

---

# Gestural-Sonorous Objects: embodied extensions of Schaeffer's conceptual apparatus

---

ROLF INGE GODØY

Department of Musicology, University of Oslo, P.B. 1017 Blindern, N-0315 Oslo, Norway  
E-mail: r.i.godoy@imv.uio.no

One of the most remarkable achievements of Pierre Schaeffer's musical thought is his proposal of the *sonorous object* as the focus of research. The *sonorous object* is a fragment of sound, typically in the range of a few seconds (often even less), perceived as a unit. *Sonorous objects* are constituted, studied, and evaluated according to various criteria, and *sonorous objects* that are found suitable are regarded as musical objects that may be used in musical composition. In the selection and qualification of these *sonorous objects*, we are encouraged to practise what Schaeffer called 'reduced listening', meaning disregarding the original context of the sound, including its source and signification, and instead focus our listening on the *sonorous features*.

However, it can be argued that this principle of 'reduced listening' is not in conflict with more fundamental principles of embodied cognition, and that the criteria for the constitution, and the various feature qualifications, of *sonorous objects* can be linked to gestural images. Also, there are several similarities between studying sound and gestures from a phenomenological perspective, and it is suggested that Schaeffer's theoretical concepts may be extended to what is called *gestural-sonorous objects*.

## 1. INTRODUCTION

Although Pierre Schaeffer is commonly associated with the emergence of *musique concrète* and its twin concept of *acousmatic* listening, I believe one of Schaeffer's most remarkable achievements is his idea of the *sonorous object* as the focus of musical research. Briefly stated, the *sonorous object* is a fragment of sound, typically in the range of a few seconds (often even less). Most importantly, the *sonorous object* enables us to have an overview of the entire fragment of sound as a shape, hence as an object, and notably so an object with several concurrent features evolving between the start and end points of the fragment (e.g. timbral, dynamic, textural, etc., evolution in the course of the fragment). *Sonorous objects* are 'raw' fragments of sound, some of which by various subsequent feature evaluations may be found 'suitable' for use in compositions, and may thus be elevated to the status of *musical objects*.

The idea of the *sonorous object* was, and is in my opinion still, four decades after the publication of

Schaeffer's monumental *Traité des objets musicaux* (Schaeffer 1966), in several ways a radical one: it allowed denominating features of musical sound previously not possible within the Western conceptual apparatus, features we today would variously regard as timbral-textural elements of musical sound. It was universal in the sense of being potentially applicable to any fragment of musical sound, regardless of source (vocal, instrumental, electronic, environmental), style, and/or musical culture. But most of all, the concept of the *sonorous object*, and the extensive apparatus for feature qualification that goes with it, provided the means for capturing and reflecting on the otherwise ephemeral or transitory nature of musical sound, i.e. allowed what we could call a mental *recoding* (to borrow an expression from G. A. Miller's famous paper on chunking [Miller 1956]) of sound to more stable images in our minds.

In fact, after extensive discussions of what the *sonorous object* is and is not, Schaeffer ends up with stating that the *sonorous object* is an 'intentional unit', constituted in our consciousness by our own mental activity (Schaeffer 1966: 263). The *sonorous object* can be inspected, explored, and progressively differentiated with regards to features, features which often evolve or have various envelopes which can be traced, hence in my opinion actually becoming more like what I would call a *gestural object*. In our present work on musical gestures (<http://musicalgestures.uio.no>) we suspect that there are gestural components in the recoding of musical sound in our minds. As indicated by the figure, we hypothesise that there is a continuous process of mentally tracing sound in music perception (and in musical imagery as well), i.e. mentally tracing the onsets, contours, textures, envelopes, etc., by hands, fingers, arms, or other effectors, when we listen to, or merely imagine, music. This means that from continuous listening and continuous sound-tracing, we actually recode musical sound into multimodal gestural-sonorous images based on biomechanical constraints (what we imagine our bodies can do), hence into images that also have visual (kinematic) and motor (effort, proprioceptive, etc.) components. Furthermore, this recoding is conceived of as a bidirectional process, i.e. that gestural images

may engender sonorous images as well (see discussion and references on this in section 6 below).

This does raise some questions about the nature of sound perception in Schaeffer's case, because of his principle of acousmatic listening, i.e. that the listener is not able to see the sound-producing gestures, and even more so because of his principle of 'reduced listening', meaning that the listener should intentionally disregard causal and/or anecdotal significations of the sound, and rather focus on the sound-features alone (more on this in section 2 below). However, it is also quite clear that Schaeffer did make use of a number of gestural concepts and metaphors in qualifying sonorous objects. Going a bit deeper into this, I shall argue that Schaeffer's use of gestural concepts and metaphors can be related to the idea of so-called *embodied cognition*, meaning that virtually all domains of human perception and thinking, even seemingly abstract domains, are related to images of movement (Gallese and Lakoff 2005). For these reasons, I shall introduce the concept of *gestural-sonorous objects* in this paper, meaning an extension of Schaeffer's thoughts on exploring sonorous objects to also include the exploration of gestures associated with the various sonorous objects.

