

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/29977> holds various files of this Leiden University dissertation.

Author: Çamci, Anil

Title: The cognitive continuum of electronic music

Issue Date: 2014-12-03

CHAPTER 5

DIEGESIS: A SEMANTIC PARADIGM

DIEGESIS

Narratologist Gérard Genette defines *diegesis* as “the spatiotemporal universe” referred to by a narrative (1969: 211). The concept of diegesis can be traced back to Plato’s dichotomization of narrative modes into imitation and narration (1985 [c. 380 BC]: 247). However, it has since yielded various incarnations that have been used for describing narrative structures in art, and situating the components of an artwork in relation to one another. On a meta-level, the resulting narratological perspectives also provide insights into the fabric of the artistic experience by delineating relationships between the artist, the artistic material and the audience.

As an artistic form of temporal nature, music too prompts narratives. The narrative mode of music is determined by its language. In instrumental music, narratives are conveyed to the listener through a culturally embedded musical language that has been established over the course of centuries. Since the material of electronic possesses a cognitive disposition which extends beyond the well-ingrained structures of a traditional musical language, electronic music can assume a mimetic role: listeners are *presented* with sounds that can *represent* extra-musical events while the medium of the recounting remains the same as that of the recounted. However, when this material meets the esthetic capacity of the listener, the physical artifact is inevitably succeeded by the manifestation of a story. Therefore, a diegesis (i.e. a spatiotemporal universe) emerges in the intellectual domain. The cognitive processing of electronic music institutes a bond between the mimetic and the diegetic. As a result, the figure and ground relations between gestures extend beyond those of physical forms, and a narrative unfolds both in the spatial domain of the concert hall and in the semantic space superimposed onto this domain by the listener.

In this chapter, I will discuss the semantic and physical dimensions of electronic music and describe how these two dimensions come in contact with each other. I will assess the explicit and implicit aspects of the listener experience while, at the same time, questioning the extents to which the listener is inside or outside the musical material. Rather than merely contrasting the mimetic and the diegetic aspects of electronic music, I will describe the role of their coalescence in actively shaping our experience of electronic music.

An Interdisciplinary Contextualization of Diegesis

Throughout history, the concept of diegesis has come to assume several meanings. Most of these meanings can be associated with the modern field of narratology. In Platonic mimesis, events “either past, present, or to come” are presented through imitation. Diegesis on the other hand, relies on narration (Plato 1985 [c. 380 BC]: 247). Therefore, under Plato’s taxonomy of narrative forms, theatre is mimetic, because actors imitate (i.e. re-enact)

situations, while poetry is diegetic because the poet, speaking in their own person, recounts events as a narrator who is external to the immediate world of the story.

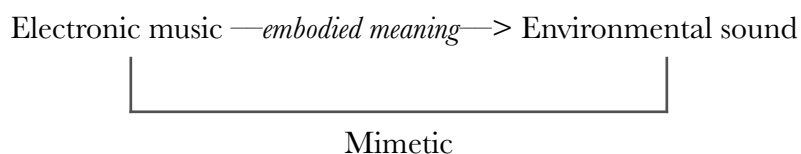
Art forms, and moreover, how artistic material is experienced, have indeed evolved considerably since Plato's delineation of these concepts. For example, in Plato's time, reading poetry was not a recreational activity one would enjoy individually. Rather, the artists themselves recited their poems to the public during gatherings. Due to changes in such practices, concepts of diegesis and mimesis have also been redefined several times to accommodate new art forms and new esthetic routines. These redefinitions have inevitably given rise to contradicting views amongst theorists who have elaborated new demarcations of the terms specific to particular art forms.

In film, for instance, the sounds that occur within a scene (e.g. a dialogue between two characters, or music coming from a radio in the scene) are considered diegetic, while the film score that emerges from outside of the universe of the story is labelled as non-diegetic sound (Taylor 2007). Diegesis serves broader functions in film theory. A prominent perspective is that all film is diegetic because the director chooses certain parts of the story's universe to be displayed on the screen. The director therefore assumes the role of a narrator, and all that is going on on-screen is illusory (Hayward 2006). But then we can question which form of art fails to meet this criterion. Even a generative work of art necessitates a moment in time when the artist initiates an algorithm, thus creating a narrative context for the piece. From this point of view, all artistic material can be deemed to display diegetic features, even when these are communicated through a representational form that is mimetic in the Platonic sense.

Genette applies the term *diegesis* exclusively to literary theory. By describing such subcategories as *heterodiegetic* and *extradiegetic* (1980: 50; 1969: 202, 212), he creates a narratological terminology, which he utilizes to situate the author, the reader and the components of a literary text (i.e. characters, venues, time) in relation to one another. Differentiating between cascading layers of a narrative by starting from the physical world of the author on the outermost level, Genette outlines the concept of diegesis as the *spatiotemporal universe* to which the narration refers. Therefore, in his terminology, a *diegetic* element is "what relates, or belongs, to the story" (translated in Bunia 2010: 681). Here we can observe a thread emerging between Genette's literary definition of the term and its usage described above in the context of film where a diegetic sound originates from a source that belongs to the scene.

Coalescence of Mimesis and Diegesis

As previously described, in Plato's categorization, tragedy and comedy (i.e. theatre) are mimetic modes of narrative, while poetry and mythology (i.e. literature) are diegetic. The distinguishing factor between the narrative actions pertinent to each mode, namely re-enactment and recounting, is whether the medium of representation remains the same as that of the represented, or, in other words, whether there is a mediation between the expression and the expressed. Electronic music, in this sense, is mimetic. While it may represent extra-musical events, it does so through connotations of sound. It is not narrated like diegetic poetry, but speaks for itself; it represents not as a mediator but as a portion, or an abstraction, of reality. The loudspeaker will detach the sound from its source but the medium of the phenomenon remains unchanged. Here, we can revisit the semantic relation between gestures in electronic music and environmental sounds described in the previous chapter:



The embodied meaning which constitutes the semantic relationship between electronic music and environmental sounds is mimetic in nature because the medium of both is sound. However, electronic music does evoke more than memories of sounds, just as an environmental sound signifies more than the physical entity which emitted it. What is auditorily perceived can be an artifact of a phenomenon but not the phenomenon itself:

Let us begin by noting that *information about* something means only *specificity* to something. Hence, when we say that information is conveyed by light, or by sound, odor, or mechanical energy, we do not mean that the source is literally conveyed as a copy or replica. The sound of a bell is not the bell and the odor of cheese is not the cheese (...) Nevertheless, in all these cases a property of the stimulus is univocally related to a property of the object by virtue of physical laws. (Gibson 1966: 187)

Furthermore, every sound we hear triggers a semiotic web, allowing us to imagine and comprehend more than what the sound immediately represents. The mimetic acting of tragedy engenders a similar reaction. What we witness on stage is just a portion of the world we imagine and situate the characters within. Here a bond between Plato's mimesis and Genette's diegesis comes into being. Although a narrative form might be purely mimetic, it will nevertheless imply a spatiotemporal space different from that which the audience inhabits. Electronic music *presents* to listeners sounds that *represent* events; it does not speculate about — or recount — sounds. Electronic music is therefore mimetic in the spatial domain of the concert hall, but it creates a diegesis for the listener in the semantic domain.



(Re)presentation

However, the listener can be absent from the artwork's universe. If we go a little further with the adaptation of dramaturgical concepts, electronic music engages with listeners in a similar fashion to *representational acting*. This type of performance ignores the presence of an audience and situates them outside the context of the unravelling universe of the story. This is unlike presentational acting, which acknowledges the audience and moreover, addresses them. Actor Constantin Stanislavski's typology of these terms, although from an entirely different perspective, reveals new threads across separate art forms. Stanislavski asserts that presentational actors "must live the part every moment that [they are] playing it" (Stanislavski 1948: 19) and expose the character through their understanding of it, *becoming one* with their role. Representational actors on the other hand, do not live the part but *play* it. The actor "remains cold toward the object of his acting but his art must be perfection" (22).

Parallel to Stanislavski's definition of presentational acting, American visual artist Sanford Wurmfeld describes *presentational art* to be "structured by a human being and presented as a statement (...) to be experienced or received by an active viewer. By its sensory nature, such art is untranslatable and the ideas or feelings transmitted by it are tied to the particular object that expresses them" (Wurmfeld 1993). The ideas of untranslatability and affixing of affect to the art-object intrinsically relate Wurmfeld's discourse to the previously described experience of a musical material expressed through a culturally embedded musical language, whether it is in the context of an instrumental or an electronic work. Furthermore, in instrumental music, the presence of an external performer *presenting* a composer's work is explicit. We could therefore align the experience of a traditional musical language with Wurmfeld's and Stanislavski's definitions and classify it as having presentational qualities. As I will discuss in the next section, presentational qualities in electronic music can manifest themselves in two ways — through the composer's presence in the work, or through the listeners' awareness of their physical selves.

