NYU Steinhardt
Steinhardt School of Culture, Education, and Human Development
Operational & Curatorial Research
LEONARDO®
THE INTERNATIONAL SOCIETY FOR THE ARTS, SCIENCES AND TECHNOLOGY
Sabancı Üniversitesi
Goldsmiths
UNIVERSITY OF LONDON

LEA
LEONARDO ELECTRONIC ALMANAC

VOL 19 NO 3 VOLUME EDITORS **LANFRANCO ACETI, STEVE GIBSON & STEFAN MÜLLER ARISONA** EDITOR **ÖZDEN ŞAHİN**

Live visuals have become a pervasive component of our contemporary lives; either as visible interfaces that re-connect citizens and buildings overlaying new contextual meaning or as invisible ubiquitous narratives that are discovered through interactive actions and mediating screens. The contemporary re-design of the environment we live in is in terms of visuals and visualizations, software interfaces and new modes of engagement and consumption. This LEA volume presents a series of seminal papers in the field, offering the reader a new perspective on the future role of Live Visuals.

LIVE VISUALS

by

Adriana Sa

adrianasa@gold.ac.uk
EAVI Research Group, Computing Dep., Goldsmiths, University of London.
http://adrianasa.planetaclix.pt

How An Audio-Visual Instrument Can Foster The Sonic Experience

Photo 1. *Windowmatter*, performance, Adriana Sa, 2012. Sho-Zyg ("Immerge," sho-zyg.com), St. James's Hatcham Church, London, September 2012. Photo © James Bulley 2012. Used with permission.

A PERCEPTUAL APPROACH

The article asks whether, and how, an audio-visual instrument can be composed in such a way that perception is primarily driven through sound organization, rather than vision. A perceptual approach to the question also exposes a scientific reasoning for artistic perspectives that seem in contradiction. Does the visual distract from the aural? Conversely, does the aural distract from the visual? Or is the presence of each, and the synchronization between, of benefit

ABSTRACT

The text formulates an understanding of how an audio-visual instrument can be composed in such a way that the experience is driven through sound organization – modulated, but not obfuscated, by a moving image. This is particularly challenging, as normally the audio-visual relationship is skewed in favor of the visual. The investigation is motivated by insights derived from artistic practice. It outlines psychophysical boundaries with the aid of existing cognition/ attention research, and it describes three principles for the creation audio-visual instruments. As an example, the article describes how they are explored in a specific audio-visual instrument, combining an acoustic zither and modified software from audio processing and video-game technologies. This instrument addresses the three principles while exploring the disparities between an acoustic and a digital output.

to both? The article will ultimately discuss and analyze how we prioritize the sensory input – how a part of the sensory information becomes conscious, and another does not.

Audio-visual experience is not simply the sum of auditory experience and visual experience. Pierre Schaeffer argued that sounds must be detached from their originating causes to be fully experienced. He wrote: "often surprised, often uncertain, we discover that much of what we thought we were hearing, was in reality only seen, and explained, by the context." [1] The act of hearing without seeing the originating cause of sounds would enable us to focus on the traits of the sound itself, independent of its cause and of its meaning. [2] To the present, acousmatic composers such as Francisco Lopez perform in visual darkness to potentiate a full sonic experience. Lopez expressively states that "the combination of visual darkness and being 'inside' the sound (...) creates a strong feeling of immersion where your own body moves into the perceptive background." [3]

Conversely, in his many silent films Stan Brakhage aspired at highlighting the musical quality of the images and their montage, without the distraction imposed by sounds. In one of his texts he recalls a certain night of his childhood: "I was in an environment silent enough to permit me to hear 'the music of the spheres,' as it's called, and visually specific enough for me to be aware of the eye's pulse of receiving image." [4] According to his wife Marilyn Brakhage, he felt that sonic music tends to dominate over the more subtle rhythms of vision. [5] In his few later audio-visual movies, the primary principle was one of non-synchronization, of breaking any direct connection between pictures and sounds.

Meghan Stevens assessed that music is pushed to the background when the audio-visual relationship is fully congruent or incongruent. [6] The only way to highlight the music would be with partial congruency. But ultimately, congruency is subjective with abstracting sounds and images, and the role of synchrony is debatable. Stevens drew insights from Michel Chion. [7] This filmmaker and composer wrote: "For the spectator, it is not the acoustical realism so much as synchrony above all, and secondarily the fac-

ISSN 1071-4391 ISBN 978-1-906897-22-2 VOL 19 NO 3 LEONARDOELECTRONICALMANAC 285

tor of verisimilitude (verisimilitude arising not from truth but from convention), that will lead him or her to connect a sound with an event or detail." [8] Yet a scientific experiment conducted by the psychologists M. Kubovy and M. Schutz led to the conclusion that at low level in information processing perceptual binding prioritizes verisimilitude over synchrony. [9] The experiment examined how a visual stimulus may effect the perceived duration of a sound. It demonstrated that this only occurs when both the auditory and the visual stimulus seem to come from the same source, and that synchronization in itself is not bound to alter how we perceive sounds.

Synchronization in itself does not obfuscate the sonic experience. It does not necessarily blind us to the specific qualities of sounds and their organization. Other factors do. In pursuing the question of how the audio-visual can favor the aural, the present investigation looks at perception, attention and intensity, binding scientific research with insights derived from artistic practice. It arrives at three principles for the creation of audio-visual instruments:

» To threshold control and unpredictability for sonic complexity and expression – the instrument is here not solely a controlled prolongation of the body, but also a means of destabilization.
» To facilitate the perceptual simplification of the moving image by applying Gestalt principles and avoiding sudden visual changes, which would automatically attract attention – minimizing the demand for visual information processing maximizes the sonic information in conscious awareness.
» To establish fungible correspondences between the sounds and the visual events – such that the audience senses an overall cause-effect relation but quits trying to understand the instrument and focuses on the perceptual experience.

A monophonic moving image fosters the auditory experience, punctuating the sonic construction while conveying a sense of total environment. Conscious awareness is invited to draw upon the complex and often surreptitious organization of sounds. The proposed three principles result from clarifying a number of artistic insights with the aid of existing research in psychology and neuroscience. They are a set of conditions, and those conditions mark out a territory that can be explored in many different ways. To facilitate a diversity of explorations, the article provides: a summary of correspondences between gestaltist principles in audition and vision; an appropriate notion of intensity within the audio-visual context; and a taxonomy relating continuities and discontinuities with intensity and attention. The applicability of these systems is also exemplified with a detailed definition of a personal audio-visual instrument.

The principles are meant to inform design strategies, and should be considered in conjunction with individual artistic idiosyncrasies. The goal then is to manage these principles in each creative exploration. My personal work has been developed with an acoustic multi-string instrument (a zither) and AG#1, a modification of software originally written by media artist and video-game programmer John Klima (2007). In this exploration, we have been careful to distinguish between potentially different notions of 'play' in music and gaming. A detailed comparison exceeds the scope of this article, but two divergences are directly related. Rather than endeavoring to maximize a transparent sense of human-computer interaction, I seek to take advantage of, and to even exacerbate the problematic issues of human-machine agency. Considering the three principles mentioned, certain disparities between the physical input and the digital output are especially suitable to explore sonic complexity and unpredictability, as well as to create fungible correspondences between the sounds and the visual events.

 ISSN 1071-4391 ISBN 978-1-906897-22-2

Moreover, rather than designing the 3D world and the dynamic navigation through it for for a 'gameful' performance, I design them to emphasize the intrinsic cohesion of sound organization.

The article is structured in three parts: foundations, extrapolations, and explorations. Part one exposes existing research on perception, multi-sensory integration, attention and intensity. Part two extrapolates from the previous, arriving at three principles for the creation of audio-visual instruments. Part 3 describes how these principles manifest in a specific instrument that articulates acoustic sound, digital sound and digital image.

PART 1 – FOUNDATIONS

Perception, structure and complexity

Perception is a process of multi-sensorial synthesis. The same principles govern across multiple sensory modalities, all being mutually influential despite our possible unawareness of such interactions. [10] The primary aim of the brain is to use the information derived from the various senses in order to detect, perceive and respond appropriately to objects and events.

From the early 20th century to the present, gestaltist psychologists have described how we organize the perceptual field in the simplest and clearest way possible, deriving the meaning of the parts from the meaning of the whole. The principles of this perceptual organization have been studied in the aural and the visual domains. Table 1 summarizes them in their most basic form. Research on their subtleties could proceed with A.S. Bregman, [11] J. Tenney, [12] B. Snyder, [13] and F. Lerdahl & R. Jackendoff [14] in the aural domain; M. Wertheimer, [15] E. Rubin, [16] K. Koffka [17] and S. Palmer [18] in the visual domain.

