# PROCEEDINGS OF THE FIFTH CONFERENCE

X · Computation Communication Aesthetics And

# xCoAx 2017

# DESIGNING MUSICAL EXPRESSION

**ADRIANA SÁ**
adrianasa@clix.pt

Deps. of Architecture &
Dep. of Electrical and
Computer Engineering
University of Coimbra, Portugal

**Abstract**

The term New Interface for Musical Expression (NIME) has been applied to a great variety of instruments and systems, since the first NIME conference in 2001. But what is musical expression, and how does an interface intended for idiosyncratic expression differ from ubiquitous interfaces? This paper formulates an understanding where the reciprocal interaction between performer and instrument is important. Drawing from research in perception science, interface design and music, the paper specifies methods that can be used to analyse interaction, attention dynamics and semantics. The methods are applicable to any technical platform and aesthetic approach, facilitating the discussion of creative strategies and the analysis of music experience. The paper uses these methods to describe a NIME that combines an acoustic string instrument and software that operates based on the acoustic sound. The software applies the difference between the detected pitch and the closest tone / half tone to the processing of pre-recorded sounds. The proposed methods help to explain how this NIME enables versatile musical forms, and prevents undesired outcomes.

**Keywords**
NIME
Musical Expression
Musical Interface
Perception
Attention
Semantics of Sound

## 1. INTRODUCTION

Our ways of interacting with the world condition how we perceive the world, and vice-versa. In the 1950s John Cage proposed that uncontrolled features can be used as musical material, and many recent artists explore related compositional approaches (e.g. Vasulka 1996 and Cascone 2000). His strong assimilation of eastern philosophies is well known. These philosophies suggest that suppressing intention is required to permeate the unity and mutual interrelation of all things, which are inseparable parts of a cosmic whole. Cage studied Indian philosophy and music. When he asked what is the purpose of music in Indian philosophy, he was answered: "to sober the mind and thus make it susceptible to divine influences" (Cage 1967, 158). This "sobering the mind" is not exclusive to Indian *raga*. It seems equally related with that which Francisca Schroeder and Pedro Rebelo called the *performative layer* (2009). They coined the term to describe how the performer's strategies in dealing with discontinuities, breakdowns and the unexpected reflect "a becoming-aware-of and awakening of unused abilities".

Understanding how perception works is crucial to certain research in music, and it also provides useful perspective upon artistic discourses that are seemingly distant from science. Bob Snyder described musical motion as the continuous oscillation between points of low intensity and high intensity (2001). He defined intensity as any change in the chain of stimuli causing an increase in neural activity. Looking at neuroscience and psychology led me to note that intensity reflects attention, activity reflects attention. Hence, attention depends not solely on the stimulus, but also on the panorama and a person's current perceptual resolution (Sá 2013).

Perception science tells us that counteracted expectations cause an increase in neural activity, and fulfilled expectations cause a decrease. It also explains that attention causes us to optimize perceptual resolution (Knudsen 2007). Optimization occurs when the eyes are directed toward a target, and / or when the sensitivity of neural circuits is modulated for an auditory target. Both automatic and deliberate attention causes perception to improve the quality of related information processing in all domains: sensory, motor, internal state and memory. Consequently, a person becomes more susceptible to any changes in the chain of stimuli. This understanding led me to create a taxonomy of continuities and discontinuities that seems very useful to analyze musical motion. It distinguishes between apprehensions automatically driven by stimuli and apprehensions under individual control.

So, how can the interaction with a musical instrument convey an expressive interplay of continuities and discontinuities?

Peter Weibel and many other researchers have based their notion of interaction on cybernetic theories: "the concept of systems is already anticipated in the concept of the environment — as an interaction between components of the system, where if one component of the system is absolutely dominant, the system can collapse" (Weibel and Lischka 1990, 67). The idea of reciprocity is important, but a useful notion of musical expression should not be restricted to digital systems. Also, the term "interaction" is not the keyword. Some well-known interactive systems are merely intended for a person to randomly move their arms or press a button, and the extent to which NIME designs allow for agency is quite variable. For example, in an article titled *Beyond Guitar Hero — Towards a New Musical Ecology*, Tod Machover declared the purpose of "diminishing the current exaggerated distinctions between celebrities and amateurs", so as to compensate for "people's limitations" (Machover 2008).[1] Such "compensation" implies that the software prescribes which output results are desirable, and which are not. In opposition, one can argue that an instrument should not compensate for limitations: an idiosyncratic, personal interface, which requires particular skills and constant practice, can be governed by very different principles than interfaces intended for non-musicians.

Joel Ryan, who pioneered digital signal processing of acoustic instruments, affirms that it is interesting to make control as difficult as possible, because effort is closely related to musical expression (Ryan 1991).[2] Andrew Johnston speaks of a type of interaction where "the musician allows the virtual instrument to 'talk back'" (Johnston 2011, 293).[3] Atau Tanaka stresses the importance of volatility in expression (Sá et al. 2015:20).[4] And Ryan, MacPherson,[5] Magnusson[6] and Tanaka agree on the importance of timing, despite their very different digital music interfaces and creative methods (Sá et al. 2015, 15-20). Ryan is particularly eloquent:

> **The fact is I know when. Before it happens, I know when a beat should come, I know after, when it didn't. (...) The time referred to here is not the objective, uniform time inferred by physics or fashioned by technology, but another, local time. It is (...) the time we make, enacted time, dense and polyvalent, the most elaborate aspect of time in music. (Sá et al. 2015, 15)**

We can say that this enacted, musical time is simultaneously personal and universal: indeed, the audience is equally sensitive to its logics, as it reflects an "awaken of unused abilities", quoting Schroeder and Rebelo 2009. It requires the musical interface to be effortful, to a certain extent. The term "effort" seems preferable to the term "virtuosity", which might not account for the creative potential of the unexpected due to its origin in classic music tradition. Effort can be quantified relatively to the amount of cognitive processing required in the construction of musical time. It manifests in the dynamics and semantics of music.