Narrativity

Deleuze and Guattari describe that the artist's greatest challenge is to make an artwork stand up on its own. For them, this requires "from the viewpoint of lived perceptions and affections, great geometrical improbability, physical imperfection, and organic abnormality" (1994: 465). The idea of the "abstract, as a negation of reality" discussed in Chapter 3 applies to such imperfection and abnormality; an improbability is intrinsically defined in relation to our preconceived notions of what is probable. In literature, a text that demands too much interpretation prompts the reader to naturalize it by "using acquired knowledge to simplify it" or resolving semantic inconsistencies by turning to the narrative structure (Mikkonen 2011: 113). An assumption of *world semantics* is transferred from the real world to the fictional world (Bunia 2010: 699). Impossible fiction is therefore considered "an ostensible oxymoron" (Ashline 1995: 215).

While there is no narrator in non-vocal music comparable to that in a literary text, listeners partly assume this role by constructing stories out of their experience of the narrative. Narration "can inform us about a universe and yet restrict its information to a small set of events and characters populating this universe" (Bunia 2010). The artwork does not need to provide every element of the diegesis, since the listener expands the physical experience with the semantic by filling in the gaps. This license on the listener's part is further apparent in Souriau's interpretation of diegesis, which he describes as "all that belongs, 'by inference,' to the narrated story" (Gorbman 1980: 195). A similar attitude is described using the term "the principle of minimal departure" (Ryan 1980). This principle states that we structure our interpretation of alternative realities as closely as possible to our own realities (403): we project things we know about the real world upon the implied reality of a story (406). A narrative can restrict the information it communicates about a universe to a small set of actors and events that populate it (Bunia 2010: 686). But by virtue of the principle of minimal departure, we are able to form "reasonably comprehensive representations" of such worlds even though they are always described incompletely. An imagined spatiotemporal universe can be insufficiently narrated, but the universe will nevertheless be logically complete.

Genette describes narrative as a statement or a discourse which signifies a story (Genette 1980: 27). Cultural theorist Mieke Bal describes *narrative text* similarly as a text in which a story is told. However, Bal divides what Genette refers to as a story into two layers, namely *story* and *fabula*. In Bal's definition, a story is "a fabula presented in a certain manner" (1997: 5). By this,

Bal means that a *fabula* consisting of a series of chronologically connected events can be communicated to cause different appraisals and therefore different stories. Furthermore, Bal emphasizes that a text can be understood as a narrative in any medium including “language imagery, sound, buildings, or a combination thereof” (5). When situated in Nattiez’s semiological model, Bal’s *text* corresponds to the trace¹⁷ and *fabula* is the outcome of the reader’s esthesis (or interpretation) of the text manipulated by the story (9). Genette’s and Bal’s approaches can be differentiated with the below adaptations of the semiological schemas discussed in Chapter 3:

Genette’s approach: Narrative (trace) —*esthesis*—> Story

Bal’s approach: Producer —*poiesis* (story)—> Text <—*esthesis* (fabula)— Receiver

Deriving from Bal’s model, Meelberg describes narrative as “the representation of temporal development”. He further emphasizes that narrative represents a succession of events in a process and it is not the process itself (2006: 39). Theorist David Herman characterizes narrative as “a basic human strategy for coming to terms with time, process, and change” (2009: 2). According to Roads, we constantly construct narratives from our sensory experiences “by anticipating the future and relating current perceptions to past” ([forthcoming]). From this point of view, it is inevitable for listeners to extract narrative structures from their musical experiences due to the simple fact that a piece of music encapsulates a series of events between a starting point and an anticipated ending in the future. Recalling the discussion on gestural intentionality in the previous chapter, the extent to which the extracted structure is concordant with the composer’s design does not jeopardize its materialization.

In instrumental music, the narrative emerges in the abstract realm of the musical sound: musical expectancy is built upon either culturally or physiologically evaluated traits of music. Music can narrate emotional experiences (Walton 1994: 60), or simply a form unfolding over time (Meelberg 2006: 1). But as stated in Chapters 3 and 4, electronic music adds to this experience a distinct form of representationality made possible by the electronic medium. A layer of meaning attribution intercepts the passage from physical phenomenon to musical appraisal. Listeners inhabiting the spatial domain of the concert hall superimpose semantic representations over their embodied experience of the sounds. The affective quality of the artwork is immanently informed by this act. Here are two general impression entries about *Birdfish* as written by two participants:

I heard robotic bugs moving around being commanded by more intelligent robotic beings. There was water, stepping into water, robotic dialogues and also progress made by the robotic bugs in their task.

This music reminded me of a cartoon I used to watch when I was in high school. I related the piece to the story of the cartoon which told the struggles of liquid-like alien creatures who on the one hand were not from this world but on the other hand had to adapt to survive.

In the real-time descriptors of the first participant, the same narrative is apparent but the robotic bugs are replaced with “bugs” as the workers who are “flying” and “walking” while

¹⁷ Although Bal provides another definition of *fabula* as “a memorial trace” of a story “that remains with the reader after completion of the reading” (1997: xv), trace, as she refers to it, has a different meaning than how Nattiez uses the term in his model described in Chapter 3.

making progress on a given task. In the second participant's real-time descriptors, instead of aliens, there are other creatures referenced as actors such as "baby bird", "huge ant", "snail" and "worm". The story, on the other hand, persists with such descriptors as "sent to earth", "can't fit in" and "struggle again". The timing of the latter coincides with that of such descriptors by the first participant as "some adjustment" and "project continues". A correspondence chart between two stories can therefore be constructed as seen below:

Character	Setting	Action
Bugs (robotic)	A project	Making progress
Creatures (aliens)	Earth	Struggle

In these cases, we can observe two distinct stories constructed from the same narrative. However, there is an apparent pairing between how these two separate diegeses are populated and how the actors populating these universes act. In the next section I will examine those physical and semantic features of electronic music which contribute to the construction of such stories from a given narrative. I will return to congruences across different stories that emerge from the same narrative when I discuss the effects of semantic context on the construction of diegeses.

DOMAINS OF EXPERIENCE

As described in Section 1 of Chapter 3, there are several mental mechanisms which underlie the evocation of musical emotion. Said mechanisms were brain stem reflexes, evaluative conditioning, emotional contagion, visual imagery, episodic memory, and musical expectancy (Juslin and Västfjäll: 559). These mechanisms do not function in a mutually exclusive manner but rather assume complementary roles in inducing musical emotion (563). Juslin and Västfjäll offer a somewhat simplified scenario to explain how these mechanisms might be activated when listening to a piece of instrumental music:

A sudden, dissonant chord induced a strong feeling of arousal (i.e., brain stem reflex), causing [the listener's] heart to beat faster. Then, when the main theme was introduced, he suddenly felt rather happy – for no apparent reason (i.e., evaluative conditioning). In the following section, the music turned more quiet . . . The sad tone of a voice-like cello that played a slow, legato, falling melody with a trembling vibrato moved him to experience the same sad emotion as the music expressed (i.e., emotional contagion). He suddenly recognized the melody; it brought back a nostalgic memory from an event in the past where the same melody had occurred (i.e., episodic memory). When the melody was augmented by a predictable harmonic sequence, he started to fantasize about the music, conjuring up visual images – like a beautiful landscape – that were shaped by the music's flowing character (i.e., visual imagery). Next, the musical structure began to build up towards what he expected to be a resolution of the tension of the previous notes when suddenly the harmonics changed unexpectedly to another key, causing his breathing to come to a brief halt (i.e., musical expectancy). He thought, “This piece of music is really a cleverly constructed piece! It actually made me reach my goal to forget my trouble at work.” Reaching this goal made him happy (i.e., cognitive appraisal) (563)

While such so-called goals are rarely a conscious part of the listener's experience, these mechanisms constitute components of how we navigate through not only musical experience, but also the experience of daily life because, as Juslin and Västfjäll conclude, “music evokes emotions through mechanisms that are not unique to music” (559). In Chapter 3, I identified survival instincts and the evolution of our perceptual mechanisms as the bases of our appraisal routines. Therefore, studies on the cognition of sound provide us with insights into how the human mind copes with external stimuli, whether these may emanate from environmental phenomena or the speakers in a concert setting. Throughout this section, I will introduce evidence from both my experiments and other empirical studies to trace out compositional strategies which exploit the semantic and the physical domains of electronic

music. I will bring these findings together with the theoretical perspectives presented in the earlier sections to portray various dimensions of narrativity in electronic music.

The Physical Domain

The physical domain represents the empirical reality which the listener inhabits as a corporeal entity. Sound is a part of this reality as an acoustic phenomenon. Time in the physical domain progresses at a steady rate. The discussion below relates to physical objects and the measurable qualities of sound, as well as the implications of such qualities on the listening experience.

Perceptual Properties of Exposure

According to Moore and Hedwig, when hearing a sequential stream of sounds we can either perceive it as coming from a single source (i.e. fusion or coherence), or from multiple sources (i.e. fission or stream segregation) (2012: 919). They describe the “degree of perceptual difference” as determining the extent to which stream segregation occurs (2002: 320). This degree is dependent on various attributes of a sound such as lateralization between the two ears, frequency, temporal envelope and phase spectrum (320): “[w]hen the differences between successive sounds are very large, fission nearly always occurs, whereas when the differences are very small, fusion nearly always occurs” (2012; 919). Differences of intermediate size bring into being a property called “bistability” during which the percept flips between one or multiple streams (919). This provides an empirical ground for the compositional strategy of articulating a gesture by shifting its spectral, dynamic or temporal characteristics. Although in visual perception figure and ground configurations establish depth relationships between concurrent elements, Moore and Hedwig’s findings pertaining to the auditory domain supports the idea that non-layered sequential sounds can be imbued with figure or ground roles through perceptual segregation.

