# Instrumental Listening: sonic gesture as design principle

**Article** *in* Organised Sound · August 2009

DOI: 10.1017/S1355771809000284



**1 author:**

Doug van nort
York University

# Instrumental Listening: sonic gesture as design principle

DOUG VAN NORT

Department of Music Research, McGill University, Montreal, QC, CANADA
E-mail: doug@music.mcgill.ca

**In the majority of discussions surrounding the design of digital instruments and real-time performance systems, notions such as control and mapping are seen from a classical systems point of view: the former is often seen as a variable from an input device or perhaps some driving signal, while the latter is considered as the liaison between input and output parameters. At the same time there is a large body of research regarding gesture in performance that is concerned with the expressive and communicative nature of musical performance. While these views are certainly central to a conceptual understanding of 'instrument', it can be limiting to consider them a priori as the only proper model, and to mediate one's conception of digital instrument design by fixed notions of control, mapping and gesture. As an example of an alternative way to view instrumental response, control structuring and mapping design, this paper discusses the concept of gesture from the point of view of the perception of human intentionality in sound and how one might consider this in interaction design.**

## 1. INTRODUCTION

In the majority of discussions surrounding the design of digital instruments and real-time performance systems, notions such as control and mapping are seen from a classical systems point of view: the former is seen as a variable from an input device or perhaps some driving signal, while the latter is considered as the liaison between input and output parameters. At the same time there is a large body of research regarding gesture that is concerned with the expressive and communicative nature of human performative action. While these views are certainly central to a conceptual understanding of 'instrument', it can be limiting to mediate one's conception of digital instrument design entirely through them. As an example of an alternative way to view instrumental response, control structuring and mapping design, this paper will discuss the concept of gesture from the point of view of the perception of human intentionality in sound and how one might consider this in the process of interaction design. I will examine and reflect upon the ways in which gestures – in the sense of dynamics that are either a result of or which suggest human action – are embedded as a trace within a sound signal, and how these might recursively be embedded within a mapping proper. To this end I introduce the notion of 'sonic gesture' in order to extend the notion of instrumental mapping design, evoking the latent embodiment in Schaefferian theory.

A primary goal of this work is to move away from discussions of mapping in isolation – as simply a connective tissue between control and sound parameters – instead viewing it as a process in larger control structuring that is very much an interplay between two complementary aspects of musical performance dynamics, namely human actions and abstracted musical dynamics. In the case of human gestures, the literature is vast and often covers communicative visual aspects of performance. In this work, my consideration of actual (as opposed to imagined) human action is similar to the gestural primitive notion of Choi (2000) in that the concern is working with the dynamic and organic nature of such control signals in data space. To make this distinction, I will refer to these particular gestural objects as *control gestures*.

At the same time, the majority of this article will focus on the gestural nature of musical dynamics and their interplay with mapping and control structuring. In designing an instrument, of course one is considering the sonic material that will be controllable. Beyond this, however, it is the *shaping* of said material as a function of human action that must be designed – and that is so often implicitly embedded within the mapping strategies employed to this end. Many researchers have explored the relationship between the experience of listening to music and creating mental images of human movement (Hatten 2004; Delalande 1988) in the context of classical music rendered from acoustic performances. One hypothesis of this research is that such *musical gestures* that evoke a sense of embodiment or 'anthropomorphic projection' (Chadabe 2004) extend to the experience of listening to electroacoustic music. Without delving too deeply into the discussion on defining electroacoustics, in the context of this discussion let the term come to mean musics primarily focused on formal development through shaping of timbral and textural sound qualities. Just as most instrument design arguably does not focus deeply on sonic sculpting, I will not pretend that this work extends to all instrumental design contexts, but rather those in which control of timbre and texture is the main concern.

What this article presents, then, is both an analysis framework and design methodology that are shaped

*Organised Sound* **14**(2): 177–187 © 2009 Cambridge University Press. Printed in the United Kingdom. doi:10.1017/S1355771809000284

by theories and views normally associated with the out-of-time electroacoustic tradition. As I wish to avoid delving too deeply into stylistic or ontological discussions about the musical aspect of this established notion of musical gesture, I will instead favour the term *sonic gesture*, which further highlights that the focus here is on signal and sub-to-note-level phenomena rather than larger elements of musical form often described using the former term.

## 2. CONTROL GESTURE

One of the constraints of the design methodology presented herein is a preservation of the immediacy of action–sound coupling, so that by control gestures I do indeed mean direct human actions that are functional in nature, so that we may consider them as existing both in control signals as well as human-input gestures. This is precisely because they are *instrumental* in nature (Cadoz and Wanderley 2000) and so are intentionally acted in such a way as to embed them in the resulting control data stream. While a complete analysis of instrumental gestures is beyond the scope of this work, suffice to say that these encompass *sound-producing* gestures (Godoy, Haga and Refsum-Jensenius 2006b), which may manifest as continuous or impulsive *excitation gestures*, the continuous or discrete modulations of *parametric modification gestures* or *structural* modification gestures that alter the makeup of the instrumental system. Note that while I focus here on traditional acoustic musical instruments, the discussion extends to the more general notion of sound-making actions performed with musical intention on physical objects.

Such performative gestures are directly transduced into control gesture as they, by definition, assume that human contact with an object is taking place, with a skilled manipulation and an 'existence of an energy continuum between the gesture and the perceived phenomena' (Cadoz and Wanderley 2000). Control in an electroacoustic context, however, is often geared towards driving elements of sound that are not directly controlled (in a parametric sense) by sound-producing actions in an acoustic instrumental context. Rather, such sounds arise from the sustain portion of a musical event – not controlled by direct energy transfer to the resonating body, but instead guided in an indirect and nonlinear fashion. For example, one may shake an object that includes a spectrally dense resonator in such a way as to influence the roughness of its output. Many other acoustic examples exist (e.g. manipulating bow/string interactions) in which one indirectly controls sustained textural elements that may be clearly perceived as separable in time – acting on a different timescale from timbral elements – and that conjure images of gestural motion.

