#5
TIMESCALES FOR SOUND-MOTION OBJECTS

Rolf Inge Godøy

5.1 Introduction

There can be no doubt that Pierre Schaeffer was one of the most influential figures in 20th century European music, both as a composer of electroacoustic music and as the developer of a new and extensive theory of music. We are much in debt to his artistic and theoretical achievements, and in spite of more recent advances in music technology, we can still find his contributions highly relevant for music making. In particular, Schaeffer's idea of *sound objects* as the basis for musical creation, perception, and music theory (Schaeffer 2017; Godøy 2021), was a remarkable change of paradigm in Western musical thought, and this sound object paradigm is still highly relevant for music theory, as will be highlighted in this chapter.

In brief, the sound object can be defined as a holistic mental image of a fragment of sound, typically in the approximately 0.5 to 5 seconds duration range, and the sound may be of any origin, vocal or instrumental, electronic or environmental, and from any musical culture. The focus on sound objects as the basis for both creation and theory, this chapter will argue, has the following advantages:

a) Universal in scope, applicable to very different kinds of music
b) Holistic, making the object-timescale features accessible for scrutiny
c) Salient musical features (style, motion, affect) can be found at the object timescale
d) Human motor control seems to function optimally at the object timescale

The main point of this chapter is that the object timescale has a privileged role in music, both in terms of intrinsic sound features and the associated body motion fea-

https:/doi.org/10.34727/2022/isbn.978-3-85448-047-1_5.

tures. The main source for this view of the object timescale is Schaeffer's monumental *Traité des objets musicaux* (Schaeffer, 1966), now also available in English and referred to as (Schaeffer, 2017). Other sources include Michel Chion's *Guide des objets sonores* (Chion 1983), an excellent overview (endorsed by Schaeffer himself) of the most important topics of Schaeffer's Traité (also available in English as Chion, 2009), and Schaeffer and colleagues' *Solfège des objets sonores* (Schaeffer et al., 1998), a collection of three CDs with sound examples and music theory topics narrated by Schaeffer. This work is very useful because of the sound examples, and can make Schaeffer's ideas accessible for a broader audience. For practical reasons, references to these sound files in this chapter will be given as follows: '*Solfège*, CD#, track#'.

5.2 Schaeffer's object focus

The idea of sound objects as the basis for a music theory emerged from practical work with concrete music (sometimes referred to with the French label, *musique concrète*), when composers, before the advent of the tape recorder, used looped sound fragment recordings, i.e. *closed groove* (*sillon fermé* in French) on phonograph discs, to combine different sounds in compositions. As the composers listened to innumerable repetitions of such sound fragments, they discovered that their attention shifted from the immediate and everyday signification of the sound to the more qualitative features of the sound, i.e. to the features that came to be the content of the so-called *typology and morphology of sound objects*, the elaborate scheme for feature classification in Schaeffer's theory. Later, Schaeffer and his collaborators came to realize that their modus operandi during the early years of the *musique concrète* had involved a phenomenological shift in focus, known as *epoché* in the writings of Husserl (Husserl, 1982). In retrospect, Schaeffer called this 'doing phenomenology without realizing it' (Schaeffer, 2017, p. 206). We can now see that Schaeffer's music theory has several other affinities with phenomenological philosophy, such as the procedure by sketches and the object-centred view of perception (see Schaeffer, 2017, p. 210).

The shift of focus from everyday significations (e.g. the squeaking door signalling that someone is coming) to the more qualitative features (e.g. the overall dynamic envelope and the upward glissando of the squeak sound), was called *reduced listening*, and it should be emphasized that this was a method for exploring sound features. This reduced listening was related to the idea of *acousmatic music*, 'acousmatic' here denoting music emanating from loudspeakers with no other visible sound source. Furthermore, a close reading of Schaeffer will show that any sound object is ontologically composite, i.e. it will usually have several different features and significations in parallel, but the overall dynamic and spectral features form the basis for the *typology* of sound objects, enabling an initial and coarse classification of sound objects based on their dynamic and spectral shapes.