This paper will examine the reciprocal interaction between performer and instrument, and look at how different types musical motion influence the semantics of music. The first part formulates an understanding of expression that embraces a great diversity of musical idioms and interface designs. The second part describes a personal NIME, which combines a *concert zither*[7] with custom strings and tuning, and 3D software written from scratch. Its description will focus on the relation between acoustic and digital sound, independent from the image and the audio-visual relationship.

## 2. A WAY TO ANALYSE SONIC EXPRESSION

This section discusses how musical interface behaviors convey expression, and how expression substantiates in music. It specifies methods to analyze interaction, dynamics and semantics. Accordingly, it also describes a general creative principle.

1 That article was written years after the *hyperinstrument project*, which Machover and the MIT Lab started in 1986. The hyperinstruments used acoustic analysis, motion sensing and pressure sensing.

2 Talk: https://vimeo.com/3392802; Performance: https://www.youtube.com/watch?v=IQ4DqRgtbHc.

3 Performance: https://vimeo.com/16190938.

4 Demo: https://vimeo.com/46091343.

5 Demo: https://vimeo.com/46091343.

6 Demo: http://www.borealisfestival.no/2017/threnoscope-thor-magnusson-2/.

7 Multi-string instrument with a resonant body and fretboard.

## 2.1. The role of effort in the interaction with musical instruments

The philosopher Alva Noë speaks of perception as "an activity of exploring the environment drawing on knowledge of sensorimotor dependencies and thought" (Noë 2004, 228). This is applicable to perception in general. To discuss expression we need to consider the difference between two modes of perception: pragmatic and non-pragmatic. The difference was examined while bridging neuroscience and philosophy (Sá 2013). It was equally demonstrated with a study on audio-visual mapping and perception (Sá et al. 2014).

In summary, usually perception prioritizes the stimuli governed by a purpose, such as discerning a cause and a meaning, or accomplishing a task.Focusing on that goal frames the mind through previous information; perception simplifies the incoming information according to unconscious presuppositions. Conforming sensory information to presuppositions and concepts requires perception to segregate the information, and prioritize the converging over the diverging. Another mode of perception is possible when we are not driven by any purpose: we can also be consciously aware of a wider sensory complexity, beyond conclusive concepts. The non-pragmatic mode of perception is driven by intentionality, rather than intention. In philosophy, intentionality is described as the distinguishing property of mental phenomena of being necessarily directed upon an object, whether real or imaginary.[8] We can say that this "being necessarily directed upon an object" respects to the primary aim of the brain: to make sense of the world in order to survive. But whereas intention frames the mind through previous information so as to convey conclusive concepts, intentionality places conscious awareness at ground level, where we can focus on perception itself. The brain makes use of assumptions to simplify and clarify the perceptual field, and at the same time, it draws upon their ambivalences.

8 *Intentionality*, n." Oxford English Dictionary Additions Series. 1993. Online. Oxford University Press. 17 Aug. 2008.

So how do interfaces convey intention or intentionality? For example, we use a text processor while driven by the intention of writing a text. Gaming is also all about intention — the goal of the game, the challenge of accomplishing increasingly difficult tasks, the social interaction. A person can focus on writing the text or playing the game because the interface behaves consistently. For example, if pressing the "W" key made the game player move forward in the digital 3D space, pressing "W" again should make him move forward once more. Predictability maximizes our control over the interface. An interface behaving in unpredictable ways would distract us from our task; it would require attention in itself. Conversely, linear behaviors can make the interface "immaterial": the interaction becomes seemingly immediate. Paul Dourish described this dematerialization of the interface as the paradigm of ubiquitous digital media (Dourish 2004); it is the paradigm nowadays as well, possibly even more (Lombard et al. 2015).

Dematerializing the interface is often not desirable with musical interfaces, which can be designed to convey intentionality. As one dispenses with intention, interface behaviors do not need to be consistent. Indeed, complex behaviors foster an engagement with creative expression, as shown in a study conducted by Tom Mudd (Mudd et al. 2015). The authors highlight a contrast between "communication-oriented attitudes to engagement that view the tool as a medium for transmission of ideas" and "material-oriented attitudes that focus on the specific sonic properties and behaviors of a given tool".

Volatile, complex interface behaviors make the musician dwell with material properties. We can say that the interaction is bi-directional; indeed the term dwelling implies some sort of separation. This is seemingly paradoxical, because the instrument must also feel like a natural extension of the body, so that the performer can focus on the musical output. Articulating these requirements is a compelling creative challenge in the design of new musical instruments, which can be examined from many perspectives.

The separation of performer and instrument seems emphasized when derived from the resistances of musical systems, be they physical, or conceptual as one might find in the design of a programming language. The papers (Magnusson and Mendieta 2007) and (Bertelsen et al. 2009) show that many musicians enjoy engaging with such resistances, which exceed the performers' control. Chris Kiefer finds that they can become excessive in live coding performances: code is symbolic, and computer keyboards are not designed for musical expression (Kiefer 2015). He uses genetic programming representations to translate the output of a multi-parametric controller into code.[9] Whilst the controller provides a sense of immediacy, conveying embodiment, the focus on code and the keyboard typing create a separation.