Following this principle of *guiding* textural elements leads to actions such as those that Godoy et al. (2006a) have called *sound-tracing* gestures, which can be thought of as outlining general contours, in response to imagined or actual sonic stimuli. I consider such gestures to be central in designing control structures for those timbral and textural sound qualities that are themselves central in electroacoustic music. In this way embodied control gestures for electroacoustic music may find inspiration in the direct actions of acoustic instrument performance as well as the indirect human response to musical stimuli in a non-performative context.

## 3. SONIC GESTURE

Through the ubiquitous experience of watching live music, we are able to clearly identify performer gestures upon hearing recordings of instrumental music performance. Going beyond identification, researchers in cognitive psychology have put forth theories of *embodied cognition* for listening in which it is suggested that perceiving sounds (e.g. vocal utterances (Liberman and Mattingly 1985)) is a process of creating mental images of the generating articulatory gestures. These thoughts have been applied towards a 'motor-mimetic' theory of music perception (Godoy 2003), positing that the perception of musical sounds is inextricably linked to the creation of mental representations of sound-producing gestures.

I assume this embodied cognitive lens in order to examine theories normally associated with 'fixed' or non-real-time approaches to sonic media. In doing this I take as a starting point work presented by Godoy (2006), in which the author provides a reading of Schaeffer (1966) that evokes the gestural and embodied nature of these writings through his proposed notion of *gestural sonorous objects*. In doing so, the author essentially elucidates an often-overlooked aspect of Schaeffer's theory of sound objects: that he in fact proposes a link from his typology/morphology of sound objects back to an appropriate sound-producing gesture suggested by a given sound object. This can be clearly seen in Schaeffer's typological notion of temporal envelope and related morphological notion of dynamical form (Schaeffer 1998), qualified in terms of human action whose imprint is left on the sound. In the English translation this is presented as 'execution' type in concordance with the notion of 'facture gestuelle' (or 'executive gesture') as the sound-producing action. The gestural quality of the sound object is then considered in terms of the overall temporal envelope that Schaeffer breaks down into impulsive, sustained and iterative, which are in turn paralleled with punctual, continuous and iterative gestures by Godoy (2006). Considering again the typology of instrumental gestures put forth by

Cadoz and Wanderley (2000), we can further draw analogy with these temporal forms and a(n):

- instantaneous excitation gesture
- continuous gesture, either excitation or modification
- periodic gesture, either discrete or continuous, excitation or modification respectively.

It is quite easy to imagine these various control gestures giving rise to the aforementioned envelope types. However, the link with overall temporal envelope, and implicitly with temporal scale, is only a first-order view of the gesture-sound object experience, which we can press further towards the construction of mapping and control strategies. In using language such as 'mapping' I don't mean to suggest that one may map directly between 'gestural objects' and note-level sonic objects, as such an association is determined by high-level functional and perceptual criteria that one cannot directly parameterise, but rather that the constant flux between gesture and sound objects must be constrained. Setting the conditions for this interplay between perceived sonic gesture and appropriate control gesture is a global view on mapping – one that paints a more complete picture of the process. It is in contrast to mapping-as-correspondence, which preserves the out-of-time, flowchart notion of mapping that neglects the *perception of gestural dynamics*. A more complete view of mapping includes the *perceived result* of a control gesture–sonic gesture causality. I propose that if the state of action and sound dynamics are considered from the point of view of gesture–sonic-object qualities, then such a *perceptual, embodied control design* can result through considering this *in a design feedback loop with the creation of mapping strategies*. Taking such an approach suggests that rather than isolating a perceptual parameter such as roughness or brightness a priori, one engage in the phenomenological approach of observing the qualitative sonic output that results from a given control structure, in order to reveal the perceptually relevant parameters that are being directly or indirectly driven. The lens through which to view this output is an analysis of sonic gestures. In this way, from a signal point of view 'perceptually relevant' is a product of perceived gestural intent.

### 3.1. From sound objects to sonic gestures

The morphological concepts arising from Schaefferian analysis can suggest new unexplored gesture/sound links that one may use to structure the instrument design process. After a continuous sonic stream has been perceptually 'cut' using the principle of stress-articulation (Chion 1983), the resultant object's *morphology* describes both its form and matter[1] (Schaeffer 1998), where the former relates to the global properties of a sound and the latter describes its internal characteristics. Having identified a sonic object, one may examine its morphological properties for traces of perceived action. Space does not allow for a complete breakdown of those qualities described by Schaeffer (1966, 1998) but a few words must be said about those most relevant to sonic gesture understanding:

- *Dynamic Profile* is directly related to a given sound's articulation and its energy envelope over time. It is the quality most obviously related to gestural control.
- *Mass*, as Chion notes, is the 'way of occupying the pitch field'. It describes the complexity of the spectrum: pitch coherence, spectral span peakedness vs. flatness, and so on.
- *Motion* is a term translated from the french *allure*. It describes fluctuations that characterise a sound during the sustain portion, and so is integrally linked to the notion of modulation gestures acted subsequent to an initial excitation. It can be considered as a generalisation of tremolo or vibrato (Thoresen 2007).
- *Harmonic Timbre* (HT) relates to the relative balance of the spectrum. It interrelates with mass (Schaeffer 1966; Chion 1983), in that the former relates to spectral placement while harmonic timbre describes the articulation or colouration of this.
- *Grain* relates to the micro-level structure of sound objects, including both spectral and temporal properties. It encompasses the notion of *sonic texture*, which describes sound events that are globally stable or stationary with local fluctuations or non-stationary elements (Strobl, Eckel and Rocchesso 2006). Grain is explicitly related to gestural actions by Schaeffer in that it is further broken down into resonance, rubbing and iteration sub-types. In a more general and signal-focused view, I introduce the terms *spectral grain* and *transient grain*. The former encompasses the notion of resonance grain and any similar phenomena wherein the causative factor is primarily a mass or spectral feature, such as roughness resulting from dissonant tones. Meanwhile, transient grain refers to those that are primarily a time-domain phenomenon, resulting from many micro-transients within the signal. It is appropriate to continue using the iteration grain terminology here in that it exists on the boundary with the perception of gestural iteration, as suggested by motion or by an overall iterative dynamic profile. Beginning from an idea suggested by Chion (1983), Schaeffer (1966) and Thoresen (2007), I further qualify grain phenomena by virtue of identifying the spectrum, weight and placement of the grain element of a sound (as separated from the sound proper).

