

# Ecological Constraints of Timescales, Production, and Perception in Temporal Experiences of Music: A Commentary on Kon (2014)

ROLF INGE GODØY[1]

*Department of Musicology, University of Oslo*

**ABSTRACT:** In trying to structure our discussions of temporal experience in music, it could be useful to look at some basic ecological constraints of timescales, production, and perception of music. This may hopefully help us to distinguish between on the one hand readily perceived features of sound and music-related body motion, i.e. concrete sonic, kinematic, and proprioceptive features, and on the other hand, more generic, amodal, and abstract elements in musical discourse, manifest in various symbolic representations such as notation, numbers, and diagrams. Given easily accessible music technologies, it is possible to experiment with different editions of musical works, i.e. concatenate fragments in a different order and then evaluate the emergent contextual effects in listening experiments. Also, given the faculties of musical imagery (defined as our ability to mentally re-experience musical sound and body motion in the absence of physically present sound and body motion), we can at will recombine chunks of music in our minds and mentally scan through large musical works. The contention here is that such recombination in actual re-editing of musical sound or in musical imagery will still be related to the basic ecological constraints of the timescales, production, and perception in music.

Submitted 2014 August 22; accepted 2014 October 8.

**KEYWORDS:** *temporal experience, ecological constraints, contexts, abstract-concrete*

## INTRODUCTION

THE topic of time in musical experience has engendered innumerable publications and a rather heterogeneous body of conceptual frameworks, altogether making discussions of time in musical experience rather challenging. The reason for this is obvious: music is, to put it plainly, a temporal art in the sense that it basically unfolds in time, in chronometric time that can be measured in seconds, minutes, and hours. Time concerns pretty much all features of music as we know it, ranging from basic acoustic features of sound, and by way of most stylistic features, all the way to high-level affective and aesthetic features, including notions of large-scale musical experience such as of narrativity and drama. Additionally, we have discussions of whether some features of music are independent of time, e.g. as suggested by Iannis Xenakis (Xenakis, 1992) with his notion of “outside time” (“hors-temps”), or implicit in various notions of static categories (of pitch, duration, timbre, modality, harmony, etc.) of mainstream music theory. Further complicating the temporal-atemporal distinction is the now generally accepted insight that most other arts, including the plastic arts, are heavily time-dependent as well, e.g. that in perceiving a painting, there is first a frenetic activity of saccades, resulting in an experience of a coherent and seemingly “instantaneous” image of the painting only after a certain time interval.

The crux of the matter is in my opinion to recognize that all perceptual and cognitive processes take time, yet that we may have subjective notions of atemporal or “instantaneous” phenomena emerging from continuous time. As argued by Spivey (2008), the challenge is to try to understand how underlying continuous neurocognitive processes give rise to more discontinuous percepts and judgments. Clearly, recognizing the basically continuous nature of biological processes in an organism does by no means preclude subjective “outside time” or “instantaneous” mental images of whatever it is that we are perceiving or imagining. Rather on the contrary: there is now mounting evidence that not just auditory perception, but also perception and cognition in general, may proceed by what is subjectively experienced as discontinuous chunks, and in the case of music, include what I have elsewhere tried to denote as “quantal elements” (Godøy, 2013). The challenge is then that of trying to understand *how* various seemingly atemporal and relatively stable percepts may emerge from the flux of lived experience. This process of emergence, which I have earlier from my readings of

classical phenomenological and more recent cognitive science work denoted collectively as a “flux-to-solid” transition (Godøy, 1997), remains a major challenge in music perception research.

Obviously (and fortunately so), this flux-to-solid transition works well for us in music and in our everyday life, yet despite significant progress in neurocognitive research, also remains profoundly enigmatic: how is it that continuous auditory, visual, haptic, etc. sensations can engender more discontinuous and solid percepts in our minds? With this grand question lingering in the background, my approach here is more modest, focusing on some observable features in sound and in music-related behavior, and trying to advance our knowledge of time in musical experience by studying rather concrete aspects of music. This is the basis for my contribution to the discussion of the topics raised by Maria Kon's paper, namely to try to understand issues of time and succession in music by way of music-related sound and body motion features. The last point is crucial, i.e. that music is understood as consisting as much of body motion sensations as of sound, hence, that considerations of temporal experience in music need to include what can be called a *motor theory* perspective (Godøy, 2001; 2003; 2004; 2010a, 2010b).