Actually, studies of human gestures reveal a strikingly similar challenge of recoding, or of extracting and qualifying stable percepts from a continuous sensory stream, as in the study of sound. Strategically, gesture research could also have the focus on short fragments as in Schaeffer's sonorous objects, and could even use a conceptual apparatus similar to that which Schaeffer used for sonorous objects. Furthermore, recent research and development in gestural control of new musical instruments (Wanderley and Battier 2000), with enhanced possibilities for both direct and indirect control of very many sound features (through various schemes for mapping), has highlighted the close relationship between sound features and gestures in general, i.e. not only as a causal and/or anecdotic relationship, but as fundamental, ubiquitous cognitive schemata, and has opened up a large territory of gestural-sonorous interactions in need of exploration.

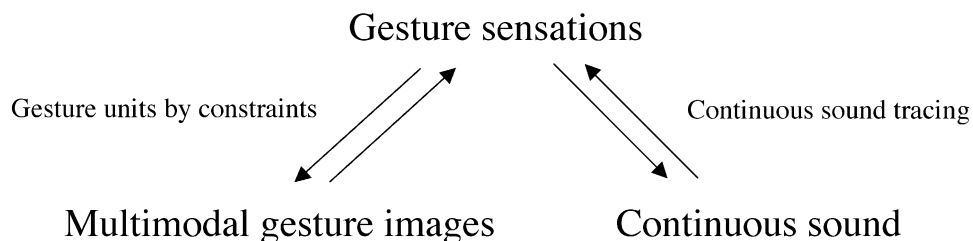
Schaeffer's universe of musical thought is very extensive and complex, and the scope and ambition of this paper must necessarily be limited to demonstrating what I believe are gestural-sonorous elements in

Schaeffer's theory, as well as some elements of his theory leading up to this. In order to illustrate some of the main ideas here, I will make a number of references to the *Solfège de l'objet sonore* (Schaeffer 1998), which on the three CDs have sound examples that better than words may demonstrate these gestural-sonorous ideas. The reader is encouraged to listen to the examples in this *Solfège* (and of course also follow the text, available in French, English, and Spanish in the accompanying booklet) and judge for herself or himself about the gestural elements in these sounds and texts.

## 2. LISTENING

The basis for Schaeffer's conceptual apparatus was the listening experience and not any kind of symbolic representation of musical sound, be that in the form of conventional notation or acoustic measurements. Furthermore, Schaeffer stated that there should be no restrictions on what kind of sonorous material could be investigated, i.e. that sounds from all sources and all cultures could be legitimate material for study. It is interesting that of the three challenges to music theory discussed by Schaeffer, i.e. non-Western music, new technology, and new aesthetics, the challenge of non-Western music was seen as most important because it questioned the nature of our Western musical system (Schaeffer 1966: 18). This clearly indicates a wish for theoretical concepts based on more universal, cross-cultural principles of human perception.

An essential element in such a new and universal theory of musical sound would be the relationship between perceptual images and the acoustic substrate of sound. On the background of some seemingly puzzling perceptual effects of manipulating the acoustic material, such as removing the attack segments and fundamentals of sounds, as well as manipulating spectral content and envelopes, Schaeffer concluded that the relationship between perception and acoustics was complex and nonlinear. This relationship was characterised as a relationship of *anamorphosis*, i.e. of 'warping', and the challenge of musical research was defined as that of establishing *correlations* between what we perceive and the acoustic substrates, rather than studying the acoustic substrates alone (Schaeffer 1998: CD1, tracks 12–92 and CD2, tracks 1–66). As Michel Chion points



**Figure.** Schematic overview of gestural-sonorous interaction and recoding.

out, the climate of musical research in the 1950s and 1960s was dominated by a kind of 'scientific' idea of the musical object 'in itself', not taking into consideration the complexities of perception (Chion 1983: 30). However, much has changed in the four decades since the publication of the *Traité*, and several of Schaeffer's ideas concerning *anamorphosis* have now become accepted within mainstream psychoacoustics and music cognition.

The consequences of *acousmatic listening*, meaning that the original sources of the sounds (e.g. musicians, various sound-producing objects, etc.) were not visible because loudspeakers transmitted the sound, were extensive, implying a general exclusion from sonorous object research of whatever means for production had been used. Related to this is Schaeffer's general distinction between generation ('faire', including so-called 'composition techniques') and perception ('entendre'), a distinction that unfortunately was all too often ignored in twentieth-century music theory (Godøy 1997). Furthermore, Schaeffer took great care to distinguish different levels and goals in listening, ranging from the basic faculty of hearing to the intentional focus on various significations of the sound (distinguishing *four modes of listening*; see Godøy 1997: 129–30 for a discussion of this). Schaeffer suggested that the listener, when doing musical research, should focus on the sound features and disregard any contextual significance, hence practise 'reduced listening' (*'écoute réduite'*).