[1]This terminology is a combination of that presented in Chion 1983 and Thoresen 2007.

### *3.1.1. Matter profile as gestural description*

Where these elements describe the overall quality of a given sound object, a consideration of the gestural nature of said object means examining the dynamics of morphological features. To that end, consider a matter profile as consisting of the time-varying aspects of all of the matter-related sound qualities: mass, HT, motion and grain. In his writings Schaeffer referred more specifically to mass and pitch-based profiles (Schaeffer 1998; Chion 1983), but the concept is extended here to all matter criteria, as these contribute to an overall image of a sound gesture and are used (particularly in electroacoustics) to convey a sense of motion. A spectrum exists between phenomena such as mass and HT or motion and grain, with the focus shifting from one to another throughout the life of a sound. Therefore, examining the co-varying nature of all matter profiles will paint a more complete picture of a given sonic gesture. Further musical implications and justification for this extension of matter profile arise naturally from the dual principles of gesture and texture, as will be discussed in section 3.3. From an instrumental point of view, a complete view of the dynamic profile of all form/matter properties may suggest a given control gesture, and a novel way to constrain the mapping design: constructing a system's 'sonic gestural response' in the spirit of analysing an acoustic body's frequency response.

## 3.2. Perceptual criteria of sonic gesture analysis

Timbre similarity studies focus on *musical listening* (Gaver 1993a, 1993b), a mode in which one is listening for properties of the sound itself, as opposed to listening oriented towards properties of the source. The latter is an assumption that dominates in the case of ecological perception studies,[2] as is evident from the nature of psychoacoustic experiments and the isolated sound events employed (e.g. Lakatos, McAdams and Causse 1997; Giordano and McAdams 2006). Though it may be that the properties of materials are most perceptually salient in a given listening context, in a complex musical situation the interaction between excitation and material might prove to be more important as it provides information about the 'expressive' nature of the performer. Such continuous variations are left out of psychoacoustic studies of musical timbre or 'non-musical' ecological studies, being attributed to such expressive performer deviations in the former case or simply not entering into discussion in the latter.

In the case of electroacoustic music, continuous variations such as modulations of intensity, frequency or density play an important role in defining musical form. Taking this position means that these phenomena – in addition to contributing to performer 'expressivity' found in the ornamentation of a gestural-sound object – can also be contributors to form-bearing elements of a musical piece. I maintain that this is highly relevant in the context of a discussion on gesture and digital instrument models in that if we are to extend these views – in particular the notion of parameter mapping – then we must reconsider our *implicit parameterisation of musical as well as perceptual phenomena*. In order to construct a mapping framework that considers gestural intentionality in sound – particularly in the abstract world of electroacoustics – one must consider the continuous interplay between sonic material, perceived human action and perceived musical form. Following these parallel roles of continuous sound modulation drive the analysis of sonic gestures, leading to a reconsideration of what a 'perceptual parameter' might mean in this context of navigating performer intent at the same time as musical meaning. This, then, takes the place of the standard practice of adopting sound features that arise in a laboratory listening context, having different musical assumptions.

In other words, while known perceptual features related to brightness, roughness, and so on clearly do arise as salient in many listening contexts, they may or may not be the main carriers of information in regards to gestural dynamics. To decide upon such perceptual relevance warrants a *phenomenological* approach – in the spirit of Schaefferian tradition – wherein it is the *perception of intentionality* (understood in terms of form/matter dynamics) that informs the design of new mapping and control structures. While certainly divergent from most presentations on instrumental design, this 'inverse mapping' approach is particularly appropriate for electroacoustics, again because of this interrelation of gesture and form.

## 3.3. Form/matter and gesture/texture: mapping considerations and structure-bearing principles

This notion of a global profile for a sonic gesture and 'internal' characteristics that generally relate to smaller-scale temporal and spectral changes (i.e. subnote level) is important for the consideration of sounds that are driven by human action. As noted previously, in the case of acoustic instruments note-level events have an overall shape that is a direct result of performer actions, with textural and timbral characteristics that result indirectly from actions such as slight bow angle/force modifications in a cello, or embouchure changes in a reed instrument. Considering the embodied nature of such a global form/internal matter breakdown from a reception/listener's point of view, Smalley (1997) has written about the

[2]Though one might argue that the work in Gygi (2001) is an exception that takes more of a 'musical listening approach'.

dual concepts of *gesture* and *texture* in the context of his work on spectromorphology. The author refers to a gesture as 'an energy-motion trajectory which excites the sounding body, creating spectromorphological life', further stating that 'when we hear spectromorphologies, we detect the humanity behind them by deducing gestural activity …'. Smalley then goes on to develop the electroacoustic music-theoretic notion of *gestural surrogacy*, differentiated by order ranging from immediate awareness of materials, through standard (acoustic) instrumental sound gestures, to abstracted shapes that cause uncertainty of source/cause, through to a so-called 'remote surrogacy' wherein the human element is lost as well as perception of source and cause. I maintain (as do many other sources, such as Emmerson 1986 and Wishart 1994) that the interplay between these types of sonic gestures – traversing different timescales as well as levels of abstraction from human action – are a primary way to create structure in an electroacoustic work. This can be related to the notion of musical gesture in instrumental music that is applied concurrently to temporal scales ranging from note-level events up to larger phrases (Hatten 2004).