At any time, we are presented with a massive amount of stimuli. According to many researchers [19] [20] incoming information transits from *sensory memory* into *short-term memory,* and subsequently to *long-term memory* from where it is constantly retrieved. Whilst long-term memory indefinitely stores a seemingly unlimited amount of information, the rapid decay of short-term memory submits the stimuli to strong competition. The term *working memory* refers to the structures and processes used for temporarily storing and manipulating information. Because perception chunks the information, working memory can handle large amounts of information simultaneously, through cues.

Bob Snyder (composer and cognitive researcher) describes how the information is chunked continuously throughout information processing, from low-level stages (e.g. a sound includes a multitude of sound frequencies) to high-level stages (e.g. identifying a sound involves retrieving memories). According to him, "the focus of conscious awareness could be thought of as being at the 'front edge' of short-term memory." [21] This 'front edge' has an even smaller capacity than short-term memory – Snyder states that it holds three pieces of information at the most. Nevertheless, each of these information pieces embeds a hierarchy of semi-conscious and unconscious information.

Snyder describes how categorizing the sensory information enables us to operate based on assumptions. Events activate parts of long-term memory that have been previously activated by similar events. Snyder named these long-term memories as *conceptual categories*. Among these long-term memories, some become highly activated and conscious, whilst others – named *semiactivated* memories – remain unconscious, forming expectations. When new situations counteract these expectations, they require new combinations of cognitive processing. Since this

ISSN 1071-4391 ISBN 978-1-906897-22-2

Phenomenon	Visual domain	Aural domain
To discern cohesiveness in a changeable form (Invariance)	A visual object is recognized independently of rotation, translation, scale, elastic deformations, lighting and component features	An auditory stream is perceived as a unit in spite of its changes over time (stream fusion)
To segregate a form from the continuum (Figure/ ground)	Visual figures are perceived as separate from their background	Discerning a sound implies segregating it from the soundscape
To group closely located elements (Proximity)	Visual elements that are closer together are perceived as a coherent object	Sounds are grouped as a single event when adequately proximal in time (sequential integration)
To group similar elements (Similarity)	Visual elements that look similar are grouped as part of the same form	Simultaneous sounds with different spectral content are grouped when holding the same fundamental tone (harmonic grouping), or when their frequency components hold similar onset time (frequency onset grouping)
To group elements that change or move together (Common Fate)	In moving image: elements that change or move together are grouped as a unit. In static image: elements with the same orientation are grouped as a unit	Sounds that change together are grouped as a unit (similarity of temporal evolution grouping)
To bias the grouping of elements that follow a consistent, lawful direction (Good Continuation)	Visual shapes are perceived to relate inextricably when a parameter changes progressively (e.g. size, brightness, location)	Sounds are perceived to hold immediate relation when a parameter changes progressively (e.g. pitch, loudness, interval)
To enclose a form despite any existing gaps (Closure)	A visual space is enclosed by completing a contour and ignoring any gaps in the figure. Furthermore, lines and colors are tied into shapes, relating at higher hierarchical level	A linear sequence of sounds, e.g. a note scale or a sound repeating at equal intervals, is still perceived as a unit if a note or interval diverges. Furthermore, musical phrases are tied together, relating at higher hierarchical level

Table 1. Gestalt principles of perceptual organization.

cognitive processing depends greatly on attention, attention is constantly being drawn to the novel aspects of situations.

Novel aspects of situations attract attention, but only if they resist perceptual simplification. In fact, as put by the media researcher Herbert Zettl, "(o)ur mental operating system encourages a considerable perceptual laziness that shields us from input overload. (...) We often see and hear only those details of an experience that fit our prejudicial image of what the event should be and ignore the ones that interfere with that image." [22] However, the term perceptual laziness should probably be put in quotes. Our mental operating system can be extremely active while simplifying the sensory information. There is no structured rational thought without abstraction, and there is no abstraction without perceptual simplification. Furthermore, one can consciously experience sensorial complexity while the brain seeks to form abstractions of that same complexity.

We can enjoy the process in music. Snyder describes how music can be playful with our psychological tendency to complete a shape despite any existing gaps. [23] This tendency is the gestalt of closure. Closure gives us the sense of cohesion in a musical shape (which in turn depends greatly on internalized musical traditions). As we experience complexity and discontinuity, closure ties musical events together, relating them at a higher level of information processing. Partial closure occurs when some parameters of musical dimension fulfill expectations, while others do not. These incompletely closed musical phrases create expectations of eventual closure. Any more completely closed musical phrase appearing subsequently will then not only close itself, but also close the less thoroughly closed musical phrases that appeared prior to it.

Jeff Pressing (composer and scientist) distinguished several types of complexity in order to define the structure of musical patterns. [24] *Hierarchical complexity* refers to the existence of a structure across many levels. Snyder's description of how musical events are tied at a higher hierarchical level would possibly fit here. *Dynamic* or *adaptive complexity* refers to a rich range of behaviors over time, or an adaptation to unpredictable conditions, or a monitoring of results in relation to a reference source, or an anticipation of changes in oneself or the environment. *Information-based complexity* involves a target problem that requires a solution; since solutions are seldom exact, one actually seeks for an approximate solution, with a certain amount of tolerance.

Intensity and musical motion

The composer James Tenney asked which "are the factors leading to the discovery of continuity" in music. [25] He proposed that the cohesion and segregation of sounds is primarily determined by the gestalts of proximity and similarity, i.e. the perceptual grouping of closely located and similar information. Intensity would be among the secondary factors.

Tenney defined intensity as "the tendency of an accented sound to be heard as the beginning of the [perceptual] grouping. The relative intensities of several concurrent elements in a clang (or several monophonic sequences in a polyphonic sequence) are also a determinant of textural focus." [26] Tenney spoke of "musical or subjective intensity," assessing that one would possibly never be in a position to describe the factor of intensity in a satisfactory way. Generally, intensity could be associated with increases in loudness, pitch, harmonic dissonance, tonal brightness, speed or temporal density.

Snyder adopted Tenney's notion of intensity [27] – a physical property of stimuli, related with preceding, simultaneous and subsequent stimuli. Snyder summarized this notion in a very interesting way: "any change in a stimulus that causes an increase in neural activity." [28] He described musical motion as a continuous oscillation between points of high intensity (corresponding to motion or tension) and points of low intensity (corresponding to rest or release). Points of high intensity indicate instability: we expect the current musical structures to move toward a resolution. Conversely, points of low intensity indicate stability: musical structures are resolved when a sense of arriving at a 'goal' has been achieved.

Not only sounds but also moving images are frequently described in terms of tension and release dynamics. Correspondences have been established between

ISSN 1071-4391 ISBN 978-1-906897-22-2
ISSN 1071-4391 ISBN 978-1-906897-22-2

speed of sound and image, timbre and color, amplitude and luminosity, [29] but these topics fall out of this article's scope.

Multisensory integration, attention and causation

The question here is how one can draw upon complex sonic constructions in audio-visual performance. The problem is the stimuli competition to access conscious awareness during "the forging of an immediate and necessary relationship between something one sees and something one hears." [30] The term multisensory integration refers to the set of processes by which information arriving from the individual senses interacts and influences processing in other sensory modalities, conveying a unified experience of multisensory events.

While we process the sensory information, divergences across the sensory modalities can produce phenomena known as multisensory illusions. For example, the *ventriloquist effect* [31] is a phenomenon in which a sound is perceived to occur at or towards the location of a spatially disparate visual stimulus that occurs at the same time. The *sound-induced double-flash illusion* [32] refers to how a single flash of light, presented concurrently with a train of various short tone pips, is perceived as two or more flashes. The *McGurk effect* [33] occurs when speech sounds do not match the sight of simultaneously seen lip movements – it leads to a perception of a phoneme that is different from both the auditory and visual inputs. Furthermore, the perceived duration of auditory events can be shortened or lengthened by conflicting visual information. [34] [35]

Researchers in psychology proposed that visual stimuli tend to dominate in the processing of spatial characteristics (as with the ventriloquist effect) and that auditory stimuli tend to dominate in the processing of temporal characteristics (as with the sound-induced double-flash illusion). [36] Michel Chion also stated that the aural supports the visual: "the eye perceives more slowly [than the ear] because it has more to do all at once; it must explore in space as well as following along in time (...) Why don't the myriad rapid visual movements in kung fu or special effect movies create a confusing impression? The answer is that they are "spotted" by rapid auditory punctuation." [37] "What we hear is what we haven't had time to see." [38]

Audition may dominate in the processing of temporal characteristics, and be crucial to visual information processing. Yet, the timings and intervals of a sonic construction may not be experienced when that sonic construction is coupled with moving images. In other words, the visual tends to obfuscate the aural. Timings and intervals are formed through relations between sonic events, and many of these events (the most subtle ones in particular) may not reach conscious awareness.