The research method of Schaeffer was that of starting out with seemingly naïve questions of what we are hearing, with a kind of Socratic approach of top-down scrutiny of what is in our minds. In the words of Schaeffer, this could be summarized as 'exploring the listening consciousness' (Schaeffer, 2017, p. 109), and with questions of the overall features of the sound objects like: What is the dynamic shape of the sound object? What is the mass (subjective sense of spectral shape and/or pitch) of the sound object? Is the sound object stable or does it fluctuate? Is it sustained or more impulsive?

A crucial point here is that the sound object is not a static entity, but first of all a mental image of a fragment of unfolding sound. Much effort in the writings of Schaeffer is devoted to what the sound object is *not*, and emphasizing that the sound object is a mental image resulting from attentive perception across multiple listening experiences, as well as being ontologically composite with a multitude of features in parallel.

Another vital point here is that a sound object may have a non-linear relationship with its acoustic basis, i.e. there may be a relationship of so-called *anamorphosis*, or warping, between the acoustic signal and the mental image. This often non-linear relationship between the acoustic features and the subjective percept is due to some perceptual-cognitive factors, primarily the following:

- The mutual influence of the parts of a sound object unfolding in time, i.e. the attack part colouring the sustain part and vice versa, or to what extent the sound object's identity is preserved or not across different cuts in its unfolding.

- Differences across the spectrum of what we perceive as a coherent instrument, e.g. if we shift the spectrum of a deep piano tone up a couple of octaves, it sounds more like a harpsichord than a piano.

The point with anamorphosis is that there may not be a one-to-one relationship between the acoustic features and our subjective perceptions. In Schaeffer's method, this means taking our perceptions as primordial and not regarding perception as flawed, but instead exploring the correlations between acoustics and perception, correlations that also take this anamorphosis into account.

What is crucial here is the internal coherence of the sound object in the sense of temporal bi-directionality, i.e. that present is tainted by past and past is tainted by present (as well as by future expectations), as was concretely documented by Schaeffer with the so-called *cut bell* experience, which showed how removing the attack segment could totally alter the sound of a bell. This past-present-future tainting is yet another reason why the *closed groove* is such a powerful tool for research, as it documents the workings of context at the sound object timescale. From this primacy of the subjective perception, the next step was to study the dynamic shapes in the so-called *typology*, extended also to pitch and spectrum-related shapes, as well as later to various internal features of the sound object in the so-called *morphology*, and the combination of these in the *typology and morphology* summary diagram (Schaeffer, 2017, pp. 464–467).

5.3 General object cognition

The focus on objects in perception and cognition is also found in other domains of thought (see e.g. Shinn-Cunningham, 2008; Starrfelt, Petersen, & Vangkilde, 2013; De Freitas, Liverence, & Scholl, 2014 for some interesting cases), and it could be useful to have a quick overview of some generic ideas on object-centered perception and cognition to see how they may contribute to the idea of sound objects.

Two strands of thought in the late 19th and early 20th century stand out: Gestalt theory, and the previously mentioned phenomenological philosophy of Husserl. From the pioneers of Gestalt theory such as von Ehrenfels, Stumpf, Wertheimer, Koffka, and Köhler (just to mention the most prominent ones) up to Bregman in the late 20th century (Bregman, 1990), the idea of holistic perception and cognition has been prominent, and principles such as *belonging* and *exclusive allocation* have contributed much to understanding what we call object cognition. Another relevant case of Gestalt theory is that of motor control (Klapp, Nelson, & Jagacinski, 1998; Klapp & Jagacinski 2011), suggesting that body motion can be understood as consisting of pre-planned chunks similar to gestalts in perception. It could also be suggested that motion chunks contribute to sound object formation, as is the point of so-called *motormimetic cognition* (Godøy, 2003) and the idea of extending Schaeffer's typology categories to body motion (Godøy, 2006).

In parallel with early Gestalt theory, the phenomenological philosophy of Husserl provided important contributions to the epistemological reasoning about objects in our experiences. According to Husserl, we need to step out of the continuous stream of sensations in order to constitute meaning, and this happens by means of a series of so-called *now-points*, i.e. points in time where we interrupt streams and lump together past, present, and expected future sensations into somehow meaningful chunks (Husserl 1991; Godøy 2009).