In adopting this gesture–texture dichotomy, it is implied that gesture as a form-bearing principle is concerned with propelling (musical) time forward, moving towards (as well as away from) a particular goal. Smalley states that 'If gestures are weak, if they become too stretched out in time, or if they become too slowly evolving, we lose the human physicality … [moving towards an] environmental scale.' In the absence of gestural motion we are left with so-called texture, which is the lack of perceived motion but rather focuses on inner details of sonic material. While this concerns structural characteristics of electroacoustic music reception, it directly relates to short-term sonic features in that the former regularly emerges from the latter in music of this sort. In this way, the gesture/texture breakdown can be drawn into analogy with the note-level Schaefferian notion of form and matter. Rather than simply equating the two – gesture is form, texture is matter – it is the dynamics of form and matter as well as the interplay between the two that define gesture. These dynamics are the substrate from which imagined (human) gestures arise, and it is in their boundary – what Schaeffer called the 'criteria of sustainment' – that Smalley's texture-as-sound resides. In this instance the signal-model definition of texture (Van Nort 2007) – directly related to perception of non-motion – converges to the spectromorphological meaning. It is in the dynamics of grain, and between grain and motion, that electroacoustics offer new boundaries of imagined human gestures – which themselves are on the boundary between sound-tracing and sound-producing gestures. There is no clear and direct path – in terms of mapping action to sound – that gives rise to such gestural dynamics. Therefore it is worth examining this concept of gesture/texture and form/matter dynamics in the larger discussion of instrument models: to view the design of a mapping/control strategy in terms of the way that it drives these dynamic qualities, realising that it can have larger implications in terms of composing for the resultant instruments. It suggests an embodied approach to constructing a *composed instrument* (Schnell and Battier 2002), one which considers idiomatic gestural dynamics arising from the often radically different musical production/reception modes of electroacoustic music while maintaining the essential concern with the human element of musical performance practice.

### 3.4. Sonic gestural shapes

In order to distinguish amongst and compare sonic gestural shapes we may use Smalley's basic scheme to define those that can be characterised as *attack alone* (a purely impulsive sound object-gesture), *attack-decay* (attack with sharp decay in the closed case, gradual decay in the open case) or the so-called *graduated continuant* (GC: focus on the continuant or sustain portion, characterised by the presence of this as well as onset and termination). Smalley notes that at least three variations exist: (GC1) a 'compressing' of the onset to include more energy near the beginning, (GC2) a lengthening of the continuant phase, thereby drawing interest away from the onset, and finally (GC3) 'increasing the spectral energy towards termination, leading towards, and creating the expectancy of, a new note-gesture' (Smalley 1997: 113).

Using the generality of these sound shapes is a means to consider them from a reception point of view as well as to understand the gestural nature of their creation. There are many spectrotemporal morphologies possible within each of these classes of sonic gesture, and the process of directing and conditioning this space of possible form/matter dynamics should be properly considered as part of the mapping process in the context of instrumental design. Certain sonic gestures are indeed suggestive of particular control gestures. Therefore, even while we've re-considered the gestural nature of sound objects from a sonic-phenomenological perspective – not pre-supposing an acoustic instrumental paradigm – it is useful in the context of an instrumental design process to reconsider the physical and perceptual nature of the control of acoustic instruments *from a sonic gesture point of view* rather than a physical gesture point of view.

## 4. FROM SONIC GESTURE BACK TO CONTROL GESTURE

The reason that I've taken the time to establish a more fluid notion of sonic gesture in the context of gesture/texture (from a musico-structural point of view) and form/matter dynamics (from a morphological point of view) is that it simply does not make sense to use sonic objects as some atomic building block from which to create formal elements in electroacoustic music in the way note events are used in acoustic music. Having said this, there are many subtle and important nuances that we may look to in the control/sound relationship of acoustic instruments, including the perceptual nature of control as well as sound for a given instrumental gesture.

This is articulated in Levitin, McAdams and Adams (2002), wherein the authors discuss the control of a musical note-event from the point of view of the required input gestures to create a given note type and the sonic perceptual parameters that may be affected. They break this performer gesture/perceptual sound feature link down into the beginning, middle and ending of a musical tone – paralleling the classic view of attack, sustain, decay – which I in turn translate into the more abstract and aforementioned notions of onset, continuant and termination. In doing so the authors create a framework that parallels the excitation/modification gesture typology of Cadoz and Wanderley (2000) and expands upon this to include psychoacoustic cues. In the process they introduce the notion that an *explicit* beginning is contrasted with a newly introduced event termed a *state change induction*, described as an *implicit* beginning that arises from some (contextually determined) spectral discontinuity. It is important to note that a key feature of an implicit beginning is that there must first be a continuous excitation from which they can emerge – think of a legato between two notes of a wind or bowed string instrument as a canonical example. Similarly, one may consider the parallel notion of explicit/implicit ending, where the former may result from damping or natural decay of resonances by stopping a gesture, while the latter exists in the same moment as the implicit beginning of another event. In regards to the middle portion of a musical event, the authors suggest that pitch modulations resulting from tension/pressure applied to an instrument's tone generator are the only relevant control actions in this context, but I argue that the 'subtle' timbral effects that can result from shaking a resonating body, for example, are particularly relevant (though not unique) to control of electroacoustic-inspired musical instruments where such timbral and textural variations are magnified in importance through amplification. This limit case of acoustic control via sound-producing gestures is an interesting meeting point with sound-tracing type gestures that might suggest novel control strategies.

In examining the typological thinking of Levitin et al. (2002) as well as Cadoz and Wanderley (2000) and Godoy (2006), we see that they consider the universe of possible gestures in a way that is mediated by the vibrating source/resonating body physical aspect of musical instruments. Levitin et al. (2002) implicitly suggest a notion of musical gesture through the consideration of separable and perceptual sound parameters that are affected, and what a player does in order to elicit a given response. Translating this into the electroacoustic realm, I propose to consider the given typologies of Cadoz and Wanderley (2000) and Levitin et al. (2002) regarding the control of the beginning, middle and ending of note events as a starting point to abstract towards control of sonic gestures: what *could* one have done in order to give rise to a given sonic gesture? This question suggests directions for control structuring. In the course of such an exploration, there is no need to presuppose an attack-resonance instrument model, but rather consider such sounds *in the larger context* of all gestural sound events that one may control.