For this reason, I shall in the following first present some considerations of timescales constraints in musical experience, followed by what I see as some essential constraints at work in music making and music perception, ending with some thoughts on exploring temporal experiences in music.

### TIMESCALES CONSTRAINTS

In musical sound, we have a very large range of timescales, extending from that of single vibrations and impulses to that of whole works, i.e. from that of milliseconds (and even shorter) to that of minutes and hours. Although we usually have several different timescales in superposition (i.e. singular vibrations in succession within tones, singular tones in succession within phrases, etc.), there are some distinct perceptual differences between various timescales in music.

As we know, we have the threshold between singular impulses and continuous sound at about 20Hz, and an approximately similar threshold in the visual domain between still pictures and moving images (the “flicker-fusion” threshold), and interestingly, also an upper speed limit for singular human body motion somewhere approaching 20 events per second (but there may be faster body motion when using the resonant properties of the body, e.g. as in vocal folds, tongue or lip vibrations). In the timescale area above this 20Hz threshold (extending in best cases up to approximately 20000Hz), we typically perceive salient musical features such as pitch, loudness and stationary timbral features, also referred to as “tone color”, i.e. not including various fluctuations that are found below this approximately 20Hz threshold. The rapid fluctuations and transients in the below 20Hz area are highly significant for our experience of timbre, and at a slightly longer timescale, we find various textural elements such as trills and tremolos. At the timescale in the approximately 0.5 to 5 seconds duration range, we typically find significant salient musical features such as envelopes (of loudness, pitch, and timbre), various style-defining rhythmical-textural as well as melodic and harmonic patterns, and importantly, also salient features of sound-producing and sound-accompanying body motion (cf. Godøy and Leman, 2010).

Additionally, we can also readily observe so-called phase-transitions between different qualitative features dependent on duration: e.g. if we have a series of rather slow tone onsets and then accelerate, this will sooner or later flip over to be perceived as a tremolo, and conversely, if we have a tremolo and gradually slow it down, it will sooner or later be perceived as a series of individual tones. Notably so, these phase-transitions are also found in the corresponding music-related body motion, e.g. between the motion of singular strokes and fused tremolo motion.

We have in our research found it useful to denote three major timescales in musical experience, the *micro timescale*, denoting the various features in the below 0.5 seconds duration range, the *meso timescale*, denoting what we perceive as coherent chunks with highly significant features in the very approximately 0.5 to 5 seconds range, and the *macro timescale*, denoting sequences of music above the approximately 5 seconds limit such as sections, movements and whole works, and usually consisting of concatenations of several meso timescale chunks. Again, this timescale classification applies to both sound and music-related body motion.

As we shall see later, we may in research zoom in and out of these different timescales at will, but for now, we should recognize that we have different qualitative features at different timescales. This is actually related to the mentioned enigmas of the flux-to-solid transition and is a source of various constraints in music in the sense that our organism (including the entire auditory pathway from the outer ear to various areas of the brain) seems effortlessly to make the transitions from continuous sound sensations to more qualitative and discontinuous percepts: we perceive pitch, timbre, dynamics, and other micro timescale features because our hearing apparatus automatically recodes continuous signals into more solid patterns in memory. Similar automatic transformations seem to be at work in our experience of body motion, including features such as amplitude, velocity, acceleration, and

various associated affective features of body motion such as calm, agitated, light, heavy, happy, sad, etc. We may refer to all these qualitative sensations as *ecological* elements of musical experience in the sense that they emerge from how our organism is made and interacts with the world. Collectively, these ecological elements could be seen as constraints on musical experience, and for this reason it may be worthwhile to look at both some production and perception constraints that condition our experiences of temporality in music.

### PRODUCTION CONSTRAINTS

At the meso timescale we have some important constraints on sound-production. The first is that musical sound and music-related body motion unfolds in time: a tone on a musical instrument or of the human voice will have an envelope, i.e. an attack segment followed by a (variably so) usually longer duration segment, and a decay segment. In our Western notation, all this is collapsed into a symbolic representation; however, dealing with temporal experience, this is something we should take into account. Furthermore, if we have several tone onsets in succession without complete damping between the tones, e.g. several tones on a piano with the sustain pedal on, we have a smearing of tones creating a reverberant context. The ecological aspects of sound production also extend to room acoustics, i.e. that the reverberant features of the place of music performance significantly shapes the sonic features and our perceptions of context, e.g. as when a choir performs in a cathedral.

