppchen uses different controllers.

In the context of real-time motion tracking (as in the project „the room of the lost memories“) we use the Optitrack visual-based system for rigidbody tracking for direct interaction with the virtual 3D world. In this project, the dancer (via action-reaction as a sequential and simultaneous method) has control over triggering the sound events and interacting with the sound processing, mainly controlled by the computer.

In the context of non-realtime full-body motion capture (projects „Lebensformen“, „Before the first flight“) the visual-based tracking technology is used for performance digitalisation purposes. In virtual post-production, the captured movement sequences and the noise produced by the motion capture system, come into creative interplay with the custom avatar design.

Using game controllers (project „U92“) as instruments for audiovisual interaction, both performers - under the control of the global radioactive decay-time algorithm - improvise cross-media sequences in real-time.

The duo rotkäppchen is using gestural controllers (in „Lebensformen“, „Floating pointers“) for direct or indirect interactions, as an added value or substitute to human sensory modalities with measuring the position, rotation, acceleration and muscle tension. However, direct interaction using gestural controllers sometimes tends to limit the freedom of the dancer, imposing additional constraints, as will be further described.

In most projects, I am using a Korg nanoKONTROL2 MIDI controller as an instrument for audio-visual HHI (projects „Grow“, „Cross-product“, „a letter to Humboldt“) for both direct and indirect interaction.

# V. FLOATING POINTERS - STUDIES ON A SOUND OBJECT

A central project for the current text is „Floating pointers“ - a real-time composition for one dancer and 4-channel electronic music, controlled by a computer-music performer (duration: ca.10-11 min). The project currently has 2 realisations, which are questioning the presence of gestural controllers and the selected method of interaction.

In realisation #1 (A) movements of the dancer are unfolding and exploring the sound objects with the help of two gestural controllers. The computer music performer's role is limited to controlling the general audio output levels and triggering the mapping presets for the dancer's interaction.

In realisation #2 (B) human-to-human interaction is used, and no sensors for the dancer are involved. Dance is derived from the real-time strategies of sound processing, the sound objects are formed as a result of choreographic sequences and simultaneous action-reaction models. The computer music performer's function is extended towards the general sound processing parameters' control.

Both realisations have the same choreographic plan (Appendix #2) and use the same digital music instruments, which receive different interactions.

The idea of artistic interpretation of the „sound object" has gone through a long evolution. I have started in 2019 with the exploration of the sound object's meaning, relevant to the choreography and sound.

Starting with the physical modelling approach, I explored sound objects as grains with changing physical properties, revealed upon collision with the container they are in.

I proceeded with exploring the sound objects within the concept of "sounds' salad" - an interactive mixture of processed recordings, which one can recognise as objects from the real world. The initial sound sources should have remained unknown for the dancer - she should have made her own interpretations of the sound objects' sources.

The next step was the concept of acoustic photography, its unfolding and transforming the environment displayed on the imagined photography into the imagined performance space. The idea of collective memory and referencing to the Ukrainian collective cultural subconsciousness is for me very important. That's why I decided to experiment with the recording of the Ukrainian folk song - lullaby as a central sound source for creating imagined environments. Additionally, the recordings of ambiences were used.

The „Floating pointers" in the current state of sound object's idea evolution is centred around the idea of the fabrication of the immaterial object with the help of movements and processing of sound recordings. It consists of 2

big parts, surrounded by an intro and outro with the transition episode in the middle.

The first part is focused on exploring a sound object, which is stationary located. In the second part, the sound object's inner structure is zoomed in, and dispersed in the space. In the second part I was thinking about treating the sound object as a structure or a soundscape, a group of distinguishable sound objects in space. Additionally, with the means of sound spatialisation, I wanted to describe in the sound what the dancer „sees".

The intro, transition and outro of the performance are formed from the processed sounds of electrical devices (mobile phone, leap motion, myo), picked up by an electromagnetic microphone (the dancer is representing a robotic operator of the data centre).