In an experimental vein, we have seen research (since the pioneering work of Miller (1956) on chunking) on the workings and effects of different kinds of chunking in human behaviour (Gobet et al., 2016), and more neurocognitive views on chunking and sensations of presence can be found e.g. in Pöppel (1997), in Varela (1999), and in Wittmann & Pöppel (1999). Additionally, we have seen work on perception and cognition of auditory objects by holistic integration of sensory data (Bizley & Cohen, 2013); lastly, there is an extensive theory of objects and shapes as fundamental elements in human reasoning to be found in *morphodynamical theory* (Thom, 1983; Petitot, 1985; Petitot, 1990; Godøy, 1997).

5.4 Sound object features

What constitutes a sound-motion object is first of all a sense of energy shape, of starting and ending within a timespan of approximately 0.5 to 5 seconds (sometimes longer, but that is more exceptional). This overall energy shape, or envelope, is one of the main features of Schaeffer's typology, the typology being a first and coarse sorting of sound-motion objects. The main categories of these energy envelopes, called *facture* in French, a term designating the way they are made, are as follows:

- *Sustained*, a continuous, quasi-stationary sound
- *Impulsive*, a brief sound, e.g. percussive or plucked
- *Iterative*, a rapid series of onsets, like in a tremolo

But also a first and coarse sorting of sound content is included in the typology with the three categories of so-called *mass*, denoting the frequency domain (and not just perceivable pitch) as follows:

- *Tonic*, more or less clearly perceivable pitch, relatively stable
- *Complex*, a composite sound, inharmonic, or noise dominated, but stable
- *Variable*, changing in perceived pitch or spectral placement

These mass and facture types are combined in a 3 x 3 matrix, and can be applied to most types of sounds within the typical duration range for sound objects, and several examples of these categories can be heard in the *Solfège*, CD3, tracks 19–42.

There may be so-called *phase transitions* between these typological categories, e.g. an impulsive sound may turn into a sustained sound if extended beyond some duration threshold, or an iterative sound may turn into a series of impulsive sounds if the sound is slowed down, etc.; in short, there are categorical thresholds here due to duration and density of events. And importantly, these musically meaningful categorical thresholds are all related to sensations of body motion, to both motor control and bodily effort. The overall energy envelopes of sound objects are usually the most perceptually salient, capable of triggering sound-accompanying body motion in listeners, such as in various cases of so-called entrainment (Clayton et al., 2013).

There are in addition a number of internal features of the sound objects in what is called the *morphology*, organized in a system of main feature dimensions, each in turn with sub-feature dimensions, and sometimes also with sub-sub-feature dimensions. These dimensions concern the frequency domain features and their various sub-features, and some dynamic features and their sub-features. The most important are those concerned with textures:

- *Grain* includes various instances of very fast fluctuations in the sound, such as in a trill, yet not so fast as to enter into the audio region (i.e. not above $\approx$ 20 hz).
- *Gait* denotes slower kinds of fluctuation, e.g. such as those found in dance, walking, and other body motion patterns.

There are also other morphology dimensions such as *mass* (overall spectral content), *dynamics* (overall loudness), *harmonic timbre* (spectral distribution), *melodic profile* (pitch-related shapes), and *profile of mass* (spectral shapes); however, the typology contains the most prominent features for sound-motion objects, because it includes shapes at the sound-motion object timescale.

Large-scale forms may also be conceived of as objects, as has often been the case in Western music theory, but Schaeffer was clear that his focus was on the materials of the sound objects and not on large-scale works (Schaeffer, 2017, p. 17). Schaeffer introduced the concept of the *suitable object*: neither too short nor too long, and in practice (as evident in the examples of the *Solfège*) typically in the 0.5 to 5 seconds duration range, as well as some other criteria of information density, i.e. neither too complex, nor too simple, to keep the attention of the listener.