In general, sound events are quite time-direction sensitive with regards to the overall envelopes: if we for instance listen to a recording of a hit cymbal played backwards, the recording will probably sound rather strange to many people because of the unusual energy envelope: a hit cymbal has an infusion of energy only at the impact point of the mallet with the cymbal, and the rest of the sound is a dissipation of this impact energy. If this sound is played back in reverse, we have in ecological terms the unusual schema of a sound building up and suddenly, at the peak, going silent. If we made a tremolo with two mallets on the cymbal, starting from very soft and then making a large crescendo to a peak, followed by an abrupt damping, we would have an envelope shape which resembles that of the reversed cymbal recording, but probably accept it as not so unusual because of the series of impulses in the tremolo. On the other hand, if we had long sustained tones with little or no evolution e.g. on wind instruments or bowed string instruments, a time reversal could perhaps pass unnoticed, as could also a rapid sustained tremolo on such instruments.

In general, we seem to be quite sensitive to most (but not all) ecological schemas of temporal evolution, and this goes also for higher level features: the order of tones, both in melodies and in chord progressions may (variably so) be critical. Knowing that Western music has shown much fascination with various temporal order manipulations, e.g. retrograde and augmentation in various instances of classical counterpoint and in dodecaphonic music, the emergent results of such manipulations are dependent on more global stylistic criteria for the music in question. For example, retrograde of harmonic progressions may violate relevant stylistic principles in some cases (e.g. in the resolution of dissonances and leading tones of functional harmonic music), in other cases not (e.g. in more modal or free tonal harmonic progressions). Likewise, retrograde and permuted orders of the 12 tones in dodecaphonic music may often be less significant than the overall harmonic content of the music. Common to such temporal order manipulations in Western music is that they are based on notation, i.e. on the abstraction of pitch into symbols that enable production of music with reversals of event sequences that otherwise would be difficult if not impossible to produce given ecological constraints on production and perception.

Furthermore, sound can be perceived as included in action trajectories of sound-producing body motion, and this is the meaning of the motor theory perspective: we tend to perceive music not only as a sequence of sounds, but just as much as body motion that produce the sounds (Godøy 2010a, 2010b). That musical sound is embedded in contexts of sound-producing body motion implies a number of constraints on temporal experience in music, including the already mentioned phase-transitions and so-called coarticulation, meaning a contextual smearing of sound-producing motion and the resultant sound, very much shaping the chunk timescale features in music (Godøy, 2014).

In brief, we seem to have a number of quite significant ecological constraints on the production of musical sound and on the associated body motion, constraints that are supplemented by some important perceptual constraints.

### PERCEPTION CONSTRAINTS

Musical features at different timescales will need different durations in order to be manifest in our minds. Some salient features of timbre and overall energy of the music may be perceived very quickly, actually down to the 250 milliseconds range in some cases as suggested by Gjerdingen and Perrott (2008), however other stylistic features require durations at the meso timescale to be perceived.

From a motor theory perspective, the mentioned constraints in the production of musical sound (phase-transitions, coarticulation) also have consequences for the chunking of musical sound. The contextual smearing of individual tone events into holistically perceived, gestalt-like chunks at the meso timescale may be so strong that these chunks become perceived as the basic units of temporal order in music (Godøy, 2013, 2014).

The combined production and perception constraints seem to unite in establishing meso timescale chunks as the most significant element for an ecological foundation of temporal order in musical experience. Breaking up such fused chunks into its lower level constituent parts will of course destroy them, and in fact create new and qualitatively different chunks. As such, we may freely create new chunks just simply by cutting out some fragment of musical sound, as was one of the original ideas of the *musique concrète* and ensuing research on sonic objects by Pierre Schaeffer and co-workers (Schaeffer, 1966; Chion, 1983; Schaeffer, 1998), and something that has been inherited by the present musical practice of DJ scratching. The point is that by repeated listening to a fragment, we accept almost any fragment of sound as a coherent chunk with a shape, i.e. as having an envelope with a start, a body, and an end. However, if a chunk of musical sound is broken up into even smaller fragments, e.g. in the below 200 milliseconds range, and then recombined, we will often have entirely new emergent perceptual features, something that in recent years has become popular as so-called *granular synthesis*.