In explaining this principle of reduced listening, Schaeffer points to the practical experience in the early days of the *musique concrète* of repeatedly listening to looped fragments of sound, to what was called 'sillon fermé', meaning 'closed groove' on a phonograph record. Initially, this was actually a pragmatic matter, as the use of these phonograph loops was the available technology for the manipulation and mixing of sounds before the arrival of the tape recorder. However, such repeated listening was turned into a research strategy, because these repeated listenings would inexorably lead the attention away from the original contextual significations of the sound and towards various features of the sound itself, hence towards what would later be called the *typology* and *morphology* of the sonorous object.

But Schaeffer also draws on Husserl's concept of *epoché*, meaning the 'bracketing' or 'suspension' of the real world outside our minds, and rather trying to perceive things 'as they are', untainted by our usual, everyday, contextual associations. This means that we may intentionally shift our focus towards features we previously were not aware of. Because of the earlier experience with 'sillon fermé', Schaeffer noted that he and his associates had actually been practising phenomenology for several years without knowing it, dryly commenting that this was better than claiming to

practise phenomenology without really doing so (Schaeffer 1966: 262).

For some people, Schaeffer's idea of reduced listening may be difficult to accept, claiming that it is just not possible to eradicate various associations of contextual significance from our minds when we listen to musical sound. However, from my reading of Schaeffer, I believe the principle of reduced listening is first of all a point of method: It is a matter of intentionally shifting focus towards the various features of the sound for the purpose of knowing more about the sound, i.e. to actively trace the evolution of the various features of the sonorous object and hence progressively build up an increasingly detailed image of the sonorous object. Discussing our intention in listening, Schaeffer wrote: 'Nothing can prevent a listener from making it waver, passing unconsciously from one system to another, or from a reduced listening to a listening which is not. One can even be pleased with that. It is by such a swirl of intentions that the connections are established, that information is exchanged' (Schaeffer 1966: 343).

### 3. SONOROUS OBJECTS

In principle, just arbitrarily cutting out a fragment from a continuous stream of sound, and making a 'closed groove' (or a tape loop or a digital sound file), could be the point of departure for exploring that fragment. However, an arbitrary cut could also result in various artefacts, and radically transform our perception of the content of the object. Schaeffer states that any new object after a cutting will have a head, a body, and a tail, comparing the cutting of a sonorous object to the cutting of a magnet into smaller parts: Each of the new magnet parts will have their respective polarisations (Schaeffer 1998: CD2, tracks 87–8). In order to avoid the artefacts of cutting, Schaeffer suggested that sonorous objects should be cut at what could be considered 'natural' discontinuities in the continuous stream of sound, by the principle of *stress-articulation* (Schaeffer 1998: CD3, tracks 17–23). 'Articulation' is here defined as 'breaking up the sonorous continuum by successive distinct energetic events' (Schaeffer 1966: 396), and 'stress' as the prolongation of the sound, similar to vowels in speech (Schaeffer 1966: 366). Applying this principle of *stress-articulation*, we actually end up with quite short and clearly shaped sonorous objects, actually so small that they could be called 'gestural primitives' (Choi 2000). To get an idea of how this works, the reader is encouraged to compare the CD3, track 13 with CD3, tracks 18–22 of Schaeffer (1998).

This 'rule' of *stress-articulation* is important, because it provides Schaeffer's theory with a universal principle for constituting sonorous objects. But also, this *stress-articulation* rule is significant because it refers to the more general, ecological principle of chunking sensory

streams according to experienced discontinuities. This may clarify our understanding of reduced listening, because whereas causal and everyday significations are to be ignored, basic schemata of perception, such as energy and qualitative discontinuities, are clearly not to be ignored in reduced listening. In our context of gestural-sonorous objects, this fits quite well with principles of embodied cognition, meaning that Schaeffer in fact applied fundamental schemata of bodily experience to sound perception (see section 6 below).

The experiences of anamorphosis between sonorous objects and acoustic substrates made Schaeffer realise that sonorous objects had to be perceived holistically in the sense that sequentially presented acoustic information (e.g. attacks before the sustain portion of a sound) would influence the perception of the sonorous object as a whole. In this way, the sonorous object is a cumulative image of a certain stretch of sound, constituted in our minds. For this reason, Schaeffer also took great care to delimit what kind of ontological status the sonorous object should have, i.e. emphasising what a sonorous object is and is not (Schaeffer 1998: CD2, tracks 66–88). As a mental image, the sonorous object may vary from one listening to the next, yet remain identifiable (Schaeffer 1998: CD2, tracks 88–9). This is in line with the abovementioned idea that the sonorous object is an ‘intentional unit’, meaning that the perception of the sonorous object proceeds by sketches, making us progressively more and more aware of its many features.