The relation between embodiment and separation is not exclusive to digital instruments. An acoustic instrument is effortful, yet the interaction becomes fluid with training, to the extent of seeming natural. At a certain point, the musician does not focus on physical gestures—their techniques became body knowledge. Unpredicted events can bring the focus back to materiality, by creating a separation between musician and instrument. For example, Pedro Rebelo speaks of a *parasitical relationship* between the grand piano and the myriad of objects used in its preparation (Rebelo 2015).[10] Introducing "parasitical" elements during performance brings unpredictability. The same can happen when audio software operates based on an audible, acoustic input: the digital sound becomes a "parasite" of the acoustic, in ways that bring unpredictability. Regardless of whether any sensors can capture the resilient nuances of the acoustic sound, software is necessarily symbolic, and physical action will always be mediated through code.

The purpose of separation is to challenge the performer's body knowledge. A study conducted in an hospital environment showed that physical movements change from *exploratory to performatory* when a person becomes skilled in the execution of a specified task: movements become fluid, with a "focus on timing" (Kilbourn and Isaksson 2007). Whilst exploratory movements imply an "initial mode of attention", with performatory movements every gesture is a "development of the one before and a preparation for the one following". We can say that exploratory movements imply effort, and performatory movements imply embodiment. As a musician embodies their techniques, effort motivates a constant return to that 'initial mode of attention'. Whereas the performatory aspect of the music entails fluency and focus on timing, the exploratory aspect makes the musical thread unrepeatable and unique; it brings a "fresh" flavor to the music.

Effort is a variable in interaction design, which can be quantified so as to distinguish this understanding of expression from others. Little effort means one of two things: either the music does not depend much on the performer's interaction, or the relationship between deliberate human agency and sonic results is linear and clearly perceivable. High effort implies particular skills and / or high cognitive demand; the interaction with the system does not feel immediate, and / or the system does not rule out undesired outcomes. Medium effort means that

9 Performance excerpt: https://vimeo.com/122340471.

10 CD excerpt and review: http://www.squidsear.com/cgi-bin/news/newsView.cgi?newsID=804.

interface behaviors are complex, i.e. the performer needs particular skills to play the instrument; yet a sense of immediacy conveys musical timing, and / or technical configurations rule out undesired outcomes. The notion of expression formulated in this paper requires medium effort, or medium-high.

This notion of expression entails what Jeff Pressing called *dynamic complexity*: a rich range of behaviors over time, an adaptation to unpredictable conditions, a monitoring of results in relation to a reference source, and an anticipation of changes in oneself or the environment (Pressing 1987). We are constantly comparing what we hear with the "grid" of expectations derived from our psychophysical processes and internalized musical traditions. Effort manifests in the constant approach and deviation from that "grid". Those deviations are often very subtle, inviting for deliberate attention. As deliberate attention increases perceptual resolution, a person becomes more susceptible to automatic attention.

## 2.2. The dynamics and semantics of music

The dynamics of musical motion can be analyzed using the taxonomy described in (Sá 2013), which relates intensity and attention with continuities and discontinuities. Intensity is the neural impact of any change in the chain of stimuli causing an increase in neural activity.

*Steady continuity* is of lowest intensity; it dispenses with attention. Conscious awareness is likely to deviate and focus upon any simultaneous stimuli, or upon internal states.

*Progressive continuity* occurs when successive, non-abrupt events display a similar interval of motion. It is of low intensity, as it fulfills the expectation that once something begins to move in a certain direction, it will continue to move in that direction.

*Ambivalent discontinuity* is simultaneously continuous and discontinuous. Perceiving discontinuity depends on deliberate attention, which causes one to optimize perceptual resolution. At lower resolution, the foreseeable logic is shifted without disruption. At high resolution, the discontinuity becomes more intense.

*Radical discontinuity* is disruptive; it violates psychophysical expectations. It is of highest intensity, prompting automatic attention.

Finally, *endogenous continuity* binds any types of continuities and discontinuities in meaningful ways. It occurs at high level in perceptual organization, corresponding to the global semantics of the music.

The perceived continuities / discontinuities depend on stimuli, panorama and perceptual resolution. The panorama also incorporates time, but the relation of time and attention is not linear. For example, a long period of continuity can make an inconsistent event more discontinuous, but that is not a given because continuity also leaves attention under greater individual control. Sustained attention will increase perceptual resolution, making the discontinuity more intense. Without attention, the inconsistency becomes less intense; it can even go unnoticed. In any case, music is the construction of time, and the taxonomy of continuities / discontinuities should be considered relatively to the timescale of experience.

Another question is how to characterize the semantics of the music. Causes, concepts and meanings are important when we listen to music, but that is solely one aspect of the experience. At points they come to the focus of conscious awareness, and other times they submerge into the perceptual background, as we focus on the experience itself. Pressing distinguished informational, expressive and environmental sounds (Pressing 1997), and we can redefine those terms so as to account for how the interplay of continuities and discontinuities influences attention.

The *expressive* dimension of music coveys a focus upon the performer's personal expression and skills. It can be associated with ambivalent discontinuities and radical discontinuities. This narrows Pressing's notion of "expressive sounds", i.e. all kinds of music and song. It relates to what he termed dynamic complexity: a rich range of behaviors, an adaptation to unpredictable conditions, a monitoring of results, and an anticipation of changes in oneself or the environment.

The *environmental* dimension of music conveys a focus upon space and context, as opposed to the performer. This extends what Pressing termed "environmental sounds", e.g. animal calls, wind sounds and the noises of machinery: steady continuities and progressive continuities convey environmental semantics as well, because they allow for attention to deviate from the performer.