In many scientific experiments, when participants were presented with audio-visual stimuli in a speeded discrimination task, they failed to respond to the auditory component significantly more often than to the visual. The phenomenon is known as *Colavita visual dominance effect.* [39] Perception seems to delay the aural when the audio-visual linkage is verisimilar. Pressing observed that, in digital 3D environments, "where a direction of causation is discernible, it operates from visual event to sound event and not the reverse. Nevertheless, sometimes the sound for a new cinematic scene can precede the visual image." [40] Possibly due to this sonic 'delay,' 3D animators often place the sound of a footstep slightly before the foot actually hits the ground.

The visual tends to dominate in conscious awareness, but recent scientific research confirms that it can also subordinate to the aural: "The visual channel is sampled before, or possibly more frequently than the auditory channel. However, because this difference in sampling rate (or bias) may be attentional in nature, it can be manipulated by focusing attention on one sensory modality or another." [41] This scientific conclusion is a steppingstone to the present investigation.

Attention dynamics are hence of major importance here. Attention is *exogenous* (or bottom-up) when automatically driven through stimuli, and *endogenous* (or top-down) when more under individual control. [42] Attention is typically drawn to events that are infrequent in time or in space – e.g. to stimuli that appear or disappear in a sudden manner. These events are of high biological relevance. They cause the nervous system to respond strongly and automatically. Conversely, attention is under greater individual control when there are no salient changes.

Attention determines whether an event possesses greater or lesser resolution in the perceptual flux. A researcher in the field explains: "At any point in time, the information that gains access to working memory is selected by a competitive process (...) Signal strength reflects the combined effects of the quality of the encoded information, top-down bias signals, and bottom-up salience filters." [43] Attention causes working memory to optimize the resolution of the information concerning its target. This optimization occurs when the sensory organs are directed toward the target, and/ or when the sensitivity of neural circuits is modulated accordingly. Working memory then improves the quality of related information processing in all domains: sensory, motor, internal state, and memory.

An important question here is whether automatic attention and deliberate attention have equal impact in multisensory integration. Endogenous (deliberate) attention has an influence when there is considerable competition between inputs to different modalities, and the attention target has to be selected deliberately. [44] Scientific experiments occur in very controlled environments. Thus one should be careful with extending certain conclusions to situations where many diverging stimuli are competing to reach conscious awareness.

That been said, a number of experiments demonstrated how automatic attention influences multisensory integration, and deliberate attention does not. For example, the ventriloquist effect has been shown to largely reflect automatic sensory interactions, with little or no influence from deliberate attention. [45] Similarly, Van der Burg et al conducted experiences showing that, given a sound and a specific target in a visual field, the target only 'pops out' if changes in the sound and image signals are both synchronized and abrupt. [46] In other words, if sound and image prompt automatic attention simultaneously.

Regardless of whether attention is automatic or deliberate, multisensory interactions are greatly driven by the primary aim of the brain. We need to make sense of the environment in order to survive. The interactions themselves consist of a reciprocal mapping between auditory and visual information. M. Kubovy and M. Schutz speak of the product of such interactions as *audiovisual objects*, stating: "even though we perceive the world and not percepts, we cannot dispense with mind-dependent concepts, and indeed entities." [47]

Kubovy and Schutz conducted an experiment showing that, above all, the "[perceptual] binding depends on the ecological fit between visible events and sounds. For example, the sound of a marimba binds with the visible impact because such an impact could have produced such a percussive sound." [48] The visible impact altered the perceived duration of the marimba sound, even when that sound was delayed (up to 700 ms). The seen moment of impact was also simultane-

 ISSN 1071-4391 ISBN 978-1-906897-22-2

ISSN 1071-4391 ISBN 978-1-906897-22-2

ous with the onset of non-percussive sounds, which kept the same perceived duration. [49] The experiment showed that multisensory interactions were not just due to synchrony. Synchrony may cause us to bind aural and visual, but it is not bound to skew one sense in favor of the other.

PART 2 – EXTRAPOLATIONS

Perception is conditioned through multisensory integration, which in turn is influenced by attention. In many aspects, vision tends to dominate over audition, but attention can be manipulated so that audition dominates over vision. One can extrapolate requirements for one sense to dominate over another, by considering how attention is driven through continuities and discontinuities, and how it simultaneously influences how these continuities and discontinuities are perceived. If an audio-visual instrument addresses these requirements specifically for the aural, it can highlight the subtle complexities of sound organization in the audience's experience.

(Re)defining intensity

How can we define intensity in cross-sensorial terms? Snyder's expression is most useful here – intensity is any change in the stimuli chain causing an increase in neural activity. His description of musical motion is appropriate – it operates through an oscillation between high intensity (tension) and low intensity (release). However, Snyder speaks of intensity as related to qualities that are specific to sounds.

The question of how the aural can dominate over the visual should consider the impact of stimuli, rather than the stimuli themselves. In truth, a stimulus is likely to be acknowledged (and potentially cause an increase in neural activity) when contrasting with panorama; yet the same stimulus may be overlooked when competing with other equally bold, or bolder, stimuli. For example, a sound may stand out in a quiet environment, and disappear in a loud environment. The color red stands out in a black-and-white environment, but not in a yellow-red environment. Moreover, a salient, unusual event is necessarily intense as it automatically attracts attention, but a minute variation on a common event can be equally intense if attention is being paid to it, and working memory optimizes its resolution.

To begin with, intensity must refer to events rather than to any physical properties of sounds or images, and these events must be considered within the sensorial context. For example, both the abrupt appearance of a stimulus, or its abrupt disappearance, can make for an event to be highly intense. Moreover, an appropriate notion of intensity must also comprise the perceiving subject: attention and multisensory integration are determinant to whether the chain of stimuli causes an increase or a decrease in neural activity.

The same event can be highly intense when under focus in conscious awareness, and not intense at all when ignored. The intrinsic qualities of sounds and images – e.g. loudness, pitch, color, timbre, dissonance, or brightness – must be prioritized in multisensory integration in order to be consciously grasped. The process depends greatly on attention.

It has been previously mentioned that the competition between stimuli to access conscious awareness evaluates 'signal strengths' derived from the combined effects of deliberate attention, automatic attention, and the quality of encoded information. [50] I propose that intensity derives from this comparison. From this angle, intensity is a psychophysical property that is directly proportional to the impact of a change. Thus, it depends on the event itself, on cross-sensorial interactions with other events, and on the current perceptual resolution, which depends on the current focus of attention (Figure 1).

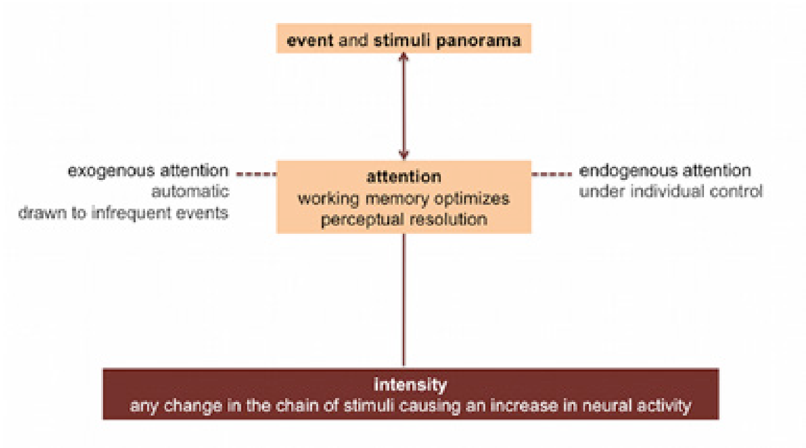

Figure 1. The intensity of an event depends on the event itself, on cross-sensorial interactions with other events, and on individual attention. © Adriana Sa, 2012. Used with permission.