On the other hand, meso timescale chunks may be recombined in innumerable alternative orders creating new macro timescale musical experiences, similar to new mosaic pictures from a collection of glass fragments. However, the macro timescale seems to be not so well studied in music perception research, although Western musical discourse has dedicated considerable attention to large-scale architectures in music such as sonatas and symphonies, often trying to document the importance of large-scale schemes for tonal, harmonic, and motivic/thematic organization. Until further notice, we should be a bit sceptical about the efficacy of such large-scale formal schemas: the little evidence we have seems to undermine some inherited notions of the perceptual significance of large scale formal schemas in Western music (Eitan and Granot, 2008).

Looking around at different musical cultures, we see that they often do not have these Western notions of large scale forms, and that the focus of the music often is on the meso timescale chunks: chunks may be recombined in different orders in performances of tunes, tunes that in spite of different concatenations of fragments are considered identical (Kvifte, 2001). Also in Western art music, we have seen chunk-based music, ranging from the *Musikalisches Würfelspiel* (perhaps wrongly) attributed to Mozart (cf. Hedges, 1978), to various 20<sup>th</sup> century music based on the concatenation of fragments as collage forms or by various principles of chance. The net effect of all this from a perceptual point of view is that we as perceivers unavoidably tend to make a context of whatever succession of events we are exposed to.

## EXPLORING TEMPORAL EXPERIENCES OF MUSIC

The considerations of timescales, production and perception constraints presented above serves to point out that music is grounded in a number of concrete experiences. In his grand project of establishing a more universal theory of music in the face of the new 20<sup>th</sup> century music and the music of non-Western cultures, Pierre Schaeffer introduced what he denoted as the *abstract-concrete* distinction (Schaeffer, 1966; Chion, 1983). The abstract is that which is based on generic categories represented by symbols, e.g. pitch and duration in Western music theory, whereas the concrete is that which relates directly to the substrate of the sound, i.e. its overall dynamic shape (envelope), its timbral content, both globally and more locally, and including its various nuances of pitch-related, dynamic, timbral, etc. fluctuations, etc., in short, all that we may perceive as the shape and content of the sound.

In our context, the concrete would include all the abovementioned ecological features of music, whereas the abstract would not. The abstract would be a symbolic representation of some ideal, unspecific generic categorical feature in Western music theory. Given this symbolic abstraction (for good and bad), the door is then open to all kinds of more abstract organizations of the tone symbols, e.g. as in the permutation of tone events into different sequences and at different timescales. In my opinion, we should be aware of the various ecological components of musical experience that may disappear in such more abstract approaches to temporal experience in music.

Given the multiple timescales and features at work in music, we also have the possibility to direct our attention volitionally to different features in the music. Again, this was one of Schaeffer's main points, namely that we may focus intentionally on different aspects in the music, relegating other features more to the background. Given present day readily available digital audio software, we can enhance very many features in recordings of musical sound in that we may scrub back and forth, may jump to any location in the music, may zoom in and out of different timescales (using suitable time stretching or compressing software, e.g. a digital phase vocoder), filtering out various frequency

regions, etc., in short, have a number of perspectives on real, physically audible musical sound, a literary speaking “hands-on” control of otherwise ephemeral sound features.

This idea of freedom of perspective in listening and research is related to another key concept of Schaeffer, what he called the *context-contexture* distinction: the term *context* denotes the overall surroundings of a sonic object in music, whereas the *contexture* denotes the internal context of the sonic object. Given our freedom of mental focus in listening and imagery (supplemented with the mentioned audio technologies), we may zoom in and out at will to whatever feature(s) we are interested in, progressively exploring what we believe are perceptually salient features of the music at different timescales.

Also with the help of readily available technology, we may by a few keystrokes and mouse movements cut, copy, paste, etc. fragments of music, in short, re-edit whatever work of music in whatever order we like, and then evaluate the results: how does an alternative version of a given musical work sound? This is something that could be done in experiments with several participants, e.g. as in Eitan and Granot (2008). In fact, this is an *analysis-by-synthesis* approach to exploring temporal experience in music, and there is nothing stopping us from experimenting with alternative concatenations of chunks in all kinds of music, producing as many alternative edited works as we like.