This more generalised context suggests a modified notion of what constitutes a singular control gesture: in the exposition of Levitin et al. (2002) there is a supposition of one single gestural action constituting a single gestural unit. However, in the case of a sonic gesture, an iterative sound object may be understood as a unit – one that suggests a continuous or iterative excitation. Among other reasons, causality is extended in electroacoustics through the potential for repetition and automation. Therefore I consider explicit beginnings to include iterative excitations as a separate category from singular impulsive or continuous excitations. Note that this is in contrast to Cadoz and Wanderley (2000), where iterative excitation is considered as a subset of continuous gesture, having discrete excitation. Similarly, the notion of a state change induction that is brought on from a spectral discontinuity is very prominent in electroacoustic music, as sudden changes in matter are often used to signal a new sonic gesture/texture context. Applying the language that we use here, such a spectral discontinuity may exist at the termination of a sonic gesture in terms of the third type of graduated continuant, in which spectral energy is increased just prior to such a state change.

Following the sound-first principle of this article, I've thus paralleled morphological descriptors, sonic gestural shapes and control gestures, as presented in figure 1, which describes the onset, sustain and termination phases for a given sonic gesture and what sort of control is possible at each phase. In the case of attack sonic gestures, an impulsive excitation is the entire control gesture, and the dynamic profile and instantaneous matter features (e.g. pitch) dominate our perception. Attack–decay gestures are also impulsively excited, but there is a chance for control of the ending through damping, and for continuous excitation of the middle through the introduction of the concept of mode changing, which in this context

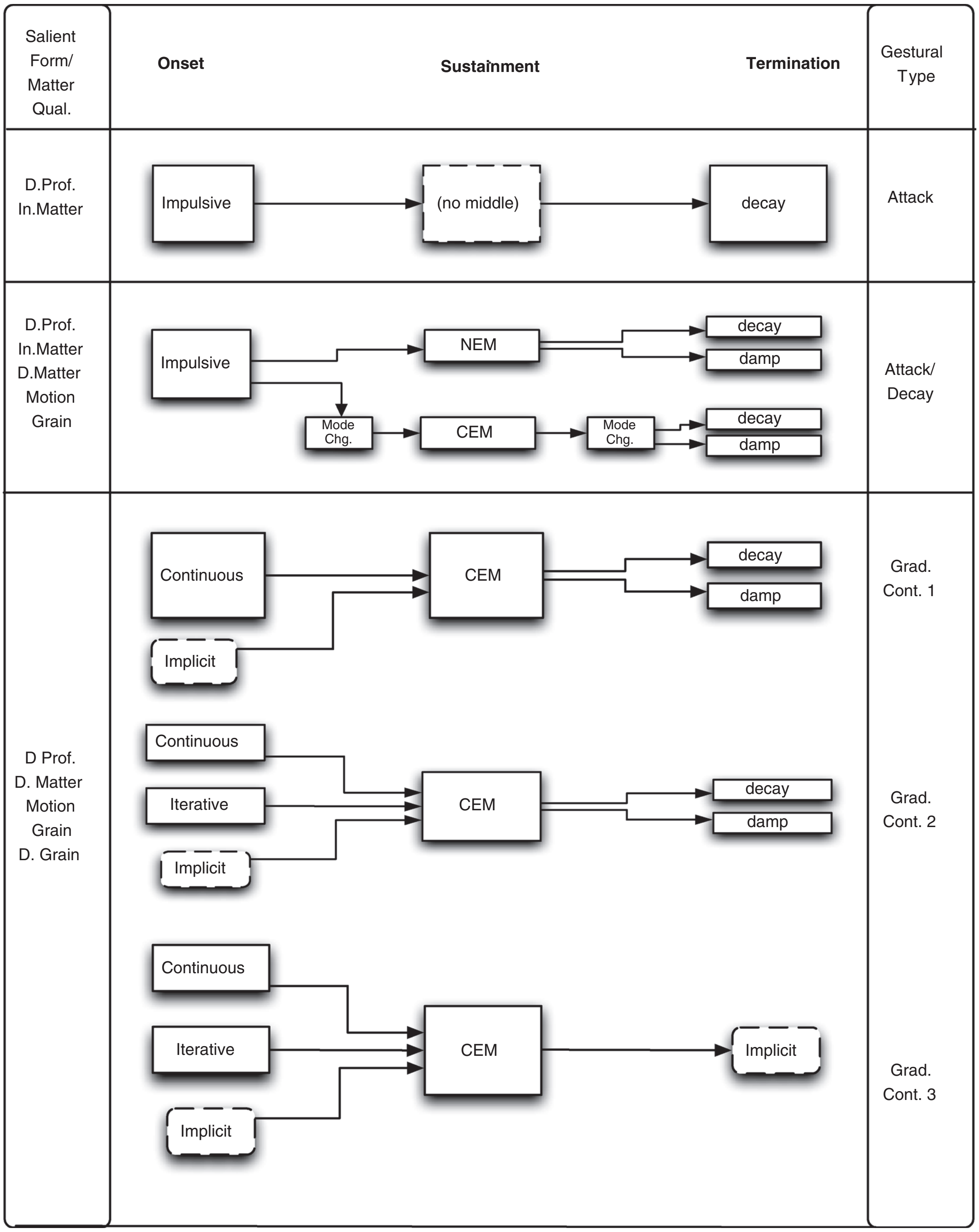


**Figure 1.** Possible control structures for different sonic gestures types (attack, attack/decay and graduated continuant 1, 2, 3), including control type for all portions of a sound's life. Also included are the form/matter features that are perceptually relevant and so may be affected (dynamic profile, instantaneous matter, dynamic matter, motion, grain and dynamic grain).

is enacted by an abstracted form of a structural modification gesture. Continuously exciting the middle of an attack decay gesture through mode change amounts to 'guiding' or otherwise affecting the decaying resonances. In this archetype there is added salience from the sonic features related to dynamic matter as well as motion and grain during the decaying sustain. In regards to the sonic gesture archetype of graduated continuant, each of its three variants may result from continuous excitation or from an implicit beginning, and by definition extend into a continuously excited middle (CEM). The swelled onset nature of the first type means it does not result from an iterative excitation – at least not in the sense of repeated application of the same gesture. However, this may be the case for the latter two.