Considering the inextricable relation between intensity and attention garners a useful perspective over musical motion, be it sonic, visual, or audio-visual. Musical motion thus appears through the interplay of discontinuities and continuities, as it is this interplay conducting the interplay of exogenous and endogenous attention.

Continuities and discontinuities

Whether musical motion is sonic or audio-visual, the interplay of continuities and discontinuities fundamentally directs attention. The intensity of each event depends greatly upon its relation with other events, as well as upon the current state of the person's attention. For example, a certain event can be more intense when preceded by a moment of low intensity, i.e. of rest. Moreover, events can increase or decrease in intensity due to changes in the perceptual resolution, rather than to actual changes in the chain of sounds or images. For example, continuities can become discontinuities – increasing in intensity – and discontinuities can become continuities – decreasing in intensity.

The complex dramaturgy of musical motion cannot be described to its fullest extent, but it is useful to systemize how attention relates to different modes of continuity and discontinuity. The following terminology proposes a distinction between a sense of continuity that is primarily driven from stimuli (exogenous), and a sense of continuity that depends more on the individual (endogenous). It also proposes a distinction between radical discontinuities, which impose disruption, and ambivalent discontinuities, whose acknowledgement depends on the perceptual resolution.

Exogenous continuity is of low intensity. It occurs when there is no significant change from one event to the following: when subsequent events seem unequivocally related. A constantly changing emission can also exhibit exogenous continuity, as change becomes the norm. Exogenous continuity can be steady or progressive. *Steady continuity* has no intrinsic motion: it is of lowest intensity. Being steady and continuous, it dispenses with attention. Conscious awareness is likely to focus on any simultaneously fluctuating chain of information, or alternately, upon internal states. *Progressive continuity* occurs when successive events display a similar interval of motion. It fulfills the expectation that once something begins to move in a certain direction, it will continue to move in that direction (gestalt of *good continuation*).

Discontinuity counteracts expectations: the sequence of elements is not foreseen. *Radical discontinuity* is disruptive. It corresponds to the sudden appearance of a bold stimulus, or the sudden interruption of cohesive motion. The event is of highest intensity, corresponding to a drastic change – it automatically attracts attention. *Ambivalent discontinuity* is simultaneously continuous and discontinuous. At lower perceptual resolution, the foreseeable logic is shifted without disruption. At high resolution, the discontinuity becomes more intense. Higher intensity implies greater attention, and lower intensity implies lesser attention.

ISSN 1071-4391 ISBN 978-1-906897-22-2
ISSN 1071-4391 ISBN 978-1-906897-22-2

Finally, *endogenous continuity* 'wraps' musical motion, binding all types of continuities and discontinuities in meaningful ways. Endogenous continuity occurs at high hierarchical level in the perceptual organization. It requires the use of long-term memory, attention and conscious awareness.

Using continuities and discontinuities

In many musical languages, sounds relate in surreptitious manners. Irregular and multi-layered, the sonic motion ties and unties many sonic qualities at many hierarchical levels, inviting perception to navigate between detail and structure. Such sonic orchestration shapes the whole range of continuities and discontinuities, fulfilling, tempting or violating expectations. Experiencing cohesion within such complexity (i.e. establishing endogenous continuity) requires attending how the aural unfolds: how a moment in the musical motion reports to a multitude of previous sounds, silences, or kaleidoscopic soundscapes.

If we close our eyes, perceptual resolution increases and decreases according to the sonic motion. If the desire is to maintain this same perceptual oscillation when a person looks at the screen, one must not employ radical visual discontinuities, such as image cuts, because these would automatically attract attention. Radical visual discontinuities would disrupt the sense of auditory continuity. [51] Synchronized or not, such discontinuities would blind us to the nuances of sounds and the logics of sound organization.

The diagram in Figure 2 shows how to articulate continuities and discontinuities instead. The sonic motion oscillates between points of lowest intensity and points of highest intensity, and apprehending its cohesion requires an individual to establish endogenous continuity. In-between steady continuities and radical discontinuities, there are gradients of ambivalence – experiencing continuity or discontinuity depends then on deliberate attention. If the desire is to invite attention to focus on the wealth of sounds, the moving image should not attract attention automatically – its discontinuities should not exceed medium intensity.

While multiple visual changes may occur at once, they should converge into a same form. This is a second condition to subordinate the visual motion to the aural. In many audio-visual performances, the moving image orchestrates a number of different shapes – musical voices – that change simultaneously forming polyphonic movements. These movements create as much, frequently more tension and release than does sound. Thus, the audio-visual skews visually: the demand for visual information processing obfuscates the sonic experience, blinding perception to the nuances of sound organization. My alternative approach is to minimize the amount of mental activity that the visual information imposes on working memory. Hence, a monophonic moving image brings an environmental context to the sonic experience, punctuating the sonic motion in non-disruptive ways (Table 2).

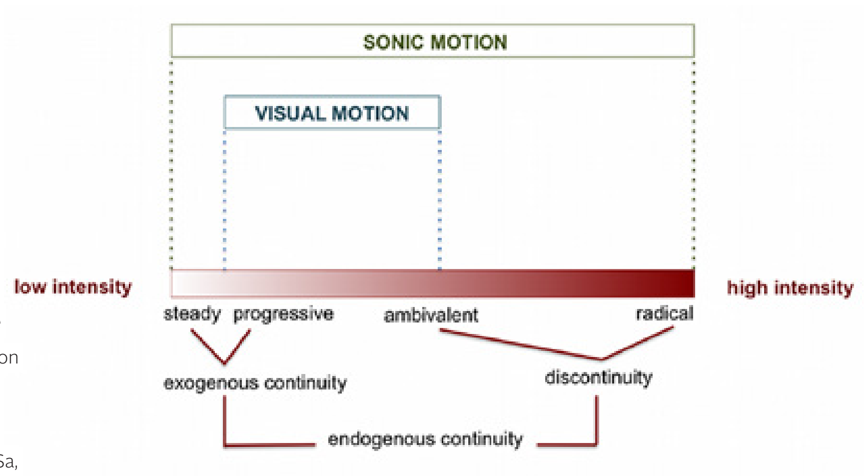

Figure 2. Articulating the sonic and the visual motion for the experience to be driven through sound organization. © Adriana Sa, 2012. Used with permission.

 ISSN 1071-4391 ISBN 978-1-906897-22-2

Sound	Moving image	Relative intensities	Results
Steady continuity	Progressive continuities and ambivalent discontinuities	The aural is less intense than the visual	Because neither sound nor image are disruptive, perception is invited to draw upon ambivalent discontinuities
Progressive continuities and ambivalent discontinuities	Progressive continuities and ambivalent discontinuities	The aural is as intense as the visual	Progressive continuities and ambivalent discontinuities
Radical discontinuity	Progressive continuities and ambivalent discontinuities	The aural is more intense than the visual	Because the aural is disruptive and the visual is not, the aural subordinates the visual within multisensory integration

Table 2. Perceptual integration of the sonic and the visual motion. © Adriana Sa, 2012. Used with permission.

Permeability and sonic expression

Perception entails a tendency to simplify information through all manner of assumptions, but simultaneously we can be consciously aware of the broad membrane of complexity formed of that same information. Perceptual simplification suits the primary aim of the brain and the limited capacity of short-term memory. It dominates when we focus on a purpose, such as accomplishing a task. Another mode of perception is possible when we are not driven by any purpose, focusing on the experience itself.

Whereas focusing on a purpose frames the mind through previous information, art invites us to another sort of experience: the brain makes use of assumptions to simplify and clarify the perceptual field, and simultaneously it draws upon their ambivalence. The work shapes a field of possible experiences, with varying complexity and non-univocal meanings. It can be accessed via multiple perspectives. It can transcend any of these perspectives, indicating their subjectivity.