## Concept and motivation

The „Floating pointers" project is aimed at exploring the sound object as fabricated by performance gestures. The exploration process via embodied listening is referred to from two different perspectives (outside and inside) with the help of HHI and HCI interaction models. In a more general sense, rotkäppchen's goal in this comparative study is to evaluate the aesthetical difference between human-to-human and human-computer interaction modes via the comparison of two versions of the same piece.

Using the concept of sound objects' unfolding into an immaterial environment, referred to as soundscape, the sound and movement media are being balanced between recognisability and non-recognisability as perceived by the dancer and audience. This balancing and variability of source recognition is defining the expressiveness of the piece.

The title "Floating pointers" is related movements-wise to the concept of data pointer existing in computer science. During the performance, the dancer is exploring the digitalised memory space or data centre. Floating memory pointers are further described as disconnected data bank slots, literally floating in space without gravitation.

## Sound object as a unit

The current and next paragraphs are describing the sound and performance aspects of the project and movement interaction with it referring to the flowchart (Appendix #1).

This flowchart is meant to create an overview of the sound object creation workflow in multilayered and multidirectional processing of sound recordings in Supercollider, which is controlled by mediated computer interaction (yellow parameters), independent computer scheduled processes (events' pattern), and variable writing to and reading from the same buffer containers.

The first part of the piece uses 4 field recordings as sound sources. In order to evoke the reduced listening, I decided to „blur" certain parameters of the sound sources spectra by working with various unit generators in Supercollider (UGens), mainly: by variable speed reading-writing to the audio buffers, FFT phase vocoder processing (bin wipe, freeze), pitch warping (granular pitch shifting, auto-tuning), time domain effects (delay, decay), filters (HPF, allpass filter), distortion (disintegrator, tanh), and dynamic effects (compressor, normalizer).

As is visible in the flowchart, there are various interaction entry points, which shape the sound production at each step of processing. That means, that the gestural part of the performance in combination with this instrument can fabricate a large variety of sound objects - complex enough to stimulate the reduced listening of the dancer (and then - of the audience), but be able under certain conditions to establish references to the unprocessed recordings.

The performative intentions of this part, according to the choreographic plan (Appendix #2), are centred around fabricating the sound objects through the gestural mediation, listening to the sound result, reacting, re-shaping or changing some aspects of the object interactively, skipping to the next one.

## Sound object as a structure

The second part of the piece, as is seen from the flowchart (Appendix #1), has fewer stages of sound processing involved.

The central aspect of sound in part 2 is working with the initial sound recording, which is selected interactively out of the same 4 recordings used in part 1. Further processing includes sound buffer writing and reading, granulation, FFT phase vocoder processing (random wipe, magnitude noise), time domain effects (echo), filter (non-linear), distortion (tanh), and dynamic effects (compressor). Afterwards, a certain stage of this sound processing is selected interactively. In the end, the sound is spatialised with the help of the Korg nanoKONTROL2.

The structure of the selected and zoomed-in sound object is explored by the dancer in relation to the stages of sound processing, mentioned above. In order to imitate the „blurred“ space in between the sound object grains or to imitate the destroyed sound objects, one extra sound source (brown noise) is used.

The performance side of this part is describing the process of moving through the dense medium of the inner structure of the sound object. The sound space and structural parts of it could be heard only upon interaction, upon movement. This movement should not only explore the medium and find the sound objects but also try to stop and concentrate on them while listening. During the moving sequences apart from sound processing, the selection and spatialisation procedures are performed. Here again, the search by listening and recognition should create imaginative choreographic objects and explore the found sound objects.

For the spatialisation strategy I have referred to the change of perspective for the sound object’s observation: in the introduction and first part the stereo perspective was chosen - 2 front audio channels, for the transition and the second part - 4 channels (placed in the corners of the audience’s listening space), the outro is composed for the rear stereo-pair - supporting spatially the gradual change of the axes in the movements’ domain: from left-right, operating always close to the ground, through the all-around space exploration towards the standing full-height and rising up in the very end.