In an fMRI-based experimental study of emotional and neural responses to music, Chapin et al. found that tempo fluctuations caused emotional arousal in both experienced and inexperienced listeners. In another experimental study on the perception of emotional expression in musical performance, Bhatara et al. concluded that timing variations alone had more impact than variations in amplitude, although the latter was also found to communicate a significant amount of expressive information (2011: 932). Huron similarly draws attention to the effects of the dynamic and temporal characteristics of music on emotional experience. For instance, frisson response is found to be correlated with loud passages in music (Huron 2006: 34). On the effects of temporal organization, Huron offers an intriguing perspective on *repetition*. In what he refers to as an *orienting response*, individuals turn their heads in the direction of an unexpected sound as a basic reflex. However, when such a stimulus is repeated, the individuals will habituate to it after a while and cease the orientation reflex. A sufficiently novel change in the sound however will cause *dishabituation* and the individuals will reorient to the stimulus (49).

The opening impact sound of *Christmas 2013* recurs twice in its unaltered form and once more in a less reverberant environment. This is an unmistakable and relatively loud gesture. Participants referred to the first instance of this sound with “bam”, “sudden start” and “percussion” which describe the abruptness of the sound. The subsequent instances of the gesture however went unmarked by these participants which indicates a habituation.

Additionally, other impact sounds of comparable amplitude appearing throughout the piece were usually noted only once by the other participants.

From an entirely different yet convergent perspective, Truax describes *habituation syndrome* as individuals' adaptation to annoying noises in their acoustic environments as a result of extended exposure (1984: 90). He further characterizes habituation as a form of desensitization (90). When a noise lacks an immediate significance, we are able to habituate to it and disregard it as background noise. However, as the level of background noises increase, so does their "information load on the brain" (23). A musical correlate of this is explored in an experimental study by the music psychologist Diana Deutsch where she found that unstructured musical sequences "imposed a much heavier memory load" on the listener than structured ones.

With the help of Stockhausen's *Unity of Musical Time*, which characterizes *noise* as event sequences of irregular durations (1962: 41), these two different research perspectives can be linked together: a lack of periodicity, which is typical of daily soundscapes, can be considered to cause a more mentally demanding experience for the listener when compared to that of an even temporal structure in music. In that respect, a clear rhythmic grid can be utilized as a device of relief, for instance, in an electronic music piece which exhibits a temporal complexity matching of everyday sounds. A powerful representation of this is heard at the end of Roads' piece *Touche pas*, where he uncharacteristically introduces looping of a 3-second segment towards the end of the piece, which had thus far evaded such patterned structures. While one participant distinguished this moment as being funny, another participant entered the number of times the loop repeated, only to immediately correct himself, almost taking pleasure in his ability to keep track of the number of instances. Many other participants used descriptors, such as "repetition", "counting", "rhythmical", "coda" and "loop", which indicate a clear engagement with the section.

Although individual events in daily life might exhibit periodic behaviors (e.g. "an engine running"), multiple periodic events almost never align to a temporal grid, meaning that environmental sounds do not share a rhythm. As a result, rhythmic patterns are almost always *crafted* and indicative of musical structure. A rhythm in the context of an electronic music piece that is composed mostly without a rhythmic grid therefore represents a powerful tool, as it is capable of disrupting the diegetic experience by inciting an immediate awareness of a musical structure. Amidst the uneven temporal configuration of *Christmas 2013*, a half-measure reprise of the *Silent Night* at 1'26", which lasts a relatively brief duration of approximately 1 second, was enough to incite tempo-related descriptors such as "rhythm", "rhythm again", and "little beat" indicating a tendency to latch onto temporal patterns.

Another example in the same vein is *Element Yon*, in which the composition of abstract sounds constantly reveals temporal and spectral contrasts. For this reason, the piece can be characterized as consisting of "unstructured musical sequences", to use Deutsch's term. The impact of this quality is apparent in the real-time descriptors, such as "exhausting", "chaotic" and "confused", submitted by participants who listened to *Element Yon*. Furthermore, in his general impressions, one participant stated that "sounds without a rhythm made [him] curious but at the same time they were really exhausting". A moment in the piece which this participant accordingly marked as "exhausting" was marked by another participant as "I repeat and repeat but you don't get it". This participant had already elaborated on this section in his general impressions when he likened his experience to witnessing a redundant argument between people. In his general impressions, another participant described a section

towards the middle of the piece as having become boring because of unclear sound structures.

Awareness of the Physical Self

The sounds of electronic music come into contact with the listener in the physical domain. The embodied experience of the listener can cause a perceptual focus on the physical characteristics of a sound. But, furthermore, a sound can make the listener aware of his or her own physical presence. For instance, boundary cases in loudness and frequency are capable of prompting such phenomena: spectral extremes at sizable amplitudes, such as very high and very low frequencies that are clearly audible, make listeners conscious of their acts of listening. As I discussed in Chapter 3, when we fail to identify a source for a sound, we tend to focus on its acoustic properties. Given the abstract nature of *Element Yon*, it was therefore expected that the listeners would be more inclined to describe the spectral characteristics of the piece. But more interestingly, the results have revealed that *Element Yon* has generated a much more articulated sense of self when compared with the results from other pieces. Participants who listened to *Element Yon* reported their general impressions mainly in prose form and using the first person (i.e. “I felt (...”). Contrastingly, participants who listened to *Birdfish* commonly assumed the role of an outside viewer who observes and reports the unfolding of certain events (i.e. “(...) happened”). In this sense, the experience of *Birdfish* can be likened to that of watching representational acting, during which the audience is situated outside the diegesis. The self-awareness apparent amongst the listeners of *Element Yon* is akin to what is experienced by audience members when they are personally addressed by the actor during a play. The music becomes a *presentational object* as the material addresses the audience by making the listener acknowledge his or her own presence.

Almost half of the participants who listened to *Element Yon* expressed a form of annoyance with the high frequencies by submitting such descriptors as “disturbing”, “annoyed” or “irritating”. Two participants described physical pain in such comments as “painful sometimes” and “it hurts”, although they later reported having had a comfortable listening experience. However, three quarters of the participants particularly noted the rests in the piece either by pointing out the pockets of silence themselves, or by describing the relief they induced. While *Birdfish* incorporates the momentary use of high frequency gestures comparable to those in *Element Yon*, only one participant used the descriptor “harsh high” to indicate a similar annoyance. Furthermore, although the silences are structured very similarly between the two pieces, none of these silences were denoted as bringing relief in the context of *Birdfish*.

This kind of self-reference is apparent also in the general impressions of the participants who have evaluated *Element Yon* more conceptually. This indicates that the aforementioned boundary cases are capable of piercing through the diegesis and bringing the listeners back to a cognizance of their physical presence in the concert hall.

The Experienced Listener

As I mentioned in Chapter 2, the participants in my experiment varied greatly in terms of musical background. Amongst this variety of participants were students of sound engineering, sonology and sonic arts. These students therefore had either been exposed to a repertoire of electronic music or were capable of discussing auditory phenomena in technical terms. A visible tendency in the results obtained from these participants was the incorporation of

technological listening (Smalley 1997: 109) in their impressions. This was mostly evident in the meta-descriptors submitted by these participants, which addressed the technology or technique behind the music rather than their experience of the music *per se*. Such impressions ranged from the description of synthesis techniques to spectral, spatial and formal analyses of the works. Descriptor characteristics did not otherwise vary significantly between experienced and inexperienced listeners: while prior exposure was of use in the characterization of abstract elements, the descriptors denoting representational forms were consistent in terms of both content and frequency across participants of varying musical backgrounds.

Presence of the Composer in the Work

Just as sounds can make listeners aware of their physical selves, they can also make them aware of the composer as an external being. Gestures which accentuate the presence of the composer are presentational. These gestures can also take place in representationally rich pieces. From a narratological perspective, the presence of a composer in a representational narrative implies an *extradiegetic narrator* who is outside the diegesis but inside the spatiotemporal universe of the *narration* “understood in terms of physics” (Bunia 2010: 683). In the experiment results, such instances manifested themselves in meta-descriptors that refer to the material being of the piece in a way that acknowledges a poietic entity. For instance, in their real-time descriptors, several participants directed questions towards the pieces, as in “where is it going?” (*Christmas 2013*), “what’s the point?” (*Element Yon*), “is it repeated?” (*Birdfish*), and “repetition, why?” (*Touche pas*). Other descriptors referred to those aspects of the piece that are conceived at the level of narration, such as “final crescendo” (*Birdfish*), “development”, “harmonic progression”, “motif” (*Touche pas*), “chaotic composition” (*Element Yon*) and “gesture” (*Diegese*). These meta-descriptors refer to the works as the conscious products of a composition process.