With too-long fragments, it would not be possible to focus on salient features, because they would change. With too-short fragments, perceptually salient features would not exist (would not have time to become manifest). In exceptional cases, sound objects in the *Solfège* last for up to 30 seconds. But the main duration criterion is that of content, i.e. that the sound object manifests a clearly perceivable salient shape. This means that the typological categories mentioned above become the main criteria of timescales, as can be heard in *Solfège*, CD3, track 19–22, as well as examples from various sources, instrumental and electronic, in the *Solfège*, CD3, track 22–42. To make this point about suitable objects, we are in *Solfège*, CD3, track 42–59 presented with other examples of sound objects deemed not to comply with suitable objects criteria by being either too long, too short, too dense, having too much redundancy, being too unpredictable, or too chaotic.

5.5 Continuity vs. discontinuity

What the preceding object principles boil down to is the relationship between continuity and discontinuity in musical experience. The question is: To what extent does our organism work by continuous or discontinuous processes and decisions? For more general points of view on this, see e.g. Miller (1982), Sergent and Dehaene (2004), Spivey (2008), and Reybrouk (2021); but in motor control, continuity vs. discontinuity has been debated for more than a century, and has engendered various models of how we plan, trigger, and control body motion in different contexts (Elliott, Helsen, & Chua, 2001).

The conundrum of continuity and discontinuity in motor control can in a sense be bypassed by seeing how triggering and control may be discontinuous, while the results may yet be continuous, i.e. the motion trigger may happen at discrete points in time, but the resultant motion may be extended over longer stretches of time. Discontinuity in motor control is often believed to be based on constraints, i.e. that our organism is too slow to enable continuous control, and that continuous motion is an emergent feature of our organism's implementation of discrete control impulses, as suggested by the following:

- Klapp and Jagacinski (2011), with discontinuous action gestalts (as mentioned above) resulting in seemingly continuous body motion.
- Grafton and Hamilton (2007), with discontinuous control through command hierarchies resulting in continuous motion and also emergent coarticulation, i.e. a contextual smearing of otherwise separate motion units into more continuous motion.
- Rosenbaum (2017), suggesting that motion is controlled by so-called goal postures with continuous body motion emerging from transitions between distinct postures.
- Sosnik et al. (2004), demonstrating how initially discontinuous, point-by-point motion may turn into more continuous smooth motion by coarticulation.
- Godøy (2014), arguing how coarticulation is at work in sound-producing body motion, resulting in a contextual smearing of both the body motion and the resultant sound.

An interesting development in motor control here is the theory of so-called *intermittent control*. This theory suggests that human body motion may be controlled discontinuously, i.e. in a point-by-point manner called *serial ballistic* control (Loram et al., 2011). The reason for this serial ballistic control scheme is again that our organism seems to be too slow to enable totally continuous control, and needs to work by anticipatory control, i.e. in a point-by-point manner by so-called *open loop* control. However, the perception of the output motion may be more continuous, hence the idea of 'observe continuously, act intermittently' (Loram et al., 2011, p. 317). Intermittent control is an ongoing research topic, but we may hypothesize that it could be a plausible way to reconcile continuity and discontinuity in sound-producing body motion as well, and hence also in our mental images of sound-motion objects.

5.6 Concluding remarks

The mentioned categories and criteria of a combined sound-motion object theory could make us aware of, and give names to, various perceptually salient features of the music. This could provide us with both analytic and generative tools, particularly useful for enhancing our grasp of the different sound-motion object types derived from Schaeffer's work (see Godøy, 2018 for details), such as:

- *Composed objects*, combining different components by juxtaposition, additively enhancing sound-motion objects.
- *Composite objects*, denoting components in a sequence, fused by coarticulation into more complex extended sound-motion objects.

Beyond the suitable objects mentioned above, there are also various other objects that can be named, and there are very many possibilities of object combinations and concatenations into more extended collage compositions. All this can have useful applications in musical analysis as well as in creative tasks, such as:

- Sound design and orchestration by systematic combinations of typological features.
- Composition and Improvisation as scripts with concatenated typological shapes.

Throughout these processes, thinking sound-motion objects and various typological and morphological features, could be a systematic approach to handling otherwise ephemeral material.