Meanwhile, the first two types may terminate from damping or stopping gestures (i.e. allowing for decay), while the increasing of spectral energy towards the end in the third type leads, perceptually, to an implicit ending. All three of these sonic gesture archetypes possess salient dynamic form/matter qualities including motion and grain, in addition to possibly dynamic grain as well.

There is certainly no need to constrain the universe of sonic gestural shapes by strict instrumental control gestures. However, just as sonic gestures give rise to an imagined embodiment, so too are they mediated by an understanding of musical control. This fact has led me to the current typology as one example of structuring control based on sonic phenomenological principles – I don't claim that it is in any way absolute or complete. Remember that these gestural archetypes are very abstract, and do not relate to instrumental sound or strict control.

### 4.1. Recapitulation

Bringing these notions together presents a unified framework through which to understand as well as construct an instrument. We may consider an instrument in terms of the ideal sonic gestural shapes that it should produce, or those that it can produce. Desired gestural shapes may constrain control structures as with figure 1, which also suggests those elements of sonic gestures that are relevant to examine. Form/matter qualities provide a framework from which to build tools to analyse sonic gestural shapes, with the archetype of gesture shapes suggesting further methods of analysis (of attack, modulation, etc.). The details of how each dynamic form/matter quality evolves depends on musical context, and the principle of gesture/texture articulates the reception strategies at play in electroacoustic music and why using this meta-framework to design instruments requires *choices made by the instrument designer while listening to sonic gestures in context*. This is why I stress this *phenomenological model of instrument design* as a 'meta' approach from which personalised and idiomatic design tools may be built. A complete discussion of my own process is beyond the scope of this paper,[3] but I will present one analysis approach that arises from these principles.

### 4.2. Instrumental morphology by examining the gestural/textural nature of sustainment

Viewing an instrumental system by virtue of its sonic gestural potential brings the question back to the ability (or lack thereof) of the instrument to compose a given sound world: for example, can it produce one type of GC-gesture with a gradual increase in spectral or transient grain? This is again a meta-level view in that it does not concern itself with a given sound model, controller or mapping type.[4]

In my own instrument design work an ongoing interest is the creation of real-time control strategies to influence textural features, and to move between different gestural contours that traverse the spectrum of phenomena between the qualities of grain and motion. In terms of a sound's morphology, these two lie on the boundary of the internal matter of a sound and the dynamic form: grain may manifest as matter properly or an element which ties matter to form, depending on its rate, regularity, spectral placement and amplitude. It can suggest the surface of an object as much as the surface of the sound itself. Similarly, motion may have an intrinsic quality in itself within a sound object, yet it simultaneously suggests the action[5] that caused it. This led Schaeffer to study these phenomena in a different chapter of his morphology, under the heading of a 'theory of sustainment'. Indeed, these qualities do arise during the sustain portion of a sound event, and in terms of sonic gesture profile are suggestive (in the case of imagined human action) of continuous excitation and modification gestures. In order to describe the grain/motion of a given sonic gesture that arises from such control gestures, I've needed to look into the deeper structure of a given sound object to extract grain and motion behaviours in order to characterise them relative to the signal. Further, I've wanted to look at the dynamics of grain in terms of its relative spectrum, weight and placement in order to understand the control-sound gesture coupling.

The notion of transient grain, in so far as it relates to the idea of a 'rubbing grain', immediately suggests the textural quality of both physical and sound object surface. It is feasible that the degree of graininess of this particular type of sound could be captured well by simple measures. However, the notions of iterative grain – such as a fast drum roll – or of spectral/resonance grain – such as shimmering and fast modulations that arise from the resonance of dense inharmonic objects – are more interrelated and thus more difficult to decouple. They again are closely tied to the perception of motion-based modulations, differing only in regards to rate and depth from a signal point of view.

[3]The reader is directed to Van Nort (2009) for a more complete exposition.

[4]Understanding an instrument in terms of its dynamic movement between sonic gestural potential is in a similar spirit as Paine's (2004) view of an instrument as a collection of soundworlds. For this he draws upon Wishart's notion of dynamic morphology (Wishart 1996) that applies to sonic objects in a constant flux between steady states. Though similar in spirit, there are many differences here with a major one being the focus on embodiment in sound.

[5]This is broken down into mechanical, living and natural in Schaeffer (1966).