Furthermore, the holistic perception and cognition of sonorous objects is necessary in order to make various qualifications of the sound, to demarcate what Schaeffer called morphological features (see section 4 below). This is true for all time-varying features, both on the micro-level and on the more superordinate level of the entire object. In general terms, this means thinking of the various features of sonorous objects as *shapes* (Godøy 1997). In Schaeffer’s thinking, this principle of holistic perception applies equally well to more traditional music: A single tone may be a sonorous object, but also a complex chord with many notes, a glissando of many notes, a rapidly played group of notes, an ornamental figure, a textural fragment, etc., may all be considered sonorous objects, as long as they are perceived holistically.

The strategic advantage of studying fragments is that it enables focus at a significant level of resolution, something Schaeffer was quite clear about, criticising the study of large-scale forms (Schaeffer 1966: 35). In studying gestures, we can make a similar choice of fragment-centred focus: Having knowledge of what happens in continuous or long stretches (be that of sound and/or gestures) is quite difficult and will usually result in global qualitative judgements, whereas setting limits or giving a time frame, makes it possible to

demarcate and qualify a number of concurrent trajectories which are perceptually significant.

The detail criteria for what is the ‘right’ duration of a sonorous object is dealt with in Schaeffer’s typology (see section 4 below), but the general principle of perception by a series of chunks is quite fundamental in phenomenological thought. As in Schaeffer’s theory, the exact size of these chunks is relative in phenomenology; however, the principle is that of proceeding by a series of discontinuous points in time where these chunks are perceived holistically and ‘in a now’: ‘... the assumption that the intuition of a temporal interval takes place in a now, in a temporal point, appears to be self-evident and altogether inescapable . . .’ (Husserl 1964: 40–1). In other words, Husserl claimed that if we would be continuously immersed in the stream of sensory impressions, we would simply not have any perceptual images at all (see Schneider and Godøy 2001 for more on this). Or as nicely stated by Ricoeur: ‘... we interrupt lived experience in order to signify it’ (Ricoeur 1981: 116). Interestingly, there seems to be support for this chunk-by-chunk type of perception and cognition as a general phenomenon in recent neurocognitive research with suggestions of rather short attention segments, typically in the less than 3-seconds range (Pöppel 1997).

Schaeffer introduced the twin concepts of *context* and *contexture*, where *context* signifies the large-scale context that the sonorous object may be included in at any time, and *contexture* signifies the internal substance of the sonorous object (Schaeffer 1966: 503). The sonorous object thus becomes a focal point where we can think of a large-scale context for the object on one side (e.g. a whole work of music), and an internal divisibility of the sonorous object (e.g. down to a single point in time) on the other side, hence Schaeffer’s term ‘two infinities’ (Schaeffer 1966: 279). It could be convenient to introduce the terms *micro*, *meso* and *macro* here, where *micro* denotes the continuous acoustic substrate, in principle divisible down to the size of a single sample, the *meso*, or ‘mid level’, denotes the sonorous (or gestural) object level, and the *macro* denotes the continuity-level of sensory impressions, so that we have a conceptual apparatus when zooming in and out of sounds as well as gestures. Focus on the meso-level of sonorous objects does not imply any denial of continuity in experience, i.e. that there is a macro-level simultaneously at work with this meso-level (and micro-level as well). As suggested by Pöppel (1997), a succession of such meso-level chunks does indeed result in an experience of continuity, even though the attentional focus may be discontinuous. Interestingly, these meso-level chunks of attention seem to apply also when perceiving stationary phenomena as in the shifts between figure and ground in bi-stable images (e.g. well-know gestalt-related figures such as the Necker cube) following this approximate 3-second duration, but

whether such attention shifts also would apply for long sounds seems to be an open question.

Although apparently much remains to be known about auditory perception of sonorous objects (Griffiths and Warren 2004), this three-level model consisting of holistically perceived chunks at the meso-level of the sonorous object (and the gestural object as well), with a concurrent micro-level of continuous sound, as well as a macro-level of cumulative memory images, seems at least not implausible in perceptual theory. But most of all, it is clearly a suitable model for studying sound and movement by allowing focus on the many significant features found on the meso-level. Although there have been some interesting projects following up Schaeffer's meso-level focus on sonorous objects, such as with the *UST* project (Delalande, Formosa, Frémiot, Gobin, Malbosc, Mandelbrojt and Pedler 1996), Smalley's *Spectromorphology* (Smalley 1997), and some similar (but apparently not directly influenced by Schaeffer's thought) projects such as the *Sounding Object* project (Rocchesso and Fontana 2003), and even parts of the *Auditory Scene Analysis* work by Bregman (Bregman 1990), this kind of object-focused research has much not yet exploited potential for research.

#### 4. TYPOLOGICAL AND MORPHOLOGICAL CONCEPTS

In order to explore and qualify sonorous objects, Schaeffer established the twin concepts of *typology* and *morphology*. Briefly stated, the *typology* is the first and approximate sorting and characterisation of sonorous objects, based on their most salient features, such as what we could call their envelopes and overall pitch and spectral content, and the *morphology* is a more detailed demarcation of the various internal features of the sonorous objects, in principle down to the most minute fluctuations in the sound, i.e. various textural and/or timbral features. The typology and the morphology should be seen as complementary, and in actual use, there is often a shift between these two, evaluating sonorous objects from different perspectives.