Art invites perception to bind the information without subordinating to any goals. It conducts attention beyond the pragmatic simplification so useful to handle the world – one draws upon elements that are usually ignored. Susan Broadhurst outlines the neural implications of such non-pragmatic experience: "all works of art that (...) frustrate our expectations of any clear resolution (...) are likely to activate a specific area of the frontal lobe which appears to deal with the resolution of perceptual/ experiential conflict." [52]

Playing with those expectations, many artistic practices relate Pressing's hierarchical complexity and his dynamic or adaptive complexity. To these artists, the adaptation to the unexpected can reveal a structure in the indeterminable. In the 1950s, John Cage proposed that uncontrolled features are creative matter, and many recent performance languages address them as such. [53] [54] Cage tossed I-Ching coins and Tarot to score many of his pieces. His strong assimilation of Eastern philosophies is well known, and these philosophies suggest that suppressing intention is required to permeate a relation between all things. Today one can also speak of this major order in scientific terms. The Chaos theory, the existence of an order underlying apparently random facts, implies what Karl Popper named *causal chance:* "It was only the incompleteness of our knowledge which gave rise to this kind of chance." [55]

ISSN 1071-4391 ISBN 978-1-906897-22-2

As a performer I feel that dealing with 'chance' is a way to permeate rather than impose a structure upon the sensory complexity. An instrument is simultaneously a controlled prolongation of the body, and a means of expanding action beyond intention. It is both a tool and an entity in itself. As such, a threshold exists between its unpredictability and the performer's control. Manipulating this threshold potentiates an artist's idiosyncratic expression. An unexpected, often minute event can produce compelling performative tension. It causes a minimal, yet graspable hesitation – a moment of suspense. Resolving the musical challenge in good time then causes a sensation of release. The performer's expression transpires openness to the Now in transformation.

Thresholding causation

Returning to the relation between audition and vision, transparent cause-effect relationships can be as problematic as radical visual discontinuities and polyphonic visual movements. Many composers [56] [57] are well aware that perception tends to bias those sonic events and qualities which help visual apprehension. It tends to subordinate other qualities such as timbre, texture, vibration, and the nuances of the performer's expression, which form the wealth of multilayered relations among the sounds themselves. On the other hand, discerning causation brings an additional layer to the experience: an audio-visual instrument can be a creative work in itself. The question is how to evidence causation, and simultaneously confound the cause and effect relationships. The audience should bind aural and visual, yet remain unsure about the actual audio-visual mapping.

The role of synchronization in this process seems debatable. Whereas Chion states that synchronization leads inevitably to perceptual binding, [58] Kubovy and Schutz state that it does not. [59] To Chion, synchrony is more relevant than verisimilitude because verisimilitude arises from 'convention' rather than 'truth.' Given the term convention, his type of verisimilitude is symbolic, which makes binding undoable. Conversely, Kubovy's and Schutz's binding occurs at lower level in information processing. As the marimba sound binds with the visual impact, which seems to have caused it, the audio-visual linkage becomes so plausible that binding becomes undoable.

Perception seems anxious to bind synchronized information, but depending on verisimilitude it may do so, or it may not. Consistent synchronization makes the mechanical connection unequivocal. We are driven to form conclusive concepts at the expense of overlooking or skewing any conflicting information. Conversely, fungible audio-visual correspondences foster ambiguity. One senses causation, but cause and effect relationships are confounding. Perception keeps acknowledging conflicting information, embracing convergences and divergences with inconclusive concepts.

With a fungible audio-visual mapping, cause-effect relationships range from transparent to opaque. Points of sensory unison provide a sense of connection, but it becomes unclear whether such unison is purely perceptual, or if it corresponds to actual instrument mappings. Whilst these unions are indicative of the instrument, confounded cause-effect relationships invite the audience to 'quit trying to understand' the instrument, and rather explore perception itself.

Principles for audio-visual instruments

The previous extrapolations can be summarized in three principles which define an audio-visual instrument that emphasizes the sonic experience:

» To confound causation in the audio-visual mapping, by creating fungible rather than consistently transparent correspondences between sounds and visual events.

» To convey sonic complexity and expressiveness, while thresholding certain unpredictability in the sonic output.

» To facilitate the perceptual simplification of the moving image, by ruling out disruptive visual changes and minimizing divergences in visual movements occurring at the same time.

PART 3 – EXPLORATIONS

The three principles for audio-visual instruments can manifest in many different ways. My personal research has been developed with an instrument whose output combines acoustic sound, digital sound and digital image.

A personal instrument

With an audio-visual instrument combining acoustic and digital, physical gesture plays an important role in whether its mechanics are transparent or not. Whereas the synchronization of gesture, sound and image corresponds to a transparent interaction with the instrument, desynchronizing any of these elements blurs its mechanics. Quoting Simon Emmerson, "we may opt for a more ambiguous relationship, mixing some directly perceivable cause-effect chains with (a) relationships between performer gesture and result which the performer may understand but the audience not, or (b) relationships of a more 'experimental' nature the outcomes of which may not be fully predictable." [60]

The instrument includes a zither and AG#1, a modified software combining video-game and audio processing technologies, in which sound and image affect each other reciprocally (Figure 3). [61] [62] A video-game engine is appropriate for rendering a digital 3D world and making the position of elements in that world have an impact on the spatialisation of related sounds. Some people also elaborate on a (debatable) connection between playing games and music. Works by John Klima, [63] [64] Mick Grierson, [65] Robert Hamilton, [66] or Florent Berthaut et al. [67] make obvious references to video-gaming while using participatory interaction, 1stperson player-paradigms and allusive iconography.

This is not the present case. AG#1, the digital component of my audio-visual instrument, is a modification of software originally written by Klima (video-game programmer and media artist). My modifications circumvented the original functions and creative guidelines. The audience does not interact with the instrument. There are no allusive icons or player-paradigms. The performer does not face the screen. The interaction with the instrument is not fully transparent. Control and unpredictability are thresholded to potentiate performative expression. The audio-visual is architected to emphasize the sonic experience.

AG#1 detects amplitude and frequency from the zither input. The zither playing drops audio-visual objects in the digital 3D space. These objects emerge as light particles, which emit pre-recorded sounds (Figure 4). The acoustic input detection also determines how the virtual camera moves, producing an undulating view over an unchangeable landscape. As the camera moves and the sound-emitting particles remain where they were dropped, the digital sounds are spatialized accordingly.

With 2D sound, each aspect of the sound environment is directed to a particular output speaker. With 3D sound, output depends upon the position of the source in the 3D world, relative to the 3D camera/listener. If the camera/listener or the sound-source moves, the audio output reflects this shift. This is achieved by constantly altering how the signal is routed to the output channels.

 ISSN 1071-4391 ISBN 978-1-906897-22-2

A R T I C L E

ISSN 1071-4391 ISBN 978-1-906897-22-2

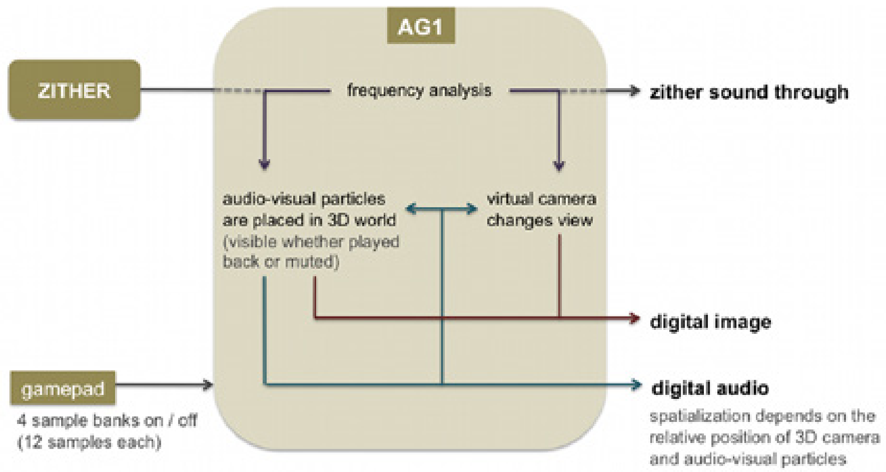

Figure 3. Technical diagram of the audio-visual instrument (2012). © Adriana Sa, 2012. Used with permission.

On modifying video-game software for an instrument combining acoustic and digital

As a rule-based software, AG#1 embeds a set of 'if-then' statements using a set of assertions, to which rules on how to act upon those assertions are assigned. Whilst these statements rule out undesired outcomes, their combination with each other and with the acoustic input generates complexity and certain unpredictability.

Combining acoustic and digital within a single instrument entails addressing their fundamental differences. As Thor Magnusson observes, the body of the acoustic instrument is physical: the interface and the sound engine are one and the same. Diversely, the body of the digital instrument is intrinsically theoretical: there is no natural mapping between physical input and digital output. Furthermore, "it is the designer who decides with clear rational arguments what is revealed and what is concealed in the use of the system (...) This activity of blackboxing, of creating abstractions of activities where bodily movements and thoughts are represented as discrete chunks in time, grounds the complexity and the non-transparency of digital tools." [68]

The coder frames and freezes his or her affordances through symbols: code. If the 'user' explores ways of unfreezing these affordances, this entails an ideological dimension. Particularly interesting is if the original coder actually collaborates by revealing what was concealed – then the process involves negotiation. Such is the present case: Klima showed me how to change several functionalities in his code. In this modification I consideredmy idiosyncrasies as a performer, my non-traditional modes of playing a specific zither, and research on how the aural can subordinate the visual. I changed the audio-visual reactions to the sonic input, the digital audio and its mappings, the 3D-world design, the images applied to the 3D-world, and the parameter configurations.