Finally, *informational* semantics embraces not solely Pressing's "informational sounds", e.g. speech, alarms and sonified data, but all situations where the sonic construction evokes something beyond itself. This means that the informational dimension of music can support its expressive and environmental dimensions. For example, a recording of singing birds evokes birds, and a piano recording leads us to imagine a piano and a pianist although we are solely hearing the sound.

## 2.3. A creative principle

The notion of expression formulated in this paper corresponds to a creative principle, which can be explored in many different ways. The principle dictates that interface behaviors should be complex, yet convey a sense of immediacy, and/or rule out undesired outcomes. Sound organization in real-time should require medium effort, or medium-high effort. This is a distinguishing factor: musical interfaces that require little effort or very high effort are not governed by this principle. The dynamics and semantics of music are not distinctive, because the principle allows for any type of continuities/discontinuities, and different semantic dimensions (Fig. 1).

The core idea of this creative principle is that the performer's interaction with the instrument is reciprocal. One can calibrate thresholds between the performer's control and the instrument's unpredictability, so as to convey idiosyncratic expression. As Anthony Gritten wrote, "while the subject is certainly performing, it is also performed" (Gritten 2006, 106). With this he meant that the experience of performing is simultaneously perceived through another type of experience.

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. 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. Once musical expression derives from addressing the unexpected resourcefully, performative action must exceed intellectual deliberation because musical timings and intervals possess biophysical logics. Dealing with non-anticipated sonic events makes me acknowledge and respond to sensory details that would otherwise go unnoticed. In a sense, creating musical meaning upon the unexpected augments my sense of control. As my capability of response is challenged, my sensitivity and resourcefulness become greater than if I was strictly executing a plan.

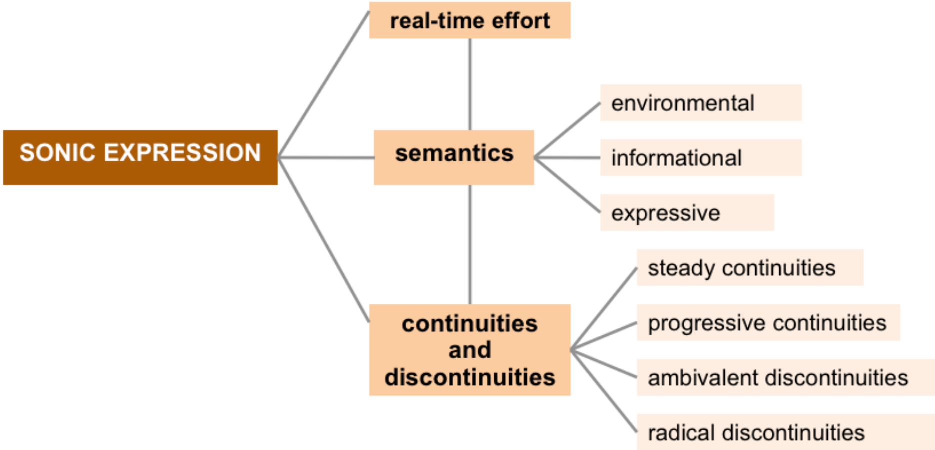


**Fig. 1**
Variables relevant to the analysis of sonic expression.

## 3. A PERSONAL INSTRUMENT AND MUSICAL IDIOM

The methods described in the previous section can be used to discuss a specific performance or a recorded music piece, whether the discussion is based on user-studies or individual assessments. I will leave that for another occasion. The methods are equally useful in the design of a versatile instrument, when the purpose is to create an open field of possibilities and rule out undesired outcomes. This section describes an instrument and its configurations. It details creative strategies regarding interaction, dynamics and semantics.

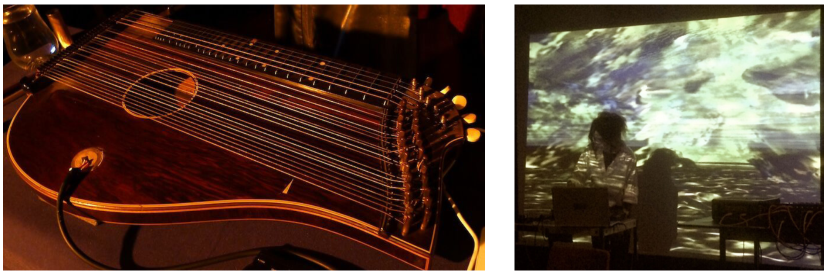

**Fig. 2**
Zither and visual projection of AG#2 (http://adrianasa.planetaclix.pt/research/AG2.htm).

The instrument combines a custom concert zither and a 3D software called AG#2, which operates based on amplitude and pitch analysis (Fig. 2). The software was developed for the audio input of a particular zither, which has aged strings and a personal tuning system. The 3D engine was written by John Klima, using an iOs / Android system from video games called Marmelade and an audio library called Maximilian. I specified the software design, created the audio, the mappings, the 3D world and the parameterizations.

### 3.1. Amplitude and pitch analysis

Both acoustic and digital music instruments can be designed so as to convey a sense of embodiment and a sense of physical separation. However, there is a fundamental difference, which can be explored so as to convey a certain amount of real-time effort in sound organization. With an acoustic instrument, the sonic output depends directly on physical interaction. In contrast, software mediates physical interaction through code. Code embeds theories informed by specific purposes and criteria, which are usually concealed to the "user". An advantage of designing software from scratch is that one can think of the process of amplitude and pitch analysis as creative material. Amplitude, or loudness, is the maximum displacement of the sound wave from a zero level position; it is unrelated to frequency, which is inversely proportional to wavelength. The amplitude of a digitized sound wave varies between -1 and 1, being 0 equal to silence, but algorithms for amplitude analysis usually consider the absolute value so that amplitude varies between 0 and 1 (Fig. 3). The software collects a number of samples from the audio input (discrete analysis) and calculates the average amplitude value. The number of collected samples corresponds to the size of the sample buffer. In the AG#2 software, the sample buffer for amplitude analysis is small, thus the average value is accurate. The software can respond inconsistently to amplitude variations, but that is because the amplitude detection threshold is high; it is a constant, which I set prior to performance.