Although Schaeffer's typological and morphological concepts, as summarised in the typo-morphological matrix, may seem quite complex (Schaeffer 1966: 584–7), the main principle is essentially a top-down exploration and qualification of the sonorous object, going from its overall shape and features downwards into progressively more detailed features and feature-values, e.g. rates and range of changes in features. Also, Schaeffer emphasised that the typo-morphological matrix of feature dimensions and their respective sub-dimensions were not to be understood as a 'balance sheet' but as a 'questionnaire', in other words, as a stimulus to explore features and not as any kind of strict classificatory system of musical sound (Chion 1983: 93).

In this paper, I will just briefly mention some main concepts from the typology and a few concepts from the morphology as these are presented in Schaeffer (1998).

A bit simplified, we could say that in the typology, the first step is that of cutting the continuous stream of sound into sonorous objects according to the mentioned principle of *stress-articulation* (Schaeffer 1998: CD3, tracks 17–22). Given these sonorous objects, the next step is to qualify objects according to what we could call their overall envelopes of duration (Schaeffer 1998: CD3, tracks 23–43):

- impulsive types
- sustained types
- iterative types

Schaeffer linked these duration envelopes to sound-producing gestures (Schaeffer's expression '*facture gestuelle*' is rendered as 'executive gesture' in the English version of the *Solfège* booklet [Schaeffer 1998: 69]):

- punctual gesture
- continuous gesture
- iterative gesture

But in this first sorting of sonorous objects in the typology, harmonic and/or pitch content, called *mass*, is also taken into consideration, with types variously having definite pitch, complex pitch, and various degrees of stability, evolution or instability in pitch. In this way, Schaeffer establishes a matrix of sonorous objects where overall envelopes, i.e. *impulsive*, *sustained* and *iterative*, are paired with pitch/spectral, i.e. *mass* features, meaning *tonal*, i.e. pitched, *complex*, i.e. fixed but indeterminate pitch, and *varied*, i.e. fluctuating in pitch, as may be heard in the same set of examples, i.e. Schaeffer (1998: CD3, tracks 23–43).

This matrix of 3 by 3 typological categories is at the centre of the typological classification scheme, and sonorous objects within this matrix are considered 'balanced' in the sense that they are of medium duration, i.e. neither too short, nor too long, and also of medium complexity. However, there is of course also the possibility that sonorous objects may be situated outside this centre, as can be seen from the overview in the *Traité* (Schaeffer 1966: 459), and from the ensuing examples in the *Solfège* (Schaeffer 1998: CD3, tracks 43–64). Furthermore, Schaeffer introduced some criteria for what he called 'suitable objects', criteria implying that the object should not be anecdotic and should be suitable for integration in a musical context (Chion 1983: 97–8). Schaeffer does emphasise though that these typological considerations are just tools for guiding our thinking, and that sonorous objects may be moved from one typological category to another, depending upon the context and the attention we give to them (Schaeffer 1998: 74).

Once a sonorous object is found suitable for musical contexts, it may be further evaluated with regards to morphological features. These morphological features are mainly concerned with the internal features of the object, such as its pitch and/or spectral content, the evolution of the pitch and/or spectral content, in short with what we often refer to as timbral features when these evolutions or fluctuations are fast and on a small scale, e.g. sub-note-level, and as textural features when they are slower and on a larger scale, e.g. note-level. Some of the morphological features are illustrated in the *Solfège* (Schaeffer 1998: CD2, tracks 90–5): The *shape*, meaning the overall envelope, the *mass*, meaning pitch and spectral features (e.g. having clear pitch, ambiguous pitch, being inharmonic, various kinds of noise, etc.), the *grain*, meaning fast/small fluctuations, the *harmonic timbre*, meaning spectral distribution (e.g. spectral envelope), and the *motion*, meaning slower/larger fluctuations (the French ‘*allure*’ is translated as ‘motion’ in the English text of the booklet in Schaeffer (1998: 59), but could perhaps also be translated as ‘gait’). These morphological features are first presented and varied one by one (tracks 90–4), and finally combined in an ‘exaggerated object’ (track 95).

## 5. SOUND-RELATED GESTURES

In the brief look at some of Schaeffer’s typological and morphological concepts in the previous section, I believe we can observe several gestural components, and we should now try to see these gestural components in the broader context of sound-related gestures in general. Although there have been suggestions made for more systematic classifications of sound-related gestures (e.g. Cadoz and Wanderley 2000), there still seem to be divergent opinions about how this should be done. For practical purposes, I shall here give a rather simple overview of sound-related gestures in music as previously presented elsewhere (Godøy, Haga and Jensenius 2006):

- *Sound-producing gestures*, including both excitatory gestures such as hitting, stroking, bowing, blowing, singing, kicking, etc., and modifying gestures such as modulations of pitch and timbre.
- *Sound-accompanying gestures*, such as dancing or marching, or more vague sound-tracing gestures such as following the melodic contours, rhythmic/textural patterns, timbral or dynamical evolutions, etc., with our hands, arms, torso, etc.
- *Amodal, affective or emotive gestures*, including movements associated with more global sensations of the music, such as effort, velocity, impatience, unrest, calm, balance, elation, anger, etc.