[W]e will have to mark the contrast between mimetic and diegetic by a formula such as: *information + informer = C*, which implies that the quantity of information and the presence of the informer are in inverse ratio, mimesis being defined by a maximum of information and a minimum of the informer, diegesis by the opposite relationship. (Genette 1980: 166)

The more the composer (i.e. the informer) articulates his or her presence in the piece, the less of a self-sustaining environment the diegesis implies. This is an aesthetic choice which impacts the listener’s experience considerably. The choice is made between the intelligibility of a human performer versus the autonomy of a diegesis, or between the listener being left unattended in the universe narrated by the piece, and the listener coming into contact with the composer.

A lack of thought concerning the implications of self-presence in a piece is another sign of the novice composer. Adorno refers to the aesthetic paradox of “making the impossible possible” when he asks how making can “bring into appearance what is not the result of making” (Adorno 2002: 107). If it is indeed the intention of an electronic music composer to suggest a spatiotemporal universe entirely separate from that which the listener inhabits, the *craft* should be extracted from the artifact. If the listener can visualize the hands of the composer as they change a parameter, the implied intentionality will amount to a presentational gesture that will take over the narrative. This becomes apparent, for instance, when creating parameter automations: if the imperfections of performing a parameter-change by hand are not smoothed out, the experience of the listener can immediately swing

from hearing a wind blowing to a visualizing a performer changing the cut-off frequency of a filter on the mixer. The latter too can be expected of electronic music, but it should nevertheless be the product of a poetic initiative rather than a lack of precision.

The Semantic Domain

The semantic domain of an electronic music piece first emerges during the composer's act of poiesis and becomes translated into the physical domain. Another semantic domain is later constructed by the listener. This domain accommodates the story extracted from the narrative. Every instance of new material primes the listener contextually for what is to come, and the flow of gestures constitute a constant semantic realignment. Even when an explicit sound element is removed from the scene, it implicitly persists. The diegesis established thus far in the piece maintains a semantic context. In the listener's mind, past diegetic actors interact with present ones, and the listener starts filling in the gaps. Through imagining the implicit world of a piece, the listener, for example, can obtain semantic polyphonies from perceptual monophonies and construct implied figure and ground relations. This act of *world making* creates an immersive listening experience.

Effects of Semantic Context

In an experimental study which examined the effects of context on the identification of everyday sounds, Ballas and Mullins found that contextual inconsistencies had a negative impact on the identification of sounds (1991: 199). In another study on semantic distance effects, Vigliocco et al. used visual representations of objects and actions to demonstrate a similar impact of semantic context on identification (2002: B61). Furthermore, Orgs et al. offered corroborating evidence of a comparable effect with a cross-modal study (i.e. between visually displayed words and sounds) on conceptual priming (2006: 267). In another study by Ballas, sound identification time and causal uncertainty were found to be highly correlated (1993: 250). Finally, Guastavino's study on environmental sound categorization affirmed the determinant role of presentation context of a sound (2007: 54).

These findings can be correlated with the concepts of contextual meaning and semantic coherence briefly touched upon in Chapter 4. According to Wishart, "[c]ontextual cues may not only change our recognition of an aural image, but also our interpretation of the events we hear" (1996: 152). Cognitive cues instigate the formation of semantic contexts for gestures. In *Birdfish*, clear references to water and organic creatures, which were the two most salient types of descriptors for this piece, caused listeners to imagine possible environments (i.e. contexts) such as "underwater", "lake", and "aquarium". When the recognition of amphibian-like sounds was evaluated within a space articulated with reverberation, such descriptors as "cave" and "dungeon" appeared with the former being one of the salient descriptors in the preliminary studies. Similarly, a combination of water-like sounds with the inference of a cave formed general impressions such as "water dripping off of a cave wall", and "slimy rocks and stalactites". Such combinations of descriptors instigate high-level semantic processes beyond what the individual components of these combinations could achieve alone. In other words, the semantic coherence between the actors can imply environments or even new actors, since, as the neuroscientist Moshe Bar states, "recognition of an object that is highly associated with a certain context facilitates the recognition of other objects that share the same context" (2004: 617). This is true "even if these objects are ambiguous when seen in isolation" (619). Furthermore, when such a coherence is present

between an actor and its context, high-level semantic processes, in which the listener embellishes the diegesis with other appropriate objects, are activated. “[A]n ambiguous object becomes recognizable if another object that shares the same context is placed in an appropriate spatial relation to it.” (619).

[E]ach context (for example, an airport or a zoo) is a prototype that has infinite possible exemplars (specific scenes). In these prototypical contexts, certain elements are present with certain likelihoods, and the spatial relations among these elements adhere to typical configurations. Visual objects are contextually related if they tend to co-occur in our environment, and a scene is contextually coherent if it contains items that tend to appear together in similar configurations. (617)

When an object is not congruent with its context, it is processed more slowly: typical items and relations will have faster processing times when compared to novel ones (618). In other words, contexts gathered from the recognition of objects feed back into how the individual objects in the extracted context will be perceived. Bar describes context frames as “prototypical representations of unique contexts” (618). This concept is reminiscent of models of experience such as *perceptual symbols* and *schemas*, already discussed in Chapter 4. Prototypical contexts imply both the likelihood of certain elements and their spatial configurations (617).

Once listeners establish a semantic context, they have a tendency to hold on to it for the remainder of the piece. For instance, a listener of *Birdfish* who established, early on in the piece, a scene implying a sea with such descriptors as “underwater”, “sand”, “water” and “waves”, described the ending of the piece with “big waves”, “sea is projected in the air”, “and explodes”. Another participant who extracted, from the same piece, a story of robotic bugs working on a project as mentioned earlier, described the ending of the piece with “workers are pleased”, “big cheers” and “project successful”. Here we can see that both participants felt a need to address the climactic ending of the piece. But how this climax was situated in the diegeses of their imaginations shows a semantic coherence with the contexts which had already been established.

Reverberation provides information regarding the static size (i.e. the spatial affordance) of an environment. However spatial movement, such as the panning of a sound object across a stereophonic panorama, not only can establish an equivalent sense of a static space (i.e. spatial context), but can also instigate in the listener’s mind a semantic gestalt as a context. This is apparent, for instance, when the “spatial activity of insects” causes the imagination of a “hive” or a “swarm”.

As discussed in Chapter 3, the contextual segmentation of everyday sounds relies greatly on those distinctions between sound events which denote sources, and ambient noises in which background sounds are blurred together (Guastavino 2007: 57). The ambient noise context affects how a listener reacts to an environmental sound (Raimbault and Dubois 2005: 342). These everyday tendencies are without a doubt reminiscent of figure and ground configurations in music, but their impact becomes even more prominent in the context of electronic music, in which everyday sounds themselves constitute a frame of reference. The blurring effect is a cognitive artifact of selective attention; however, the composer can predetermine this selection by actually obscuring sound events, and therefore building this cognitive artifact into the work. This phenomenon is exercised in the second movement of *Birdfish*: the organic gestures from the first movement which there assumed figure roles are

subjected to such processes as low-pass filtering, reverberation and frequency shifting in the second movement. Although the temporal density of these gestures remains the same, by pushing them backwards in the spatial structure of the piece, I establish a stage for new figure elements. These gestures, which now serve a ground function, are no less evident yet the contrast in spatial configurations between the two layers imposes a sense of blurriness upon the background, since while we pay analytical attention to short-term details in the foreground, we tend to group background elements in gestalt patterns (Truax 1996: 58).

Playing with Anticipation

As Huron states, the contrast between the predicted and the actual outcome of an event amplifies the emotional response to it (2006: 22). On the other hand, we are better at detecting and perceiving events when they conform to our expectations (43). As composers design new worlds, they also establish a framework of expectations. Roads states that a sense of causality between sound events will imply predictability: if listeners fail to associate these events with an underlying syntax, the piece turns into “an inscrutable cipher”. This, however, does not deny the power of strategically used juxtapositions ([forthcoming]).

[Research has] provided evidence that one product of readers’ narrative experiences are *causal networks* that represent the relationships between the causes and consequences of events in a story. Some story events form the main *causal chain* of the story whereas others, with respect to causality, are dead ends. When asked to recall stories, readers find it relatively more difficult to produce details that are not along that main causal chain. (Gerrig and Egidio 2003: 44)

The final section of *Element Yon*, which I have described in Chapter 1 as being composed “to obfuscate any motivic closures to the piece”, generated several descriptors and impressions relating to listener anticipation. One participant in her general impressions wrote: “Unpredictable. I liked that a lot”. Later, in her real-time descriptors, this participant marked the final section of the piece with the word “unpredictable”. At an approximate moment in the piece, another participant submitted “when it’s over, I can’t tell” as a descriptor.

In *Christmas 2013*, the juxtaposition of a Christmas carol with causally unfolding electronic gestures was intended to establish a sense of nostalgia in a vast post-apocalyptic environment devoid of human beings. An inexperienced listener wrote in her general impressions that the opening was familiar, but as the melodic component dissipated, the piece took a turn to what she would later refer to in her real-time descriptors as causing “suspense”:

It started to sound like bits and pieces of sounds and noises that I failed to make sense of. Bu these sounds, when they are together, they gave me this tense, mysterious feeling I don’t know why.