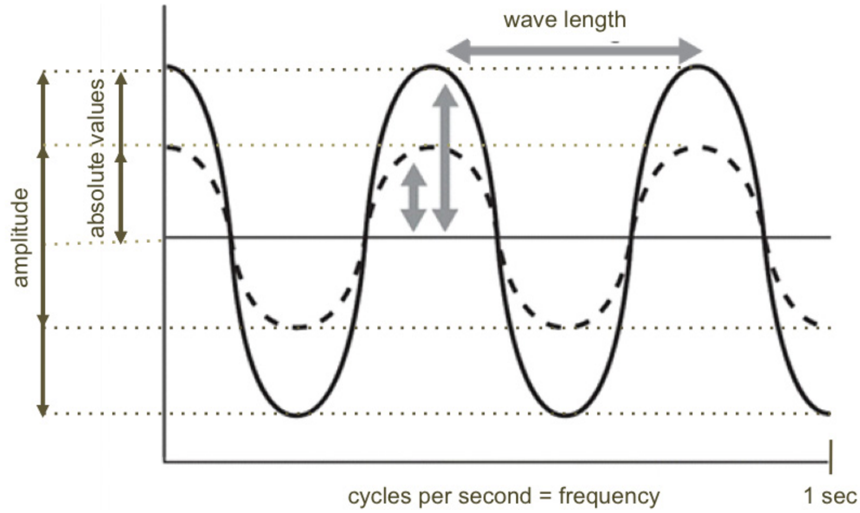


**Fig. 3**
Two sine waves with different amplitude, but same frequency and wavelength.

Pitch, i.e. frequency analysis is far more complex than amplitude analysis. Sounds include a range of frequencies (spectral shape), from which the software must extract the fundamental based on mathematical calculations. The question is what the notion of "fundamental" entails.

Usually, pitch analysis involves windowing the audio signal, as happens with amplitude analysis. A Fast Fourier Transform (FFT) function produces a series of values that represent the amount of energy, in a range of equally spaced frequency bands over a given number of samples. The frequency resolution depends on frequency bands (called FFT size). There are important limits on the information we can get from the Fourier transform of a time-limited signal extract, particularly with a small FFT size. Pitch is audible frequency, and frequency is the rate per second of a vibration constituting a wave. A very short fragment of signal does simply does not contain sufficient information about periodicity. Yet, a large window creates latency, and even a little amount of latency can break the sense of immediacy in interactive systems. Musicians can detect reaction latencies of 20-30ms in musical instruments (Mäki-Patola and Hamaläinen 2004; Adelstein et al., 2003), and the accepted target with interactive audio systems is latency under 10ms (Freed et al. 1997).

Usually, the design of interactive digital systems endeavors to negotiate frequency resolution vs. latency so as to provide a sense of immediate interaction. The construction of musical time requires a sense of immediate interaction, but that is not a concern in my software design; immediacy comes from the acoustic zither. Rather than attempting software to create a sense of immediacy, my creative strategies seek to emphasize, i.e. take advantage of the disparities between human perception and digital analysis. In fact, these disparities are unavoidable: whereas software operates based on mathematical calculations, humans sample and process the information based on attention, cognitive principles, and cross-sensorial context. Our percepts are always informed by expectations and concepts derived from past events. Conversely, software isolates the momentary input, and it responds accordingly.

The discrepancies are tangible. For example, a sound may vary in pitch during attack, sustain and release, and nevertheless, we group those pitch variations and segregate the sound from the soundscape. In contrast, the software slices the spectrum according to a buffer size, which may lead to overtones or resonance frequencies to be extracted as fundamentals. Also, an overtone can become intense as we optimize perceptual resolution, without the pitch being fundamental according to mathematical formulas.

## *3.2. The Zither Tuning, the Arpeggio-Detuning & the Interfaces*

My zither strings are from bass guitar, electric and acoustic guitar, and Portuguese guitar. Some are purposefully aged, which makes their timbre less shiny. I adopted a consistent, personal tuning to B 440Hz: around Bb, B, D, Eb, E, F and F#, but never exactly. The strings are played in any combination with hands, bottleneck, pick or bow. If the zither were plugged into a guitar tuner, the tuner would display a succession of different values upon a single string or chord.

The AG#2 software uses two streams of data from pitch analysis, as well as their mathematical difference. One data stream corresponds to the detected, fundamental frequency, calculated with a Fast Fourier Transform. The other corresponds to the nearest tone / half tone—A, A#, B, C, C#, D, D#, E, F, F#, G and G#. These tones/half tones are not played back. They are further mapped to pre-recorded sounds (octaves are disregarded), so that each audio input detection causes a corresponding pre-recorded sound to playback two times. The result is not repetitive because the second play back is pitched down, i.e. detuned according to a variable value—namely the difference between the detected pitch and the closest tone or halftone. We can think of this as a tuner: the difference between the detected pitch and the closest tone/ half tone is "displayed" in audible ways.