In our context here, the most relevant gestures are those that follow the sound closely, i.e. the sound-producing and sound-tracing gestures. The distinction

between these two types of gestures may often be not so clear; however, the main difference is that sound-producing gestures have an energy transfer from the performer to the instrument, whereas the sound-tracing gestures may mimic excitatory gestures as well as trace the evolution of the resonance of sounds, i.e. the ‘passive’ or energy-dissipating phase of the sound, hence not transferring energy to a resonating body.

The sound-producing gestures can be subdivided into discontinuous, continuous, and iterative excitatory gestures, exactly matching Schaeffer’s typological duration envelope categories, i.e. *impulsive*, *sustained* and *iterative*. Excitatory sound-producing gestures are in addition obvious in what Schaeffer called *compound* (i.e. several sounds starting together) and *composite* (sounds fusing together into one object) objects (Schaeffer 1998: CD3, tracks 2–3). As for the *composite* object, it actually demonstrates the phenomenon of coarticulation, known from both linguistics and movement sciences, meaning that smaller movements fuse into more superordinate gestures, a phenomenon at work in very many sonorous objects (e.g. rapid group of tones fusing to one gesture, as mentioned earlier).

Furthermore, excitatory sound-producing gestures also match several of Schaeffer’s morphological categories inducing changes in the sonorous object, e.g. changes in *mass* and *harmonic timbre*, but also in *dynamics* (or *shape*), *melodic profile* (overall changes of pitch/spectral content), and *profile of mass* (internal changes in spectral content), brought about by changes in e.g. the speed, pressure, direction, etc., of the excitatory gesture. Also sound-modifying gestures, i.e. modulatory gestures, can be matched well with Schaeffer’s morphological categories, first of all with those of *motion* and *grain* as when applying vibratos or tremolos at different speeds and amplitudes, but also to changes within the other morphological categories, i.e. those of *mass*, *dynamics*, *harmonic timbre*, *melodic profile* and *profile of mass*, for instance as mute changes (e.g. going from open to closed mute) or bow-position changes (e.g. going from *sul tasto* to *sul ponticello*), wind pressure changes, bow pressure and bow speed changes, etc., in short by a number of sound-modifying gestures musicians know very well.

Actually, pretty much everything in Schaeffer’s typology and morphology may be matched to various sound-producing, i.e. excitatory and modulatory, gestures. My purpose with pointing this out is to demonstrate that there is a gesture component embedded in Schaeffer’s conceptual apparatus which is on a more general and basic level than that of everyday causal listening, i.e. not on a level that the principle of reduced listening is supposed to lead us away from. The implicit gestural components I see in the typology and morphology are general in the sense that they may be applied to many rather different sounds, as well as be carried out with rather different

effectors, and hence actually demonstrate what is called *motor equivalence* in the motor control literature (Rosenbaum 1991). This means that the gestural categories have a certain degree of abstraction in the sense that they are transferable from one setting to another, both with regard to effectors (i.e. hand, fist, finger) and instrument (drum, string, metal sheet, computer), hence in fact be what we could call ‘reduced gestures’ (as suggested by Leigh Landy, personal communication), or in more general terms, become image schemata (Johnson 1987) which we use in our perception of known as well as unknown, previously heard as well as unheard, hence practising what could be called ‘anthropomorphic projection’ (Joel Chadabe, personal communication).

## 6. EMBODIED COGNITION

The idea of mental re-coding of sound into multi-modal gestural images mentioned at the beginning of this paper rests on the idea of embodied cognition. Embodied cognition means that there is an incessant mental simulation going on in our minds of whatever we perceive, so that perception is not a matter of abstract processing of sensory data, but rather a process of re-enactment of whatever we perceive (Wilson and Knoblich 2005). Remembering Schaeffer’s affiliation with phenomenology, it is also quite interesting to note recent convergences between embodiment theory within neurocognitive research and classical phenomenology (Gallese 2005).

More specifically with regard to sound perception, the so-called motor theory in linguistics (Lieberman and Mattingly 1985) claimed that speech perception is not just a matter of processing the auditory signal for certain acoustic cues, but just as much a matter of the listening subjects mentally re-enacting the articulatory gestures necessary for producing the sounds. In other words, the articulatory gestures were seen as integral to the mental image of speech sounds. This theory has often been criticised; however, more recent research seems to support the idea that there are indeed close links between perception and motor elements in our neurocognitive apparatus (Fadiga *et al.* 2002).