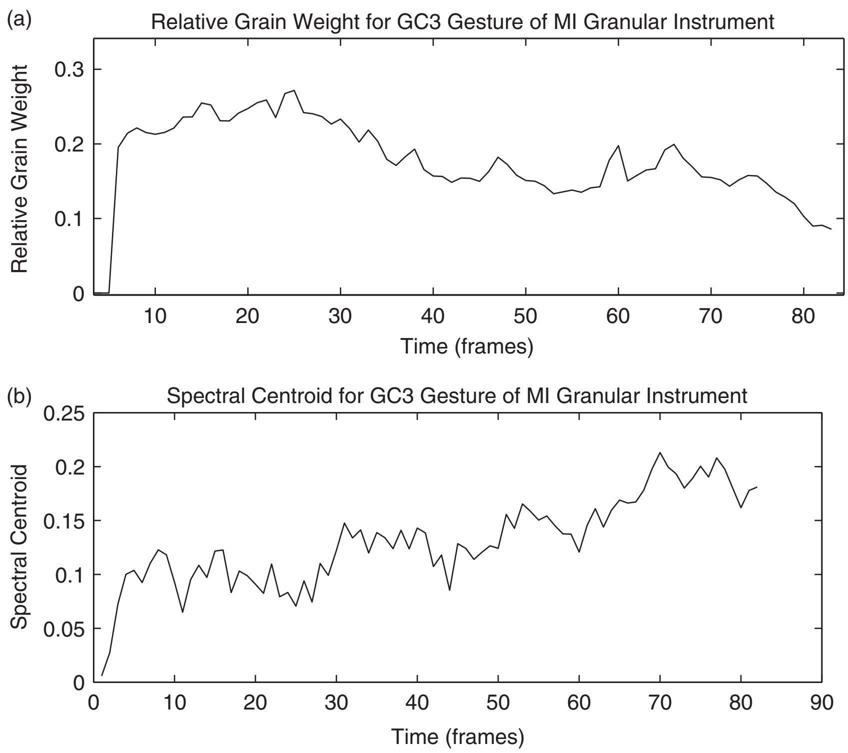


**Figure 2.** Relative grain weight (a) and spectral centroid (b) of sonic gesture. Frame size is approximately 100 ms.

Constructing an analysis that adequately represents listening to sonic gestures in musical context requires a deeper analysis of the system's output in order to decouple these phenomena, to study them separately and more fully characterise them. The fact that grain and motion are so closely tied makes them very deeply intertwined within the signal, with interactions that are often nonlinear and nonstationary in nature. Because of this, one could not hope to directly extract them using classic methods from time-frequency analysis such as Fourier or Wavelet techniques, as these are indeed linear transforms. Wavelets may well represent non-stationarity within a signal – important in this context as this is a defining characteristic of transient grains – yet this is not enough as spectral grain phenomena such as roughness and rapid beatings may arise from nonlinear interaction of modes.

With these requirements determined by the gestural context, it has been the criteria of sustainment – and grain in particular – that has prompted me to utilise a particular time-frequency analysis technique for nonlinear and nonstationary signals first presented by Huang, Shen, Long, Wu, Shih, Zheng, Yen, Tung and Liu (1998), with further development notably by Flandrin, Rilling and Goncalves (2004). This technique, known as *empirical mode decomposition* (EMD), breaks a signal down into amplitude and/or frequency modulation (AM–FM) components in the time domain, from the coarsest modulations to the finest noise components. It is defined algorithmically (rather than analytically) by virtue of an iterative process that subtracts more-to-less fine modulation functions from the signal (Flandrin et al. 2004).

At the conclusion of the EMD process, the input signal is decomposed into a set of *intrinsic mode* functions (IMFs) that express its amplitude or frequency modulation behaviour. A 'spectral' interpretation only makes sense in so far that low vs. high modes relate to more vs. less temporal detail and so high vs. low frequency content. However, there is no direct sub-band filtering – rather this method is an adaptive time-varying filter. This coarse vs. fine decomposition is precisely what motivated the exploration of EMD for this current study, so that signal qualities of grain – related to the lowest (i.e. finest) IMFs – could be decoupled from large-scale trends related to form, and more gross motions related to AM–FM components that comprise motion. This is a clear example of where compositional interests – informed by experience of listening to gestural/textural development – led to signal processing requirements articulated by dynamics of morphology, finally leading me to choose this particular signal technique. It is further appropriate for this phenomenological approach to instrument design in that it is signal-adaptive, so that dynamic shifts between matter features actually drive the analysis.

I have applied this technique to the output of granular synthesis instruments that I have built, in order to characterise the sonic gestural output. The use has been adapted to my interest in composing

**Table 1.** Mapping of morphological qualities to sound features, utilising EMD analysis, autocorrelation functions (ACF), my own temporal fine structure (TFS) measure and other more classical features.

| Property | Description | Sub-Type | Features |
|---|---|---|---|
| Matter Proper: | | | |
| Mass | Spectral Complexity | Pitch Salience | First Peak from ACF |
| | | Line-likeness | Spectral Flatness Measure |
| | | Expanse | Spectral Spread |
| HT | Spectral Balance | – | Spectral Centroid |
| Matter/Form Boundary (Sustainment): | | | |
| Grain | Microstructure | – | Temporal Fine Structure |
| | | Transient/Spectral/Iteration Grains | IMFs from EEMD |
| | | Grain Classification | Roughness, TFS |
| | | RGM | IMF to Signal Mass Ratio |
| | | RGP | IMF to Signal Centroid Ratio |
| | | RGW | IMF to Signal Power Ratio |
| Motion | Modulation | Rate | ACF of Envelope EMD |
| | | Depth | Peak of Envelope ACF |

gesture dynamics. For example, in order to examine the relative motion of grain for a given sweeping control gesture, I have examined the *relative grain weight* (RGW) defined by the power ratio of the textural grain signal – determined by a set of extracted IMF functions representing the appropriate (transient, spectral or iteration) grain – to the entire signal over an appropriate window of observation. An example graduated continuant (type 3) gesture in which a transient grain suddenly emerges and fades slowly over the course of the sound is given in figure 2. The transient grain prominence fades as other spectral mass features increase towards termination, which by definition is an implicit ending that leads, in this instrument, to a new sonic gesture.

This entire process brings the signal processing aspect of instrument design into the phenomenological loop – even the process of verification consists of listening to the output of the EMD analysis to determine whether the grain is captured properly (i.e. in the proper IMF signal). Analysis on the separated grains can suggest automatic ways of partitioning the basis functions. For example, I have constructed a measure of the *temporal fine structure* for examining transient grains, or roughness (Daniel and Weber 1997) for spectral grains. Those that are high in a given measure may be grouped for further analysis, such as relative grain weight computation. While there is no space for a complete exposition, I have included my current mapping from morphological qualities to signal features in table 1. Rather than championing this particular computational approach out of context, my example serves to illustrate a system that has arisen not directly from the literature but *from my own personal analysis of the sonic gestural response of my instrumental system.* Looking at the dynamics of these features has allowed me to describe the gestures that are produced, and how changes to the control structure of the instrument influence the gestural output. That said, the EMD technique is quite novel in application to sound analysis, and suggests an interesting new direction in general for analysis of textural sonic qualities.