The software collects 4096 samples (buffer size) from the zither input, at 8000 Hz. It calculates the fundamental frequency—the detected pitch. The mapping between the detected pitch and the closest tone / half tone narrows the input to 128 values. The next step is called a chromagram. It gives the "name" of the closest tone / half tone, establishing twelve "frequency groups". Once one of the 12 notes is detected, the software plays a corresponding audio file. Again, the file has no frequency correspondence to the note—rather, based on compositional criteria, a sample is carefully selected to correspond to the note. The sample is then played a second time, pitching the sample down by the difference in frequency of the "pure" note and the result of the FFT analysis. It is a simple file playback modulation, thus the playback duration of the sample is extended as well. The digital sounds are stored in three banks, each containing twelve pre-recorded sounds. There are two interfaces for digital audio when I play the instrument: the pitch analysis from the zither input and the computer keyboard. Pitch analysis brings unpredictability, namely regarding which specific digital sound is played back upon detection, and how much pitch down is applied to its second playback. Yet the keyboard enables me to choose musical sections: input detection will only trigger sounds from the currently activated audio sample bank.

## *3.3. Musical forms*

The combination of acoustic and digital components conveys a musical language formed of surreptitious chromaticisms and timings, where expression comes from avoiding easy developments. While the digital sounds create certain unpredictability, the acoustic immediacy of the zither enables the music to shift in good time and direction.

Unpredictability must be calibrated so as to convey expression without being disruptive. Radical discontinuities attract attention automatically, creating points of high intensity in musical motion. They need to be sparse and very precise, which requires me to have direct control over the instrument; hence, I solely create radical discontinuities with the zither. The digital sounds never create radical discontinuities, because they are always preceded by a zither sound, and the amplitude detection threshold is high.

In total, the audio sample banks contain thirty-six pre-recorded sounds. A digital sound can be short or long, simple or complex. It can also have a shorter or a longer attack (i.e. the time it takes to reach the maximum amplitude), sustain, and release (i.e. decay time). Short attacks and/or releases create greater discontinuity than long ones.

I do not have direct control over which sound is triggered upon detection, which means that each audio sample must combine well together with the other eleven samples in the same bank. Each sample can become part of many different musical shapes. As the samples combine with each other and the acoustic sound, perception will stream and segregate the component parts depending on the musical motion as a whole.

The musical shapes can become quite complex, and that is likely to create density. A single amplitude detection point can activate several digital sounds, and each is played back twice. Moreover, the pitch shift applied to the second playback stretches the original length of the sound. A digital constraint is implemented, which neutralizes audio activation whenever the detected pitch is mapped to the same sound than the previous. As such, the density of the sonic construction depends on: a) the intrinsic density of each audio sample, b) the time length of the sample, c) the loudness of the zither relatively to the detection threshold, and d) the speed of my zither playing; usually I do not play loud at high speed, so as to leave space for musical details.

I developed specific zither playing techniques with each three audio sample banks, creating a set of versatile musical forms — musical vocabulary, which can be used to create different works. Solo and collaborative audio recordings are at http://adrianasa.planetaclix.pt/research/AG2.htm#SOUND.

*Combination 1*. With audio sample bank 1 the zither is usually dribbled and played with the bow. As an input to the software it activates sounds of bass guitar, ocean waves, water drops, thunder and wind. The combination creates a changing continuum where the chromatic and textural variations from the zither merge into sounds of nature, submerging and emerging from the landscape. In the merging of acoustic and digital sounds, at times one shapes the attack, and the other shapes the release. Long-lasting bass guitar samples modulate the landscape with a dense sonic body. Their continuity allows for attention to focus on other streams of sound. But their appearance and disappearance attracts attention, creating points of higher intensity. Their disappearance is often associated with the resolution of a musical phrase, be it resolved with the zither, or the sound of a single water drop.

At times the music can become very dense and complex with bank 1. Attention is then likely to shift away from details, as perception decreases resolution in order to embrace the whole. At lower perceptual resolution one perceives progressive continuities, where successive events seem to display a similar interval of motion, fulfilling the expectation that once something begins to move in a certain direction, it will continue to move in that direction (Gestalt of good continuation). When density and complexity then faint away, attention shifts to detail. The listener increases perceptual resolution, due to attention. At high resolution, ambivalent discontinuities become intense. The overall semantics of this acoustic / digital combination are environmental, but they entail an expressive semantic dimension as well.

*Combination 2*. With sample bank 2, the zither is played with hands, bottlenecks and pick. The samples are from bass guitar, dobro (metal body guitar), and zither (played with metal pick and bottleneck). All samples have short durations and short attacks. Similarly, I play the zither with interruptions and silences. The combination of acoustic and digital sounds leads to an expressive pathos, as I avoid easy developments in the musical phrasing. In contrast with Combination 1, where the musical motion seems driven by nature, now the musical motion seems definitely human-scaled.

Silence is intense if it emerges unexpectedly, and sounds gain intensity when preceded by silences. In both cases, the discontinuity attracts attention, leading to an increase of perceptual resolution. With bank 2, sometimes the complex musical phrasings lead perception to decrease resolution, so as to embrace the whole. That is unlikely to last though, because attention is constantly being attracted by rapid musical developments. The overall semantics of the musical motion is expressive, as attention is drawn to the performer's expressivity.

*Combination 3*. With sample bank 3 the zither is played with bow and bottleneck, activating piano notes and digital timbres. The digital timbre samples are quite long, rich in bass, with gradual attacks and releases; their layering sustains sonic density. The body of the sound is inlaid with a wealth of surreptitious chromaticisms, unfolding at a textural level. The zither bow plays in unison, emerging from the sonic stream, so as to submerge once more. The piano samples — low keys — create soft punctuation.