Hopefully, sound-motion objects, combining the remarkable insights of Schaeffer and his collaborators from more than half a century ago with current research on music-related body motion, could be a vehicle for further exploration of the holistic nature of musical experience. This could be expressed as a theory of *musical quanta* (Godøy, 2013), denoting a holistic object-centred approach combining the overall sound and motion features in a way that also reconciles continuity and discontinuity in musical experience.

Notably, such an object-centred approach would not exclude more macro-level approaches as can be seen, for instance, in projects of music information retrieval. Exploring features of larger collections of sound-motion objects or of more extended, macro-level works of music, e.g. *spectral centroid, spectral flux, harmonicity,* etc. using available software e.g. the *MIRtoolbox* (Lartillot & Toiviainen, 2007), could be combined with the more local sound-motion object feature studies. This could also provide us with quantitative information about various typological and morphological features, providing the acoustic correlations of these subjective features as was indeed the stated long-term aim of Schaeffer's music theory.

REFERENCES

1. Bizley, J. K. & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience, 14*, 693–707.
2. Bregman, A. (1990). *Auditory Scene Analysis*. Cambridge, Mass., and London: The MIT Press.
3. Chion, M. (1983). *Guide Des Objets Sonores*. Paris: Éditions Buchet/Chastel.
4. Chion, M. (2009). Guide to Sound Objects (English translation by John Dack and Christine North). ElectroAcoustic Resource Site (EARS), http://ears.huma-num.fr
5. Clayton, M., Dueck, B., & Leante, L. (Eds.). (2013). *Experience and Meaning in Music Performance*. New York: Oxford University Press.
6. De Freitas, J., Liverence, B. M., & Scholl, B. J. (2014). Attentional rhythm: a temporal analogue of object-based attention. *Journal of experimental psychology. General, 143*(1), 71–76. https://doi.org/10.1037/a0032296
7. Elliott, D., Helsen, W. F., & Chua, R. (2001). A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychological bulletin, 127*(3), 342–357. https://doi.org/10.1037/0033-2909.127.3.342
8. Gobet, F., Lloyd-Kelly, M., & Lane, P.C.R. (2016). What's in a Name? The Multiple Meanings of "Chunk" and "Chunking." *Frontiers in Psychology 7*(102). doi:10.3389/fpsyg.2016.00102
9. Godøy, R. I. (1997). *Formalization and Epistemology*. Oslo: Scandinavian University Press.
10. Godøy, R. I. (2003). Motor-mimetic Music Cognition. *Leonardo, 36*(4), 317–319.
11. Godøy, R. I. (2006). Gestural-Sonorous Objects: embodied extensions of Schaeffer's conceptual apparatus. *Organised Sound, 11*(2),149–157.
12. Godøy, R. I. (2009). *Thinking Now-Points in Music-Related Movement*. In Bader, R., Neuhaus, C., & Morgenstern, U. (Eds.). *Concepts, Experiments, and Fieldwork. Studies in Systematic Musicology and Ethnomusicology* (pp. 241–258). Frankfurt, Bern, Brussels: Peter Lang.
13. Godøy, R. I. (2010). Gestural Affordances of Musical Sound. In Godøy, R. I. & Leman, M. (Eds.) *Musical Gestures. Sound, Movement, and Meaning*. New York: Routledge.
14. Godøy, R. I. (2013). Quantal Elements in Musical Experience. In Bader, R. (Ed.), *Sound – Perception – Performance. Current Research in Systematic Musicology, Vol. 1*. (pp. 113–128). Berlin, Heidelberg: Springer.
15. Godøy, R. I. (2014). Understanding Coarticulation in Musical Experience. In Aramaki, M., Derrien, M., Kronland-Martinet, R. & Ystad, S. (Eds.): *Sound, Music, and Motion*. (pp. 535–547) Berlin: Springer.
16. Godøy, R. I. (2018). Sonic Object Cognition, In Bader, R. (Ed.), *Springer Handbook of Systematic Musicology*. (pp. 761–777). Springer Nature,.
17. Godøy, R. I. (2021). Perceiving Sound Objects in the Musique Concrète. *Frontiers in Psychology*, ISSN 1664-1078. doi:10.3389/fpsyg.2021.672949

18. Grafton, S. T. & Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26, 590–616.

19. Husserl, E. (1982). *Ideas Pertaining to a Pure Phenomenological Philosophy, First Book*. London: Kluwer Academic Publishers.