## Technical realisation A - sensor-based HCI

As mentioned before, in „Floating pointers“ (A) HCI is experimentally given a central role as the sound-controlling system. That means, that interactions between gestures and controllers create sound objects within the controllers' affordances.

Still, independent computer scheduling has its place in this process, and the computer music performer establishes HCI mediation on a global level - controlling the level of general output and triggering sequentially each section of the piece's structure, changing the mapping of gestural controllers.

Initially thought to be performed inside of the Optitrack motion capture volume, because of practical reasons the performance space was scaled down to the small hemispheric interactive space for hand gestures, suggested by the Leap Motion controller (in the first part) and to the forearm's position and rotation changes, sensed with Myo armband (second part).

Leap Motion is chosen for the small-space detailed hands' movements aimed at sound objects' creation. Initially planned as a space for identification and studying gestures, because of the time constraints, imposed by the need for more detailed learning of the Leap Motion operation, this idea was postponed to the next realisation. For the first realisation of the project, the 3 degrees-of-freedom (DoF) position and acceleration were used. Those were more than enough for a dancer to become overwhelmed by the changes, which she can control, and those, which are controlled independently from her. Performance-wise the small interaction volume was placed on the ground, and the dancer was operating it with the hands from behind the back, which imposed other physical constraints on the controller's operation.

The Myo controller was chosen for imitating the space's manipulation by removing the fixed position of the volume and relating to its 9 DoF changes (3 DoF acceleration, 3 DoF gyroscope, as well as EMG, were assigned to the direct sound control). Attached to a forearm, the built-in accelerometers and gyroscopes are reacting to the position and rotation changes. An absence of change (fixed forearm's position in the space) is especially making sense in the context of finding the sound object. These changes and the absence of change came out to be easier for a dancer to learn in comparison to the Leap Motion

interaction. Incorporating forearm movements into her choreography allows interaction with the resulting space without thinking much about the controller's affordances. Furthermore, this allows the dancer not to concentrate on the arm and body relation as an instrument to create or find the sounds, but also to externalise the interaction space as an existing everywhere volume.

The Leap motion controller functions with the help of a devoted Leap motion standalone application, sending the visually sensed and interpreted data to GECO[11] - an open-source software for analysing, smoothing and outputting the OSC streams of the selected axes. These streams are received by Supercollider (a software and programming language for audio synthesis and algorithmic composition), where the strategies of sound mapping described below are further applied.

The Myo armband uses the „Myo connect"[12] standalone application, which connects the bio-sensing device via Bluetooth to the computer. Then as an OSC bridge for analysing and sending the data to Supercollider - the open-source Myo-OSC[13] C# script is used. In Supercollider controller-to-sound mappings are allocated.

The creation of sound objects for the „Floating pointers" (A) implies the creative interpretation of connections between the incoming sensing data and parameters of sound control, referred to as interactive system mapping. The strategies for those mappings provide an interface for HCI mediations in rotkäppchen performances.

The mapping strategy for the first part of „Floating pointers" (A) was concerning the preservation of the multidimensional interaction with the sound while using only 3 axes, understandable by a performer. The three-dimensional position and acceleration of one hand (independently from either it is left or right hand) were used to control many parameters at once, establishing the many-to-one mapping strategy.

For the second part of the project, in contrast, more axes were selected for mapping to match with expressive qualities of the main movements by the

---

[11] https://uwyn.com/geco/ . Accessed November 1, 2022.

[12] https://myo-connect.software.informer.com/0.9/ . Accessed November 1, 2022.

[13] https://github.com/samyk/myo-osc/ . Accessed November 1, 2022.

performer, still using a many-to-one strategy, but with fewer parameters controlled than in the first part of A.