The music unfolds at a slow pace, as if it were driven by forces greater than human time. There are no radical discontinuities. While sustained sonic density creates a space of overall continuity, conscious awareness is invited to focus upon ambivalent discontinuities, which makes the chromaticism of harmonics intense. At times attention may also shift away from details, and drift to internal states. The overall semantics of musical motion are environmental.

The graphics in Fig. 4, Fig. 5 and Fig. 6 represent the relation between interaction, dynamics and semantics in each described musical form. The dot on the top axis shows that sound organization requires medium real-time effort; this is a constant. The gradient stripe on the axis bellow represents steady continuities (SC), progressive continuities (PC), ambivalent discontinuities (AD) and radical discontinuities (RD). The darkest part of the stripe shows the dominant type of continuities / discontinuities. And the dots on semantics represent the informational, expressive and environmental dimensions of each musical form. A single axis suffices to represent expressive and environmental semantics, because they are inversely proportional. Informational semantics can reinforce one or the other, hence the additional axis.

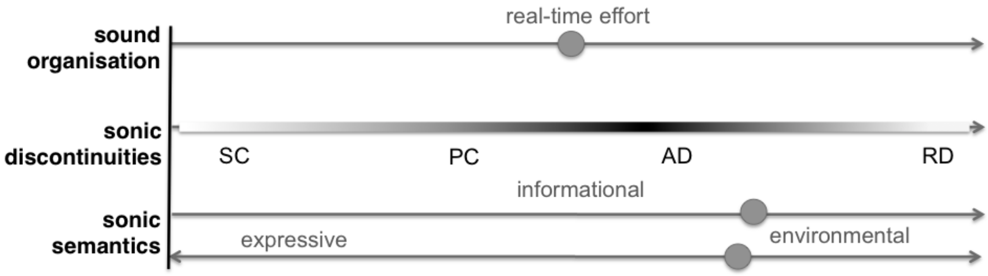


Fig. 4
Interaction, dynamics and semantics with Combination #1.

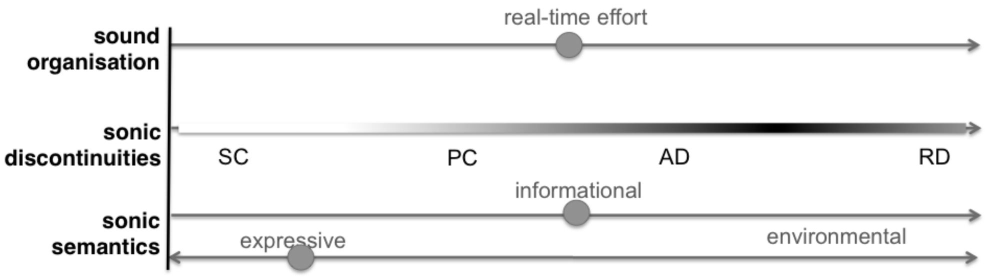


Fig.5
Interaction, dynamics and semantics with Combination #2.

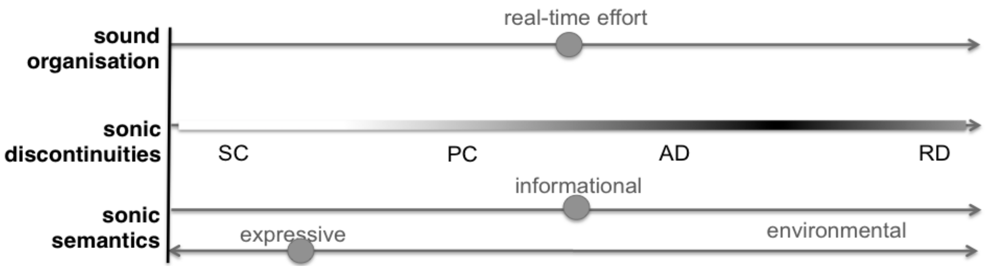


Fig.6
Interaction, dynamics and semantics with Combination #3.

My solo performances are never equal, but the musical forms maintain their dynamic and semantic characteristics. The same occurs when I play with musicians who comply with those characteristics. I also find it compelling to play with 'non-complying' musicians. For example, a great amount of radical discontinuities can cancel the environmental semantics of combination #1 and #3. And recognizable electronic devices can increase the informational semantics of the music. In any case, the interaction with my instrument requires medium effort, whether I play solo or collaboratively. Conversely, the semantics and dynamics of the music are variable, and the methods proposed in this paper can be further used to analyze the instrument's versatility.

## 4. CONCLUSION

My perceptual approach to instrument design and composition grounds an understanding of musical expression that embraces a diversity of musical idioms. Its distinguishing factor respects to interaction design: sound organization must require certain real-time effort, so that expression emerges from the reciprocal interaction between performer and instrument. This diverges from the dominant paradigm in ubiquitous interface design, which aims at dematerializing the digital interface. When designing software that operates based on an audible, acoustic input, one can face the disparities between human perception and digital analysis as creative material. Unpredictable digital behaviours create tension, increasing neural activity; they create points of intensity. Musical motion can then also shift to a release, as digital constraints rule out undesired outcomes, and the acoustic interface enables the performer to shift the music in good time and direction.

The paper specified three complementary methods that can be used to analyze any musical instrument, composition and performance. 1) Regarding interaction, the distinction between little, medium and high effort is useful to analyze whether an interface conveys the present notion of expression. 2) Regarding the dynamics of music, the taxonomy of continuities and discontinuities is useful to analyze how musical motion drives attention. 3) Regarding the meaning of music, the proposed semantic characterization is useful to describe how the music draws attention to the performer or the environment.