Human or Not

In a study on environmental sound categorization, a free-sorting task conducted with 26 participants revealed that the presence of human activity is one of the most salient features for the hedonic judgement of urban soundscapes (Guastavino 2007: 54): “Soundscapes where human sounds are predominant tend to be perceived as more pleasant than soundscapes consisting of mechanical sounds predominantly” (61). Correspondingly, based on semantic criteria proposed by Murray Schafer and Bernard Delage, Raimbault and Dubois construct a

soundscape classification consisting of “road traffic (car–truck–motorcycle), other transportation (railway, aircraft), working machines (street cleaning, working site), music, people’s presence (speech, walking), and nature (wind, animals)” (2005: 343).

The preference for a human presence in soundscapes can be broadened to a distinction between the presence of animate and inanimate objects. Based on the results of their neuroscientific study on domain-specific knowledge systems, Caramazza and Shelton assert that animate and inanimate conceptual categories are products of evolution and are “subserved by distinct neural mechanisms” (1998:1). It is suggested that dedicated neural substrates¹⁸ have evolved for the categorical representation of animate objects in semantic memory, owing to the survival advantages of being able to identify such objects (Vigliocco 2002: B62). These claims imply that humans display a particular ability (supported by a dedicated system) to distinguish semantically between animate and inanimate objects. However, unlike Guastavino’s cognitive approach, Caramazza and Shelton’s findings deal with performance rather than appraisal.

Accordingly, in a study on affective appraisal, Tajadura-Jiménez and Västfjäll found that “self-representation sounds”, which are associated with bodily functions such as breathing and heartbeat sounds, increased the corporeal awareness of the listener and might have “a stronger potential for inducing an emotional experience” (2008: 66). In another study on the categorization of environmental sounds, harmonic content has been found to indicate for participants either a vocalization or a signaling (Gygi et al. 2007: 852). This also coincides with the assertion made in Chapter 3 that harmonic sounds do not occur in nature but are rather fabricated.

Gliding pitch variations in intonation are expressive of not only meaning (Gussenhoven 2002: 47) but also personality and emotion (Scherer and Oshinsky 1978: 332). Furthermore, this is true not only of humans but also of vocalizing animals in general (Amador and Margoliash 2013: 11136). The gestures consisting of rapid frequency modulations of monophonic lines in *Element Yon* were therefore suggestive of an organic origin, as evidenced in descriptors such as “I guess he is trying to tell us something”, “communication”, “conversation”, “crying”, “scream”.

Contacts between the Physical and the Semantic Domains

The two domains previously discussed are immediately and intrinsically attached to each other. The physical domain constantly informs the semantic domain with new material. In return, the semantic content of gestures contribute to the listener’s selective focusing on the material. Contacts between these two domains can occur in various ways. Below I will offer a number of examples of such contacts supported by experimental data.

Meyer explains that the connotative capacity of a phrase in instrumental music is intrinsically connected to how much it diverges from a “neutral state”:

A tempo may be neither fast nor slow; a sound maybe neither loud nor soft; a pitch may seem neither high nor low, relative either to over-all range or the range of a particular instrument or voice. From the standpoint of connotation these are neutral states. Connotation becomes specified only if

¹⁸ Memory, language acquisition and facial recognition are other examples of neural substrates.

some of the elements of sound diverge from such neutral states. (Meyer 1961: 263)

We can assume that a ground element in music, such as an accompaniment texture, sets a neutral state, in terms of spatial attributes, for a melody to diverge from. Taking the cognitive idiosyncrasies of electronic music into consideration, we can think about semantic dimensions of figure and ground. Meyer's rationalization of connotative capacities, which come into being through contrast, can also be applied to the semantic domain. Our auditory systems allow us to perform the acts of foreground and background listening simultaneously. This way, we can achieve a gestalt perception of our daily soundscapes with certain sonic phenomena highlighted as figures, while others remain out of focus.

Spatialization and loudness determine the perceived physical proximity of a figure. In tandem, these two parameters help establish the semantic concept of *motion*. Sounds from stationary speakers follow choreographies designed by the composer and imply for the listener an animation of objects, albeit detached from any actual moving sound source. These objects can be cognitively abstract or concrete; regardless, the listener hears — and furthermore imagines — beyond the mere changes in parameters and extracts the gestalt (i.e. the motion) emerging from the interplay between them.

As for ground elements, spatialization and amplitude, along with spectral dynamics, can set the reverberant characteristics of a sound, which in return establishes another semantic concept, that of *location*. This highly representational concept transcends the metaphor of musical ground: once a location is semantically associated with a texture, a scene is established for successive figure gestures, which will then be evaluated by the listener in reference to *where* they occur. This conditioning does operate both ways, since the semantic content of a figure gesture will inevitably feed back into how a consecutive ground gesture will be received as I have previously discussed.

Inside and Outside the Diegesis

The interplay between the semantic and the physical attributes of a sound implies links between the diegesis and the concert hall. A sound can travel from an alien territory into the concert hall and weave a contact between the representational and the presentational. A stark example of this phenomenon is evident in Luigi Nono's *La Fabbrica Illuminata*, a 1964 piece for voice and 4-channel tape. The piece exhibits a mixture of live and recorded voices in multi-channel audio accompanied by electronic sounds, as it narrates a story about textile workers who had been trapped in a factory fire. For the fixed sounds, Nono made location recordings at the factory in which this event originally took place. The voices on tape transform from quiet speech into loud vocal lines and mix with the live singing. Quiet sections of the recorded voices create the illusion of a mumbling crowd, which can easily be mistaken for the audience at the concert hall where the piece is being performed. In an interview, Nuria Schönberg-Nono describes that, to achieve such surround affects, the spatial configuration of loudspeakers for Nono's pieces would be adapted to each particular performance space. In the same interview, Schönberg-Nono recounts the composer's particular focus on the spatial aspects of auditory experience:

The Basilica of San Marco in Venice was, from his early creative days, a great influence – the idea that music should come from all different directions and that you were in the centre, instead of having all the sound coming to you just from one single source. There are some wonderful films

that we have in which he explains these things about how in Venice, when you walk around, you hear so many things coming from different places and he believes that the capacity to hear all these things is in us, but that it has been shut out and it needs to be developed. (quoted in Souse 2008)

In *La Fabbrica Illuminata*, while the performance of the singer embodies a more traditional musical act, it also anchors the experience in the physical domain by serving a presentational function. This amplifies the disorientation when the recordings of the mumbling voices suddenly turn into roaring vocal phrases that are clearly in a space different from that which the audience inhabits. The listener travels back and forth between the concert hall and the burning factory, and the journey amounts to an immensely eerie experience through the interplay between the explicit and the implicit worlds.

The physical and semantic aspects of a piece can complement each other or one can overpower the other. The auditory attributes of a sound can alter the imagined universe, or the semantic content of a sound (or lack thereof) can draw attention to its physical characteristics. In *Diegese*, between 0'25" and 0'50", during which several sonic layers populate the scene, a gesture design inherited from *Birdfish* embodies the former case: the spatial configuration of these organic gestures relative to the remainder of the concurrent layers causes them to be highly noticeable. While several participants identified these sounds as "bugs" and "insects", one participant submitted "take out these bugs from my ears" as a real time descriptor. This indicates a distinct physical engagement with the semantic content of the gestures.

Such interactions will determine where listeners will situate themselves in the narrative. Listeners can witness the unfolding of the musical narrative from both inside and outside the diegesis. When a piece weaves a web between the physical and the semantic domains, gestures that are presentational on account of their auditory qualities draw listeners into the diegesis by making them aware of their physical selves. A participant, who visualized "insects flying in a cave" when listening to *Birdfish*, observed the diegesis from outside. Several participants who listened to *Christmas 2013*, on the other hand, described themselves as the subject of a similar action by expressing such impressions as "flying over a city", "makes you feel as in space", "brings you to the air", "[I imagined] open space, empty or a plain, sea (but still open space behind)".

The articulation of space and spatial activity enhances the immersion and creates a contact between the physical and the semantic domains. The listeners describe a bodily experience, but this experience is contextualized in the diegesis rather than the concert hall. Practically speaking, such impressions can be attributed to the fact that the spatial design of *Christmas 2013* consists of a stable reverberant field and intermittent low-frequency rumbles which were intended to create the illusion of a vast space. But the piece also exploits immediate and intermediate spaces which were constantly articulated with impact sounds traveling *around* the listener. Furthermore, the transitory objects displayed spectral behaviors reminiscent of Doppler shifts¹⁹ which may have established a more convincing image of a sound source moving in relation to the spatial position of the listener. Other participants described their corporeal involvement with the piece in phrases such as "inside the brain", "being inside a drum set randomly playing itself" and "music in the air and deep inside the body". All these

¹⁹ For a comprehensive overview of such localization cues, please refer to the pioneering composer John Chowning's paper *The Simulation of Moving Sound Sources*, which was of guidance in the design of various gestures in *Christmas 2013*.

impressions denote an internalization of the material which in return causes the participant to situate themselves inside the diegesis. How much of the material is internally (or corporeally) evaluated can therefore be a determinant of the extent to which the listener is inside the implicit world of the story.