Whereas software configuration requires looking at the screen, in performance my interfaces are solely the zither and the game pad. My graphic scores sketch musical sequencing, density and texture. But real-time decisions and timings are primarily visceral, as my zither playing dwells with digital mappings and constraints.

Fungible correspondences between sound and image

The moving image is projected over me, and I dress white. This reinforces its environmental role, and it brings the 3D world into the physical space – the performative arena includes both. Standing in front of the screen, I sense the image solely as light. The image has an impact upon the audience's sonic experience, but disparities between their experience and mine are limited if the visual subordinates to the aural.

3D sound with multi-channel diffusion would inevitably link digital sound and image. AG#1 outputs 3D sound, but I use inverted stereo instead. As digital sounds move through the speakers, their speed equals the

ISSN 1071-4391 ISBN 978-1-906897-22-2

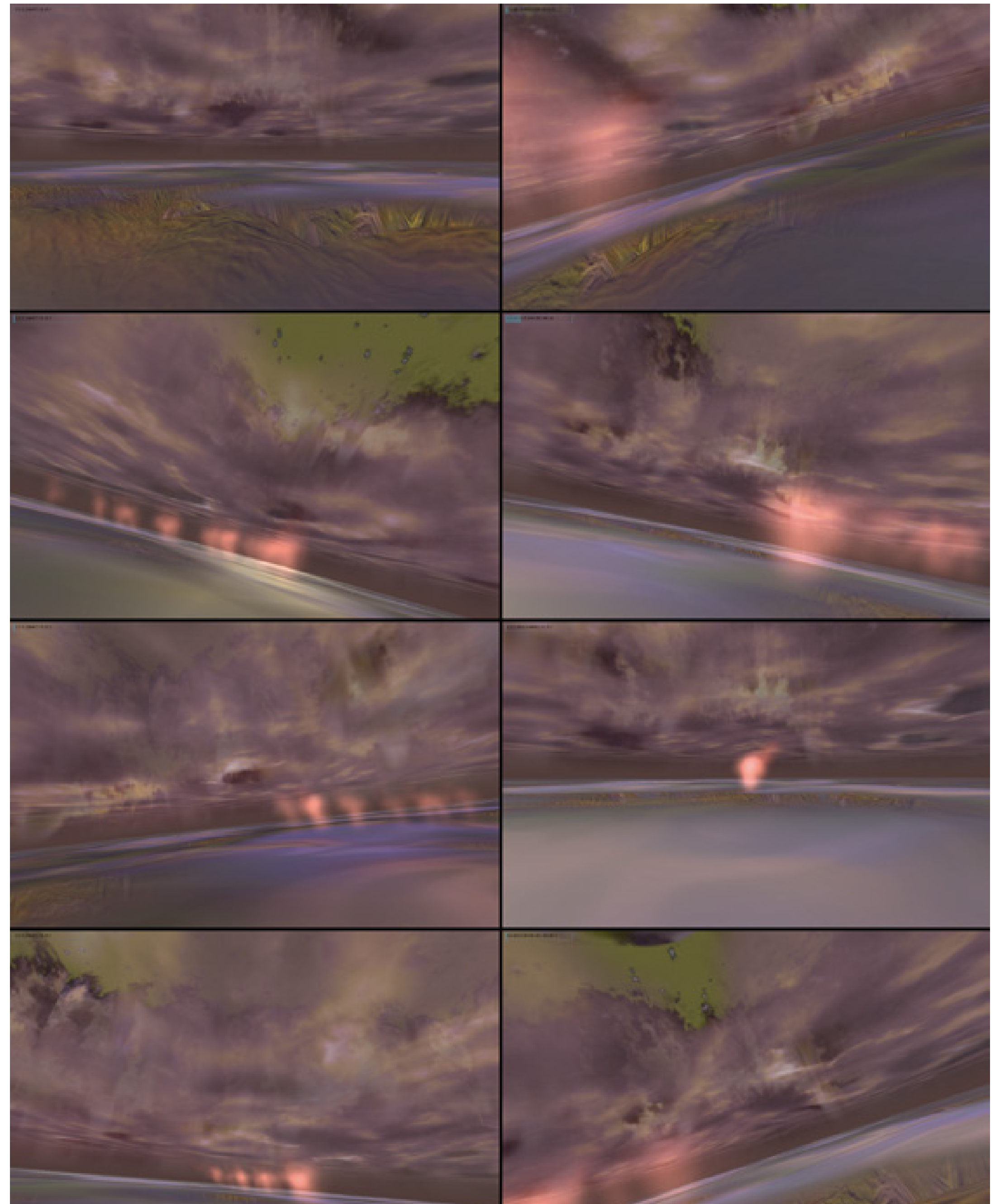

Figure 4. Screenshots of the AG#1 software (2012). © Adriana Sa, 2012. Used with permission.

visual. But their spatialization does not simulate the 3D environment. Thus, cause-effect relationships can be confounded. Furthermore, the audience's position is considered carefully. The acoustic input sounds are caused by gestures, but if gestures are not seen in detail, the cause-effect mechanics can be confounded.

Perception binds sonic and visual shapes that change or move simultaneously (gestalt of common fate), as well as those that change adequately proximal in time (gestalt of proximity/ sequential integration). Yet, the sense of sensory unison may not always correspond to the digital mappings, which range from transparent to opaque. Whilst moments of mechanical transparency indicate the instrument, moments of opaque mapping

ISSN 1071-4391 ISBN 978-1-906897-22-2

counteract its understanding – an invitation to explore perception itself.

A few connections between sound and image are clearly perceived. Bigger visual changes occur immediately upon sound detection (i.e. light particles and 3D camera direction shifts). In addition, the speed of digital sound spatialization corresponds to the speed of image.

Other factors bring inconsistency, confounding cause and effect relationships. Whereas the zither is always audible, the software only responds above certain amplitude. Bigger visual changes are consistent with zither input detection, but digital sounds are inconsistent with that same detection (I can stop digital sound activation, and several audio recordings start with silence). Furthermore, using digital recordings of the zither makes acoustic input and digital output at times undistinguishable.

Sonic control and unpredictability

AG#1 operates through diatonic pitch tracking. Each of twelve pitch detection groups is mapped to four digital audio samples – the samples are organized in four banks assigned to game pad buttons. The mechanical relation between acoustic sound and digital sound may at times be opaque to the audience, and nevertheless transparent to myself as a performer. At other times, it becomes quite opaque to me as well.

AG#1 operates according to mappings, but in performance the outcome is not fully predictable. Firstly, the actual resilience of the zither playing exceeds its codified terms – the digital processing does not handle all the acoustic information, since it is based on sampling the input. Secondly, the software at times considers elements that one is not perceptually aware of, and it responds accordingly. The incoming information is processed based on mathematical calculations and average values. Diversely, humans sample, chunk and prioritize the sensory information based on personal states of attention, on cognitive principles, and on the cross-sensorial context. When an instrument combines acoustic and digital sound, one can explore disparities between immediate and mediated sonic agency. They foster performative tension (Figure 5). As the instrument thresholds control and unpredictability, it potentiates expression.

Several technical aspects convey performative control: the zither playing is audible regardless of whether the software detects it or not; the sample banks (not the samples themselves) are activated or muted through the game pad; and clear notes generate predictable response.

Other technical aspects convey performative instability. The zither tuning is somewhat ambivalent and a few of its strings are purposefully aged. Thus I do not fully predict which audio sample will be activated, and a single string sound may activate several digital sound recordings mapped to different pitches. These may overlap with the acoustic sound, or appear immediately after, or overlap with subsequent sounds. Moreover, when the acoustic sound has no particular tonality – e.g. when zither strings are plucked, picked or strummed – the software response is unpredictable. And finally, whereas playing *mezzo forte* or *forte* activates digital audio samples, playing *piano* does not – while fluidly playing the zither, I do not fully predict where that threshold is.