## 5. CONCLUSION

Listening to dynamics, even in the case of more abstracted forms of electroacoustic musics, suggests mental images of embodied actions that may have given rise to them. Electroacoustic music further uses gesture as a form-bearing element, meaning that instruments designed to create music of this sort should consider the gestural dynamics of sound in context. Looking at the general shapes of sonic gestures – defined by the dynamics of form/matter properties – can suggest control gestures for the purpose of mapping, as with the example typology given. Finally, extracting actual features to describe sonic gestures is a personal and aesthetic decision which brings signal models into the sonic-phenomenological design loop: understanding perceived gestural shapes through dynamics of said features, thereby influencing control/mapping design and so on in a recursive loop. Following such a process – composing gestural dynamic potentials – is not only a more embodied view of mapping and control design, but in my opinion is the best way to arrive at real-time instruments that can create sound-focused electroacoustic musics in a nuanced way, as has been happening out of real-time for sixty years.

## ACKNOWLEDGEMENTS

Special thanks to Stephen McAdams, Marcelo Wanderley and the reviewers for their comments and feedback. Additional thanks to Philippe Depalle for comments and for introducing me to the EMD technique.

## REFERENCES

Cadoz, C. and Wanderley, M. M. 2000. Gesture-Music. In Marcelo M. Wanderley and Marc Battier (eds.) *Trends in Gestural Control of Music*. Paris: IRCAM – Centre Pompidou, 315–34.

Chadabe, J. 2004. Electronic Music and Life. *Organised Sound* **9**(2): 3–6.

Chion, M. 1983. *Guide des objets sonores*. Paris: Buchet/Chastel/INA-GRM.

Choi, I. 2000. Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. In Marcelo M. Wanderley and Marc Battier (eds.) *Trends In Gestural Control of Music*. Paris: IRCAM – Centre Pompidou.

Daniel, P. and Weber, R. 1997. Psychoacoustical Roughness: Implementation of an Optimized Model. *Acustica* **83**: 113–23.

Delalande, F. 1988. La gestique de Gould: Elements pour une sémiologie du geste musicale. In G. Guertin (ed.) *Glenn Gould Pluriel*. Quebec: Louise Courteau Editrice, 85–111.

Emmerson, S. (ed.) 1986. *The Language of Electroacoustic Music*. London: Palgrave Macmillan.

Flandrin, P., Rilling, G. and Goncalves, P. 2004. Empirical Mode Decomposition as a Filter Bank. *IEEE Signal Processing Letters* **11**: 112–4.

Gaver, W. 1993a. How do we Hear in the World? Explorations in Ecological Acoustics. *Journal of Ecological Psychology* **5**(4): 285–313.

Gaver, W. 1993b. What in the World do we Hear? An Ecological Approach to Auditory Event Perception. *Journal of Ecological Psychology* **5**(1): 1–29.

Giordano, B. and McAdams, S. 2006. Material Identification of Real Impact Sounds: Effects of Size Variation in Steel, Glass, Wood, and Plexiglass Plates. *Journal of the Acoustical Society of America* **119**(2): 1171–81.

Godoy, R. I. 2003. Motor-Mimetic Music Cognition. *Leonardo* **36**(4): 317–9.

Godoy, R. I. 2006. Gestural-Sonorous Objects: Embodied Extensions of Schaeffer's Conceptual Apparatus. *Organised Sound* **11**(2): 149–57.

Godoy, R.I., Haga, E. and Refsum-Jensenius, A. 2006a. Exploring Music-Related Gestures by Sound-Tracing – A Preliminary Study. *2nd ConGAS International Symposium on Gesture Interfaces for Multimedia Systems*, Leeds, UK.

Godoy, R. I., Haga, E. and Refsum-Jensenius, A. 2006b. Playing 'Air Instruments': Mimicry of Sound-Producing Gestures by Novices and Experts. In S. Gibet, N. Courty and J.-F. Kamp (eds.) *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop (GW 2005)*. Berlin, Heidelberg: Springer-Verlag, 256–67.

Gygi, B. 2001. Factors in the Identification of Environmental Sound. PhD thesis, Indiana University.

Hatten, R. S. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington: Indiana University Press.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C. and Liu, H. H. 1998. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis. *Proceedings: Mathematical, Physical and Engineering Sciences* **454**: 903–95.

Lakatos, S., McAdams, S. and Causse, R. 1997. The Representation of Auditory Source Characteristics: Simple Geometric Form. *Perception and Psychophysics* **59**(8): 1180–90.

Levitin, D., McAdams, S. and Adams, R. L. 2002. Control Parameters for Musical Instruments: A Foundation for New Mappings of Gesture to Sound. *Organised Sound* **7**(2): 171–89.

Liberman, A. and Mattingly, I. 1985. The Motor Theory of Speech Perception Revised. *Cognition* **21**: 1–36.

Paine, G. 2004. Gesture and Musical Interaction: Interactive Engagement Through Dynamic Morphology. *Proceedings of the 2004 Conference on New Interfaces for Musical Expression (NIME)*, Hamamatsu, Japan, 80–6.

Schaeffer, P. 1966. *Traité des cbjets musicaux: Essai interdisciplines*. Paris: Éditions du Seuil.

Schaeffer, P. 1998. *Solfège de l'objet sonore*. Paris: INA/GRM.

Schnell, N. and Battier, M. 2002. Introducing Composed Instruments, Technical and Musicological Implications. *Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME 02)*, Dublin, Ireland, 1–5.

Smalley, D. 1997. Spectromorphology: Explaining Sound-Shapes. *Organised Sound* **2**(2): 107–26.

Strobl, G., Eckel, G. and Rocchesso, D. 2006. Sound Texture Modeling: A Survey. *Proceedings of the 2006 Sound and Music Computing (SMC) International Conference*, Marseilles, France, 61–5.

Thoresen, L. 2007. Spectromorphological analysis of sound objects: an adaptation of Pierre Schaeffer's typomorphology. *Organised Sound* **12**(2): 129–41.

Van Nort, D. 2007. Texture Modeling: Signal Models and Compositional Approaches. *Proceedings of the 2007 Conference of the Society for Music Perception and Cognition (SMPC 07)*, Montreal, Canada, 71–2.

Van Nort, D. 2009. Modular and Adaptive Control of Sound Processing. PhD thesis, McGill University.

Wishart, T. 1994. *Audible Design*. York: Orpheus the Pantomime Ltd.

Wishart, T. 1996. *On Sonic Art*. Amsterdam: Harwood Academic Publishers.