The computer-music performer uses many-to-one mapping of the GUI to switch the controller's mapping presets and one-to-one strategy of the GUI sliders to control the main output sound levels.

## Technical realisation B - action-reaction HHI

As an experiment for comparing the interaction models, another realisation was created.

Because it was an experiment (that's different from the performance's preparation - less time to fine-tune the reactions and gestures, less time to select certain successful patterns), the mutual evaluation of the video recording was excluded. We have only made 2 video takes (full performance, and the first part again as a separate take) and just spoke in between about our own experiences of this experimental form.

For that, the initial interaction models were adapted. During the process of adaptation, the mapping strategy of many-to-one was rejected. The overall structure of the performance and its choreographic plan is preserved. A Korg nanoKONTROL2 was used as the only controller instead of Leap Motion, Myo and partly the Graphical User Interface (GUI). The GUI was preserved only in order to display the currently controlled parameters, mapped to specific ranges.

The sound processing algorithms in Supercollider were preserved, as well as the controlled ranges. In the first part of the adapted project, 8 Sliders and 2 knobs were assigned to the sound processing control, and the other 6 knobs were used for the amplitude control of the layers. Buttons were used for activation of the mapping presets (as in realisation A - there are 6 mapping presets for part 1 and 2 presets for part 2). Additional controller parameters were used for the filter controls in part 2.

The absence of the movements' noise and no need for lagging of the input values made the sound instrument to be more responsive. That has become a benefit for the sound objects in the first part, and a drawback for the second part. In particular, it became possible to show much smaller details of the sound processing in the first part, which opened more perspectives for the

human-to-human interaction model, and the second part, on the contrary, became more complicated to perform, because following the initial logic, without the controllers' movement, there is no sound preset.

## Comparison of the sound objects from Floating pointers (A) and (B)

To evaluate the aesthetical difference between human-to-human and human-computer interaction frameworks, a comparison between the two versions of the project should be made by analysing the video recordings from two rehearsals.

First I will analyze the differences in the sound component, and then I will proceed with the differences in dance improvisations, both of which equally contribute to the sound object's fabrication. Before the evaluation of frameworks, I will mention the technical experience gained while working on those projects and analyze my perspective of performing within the selected framework.

Comparing the sound components in the first part, the realisation A has more chirping-like sounds, voice-like ambience, clearly recognisable voices and other sound sources smoothly combined into the sound textures. The B-realisation includes more chopped-up granular sounds, sharper sound profiles, sound sources have less source-bonding related ambiguity, and much more rhythmical development of the structure is involved. The lullaby recording here is the most prominent material but is less stable in keeping its sonic state.

The second part of the performance in version A sound-wise reminds of the filtered wind with the grains of the dispersed song- and ambience-like processed recordings. The same part in the B realisation includes more chirping-like frequencies and timbres. The looped sequences of processed recordings sound in the B version is cleaner and more stable.

The performance in both versions of the first part contains the same gestural vocabulary. However, in version B the performed listening activity is more included, than in version A. Also in version B, one movement episode with the direct head-hands interaction is skipped. The performer is interacting by

detaching objects from the mass, and then observing and listening to the detached ones.

In the second part, there are almost no stops in the movement for the listening activity, and as a result, more continuous movement sequences are present. Scrubbing through the sound-mass movements in version B is not involving the use of the vertical dimension, the hand without the motion controller does not go up alone as a part of the choreographed sequence.

In the first part of A the dancer's movements are causing the sounds, affecting their processing, so the dancer decides the textural/structural music development with her interaction. She is aware of the volume limits, about the axes' dimensions and her position, which is tracked together with the movements' speed. She makes mistakes sometimes, which don't allow the sound to appear or change the way she expects. In the first part of B, the dancer's movements propose to me to react. If there is no sound reaction to the performer's action, the dancer in A repeats the sequence, and in B proceeds with her movement trajectories without repeating.