In music, we often see people making sound-accompanying gestures such as moving their bodies, shaking their heads, gesticulating with their arms, etc., to the music, and we may also see people making sound-producing gestures such as playing air drums, air guitar, or air piano when listening to music. Our observation studies of people with different levels of expertise, ranging from novices with no musical training to professional musicians, playing air piano, seem to suggest that associations of sound with sound-producing gestures is common and also quite robust even for novices (Godøy, Haga and Jensenius 2006). Also in cases when people are not making overt

sound-producing gestures, some studies have shown that there are quite strong links between listening and activations of motor-related areas in the brain (Haueisen and Knösche 2001), and conversely, just observing silent finger movements on a piano keyboard may activate auditory areas of the brain in pianists (Haslinger, Erhard, Altenmüller, Schroeder, Boecker and Ceballos-Baumann 2005), cf. the idea in the figure that images may be triggered both ways, i.e. from sound to gestures and from gestures to sound. Also, in the case of musical imagery, i.e. when people are merely imagining music with their ‘inner ear’, there seems to be activations of certain motor-related areas in the brain (Zatorre and Halpern 2005).

On the background of evidence from different sources, it would not seem unreasonable to suggest that there is what I have called a *motormimetic* component in music perception and cognition (Godøy 2003). The idea of motormimetic cognition implies that there is a mental simulation of sound-producing gestures going on when we perceive and/or imagine music; hence, that motor imagery (Jeannerod 2001) may actually be considered a component of musical imagery (Godøy 2004). Furthermore, this motor imagery draws on knowledge of various biomechanical and motor control constraints, meaning that we also have included kinematic and dynamic images, hence also images of effort, of chunking, of coarticulation, etc., in short, images of real-world movement elements, in our images of musical sound. Lastly, there may also be even more fundamental links between sound and gesture, a kind of auditory-motor loop in the sense of ‘low-level’ or ‘hard-wired’ interaction and cooperation of the senses (Hickok, Buchsbaum, Humphries and Muftuler 2003). Hopefully, neurocognitive research will give us important insights on the bodily basis of gestural-sonorous interaction in the coming years.

## 7. STUDYING GESTURAL-SONOROUS OBJECTS

There are different elements that converge in our studies of gestural-sonorous objects, and I shall here give a brief overview of viable methods as well as challenges:

- Analysing the conceptual apparatus, including the use of gesture-related metaphors, in various music theory research (and other music-related texts for that matter), as I have tried to show here in the case of Schaeffer’s work.
- Observation studies and analysis of what kinds of gestures people actually make when listening to music, e.g. in air-instrument playing (Godøy, Haga and Jensenius 2006). This includes also studying what features of the sonorous objects are reflected in the gestures, including onsets, pitch-space, envelopes, textures, articulation, etc., and

biomechanical and motor control constraints such as chunking, coarticulation, etc.

- Sound-tracing studies (in progress) where listeners are asked to draw various typological and morphological features of sonorous objects, such as those in Schaeffer (1998, CD3, tracks 18–22), on a Wacom digitising tablet and bimanually in three-dimensional space using the Polhemus electromagnetic tracking system.
- Compiling information from neurocognitive research on auditory-motor interaction.

The last mentioned point concerns the fact that although we may observe people's sound-related gestures (both those done spontaneously and those done according to more specific instructions), the covert sound-related gestures in people's minds are of course not directly accessible. We do hope that neurocognitive research in the coming years will give us more insight on the workings of gestural-sonorous imagery, but there are also other major challenges here:

- Better means for analysing and representing both gestures and sound. There are many and substantial challenges here of a technical nature, such as in tracking, preprocessing, and representation of data, and of a more conceptual nature, such as in categorising and interpreting gestures.
- Better understanding of the kinematics and dynamics of gestures, as well as of biomechanical and motor control constraints, assuming that these constraints condition gestures and hence are also reflected in gestural-sonorous objects.
- Better synthesis tools for the generation of incrementally different variants of sounds, allowing systematic exploration of morphological features, e.g. minute control of various aspects of *grain* and *mass* (cf. Schaeffer's 'exaggerated object', CD2, tracks 90–5), and tracking listeners' gestural responses to these variants.

## 8. CONCLUSIONS

Schaeffer's idea of focusing on the sonorous object in musical research still has great potential, and could be extended to include gestural components. The main elements of such a Schaeffer-inspired research on gestural-sonorous objects are the following:

- Focusing on fragments of musical sound at the meso-level of the sonorous object allows explorations of highly significant features such as dynamic shapes, various concurrent feature-trajectories of pitch, spectral, and textural content, and micro-features such as *grain* and *motion*.
- Sonorous objects emerge by ecologically grounded image schemata of *stress-articulation*, i.e. ecologically founded qualitative and energetic discontinuities,

but also our perceptive-cognitive apparatus seems to proceed by discontinuous chunks.

- The reduced listening strategy is not an eradication of fundamental embodied schemata in music but rather a matter of focusing on typological and morphological features, proceeding top-down from the object as a whole towards successively more detailed qualifications of significant features.
- Studies of gestures could profit from a similar focus on meso-level gestural objects perceived holistically as chunks 'in a now', as well as a similar top-down scheme for progressively finer feature-qualifications as in Schaeffer's typology and morphology.
- Sonorous objects clearly have gestural components, and the idea of gestural-sonorous objects is particularly useful for studies of musical texture and timbre (actually two overlapping domains), as well as of other entities of musical sound previously inaccessible in Western music theory.
- Studies of gestural-sonorous objects enhance images of sound in our minds and, besides helping us in the explorations of sound features, can also have several practical applications in improvisation, composition, performance, music education, and in gestural control of new musical instruments.