Sense of Time

A narrative is temporal and the time needed to consume it is the time needed to traverse the narrative (Genette 1980: 34). In literature, this time is borrowed from the pace of the reading. In music, the physical time needed for traversing a narrative is set in advance by the composer. However, our understanding of time is a result of the “experience of successions” (Fraisse 1963: 1). As mentioned in the previous chapter, the perceived time ticks in events. The time experienced by the listener can therefore speed up or slow down relative to the unremitting progression of seconds. This temporal relativity constitutes another point of contact between the physical and the semantic domains.

Participants listening to *Element Yón* referred to a lack of or a slowing down in movement with such general impressions as “something still and stable, not dynamic, not moving” and “slow movement; heaviness”. Real-time descriptors such as “heavy”, “static”, and “waiting” also point to this quality. This can be a correlate of the previously discussed increase in memory load due to a lack of structure in the piece.

In *Birdfish* a participant stated in his general impressions that “the sense of time changes over the piece”. The same participant provided the real-time descriptor of “time” after the second movement of the piece commenced. In this part, the ongoing textural density of the piece is pushed into the background. New figure gestures, which exhibit pulsations at slower frequencies, are distributed more sparsely in the foreground. Yet, the event-based pace of the background continues to move forward at the pace of the first movement. This was indeed intended to create a contrast between coexisting timescales.

In *Christmas 2013* another form of temporal shifting is observed. Different participants referred to their experience of time in their general impressions with comments such as “trying to stop time by going ultra slowly”, “objects in slow motion” and “the piece made my brain slow down for a moment”. This can be attributed to the general decaying behavior in the spectral content of the gestures. For instance each transient of the causally unfolding impact gestures exhibit a gradually decreasing frequency. Additionally, more ambient elements also exhibit a similar behavior. This is different from the aforementioned Doppler shifts in the piece which exhibit an increase in frequency followed by a decrease, in conjunction with the spatial displacement of the gestures. The decaying behavior can be reminiscent of “slowing down” by physical law. If a cyclically operating object emits sounds at a sustained pitch, this pitch gradually falls as the object slows to a halt. This behavior can be observed in most mechanical devices used in our daily lives such as washing machines, cars and vacuum cleaners. Another highly pertinent yet somewhat antiquated example is the tape machine. If the perceived slowing down in time reported by the participants is indeed correlated to the decays in spectral content, then this implies an intriguing metaphorical link between the functional pace of objects and the perceived pace of time.

Diegetic Affordances and Affect

Emerging from two distinct fields of study, namely philosophy and psychology, the concepts of affect and affordance, discussed in Chapters 3 and 4 respectively, have significantly convergent characteristics. Recalling previous discussions of these concepts, the following correspondence chart can be formed:

Affordance	Affect
Pre-personal, structured information available in the (material) environment	Pre-personal intensity
Precedes cognitive processes	Unqualified experience
Action possibility	Affective potentiality
Relative to the observer's form	A corporeal phenomenon

As seen above, these two concepts, by their definitions, are contiguous with each other. Both represent capacities, one pertaining to the perceived object and the other to the perceiver. If a link is therefore to be formed between the two, an affordance can be characterized as inductive of affect. While Massumi characterizes emotion as a sociolinguistic fixing of the experiential quality that is *affect*, he later dilutes the one-way succession of affects into emotions by stating that affect also includes social elements and that higher mental functions “are fed back into the realm of intensity and recursive causality” (2002: 30). Affects, anchored in physical reality, are therefore both pre- and post-personal (217). He further states that what he terms affect is “the simultaneous participation of the virtual in the actual and the actual in the virtual, as one arises from and returns to the other” (2002: 35). This transposable take on affect is also apparent in Freud’s interpretation of the concept: unconscious affects persist in immediate adjacency to conscious thoughts and they are practically inseparable from cognition (Seigworth and Gregg 2010: 2).

[The] meaning that listeners attribute to sound, the spatial dimension, or the interactions with other sensory modalities, are as important as the physical properties of sound in evoking an affective response. (Tajadura-Jiménez and Västfjäll 2008: 63)

In Chapter 3 I quoted from the article *Percept, Affect, and Concept*, in which Deleuze and Guattari elegantly described “how the plane of the material ascends irresistibly and invades the plane of composition of the sensations themselves to the point of being part of them or indiscernible from them” (1994: 466). Affect, as I would like to therefore interpret it, represents a landscape of experiences from which emotions sprout. This landscape is superimposed on the material. The affordances of the material evokes affects with the perceiver. An object represented in electronic sound constitutes a material of second order which induces a virtual affective experience. Simultaneously with the ascension of the embodied sound into affect, the representation ignites an affective thread of its own. The imagined spatiotemporal universe of the story will have its own dimensions, landscapes, surfaces and objects.

However, such landscapes and surfaces will only afford *diegetic* action possibilities to the listener. Gibson describes a similar behavior for surrogate objects in the visual domain, such as a photograph or a motion picture (1986: 294). While these objects also specify invariants,

they instigate indirect awareness and provide “information about” (Gibson 1966: 245). The electronic music listener can too make out acoustic invariants characteristic of a certain object. While a representation in electronic music will be “a structured object in its own right” (Nussbaum 2007: 24), the action possibility will nevertheless remain virtual for the listener since the imagined object is an external representation: “[t]he perception or imagination is vicarious, an awareness at second hand” (Gibson 1986: 295).

Affects are semantically processed, fed back into the established context and experienced as the result of *diegetic affordances*. When watching a horror movie for instance, the viewers are aware that they are in a theatre. But once they have been acculturated into the story of the film, a mundane and seemingly non-affective act, such as switching on the lights in a room, becomes loaded with affect. Because threat, as an affect, “has an impending reality in the present” (Massumi 2010: 54). Listeners of electronic music, concoct diegeses from the poietic trace left by the composer. In semantic consistency with these diegeses, listeners populate the landscapes of their imaginations with appropriate objects, situated in various configurations based on cognitive or perceptual cues. As they do so, they also *experience* this environment with implied affordances true to the objects of their imagination, and affects attached to these diegetic possibilities.

The beholder [of a film] gets perception, knowledge, imagination, and pleasure at second hand. He even gets rewarded and punished at second hand. A very intense empathy is aroused in the film viewer, an awareness of being in the place and situation depicted. But this awareness is dual. The beholder is helpless to intervene. He can find out nothing for himself. He feels himself moving around and looking around in a certain fashion, attending now to this and now to that, but at the will of the film maker. He has visual kinesthesia and visual self-awareness, but it is passive, not active. (Gibson 1986: 295)

Accordingly, the listener of electronic music experiences passive aural kinesthesia. Earlier in this section, I have provided several examples from listener impressions which display such forms of involvement. An inexperienced participant, who listened to *Diegese*, narrated a highly visual story of her experience in her general impressions:

Glass/metal ping pong balls are constantly being dropped on the floor as we walk through an empty salon with bare feet; we leave this room and go out in a jungle, moving through the grass stealthily; passing through cascading rooms; we arrive in another salon.

While many of the objects in her narrative also appear in descriptors provided by other participants, details like “walking with bare feet” and “moving stealthily” are indicative of the participant’s individual affective experience of the diegetic environments of her imagination.

Physical Attributes of the Imagined Source

In Chapter 3 I argued that even purely synthesized sounds can instigate a mental association to a sound source and, in Chapter 4, that the poietic intent of the composer and the esthetic construction of the listener do not necessarily have to match. The principle of minimal departure discussed earlier in this chapter implies for the experience of electronic music that when a source for a sound object is imagined, the mind will bridge the gaps as necessary to achieve a base level of consistency by attributing featural qualities to the source. This ability,

as discussed in Chapter 3, is informed by our mental catalogue of auditory events we extract from our daily environments: we possess a sophisticated understanding of how a certain object in action will sound in a certain environment. As Gaver states, the material, the size and the shape of a physical object will intrinsically determine how the object vibrates and therefore produces sounds: for instance, vibrations in wood damp much more quickly than in metal, “which is why wood “thunks” and metal “rings” and “big objects tend to make lower sounds than small ones” (1993a: 7).

Therefore, even the most elementary attributes of a sound can indicate a physical causality. In that respect, granular synthesis bears a significant capacity. In granular synthesis, the metaphorical relationship between a microsound and a particle can be extended to a physical model in which a microsound *behaves* like a particle²⁰. In the experiment results, gestures produced using granular synthesis were described by various participants as particles (pieces, cells, glass, metal) dividing (breaking) and merging (coming back together, colliding). These reports highlight the implication of physical causality inherent to granular synthesis. The frequency and the amplitude envelope of a grain can be altered to specify a particle’s size. *Touche pas* is particularly rich in similarly shaped objects of various sizes, as evidenced in the real-time descriptors referring to spherical objects of diverse proportions. Furthermore, the timbral characteristics of grains can be altered in order to imply different surface materials. In *Diegese*, which quotes a particular granular texture from *Touche pas*, listeners differentiated between timbral varieties by defining different material types and objects. Separate participants described imagining “glass/metal balls”, “ping pong balls”, a “pinball machine”, “champagne” (cork sound), a “woodpecker” and “knocking on the door”. Here the materials vary from metal and plastic to wood. For *Touche pas* both “coins”, “marbles”, “ping pong balls”, “bowling ball” and “xylophone” were provided as descriptors, indicating a similar spectrum of materials.