Analysing our rehearsal process for the „Floating pointers", I would like to mention several aspects. We have spent too much time with adapting the technical solutions and finding the limitations of the sensors (Myo is not operating under a certain temperature or the Bluetooth connection breaks when a certain part of the controller is being covered).

Choreography with the cables connected to the controllers, turned out to be very limited, while the dancer is thinking mostly about avoiding being trapped in cables. Located on the floor, the movement amplitude for part 1 (A) turned out to be dramatical for the sake of performance visibility, as a result - the proposed instrumental mapping was not explored much in detail.

Having not enough rehearsals with the spatialisation aspect turned out to be critical for the dancer's involvement in that. She confessed that she would like to work more just with the spatialisation concept.

Seeking for always the same straightforward sound changes, the dancer's ear still could not adapt to the sequential remapping strategy during the piece. Because of the random component involved, internal delays, even after several rehearsals aimed specifically to study the Leap motion while

performing the same interaction, the dancer was still feeling unsure which sound parameters are scheduled by the script, and which are the result of her direct interaction.

While in version B, I was performing with the Korg nanoKONTROL2 instead of the dancer's Leap Motion (first part), it appeared to me that with the help of Korg the richer sound-movements correspondences are created (even with the limited affordances of sliders-knobs operation), which lead to the idea to make the first part longer for the next performances. Using indirect action-reactions, the delayed body movements are contributing to the general flow as voices in imitation polyphony. The second part feels a bit poorer sound-wise because of the absence of the noise of the movement, which is creating the background sound textures in realisation A. However, the mapping strategy one-to-one turned it to be a complex instrument to learn for the usual rotkäppchen „blind" co-following performance, when I don't watch at the controller 99% of the performance time, so for realisation B only selective interaction was performed. Also, it's important to mention, that the dancer in the B version without having a direct impact on the sound was concentrated rather on her movements than sound operation.

Subjectively evaluating the aesthetic differences between the HCI and HHI frameworks in the two versions of the „Floating pointers" project, I would like to refer to the audience's perspective as a target. Fabricating the sound objects is an interactive process, which consists of gestural and auditory components, which produce different results depending on the framework selected. What is important for the audience from my point of view - is to be immersed in the experience and through empathy to relate to the immaterial object, to allow forming an allusion to a sound object, which rotkäppchen fabricates.

This immersion could be gained as a result of clear interaction and expressive qualities of its changes. Therefore, I find the HHI approach more persuasive for the musical dramaturgy of rotkäppchen projects. Furthermore, HHI opens more possibilities for establishing a variety of action-reaction (direct and indirect) interactions between sound objects and related movements.

Although the additional interactivity, which could be provided by the HCI mediation, gives a possibility for the performance to gain a more expressive quality of direct interactions, the artistic and technological adaptation of the HCI framework needs much more time and technical experience to become necessary component of the performance.

To sum up, within both approaches the persuasive sound objects are fabricated, and the ways for the HCI mediation require further investigation to define the productive ways of the meaningful controllers' incorporation.

## VI. CONCLUSION AND FURTHER WORK

In the current text, I have described and evaluated the potential methods of fabricating the allusion of immaterial objects through dance movements and their technology-mediated sound reorganizations, as well as revised the framework for collaboration in the duo rotkäppchen. I have shared the details of the subjective artistic process and my subjective adaptations of the relevant theoretical statements. Then I have described many aspects of two realisations of the same project as an experiment on evaluating the current artistic approaches and interaction models of the artistic practice of the rotkäppchen duo. Still, the question of the degree of HCI involvement in the process of sound objects' fabrication remains open for further versions of the „Floating pointers" project.

The next realisation of the „Floating pointers" is planned as a real-time multimedia project for one dancer, 8 channels of electronic music and video art. Further experimenting with the interactive co-creation is planned, such as combining both interaction models in the same piece. The concept of the virtual camera operator in the virtual world will be shaping the user's perspective by creating visual sequences, based on the performers' choices.