Visual continuity

Sonic complexity can be fully experienced because the demand for visual information processing is low, and visual discontinuities do not automatically attract attention. Yet, the image is not really continuous. The camera view over the 3D world changes, light particles appear and faint, either isolated or drawing paths

ISSN 1071-4391 ISBN 978-1-906897-22-2

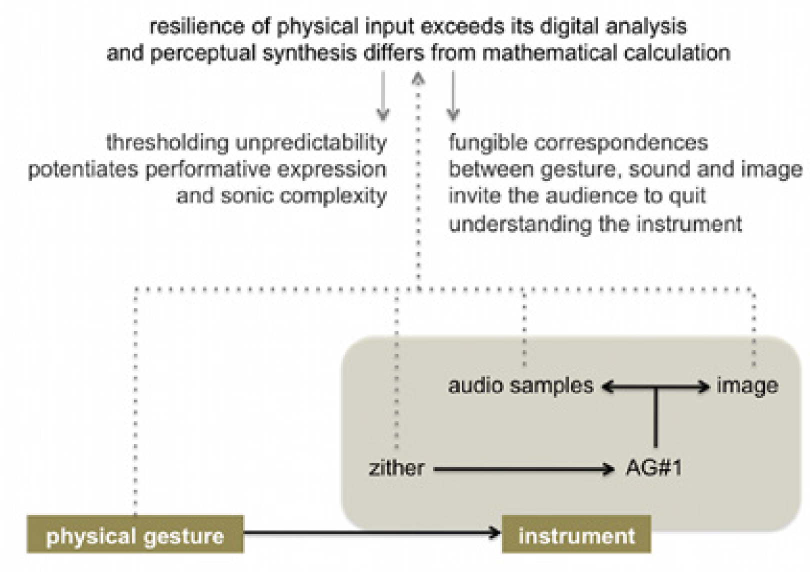

Figure 5. Exploring disparities between acoustic and digital. © Adriana Sa, 2012. Used with permission.

through the world. Such variations subordinate to an overall continuity, as a few gestaltist principles facilitate visual simplification.

There are no changes of scenario in the 3D world. The emerging light particles are monochromatic and similarly shaped. This prompts the gestalt of *invariance* (an object is recognized independent of rotation, translation, scale, elastic deformations, different lighting, and different component features) and the gestalt of *similarity* (similar elements are grouped as part of the same form). Any accumulating particles create visual paths, which may draw subtle shifts in the foreseeable logic of events, producing ambivalent discontinuities – at low resolution such discontinuities become continuous.

Adapting functionalities from a hovercraft racing video-game, the 3D camera creates an undulating view over the 3D world. Pitch detection shifts the camera view in a corresponding direction. Given the acoustic complexity, camera movements are not linear, but they are also never abrupt (progressive discontinuity).

Perceptual synthesis in performance

The sonic motion oscillates between points of highest intensity (radical discontinuity) and lowest intensity (steady continuity). It is of maximum amplitude. The visual motion is much narrower. It produces an overall sense of continuity, with progressive and ambivalent discontinuities.

As the performative arena links the physical and the digital space, the visual conveys a sense of total environment. The visual never reaches high intensity in itself; monochromatic light particles and smooth camera shifts are the most discontinuous changes. Those visual changes are synchronized with amplitude detection. Since the detection threshold is high, they are unlikely to skew surreptitious sounds.

Because the audience's attention to the image is deliberate rather than automatic, thus the ventriloquist effect does not occur (perceptual dislocation of the sound source toward the visual target). [69] The spatialization of sounds puts the focus in the physical space, rather than in the digital space beyond the screen.

Perception is invited to follow sound organization, in its fluid oscillation between tension and release. Except at points of radical sonic discontinuity, attention is predominantly under individual control. As one focuses on sonic complexity, that complexity causes changes in perceptual resolution. Because audio-visual percepts are synthesized at that resolution, sonic complexity also affects vision.

ISSN 1071-4391 ISBN 978-1-906897-22-2

CONCLUSION

The article exposes a frame of evidence in the question of whether and how an audio-visual instrument can convey a sense of total environment without obfuscating the sonic experience. It hopes to bring the subtleties of a fundamental artistic concern into debate: conscious awareness can trace unusual paths through the unconscious, but it is conditioned by the narrow time limit of short-term memory and by the primary aim of the brain. The article describes a way to emphasize sonic complexity, and extend that complexity to vision.

Building upon existing research in neuroscience and psychology the article proposed that intensity derives from the combined effects of automatic and deliberate attention. In this way, intensity depends on the relative impact of change, rather than on qualities that are specific to sounds or to images. I proposed a taxonomy that relates attention and intensity with different types of continuity and discontinuity. The corresponding terminology distinguishes whether apprehensions are primarily driven through stimuli, or if they are more under individual control. I also elaborated on the role of unpredictability in sonic expression. Furthermore, I discussed the role of synchrony, proposing that the audio-visual mapping should foster a sense of causation, but confound the cause-effect relationships.

The investigation found three principles for the creation of audio-visual instruments. One, to dispense with radical visual discontinuities and minimize the demand for visual information processing. Visual changeability can rather be explored at detail level. Two, to threshold sonic control and unpredictability, in such a way that the instrument affords sonic complexity and expression. The wealth of sound organization will be fully experienced when combined with overall visual continuity. And three, to create fungible audio-visual correspondences. Cause and effect relationships should range from transparent to opaque, at once highlighting the instrument and inviting attention to explore perception itself.

The three principles enable a diversity of creative strategies. They must be specified in each framework and technical platform. As an example, I described a personal instrument. It combines a zither and 3D software, where sound and image affect each other reciprocally. The instrument addresses the three principles while reverting functionalities from video-game technologies, and exploring disparities between acoustic and digital outputs.

Whilst the three principles suit a particular artistic sensibility, the perceptual approach they are driven from also puts in evidence the unison of apparently contradictory perspectives. It shows that whether moving images distract from audition, or sounds distract from vision, or synchrony is of benefit to both, in fact the question is how perception is driven to handle the sensory input. ■

 ISSN 1071-4391 ISBN 978-1-906897-22-2

ACKNOWLEDGEMENTS

This research is funded by FCT – Foundation of Sciences and Technology (SFRH/BD/68632/2010). John Klima supported my modification of his software. Atau Tanaka, Steve Gibson, Brigitta Zics, Rajele Jain, Alan Klima, and Baptiste Caramiaux made important observations along iterative drafts of the article. A performance series named *Windowmatter* has been developed along with the instrument. Started in 2008, the series was funded by former Direcção Geral das Artes / Ministry of Culture Portugal. I made graphic scores for solos, duos, and a group of ten. Collaborating artists included: John Klima, Toshio Kajiwara, Sei Miguel, Fala Mariam, Cesar Burago, Andre Gonçalves, O.blaat, Manuel Mota and David Maranha.

REFERENCES AND NOTES

1. Pierre Schaeffer, *Traité des Objets Musicaux: Essai Interdisciplines* (Paris: Éditions du Seuil, Paris, 1966), 93, quote translated in B. Kane, "L'Objet Sonore Maintenant: Pierre Schaeffer, Sound Objects and the Phenomenological Reduction," *Organised Sound* 12, no. 1 (2007): 15–24.
2. Ibid, 270-2.
3. Francisco López's official Web site, "Against the Stage," February 2004, http://www.franciscolopez.net/stage.html (accessed September 19, 2012).
4. Stan Brakhage, "Film and Music," in *Brakhage Scrapbook: Collected Writings 1964-1980*, ed. R. A Haller (NY: Documentext, 1982), 19, quoted in Merilyn Brakhage, "On Stan Brakhage and Visual Music," *Vantage Point* (January 31, 2008), http://vantagepointmagazine.wordpress.com/2008/01/31/on-stan-brakhage-and-visual-music (accessed September 23, 2012).
5. Merilyn Brakhage, "On Stan Brakhage and Visual Music," *Vantage Point* (January 31, 2008), http://vantagepoint-magazine.wordpress.com/2008/01/31/on-stan-brakhage-and-visual-music (accessed September 23, 2012).
6. Meghan Stevens, *Music and Image in Concert* (Australia: Music and Media, 2009).
7. Michel Chion, *Audio-Vision: Sound on Screen*, trans. Claudia Gorbman (New York: Columbia University Press, 1994).
8. Ibid., 23.
9. Michael Kubovy and Michael Schutz, "Audio-Visual Objects," in *Review of Philosophy and Psychology* 1, no. 1, ed. N. J. Bullot and P. Egré (Springer Science and Business Media B.V., 2010): 57.
10. Gemma A. Calvert, Charles Spence and Barry E. Stein, eds., *The Handbook of Multisensory Processes* (Cambridge, MA: MIT Press, 2004).
11. Albert S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound* (Cambridge, MA: MIT Press, 1990).
12. James Tenney, *Meta Hodos and Meta Meta Hodos: A Phenomenology of 20th Century Musical Materials and a Study of Form* (Oakland: Frog Peak Music, 2nd edition, December 1988).