Lastly, *musical imagery*, defined as “our mental capacity for imagining musical sound in the absence of a directly audible sound source, meaning that we can recall and re-experience or even invent new musical sound through our ‘inner ear’” (Godøy and Jørgensen, 2001), gives us the freedom to make “armchair” studies of temporal elements in musical experience. A quick glance through a score is a way to have a highly compressed overview of an entire work, but what kinds of features will such a scan include? This probably very much depends on the musical imagery skills of the musician, as reflected in this famous quote by Paul Hindemith: “If we cannot, in the flash of a single moment, see a composition in its absolute entirety, with every pertinent detail in its proper place, we are not genuine creators.” (Hindemith, 2000, p. 61). However, also for non-experts, musical imagery seems to mirror very many of the “original” features of musical performance, including the sound-producing body motion (cf. Godøy and Jørgensen, 2001; Godøy, 2001; 2004; 2010b for overviews). This is very much in line with theories of mental imagery in general, i.e. that our internal mental images are intimately linked with actual experiences, both of the world and of our own bodies and body motion. Establishing such an “ecological” basis for the contents of our minds as opposed to more abstract modes of reasoning (sometimes referred to as “mentalese”) was in fact one of the main issues of the so-called mental imagery debate of the 1980s and 1990s (Kosslyn, 1994).

In summary, I believe we need to recognize that there are ecological constraints on musical features and on the production and perception of musical sound. These constraints should be taken into consideration when we reflect on temporal aspects of musical experience, including the issues of *experience of succession* (EoS) and *succession of experience* (SoE) mentioned in Maria Kon's paper. Fortunately, we now have readily available technologies enabling us to experiment systematically with temporal order in music, by an analysis-by-synthesis exploration of concatenations of musical events and judge the effects of these various alternative concatenations. More generally so, we may cultivate our capacities for experimenting with different concatenations of musical chunks in our minds, practicing a mental analysis-by-synthesis exploration of temporal aspects of musical experience.

## NOTES

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

## REFERENCES

- Chion, M. (1983). *Guide des objets sonores*. Paris: INA/GRM Buchet/Chastel.
- Eitan, Z., & Granot, R. Y. (2008). Growing oranges on Mozart's apple tree: “Inner form” and aesthetic judgment. *Music Perception*, 25(5), 397-417.
- Gjerdingen, R., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research*, 37(2), 93–100.
- Godøy, R. I. (1997). *Formalization and epistemology*. Oslo: Scandinavian University Press.
- Godøy, R. I. (2001). Imagined action, excitation, and resonance. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp.237-250). Lisse: Swets & Zeitlinger.
- Godoy, R. I. (2003). Motor-mimetic music cognition. *Leonardo*, 36(4), 317-319.

- Godøy, R. I. (2004). Gestural imagery in the service of musical imagery. In A. Camurri & 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: Springer.
- Godøy, R. I. (2010a). Gestural affordances of musical sound. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp.103-125). New York: Routledge.
- Godøy, R. I. (2010b). Images of sonic objects. *Organised Sound*, 15(1), 54-62.
- Godøy, R. I. (2013). Quantal elements in musical experience. In R. Bader (Ed.), *Sound – perception – performance. Current research in systematic musicology, Vol. 1* (pp. 113-128). Berlin, Heidelberg: Springer.
- Godøy, R. I. (2014). Understanding Coarticulation in Musical Experience. In M. Aramaki, M. Derrien, R. Kronland-Martinet & S. Ystad (Eds.): *Sound, Music and Motion. Lecture Notes in Computer Science*. Berlin: Springer, 535-547.
- Godøy, R. I. & Jørgensen, H. (Eds.). (2001). *Musical imagery*. Lisse: Swets & Zeitlinger.
- Godøy, R. I. & Leman, M. (Eds.). (2010). *Musical gestures: Sound, movement, and meaning*. New York: Routledge.
- Hedges, S. A. (1978). Dice music in the eighteenth century. *Music & Letters*, 59(2), 180-187.
- Hindemith, P. (2000). *A composer's world: Horizons and limitations*. Mainz: Schott.
- Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, Mass.: MIT Press.
- Kvifte, T. (2001). Images of form: An example from Norwegian Harding fiddle music. In R. I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp.219-235). Lisse: Swets & Zeitlinger.
- 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.
- Spivey, M. (2008). *The continuity of mind*. New York: Oxford University Press.
- Xenakis, I. (1992). *Formalized music: Thought and mathematics in composition* (2<sup>nd</sup> ed.). Stuyvesant, NY: Pendragon Press.