## VII. REFERENCES

Altavilla, Alessandro, Baptiste Caramiaux, and Atau Tanaka. 2013. *Towards Gestural Sonic Affordances.* Poster presented at 13th International Conference on New Interfaces for Musical Expression, KAIST, Daejeon, Korea, May 27-30, 2013.

Barrett, Natasha. 2002. *Spatio-musical composition strategies.* Organised Sound, 7, 313-323.

Caramiaux, Baptiste, Jules Françoise, Norbert Schnell, and Frédéric Bevilacqua. 2014. *Mapping through listening*. Computer Music Journal, 38(3), 34-48.

Chion, Michel. 2009. *Guide To Sound Objects.* Translated by John Dack, Christine North. Pierre Schaeffer and Musical Research.

Dobrian, Christopher. 2003. *Aesthetic Considerations in the Use of 'Virtual' Music Instruments.* SEAMUS Journal*,* 30.

Gaver, William W. 1993. *What in the world do we hear?: an ecological approach to auditory event perception.* Ecological Psychology, 5, 1-29.

Gaver, William W. 1993. *How do we hear in the world?: explorations in ecological acoustics.* Ecological Psychology, 5, 285-313.

Godøy, Rolf Inge. 2006. *Gestural–Sonorous Objects: Embodied Extensions of Schaeffer's Conceptual Apparatus.* Organised Sound 11(2), 149–157.

Godøy, Rolf Inge, Egil Haga, and Alexander Refsum Jensenius. 2006. *"Exploring Music-Related Gestures by Sound-Tracing: A Preliminary Study."* Proceedings of the International Symposium on Gesture Interfaces for Multimedia Systems, 27–33.

Huron, David. 2002. *A Six-Component Theory of Auditory-Evoked Emotion.* In: Proceedings of the International Conference on Music Perception and Cognition, 673-676.

Miranda, Eduardo Reck, and Marcelo M. Wanderley. 2006. *New Digital Musical Instruments: Control and Interaction beyond the Keyboard.* Middleton, Wisconsin: A-R Editions.

Neri, Louise, and Eva Respini. 2018. *William Forsythe: Choreographic Objects.* Prestel Publishing.

Schaeffer, Pierre. 2017. *Treatise on Musical Objects: An Essay across Disciplines.* Translated by John Dack, Christine North. University of California Press.

Schafer, R. Murray. 1993. *The Soundscape: Our Sonic Environment and the Tuning of the World.* Destiny Books.

Sharma, Rajeev, and Vladimir Pavlovic. 1998. *Toward Multimodal Human-Computer Interface.* Proceedings of the IEEE.

Smalley, Denis. 1997. *Spectromorphology: explaining sound-shapes*. Organised Sound 2(2). 107–26. Cambridge University Press. Printed in the United Kingdom.

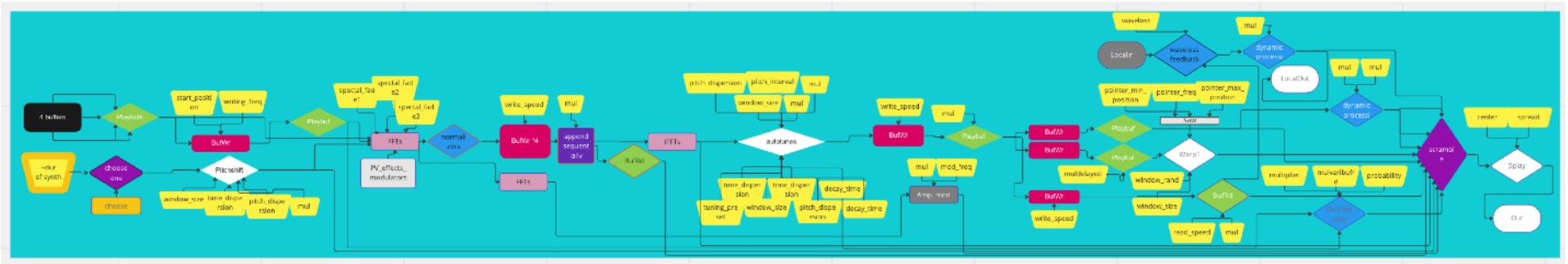
Toenjes, John. 2007. *Composing for interactive dance: Paradigms for perception*. Perspectives of new music 45/2. United States. 28-50.