ISSN 1071-4391 ISBN 978-1-906897-22-2

13. Bob Snyder, *Music and Memory: An Introduction* (Cambridge, MA: MIT Press, 2001).
14. F. Lerdahl and R. Jackendoff, *A Generative Theory of Tonal Music* (Cambridge, MA: MIT Press, 1983).
15. Max Wertheimer, "Laws of Organization in Perceptual Forms," in *A Source Book of Gestalt Psychology,* ed. W. Ellis, 71-88 (London: Routledge & Kegan Paul, 1938).
16. Edgar Rubin, *Visuell Wahrgenommene Figuren: Studien in Psychologischer Analyse* (Copenhagen: Gyldendalske Boghandel, 1921).
17. Kurt Koffka, *Principles of Gestalt Psychology (*London: Lund Humphries, 1935).
18. Steve Palmer, *Vision Science, Photons to Phenomenology (*Cambridge, MA: MIT Press, 1999).
19. Snyder, *Music and memory.*
20. Eric I. Knudsen, "Fundamental Components of Attention," *Annual Review of Neuroscience* 30, no. 1 (2007): 57-78.
21. Snyder, *Music and memory*, 50.
22. Herbert Zettl, *Sight Sound Motion: Applied Media Aesthetics*, 3rd edition, (Belmont, CA: Wadsworth Publishing, 1998), 6.
23. Snyder, *Music and memory*, 59-60.
24. Jeff Pressing, "Cognitive Complexity and the Structure of Musical Patterns," in *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*, ed. J. Sloboda, 129-178 (New York: Oxford University Press, 1987).
25. Tenney, *Meta Hodos and Meta Meta Hodos*, 17.
26. Ibid., 90.
27. Ibid., 33-41.
28. Snyder, *Music and memory*, 62.
29. Wilfried Jentzsch and Hiromi Ishii, "Multimedia Visual Music: Analogical Processes between Images and Sounds," (paper presented at the third EMSAN-CEMC Colloquium: Round Table Discussion and Forum of Development about Asian Electroacoustic Music*,* Beijing, 2010).
30. Chion, *Audio-Vision*, 5.
31. Herbert L. Pick, David H. Warren, and John C. Hay, "Sensory Conflict in Judgments of Spatial Direction," *Perception and Psychophysics* 6, no. 4 (1969): 203–205.
32. Ladan Shams, Yukiyasu Kamitani, and Shinsuke Shimojo, "Visual Illusion Induced by Sound," *Cognitive Brain Research* 14, no. 1, (2002): 147-152.
33. Harry McGurk and John MacDonald, "Hearing Lips and Seeing Voices," *Nature* 264 (December 23, 1976): 746–748.
34. Maori Kobayashi, Yoshihisa Osada, and Makio Kashino, "The Effect of a Flashing Visual Stimulus on the Auditory Continuity Illusion," in *Perception & Psychophysics* 69, no. 3 (2007): 393-399.
35. Virginie V. Wassenhove, Dean V. Buonomano, Shinsuke Shimojo, and Ladan Shams, "Distortions of Subjective Time Perception Within and Across Senses," *PLoS ONE* 3, no.1 (2008), http://www.plosone.org/article/info:doi/10.1371/journal.pone.0001437 (accessed September 23, 2012).
36. R. B. Welch and D. H. Warren, "Immediate Perceptual Response to Intersensory Discrepancy," in *Psychology Bulletin* 88, no. 3 (November 1980): 638–667.
37. Chion, *Audio-Vision,* 11.
38. Ibid., 61.
39. Francis B. Colavita, "Human Sensory Dominance," in *Attention, Perception & Psychophysics* 16, no. 2 (1974): 409-412.
40. Jeff Pressing, "Some Perspectives on Performed Sound and Music in Virtual Environments," *Presence* 6, no. 4 (1997): 10.
41. Scott Sinnett, Charles Spence, and Salvador Soto-Faraco, "Visual Dominance and Attention: The Colavita Effect Revisited," *Attention, Perception & Psychophysics* 69, no. 5 (2007): 684.
42. Knudsen, "Components of Attention."
43. Ibid., 58.
44. Durk Talsma, Daniel Senkowski, Salvador Soto-Faraco, and Marty G. Woldorff, "The Multifaceted Interplay between Attention and Multisensory Integration," *Trends in Cognitive Sciences* 14, no. 9 (2010): 400-410.
45. P. Bertelson, J. Vroomen, B. de Gelder and J. Driver, "The Ventriloquist Effect does not Depend on the Direction of Deliberate Visual Attention," *Perception and Psychophysics* 62, no. 2 (February 2000): 321–332.

ISSN 1071-4391 ISBN 978-1-906897-22-2

ARTICLE

46. E. Van der Burg, J. Cass, C. N. L. Olivers, J. Theeuwes, and D. Alais, "Efficient Visual Search from Synchronized Auditory Signals Requires Transient Audiovisual Events," *PLoS ONE* (May 2010), http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0010664 (accessed September 23, 2012).
47. Kubovy and Schutz, "Audio-Visual Objects," 58.
48. Ibid., 55.
49. Ibid., 57.
50. Knudsen, "Components of Attention."
51. See Kobayashi, Osada, and Kashino, "Auditory Continuity Illusion."
52. Susan Broadhurst, *Digital Practices: Aesthetic and Neuroesthetic Approaches to Performance and Technology* (New York: Palgrave Macmillan, 2007), 58.
53. See for example Kim Cascone, "The Esthetics of Failure: Post-digital Tendencies in Computer Music," *Computer Music Journal* 24, no. 4 (2000): 12-18.
54. And also Andrews, "Recent Audio-Visual Art."
55. Karl R. Popper and William W. Bartley, *The Open Universe: an Argument for Indeterminism* (Totowa, NJ: Rowland and Littlefield, 1982), 125.
56. Pressing, "Music in Virtual Environments," 10.
57. Schaeffer, *Objets Musicaux.*
58. Chion, *Audio-Vision*, 23.
59. Kubovy and Schutz, "Audio-Visual Objects," 57.
60. Simon Emmerson, *Living Electronic Music* (Aldershot, Hampshire: Ashgate Publishing, 2007), 91.
61. AG#1 combines the video-game engine Torque and the audio library FMOD. It extends the zither, but I hesitate in using the term 'extended instrument' because this term seems to connote different creative criteria – e.g. an absolute control over the instrument and a clearly understandable interaction with it (see Poepel and Overholt 2006).
62. C. Poepel and D. Overholt, "Recent Developments in Violin-related Digital Musical Instruments: Where Are We and Where Are We Going?," in *Proceedings of the International Conference on New Interfaces for Musical Expression,* 390–395 (Paris: NIME, 2006).
63. M. Mirapaul, "Three Dimensional Is the New Frontier for the Internet," *The New York Times* (October 5, 2000).
64. Christiane Paul, *Digital Art* (London: Thames & Hudson, 2008).
65. Mick Grierson, "Noisescape: A 3d Audiovisual Multi-user Composition Environment," in *The World as Virtual Environment*, ed. J. Birringer, T. Dumke and K. Nicolai, 160-168 (Dresden: Trans-Media-Akademie Hellerau, 2007).
66. Robert Hamilton "Maps and Legends: Designing Fps-based Interfaces for Multi-user Composition, Improvisation and Immersive Performance," in *Computer Music Modeling and Retrieval* (4th International Symposium, August 27-31, 2007), 478-486 (Copenhagen: CMMR, 2008).
67. F. Berthaut, H. Katayose, H. Wakama, N. Totani, and Y. Sato, "First Person Shooters as Collaborative Multiprocess Instruments," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 44–47 (Oslo: NIME, 2011).
68. Thor Magnusson, "Of Epistemic Tools: Musical Instruments as Cognitive Extensions," *Organised Sound* 14, no. 2 (2009): 174.
69. Bertelson, Vroomen, de Gelder, and Driver, "Ventriloquist Effect."

ISSN 1071-4391 ISBN 978-1-906897-22-2