## REFERENCES

- Cadoz, C., and Wanderley, M. 2000. Gesture-Music. In M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Paris: Ircam.
- Chion, M. 1983. *Guide des objets sonores*. Paris: INA/GRM Buchet/Chastel.
- Choi, I. 2000. Gestural primitives and the context for computational processing in an interactive performance system. In M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Paris: Ircam.
- Delalande, F., Formosa, M., Frémiot, M., Gobin, P., Malbosc, P., Mandelbrojt, J., and Pedler, E. 1996. *Les Unités Sémiotiques Temporelles: Éléments nouveaux d'analyse musicale*. Marseille: Éditions MIM - Documents Musurgia.
- Fadiga, L., Craighero, L., Buccino, G., and Rizzolatti, G. 2002. Speech listening specifically modulates the excitability of tongue muscles: a TMS Study. *European Journal of Neuroscience* 15: 399–402.
- Gallese, V. 2005. Embodied Simulation: from neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences* 4: 23–48.
- Gallese, V., and Lakoff, G. 2005. The Brain's Concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 22(3/4): 455–79.
- Godøy, R. I. 1997. *Formalization and Epistemology*. Oslo: Scandinavian University Press.
- Godøy, R. I. 2003. Motor-mimetic music cognition. *Leonardo* 36(4): 317–19.



- Godøy, R. I. 2004. Gestural imagery in the service of musical imagery. In A. Camurri and G. Volpe (eds.) *Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15–17, 2003, Selected Revised Papers, LNAI 2915*, pp. 55–62. Berlin Heidelberg: Springer-Verlag.
- Godøy, R. I., Haga, E., and Jensenius, A. 2006. Playing ‘Air Instruments’: mimicry of sound-producing gestures by novices and experts. In S. Gibet, N. Courty and J.-F. Kamp (eds.) *GW 2005, LNAI 3881*, pp. 256–67. Berlin Heidelberg: Springer-Verlag.
- Griffiths, T. D., and Warren, J. D. 2004. What is an auditory object? *Nature Reviews Neuroscience* 5: 887–92.
- Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., and Ceballos-Baumann, A. O. 2005. Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience* 17(2): 282–93.
- Haueisen, J., and Knösche, T. R. 2001. Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience* 13(6): 786–92.
- Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T. 2003. Auditory-Motor Interaction Revealed by fMRI: speech, music, and working memory in Area Spt. *Journal of Cognitive Neuroscience* 15(5): 673–82.
- Husserl, E. 1964. *The Phenomenology of Internal Time Consciousness*. ed. Martin Heidegger, trans. J. S. Churchill, Bloomington, IN: Indiana University Press.
- Jeannerod, M. 2001. Neural Simulation of Action: a unifying mechanism for motor cognition. *Neuroimage* 14: 103–9.
- Johnson, M. 1987. *The Body in the Mind*. Chicago: The University of Chicago Press.
- Liberman, A. M., and Mattingly, I. G. 1985. The motor theory of speech perception revised. *Cognition* 21: 1–36.
- Miller, G. A. 1956. The magic number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63: 81–97.
- Pöppel, E. 1997. A hierarchical model of time perception. *Trends in Cognitive Science* 1(2): 56–61.
- Ricoeur, P. 1981. *Hermeneutics and the Human Sciences*. Cambridge/Paris: Cambridge University Press / Éditions de la Maison des Sciences de l’Homme.
- Rocchesso, D. and Fontana, F. (eds.), 2003. *The Sounding Object*. Firenze: Edizioni di Mondo Estremo.
- Rosenbaum, D. 1991. *Human Motor Control*. San Diego: Academic Press.
- Schaeffer, P. 1966. *Traité des objets musicaux*. Paris: Éditions du Seuil.
- Schaeffer, P. (with sound examples by Reibel, G., and Ferreyra, B.). 1998 (first published in 1967). *Solfège de l’objet sonore*. Paris: INA/GRM.
- Schneider, A., and Godøy, R. I. 2001. Perspectives and challenges of musical imagery. In R. I. Godøy and H. Jørgensen (eds.) *Musical Imagery*, pp. 5–26. Lisse (Holland): Swets and Zeitlinger.
- Smalley, D. 1997. Spectromorphology: explaining sound-shapes. *Organised Sound* 2(2): 107–26.
- Wilson, M., and Knoblich, G. 2005. The case for motor involvement in perceiving conspecifics. *Psychological Bulletin* 131(3): 460–73.
- Zatorre, R. J., and Halpern, A. R. 2005. Mental Concerts: musical imagery and auditory cortex. *Neuron* 47: 9–12.