An important determinant of such descriptors is the motion trajectory of a grain. The particular motion trajectory used in *Diegese* is inspired by the concurrent loops of unequal durations heard in Subotnick’s seminal piece *Touch*, a behavior which is also apparent in *Touche pas*²¹. When multiple loops are blended together, the resulting texture implied for most participants a sense of “bouncing” (i.e. “marbles bouncing”) or “falling” (i.e. “rocks falling together”). One participant wrote: “the clicking sounds (...) resembled a dropped ball bouncing on a surface, since each sound came in slightly quicker than the previous one”. Another participant described *Touche pas* as displaying a “convincing physicality”. Once a motion trajectory is coupled with the imagined material of the object, higher semantic evaluations occur: while one participant described “bouncing on wood” followed by “marimba”, another participant wrote that marbles made her think about “childhood”, “fun” and “games”.

The cognition of motion trajectories can be a function of temporal causality. VanDerveer draws attention to temporal coherence as “possibly the primary basis for the organization of the auditory world into perceptually distinct events” (1979 : 41). To examine the effects of temporal factors in the identification of environmental sounds, Gygi et al. used event-

²⁰ This is discussed extensively in Roads 2001.

²¹ For a further explication of the relationship between these three pieces please refer to Chapter 1, Section 2.

modulated noises²² (EMN) which exhibited extremely limited spectral information (2004: 1252). By vocoding an environmental sound recording with a band-limited noise signal, the event-related information was reduced to temporal fluctuations in dynamics of a spectrally static signal. From experiments conducted with EMN, researchers concluded that, in the absence of frequency information, temporal cues can be sufficient to identify environmental sounds with at least 50% accuracy (1259). The idea of a sonic Rube Goldberg machine, which features in the compositions *Shadowbands* and *Hajime*, relies heavily on temporal articulations of sound events in order to reveal physical causality. Such articulations are apparent in most of my pieces, and particularly in gestures that form connections between consecutive sections (e.g. 0'33" to 0'39" in *Christmas 2013* and 1'27" to 1'30" in *Diegese*).

In *Birdfish* I utilized short-tailed reverberation and low frequency rumbles to establish the sense of a large but enclosed environment. These were reflected in the real-time descriptors with such entries as "cave", "dungeon" and "big spaceship". Similar cues in *Christmas 2013* prompted listeners to submit "open sea", "open space" and "sky" as descriptors. The spectral and reverberant attributes of the sound specify environments in various spatial proportions with the listener. This information implies, for instance, the affordance of locomotion (which in several cases manifested itself as that of "flying").

In *Element Yon*, which inhabits a strictly abstract sound world, the frequency and damping characteristics of certain gestures instigated such descriptors as "metal balls getting bigger and smaller", "high tone falls and hits the ground". Here, distinctly perceptual qualities are situated in metaphors, while retaining their embodied relationship with the listener. Another similar example is observed in the responses to high frequency gestures in *Birdfish*, which listeners characterized with such descriptors as "ice", "glass", "metal", "blade" and "knife". These descriptors imply both a metaphorical association and an affordance structure between high frequencies and a perceived sense of sharpness.

Earlier in this section, I referred to various aspects of recognizing an organic sound source. In relation to these aspects, linguist John Ohala points to the cross-species association of high pitch vocalizations with small creatures, and low pitch vocalizations with large ones. He further delineates that the size of an animal, as implied by the fundamental frequency of its vocalizations, is also an indicator of its threatening intent. Many descriptors submitted by the participants of the experiment denoted living creatures. However, a portion of these source descriptors were augmented by featural descriptors to form such noun phrases as "tiny organisms", "baby bird", "little furry animal", "huge ant" and "huge animal". Here, featural descriptors signify the proportions of the perceived organisms. In these cases, featural information available in the sounds afforded the listeners a spatial hierarchy between the imagined creatures and themselves. Based on Ohala's deductions, the spatial extent of an organism communicated in its vocalization characteristics, which would possess a survival value in a natural environment, is a diegetic affordance of threat. Featural descriptors can therefore be viewed as indicative of affect.

²² A similar vocoding technique can be heard in the introduction section of my composition *Shadowbands*, in which the semantic content of a voice recording vocoded with filtered noise is intelligible for the most part.

Music as a Diegetic Actor

The layer of meaning attribution in electronic music has a peculiar effect on how the listener engages with more traditional forms of musical material in the context of an electronic music piece. Such forms could include a tonal melody, a discernible rhythm, or even a gesture that displays a timbral similarity to a physical instrument. While these forms would be expected to cause an immediate affective appraisal in an instrumental music context, experiment results revealed a meta-evaluation of such forms when encountered in an electronic music piece. Prior to an affective appraisal, the listener identifies the phenomenon as the musical form that it is, situated in the universe of the piece. That is to say, abstract musical elements effectively turn concrete and become diegetic objects in the context of the piece, almost like *a television in a movie scene*.

As described in Section 2 of Chapter 1, one gesture in the representational sound world of *Birdfish* is an abstract leitmotif consisting of three notes in octaves played legato. This gesture is at times supplemented by almost imperceptible resolutions in the lower spectrum that remain in a territory between underwater rumblings and a significantly low-pitched pedal point. The piece is otherwise almost devoid of any material that could be aligned with structures common to instrumental music. Amongst all the narrative components of the piece, which are conspicuously set underwater, this leitmotif creates a moment of stark contrast. One participant of the preliminary studies described the final recurrence of this leitmotif at the very end of the piece as a “musical climax”. Out of all the impressions provided by this participant, such as “water dripping off of cave walls” and “factory noises”, this is the first and only reference to a musical form. It is also interesting that the leitmotif inspired a need to pronounce the *musicality* of a gesture in a piece of music. An audience member from a performance of *Birdfish* characterized this leitmotif as a “musical souvenir”. This expression appropriately illustrates the diegetic quality assumed by a traditionally musical form in the context of an electronic music piece.

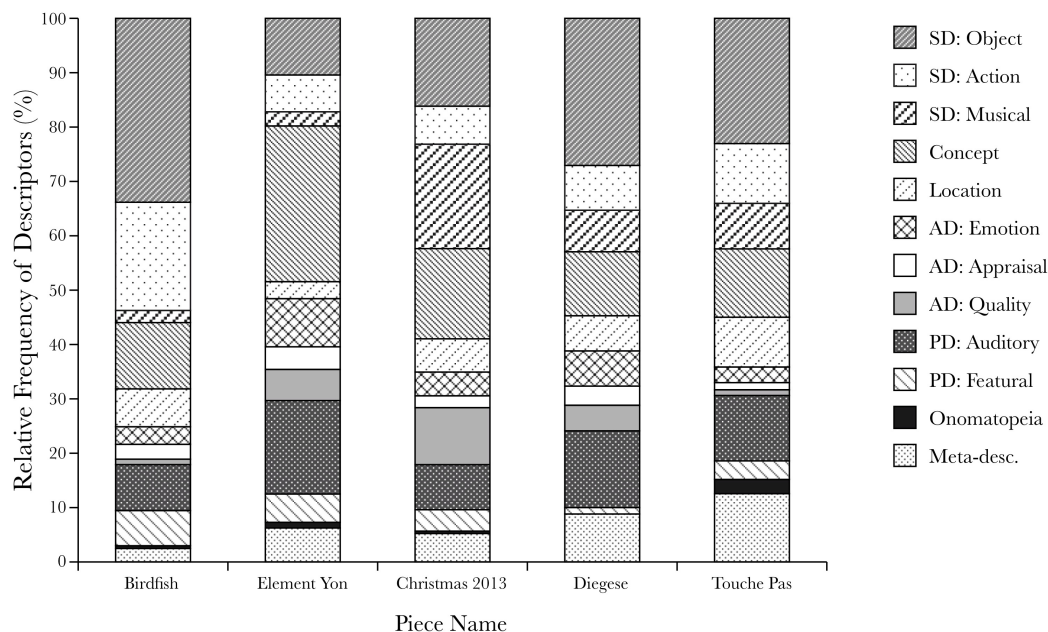
Quoting Music within Music

Quotation has a long history in music. Given its inherent capability to capture and reproduce sounds, the electronic medium grants the composer an “unprecedented ability to include exact quotations from pre-existing sources” (Beaudoin 2007: 149). Notable examples of such quotations are found in Vladimir Ussachevsky’s *Wireless Fantasy*, which cites Richard Wagner’s *Parsifal*, Pierre Henry’s appropriation of Anton Bruckner’s symphonies in *Comme une symphonie envoi a Jules Verne*, and Karlheinz Stockhausen’s *Hymnen*, which is composed of recordings of various national anthems from around the world.