20. Husserl, E. (1991). *On the Phenomenology of the Consciousness of Internal Time, 1893–1917*. English translation by John Barnett Brough. Doredrecht/Boston/London: Kluwer Academic Publishers.

21. Klapp, S. T., and Jagacinski, R. J. (2011). Gestalt Principles in the Control of Motor Action. *Psychological Bulletin*, 137(3), 443–462.

22. Klapp, S. T., Nelson, J. M., & Jagacinski, R. J. (1998). Can people tap concurrent bimanual rhythms independently? *Journal of Motor Behavior, 30*(4), 301–322. https://doi.org/10.1080/00222899809601346

23. Lartillot, O., & Toiviainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction From Audio. *International Conference on Digital Audio Effects*, Bordeaux, 2007.

24. Loram, I. D., Golle, H., Lakie, M., & Gawthrop, P. J. (2011). Human control of an inverted pendulum: Is continuous control necessary? Is intermittent control effective? Is intermittent control physiological? *J Physiol 589*(2), 307–324.

25. 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.

26. Miller, I. (1982). Husserl's account of our temporal awareness. In Dreyfus, H. (Ed.), *Husserl, intentionality, and cognitive science* (pp. 125–146). Cambridge and London: The MIT Press.

27. Petitot, J. (1985). *Morphogenèse du sens. I. Pour un schématisme de la structure*. Paris: Presses Universitaires de France.

28. Petitot, J. (1990). 'Forme' in *Encyclopædia Universalis*. Paris: Encyclopædia Universalis.

29. Pöppel, E. (1997). A hierarchical model of temporal perception? *Trends in Cognitive Sciences, 1*, 56–61.

30. Reybrouck, M. (2021). *Musical Sense-Making: Enaction, Experience, and Computation*. New York: Routledge.

31. Rosenbaum, D. A. (2017). *Knowing Hands: The Cognitive Psychology of Manual Control*. Cambridge: Cambridge University Press.

32. Schaeffer, P. (with sound examples by Reibel, G., and Ferreyra, B.) (1998) (first published in 1967). *Solfège de l'objet sonore*. Paris: INA/GRM.

33. Schaeffer, P. (1966). *Traité des objets musicaux*. Paris: Éditions du Seuil.

34. Schaeffer, P. (2017). *Treatise on Musical Objects* (English translation by Christine North and John Dack). Oakland, Calif.: University of California Press.

35. Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science, 15*(11), 720–728.

36. Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in cognitive sciences, 12*(5), 182–186. https://doi.org/10.1016/j.tics.2008.02.003

37. Sosnik, R., Hauptmann, B., Karni, A., & Flash, T. (2004). When practice leads to coarticulation: the evolution of geometrically defined movement primitives. *Experimental Brain Research, 156*, 422–438 DOI 10.1007/s00221-003-1799-4

38. Spivey, M. (2008). *The Continuity of Mind*. New York: Oxford University Press.

39. Starrfelt, R., Petersen, A., & Vangkilde, S. (2013). Don't words come easy? A psychophysical exploration of word superiority. *Frontiers in human neuroscience, 7*, 519. https://doi.org/10.3389/fnhum.2013.00519

40. Thom, R. (1983). *Paraboles et catastrophes*. Paris: Flammarion.

41. Varela, F. (1999). The Specious Present: A Neurophenomenology of Time Consciousness. In Jean Petitot, Francisco Varela, Bernard Pachoud, & Jean-Michel Roy (Eds.), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. (pp. 265–314). Stanford, CA: Stanford University Press.

42. Wittmann, M & Pöppel, E. (1999). Temporal mechanisms of the brain as fundamentals of communication - with special reference to music perception and performance. *Musicae Scientiae*, Special Issue, 13–28.