The paper used these methods to describe an instrument that enables a set of versatile musical forms, with characteristic types of continuities / discontinuities, and multiple semantic dimensions. The methods were useful to describe how these forms can draw attention to the performer, shape an environment, and extend one's sense of presence beyond the physical performance space.

# REFERENCES


__*Immersed in Media: Telepresence Theory, Measurement & Technology*; ed. M. Lombard, F. Biocca, J. Freeman, W. IJsselsteijn, R. Schaevitz. Springer International Publishing (2015) :73-94

**Adelstein, B. D.; Begault, D. R.; Anderson, M. R. and Wenzel, E. M.** 'Sensitivity to Haptic-Audio Asynchrony'. In *Proceedings of Multimodal Interfaces Conference* (2003):73–76.

**Bertelsen, Olav W., Morten Breinbjerg, and Søren Pold,** 'Emerging Materiality: Reflections on Creative Use of Software in Electronic Music Composition.' In Leonardo 42, no. 3 (2009): 97–202

**Cage, John, Silence,** 'Lectures and Writings 1939 - 1961', Wesleyan University Press, Connecticut, 1961.

**Cascone, Kim,** 'The Esthetics of Failure: Post-Digital Tendencies in Computer Music'. In *Computer Music Journal* 24 no. 4 (2000): 12-18.

**Dourish, Paul,** 'Where the Action Is: The Foundations of Embodied Interaction', MIT Press, 2004.

**Freed, A., Chaudhary, A. and Davila, B.,** 'Operating Systems Latency Measurement and Analysis for Sound Synthesis and Processing Applications'. In *ICMC Proceedings* (1997): 479–81.

**Gritten, Anthony,** 'Drift'. In *Music and Gesture, ed. Anthony Gritten and Elaine King*, Ashgate Publishing Limited (2006):104-125.

**Johnston, Andrew,** 'Beyond Evaluation: Linking Practice and Theory in New Musical Interface Design'. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (2011), 280–283.

**Kiefer, Chris,** 'Interacting with Text and Music: Exploring Tangible Augmentations to the Live Coding Interface'. In *Proceedings of ICLI 2014 — International Conference on Live Interfaces* (2015) :251-259.

**Kilbourn, K., and Isaksson, J.** 'Meaning through doing: The role of affordances over time' In *Sixth Nordcode Seminar & Workshop*, Design Semiotics in Use (2007): 6-8.

**Knudsen**, 2007?
**Machover**, 2008?

**Mäki-Patola, T and Hämälainen, P.,** 'Latency Tolerance for Gesture Controlled Continuous Sound Instrument Without Tactile Feedback'. In *ICMC Proceedings* (2004).

**Magnusson, Thor, and Mendieta, Enrike H.,** 'The Acoustic, the Digital and the Body: a Survey on Musical Instruments.' In *NIME '07: Proceedings of the 7th International Conference on New Interfaces for Musical Expression*, New York, NY, USA, ACM (2007): 94–99

**Mudd, T., Holland, S., Mulholland, P. and Dalton, N.,** 'Nonlinear Dynamical Systems as Enablers of Exploratory Engagement with Musical Instruments'. In *Proceedings of ICLI 2014 — International Conference on Live Interfaces* (2015): 172-182.

**Pressing, Jeff,** 'Cognitive Complexity and the Structure of Musical Patterns'. In *Generative Processes in music — The Psychology of Performance, Improvisation, and Composition*, ed. by J. Sloboda, New York, Oxford University Press (1987), 129-178.

**Pressing, Jeff,** 'Some Perspectives on Performed Sound and Music in Virtual Environments'. In Presence 6, no. 4 (1997).

**Rebelo, Pedro,** 'Instrumental Parasites: Interfacing the Fragile and the Robust'. In *Proceedings of ICLI 2014 — International Conference on Live Interfaces* (2015): 241-250.

**Ryan, Joel,** 'Some Remarks on Musical Instrument Design at STEIM', *Contemporary music review*, Vol. 6 issue 1 (1991), 3-17.

**Sa, Adriana,** 'How an Audio-Visual Instrument Can Foster the Sonic Experience'. In *Live Visuals*, eds. L. Aceti, S. Gibson, S. M. Arisona, O. Sahin, Leonardo Almanac vol. 19 No. 3, MIT Press (2013): 284-305.

**Sa, Adriana, Caramieux, Baptiste, and Tanaka, Atau,** 'The Fungible Audio-Visual Mapping and its Experience'. In *Journal Of Science And Technology Of The Arts* vol. 6 No. 1 (2014): 85-96.

**Sa, A.; Ryan, J.; McPherson, A.; Tanaka, A.; Magnusson, T.; Van der Heide, E.; Grierson, M.; Carvalhais, M; McLean, A.** 'Live Interfaces: Problematising Foundations'. In *Proceedings of ICLI 2014 — International Conference on Live Interfaces* (2015): 14-28 http://www.liveinterfaces.org/ (accessed January 2017)

**Schroeder, Fraziska and Rebelo, Pedro,** 'The Pontydian Performance — The Performative Layer'. In *Organised Sound* 14, 2 (2009)

**Snyder, Bob, Music and Memory,** *'An Introduction'.* Cambridge, MIT Press, 2001.

**Vasulka, Steina, and Vasulka, Woody,** 'Machine Media'. New York, MoMA, 1996.

**Weibel, Peter, and Gerhardt J. Lischka,** 'Polylog. Für eine interaktive Kunst' (Polylogue: For an Interactive Art). In *Im Netz der Systeme*. Ed. Peter Weibel and Gerhard Johann Lischka. (1990): 65–86.