Truax, Barry. 1974. *Soundscape Studies: An Introduction to the World Soundscape Project.* Numus West, no. 5.

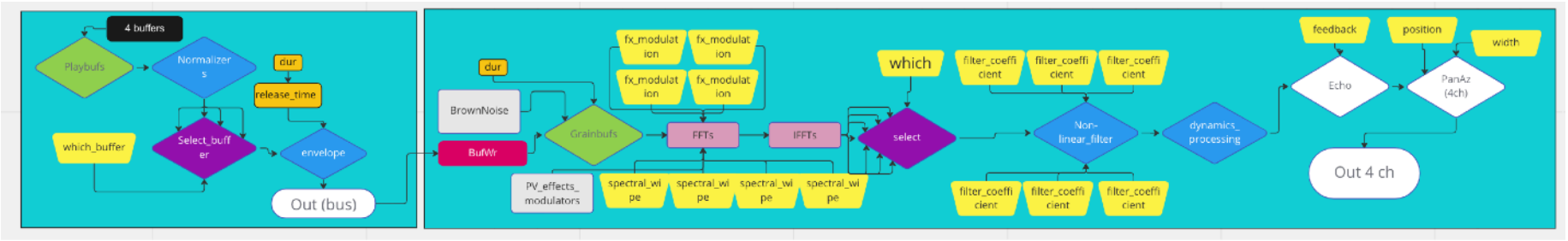
Tuuri, Kai, and Tuomas Eerola. 2012. *Formulating a Revised Taxonomy for Modes of Listening.* Journal of New Music Research 41(2).137–152.

Wishart, Trevor. 1985. *On Sonic Art*. York: Imagineering Press.

## Part 1



## Part 2




Colour-coding:

black - initial recordings
yellow - parameters for controllers
orange - parameters, controlled by a pattern
yellow with a thick orange border - in (A) it was a pattern controlled value, in (B) - value, controlled by a midi controller
light-pink - FFT processing
dark pink - Buffer Write
light green - reading of the buffer (Playbuf, BufRd, GrainBuf )
blue - dynamics processing (compressor, expander, leakDC, normalizer) and filtering; also amplitude envelope and mixing.
grey - additional UGen audio or control source
violet - selecting one from the list (also crossfades)

# IX. APPENDIX 2 (PERFORMANCE-ORIENTED NARRATIVE PLAN)

The performer takes a role of a highly-responsible and motivated operator of the data centre, which preserves information on digitised cultural artefacts. The operator works alone. In the first part, she is exploring and categorising the incoming information. The operator does not have a personal attachment to this information, but she is attached to this work. Operator checks which information is coming in by unfolding it. At some point in time (transition between the first and second part) the operator switches the laboratory space, and as a result the gravitation changes, the atmosphere gets a higher density,  and it's becoming hard to navigate normally for her - like in water, where an extra resistance of the medium is present. The connection cables from the storage systems get plugged out, the network connection disappears and the operator can not find the required information. During the second part, the performer is trying to bring storage systems back to functioning by recovering connections with hand movements. Because of the change in the medium, digital information is leaking from the servers and „floating" in space. The formed by the leakage „clouds" of information partly mix up, like the paint in the water, but not completely diffuse, allowing the „substances" or memory-leaks to keep their granular form. The performer explores the space around her by walking in order to reveal/catch/ collect the leakages she finds. To collect the leakages of different densities she uses different speeds and axes of hand movement.