Two of my pieces used in the experiments explored the idea of quoting music within music. As described in Chapter 1, in *Diegese* there are two quotations: the first is a recreation of the granular texture from *Touche pas*, and the second is a recorded snippet of Beethoven’s op. 90. Since the former is an emulation of a texture rather than an exact quotation, its effect appears in the experiment results only in the form of cognitive similarities between the descriptors submitted for the two pieces. These similarities were extensively interpreted above. However, in the latter quotation the listeners can clearly discern the piano segment even if they could not point out a particular piece as a source. Two thirds of the participants made note of the piano sound in their real-time descriptors. One participant included an appraisal descriptor and referred to the quotation with “nice piano” followed by “lovely” at 1’15”. Another participant made note of the quotation in his general impressions as “the frantic

piano sound”. Other participants did not use such appraisal adjectives when referring to the quotation.

In *Christmas 2013*, a similar quotation is from the Christmas carol *Silent Night*, as played by a jazz trio. While the entirety of the piece was composed out of sounds extracted and processed from this recording, between 0’4” and 0’20” the quotation can be clearly made out. Furthermore, brief references to the quotation are sparsely distributed through the remainder of the piece. As a result musical source descriptors constituted the most salient descriptor category for *Christmas 2013* as seen in Graph 5.1.



Graph 5.1: Categorical distribution of real-time descriptors by piece

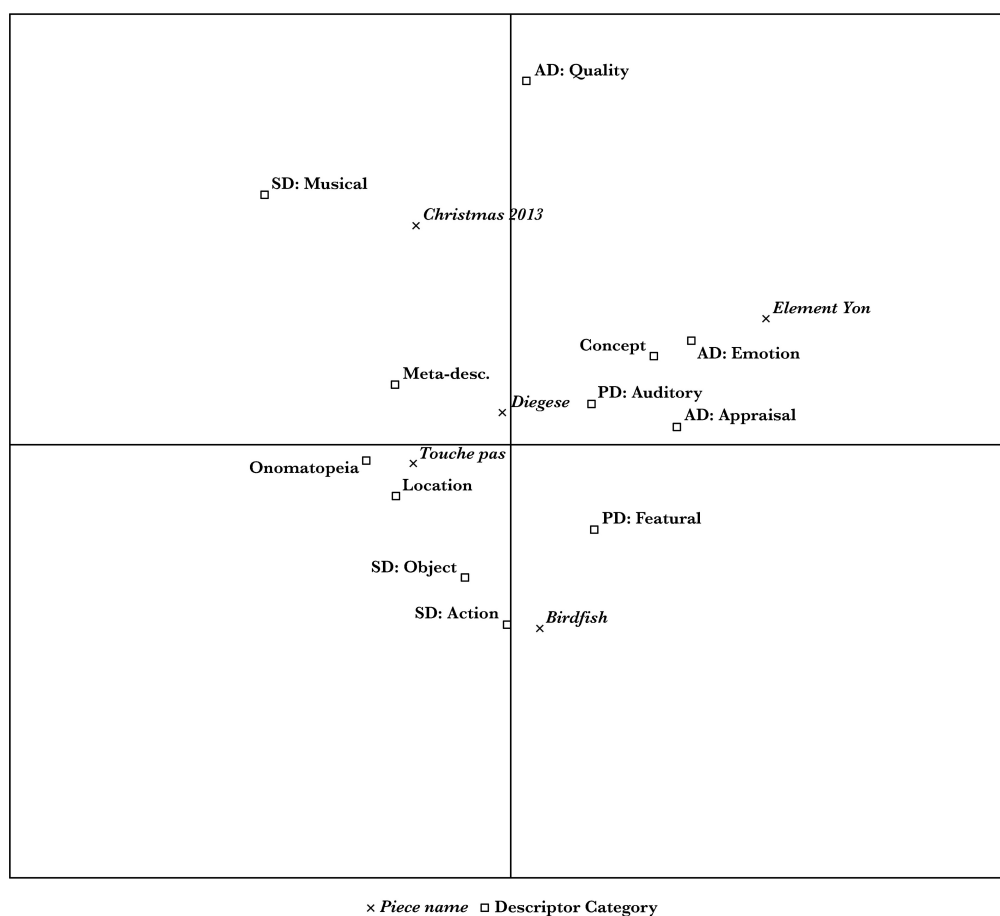
Three quarters of all participants who listened to *Christmas 2013* used the real-time descriptor “piano” at least once. Participants who did not directly refer to the piano, alternatively submitted “music I know”, “bar, dance” and “ballet”. Participants who attended an individual general-impressions session on *Christmas 2013* provided concurring feedback. The quotation was referred to with “real instruments”, “acoustic instrument”, “a phantom harmonium” and “cosy trio”. Unsurprisingly, “Christmas” was another salient descriptor. The final ambient crescendo was described by two participants as “Pink Floyd”²³. Other than the descriptors denoting musical instruments or forms, participants also submitted quality descriptors such as “familiar” and “cliché”. One participant, a Sonology bachelor student, offered an intriguing perspective on the structuring of the piece by defining “planes in a dimension”. He described the impact sounds, which indicate the temporal and spatial unfolding of the piece, as establishing a plane. According to this participant, “rooms” represent another plane. More interestingly, he refers to the quotation also as a “cliché” and explains that this too becomes a plane on its own. I find this *objectification* of the cliché to be an apt description of my poetic intent, which was to assign a diegetic function to the quotation.

²³ In hindsight, I can clearly hear the brief moment that is reminiscent of the band’s *Shine On You Crazy Diamond* from the album *Wish You Were Here*, which was a defining piece for my taste in music as a child. Although this “quotation” was entirely unintended, I find the comparison gratifying.

As I explained extensively in Chapter 1, my aim with contrasting the two sound worlds of *Christmas 2013* was to establish a sense of *future nostalgia*. Congruently, one participant described his experience as a “memory of an event” and related his sense of “something recalled from the back of the mind” to the tonal thread of the piece. One of the participants denoted that besides “flight” and “movement”, “nostalgia and space” dominated his experience and the open ending with the piano enhanced his feeling of “nostalgia/longing”. Another participant wrote:

I had the impression of being in the air (like an angel) and moving over a city on Christmas evening. The sounds escaped from human festive activities, some became distorted and merged over with other ones, others still referred to remembrances (nostalgia) of childhood (acoustic instruments).

Another notable outcome of the quotation was observed in the affective quality descriptors. Some of these descriptors were “smooth”, “mellow”, “familiar”, “childish”, “relaxing”, “incongruent”, “creepy”, “dark” and “weird”. Moreover, *Christmas 2013* yielded the highest number of affective quality descriptors (as well as musical source descriptors) across all the pieces used in the experiment. This result can be clearly observed in the correspondence analysis as seen in Graph 5.2.



Graph 5.2: Correspondence analysis between pieces and descriptor categories

These descriptors reveal a significant duality which was expressed in various ways in most of the general impressions (e.g. earth/nostalgia versus space, imaginary versus real, childhood versus distance/melancholia). This might be interpreted as the outcome of a possible priming caused by the musical quotation early on in the piece. The precedent this section sets with a relaxing, mellow and familiar feeling, amplifies the disorienting sensation of the ensuing diegesis.

Regarding the final piano part in the piece, a participant with no musical background submitted the descriptor “sounds like music”. In his general impressions, another participant recounted that although most of the sounds caused him to feel like being in “a place not on this earth”, the piano sound made him “come back to earth and reminded [him] that it was music [he] was listening to”. Along a similar line, one participant wrote “in an imaginary world, suddenly something real begins to move”. Another participant referred to the quotation as “something to hold onto in the insecure environment”.

The difference in the relative frequency of musical source descriptors between *Christmas 2013* and *Diegese* can be a result of the extents and the forms of the quotations. In *Christmas 2013*, the listeners can not only recognize a multiplicity of instruments, but they can furthermore identify a musical form in the quotation. That’s why such descriptors as “song”, “soundtrack” and “music” were exclusively apparent in the feedback for *Christmas 2013*.

A Diegetic Actor as Music: Electronic Music and Science Fiction

[The] sense of disorientation produced in some listeners by the impact of electronic sounds was the basis of the early use of electronic sound-materials for science fiction productions. The inability of the listener to locate the landscape of sounds provided the disorientation and sense of strangeness which the producer wished to instil in the listener. (Truax 1996: 139)

The choice made by film and game studios to incorporate electronic sounds to evoke various emotions in the audience has a discernible impact on the experience of electronic music, particularly with inexperienced listeners. With *Birdfish*, science fiction was a prominent point of reference for the participants as evidenced in such descriptors as “star wars”, “R2D2”, “Starcraft”, “spaceship” and “robot”. Various general impressions for other pieces also delineated similar concepts: “I get images of science fiction: spaceships etc.”, “[synthetic sounds] reminded me of a world that you may find in a movie like Tron” (*Element Yon*); “sounded like a soundtrack for a horror or science fiction movie” (*Diegese*); “astronomical documentary, museum of science” (*Christmas 2013*). Some participants associated their experience more loosely with movie soundtracks in general: “It reminded of sound effects used for tense moments in thrillers”; “I feel like they would fit to a dramatic tense moment of a film” (*Christmas 2013*). In these cases, memories of diegetic film sounds acted as a reference for the musical experience of the listener. Two participants provided more specific descriptions of this relation: “[I imagined] sound design people working on a sci-fi film, enjoying their work”; “the brief tonal sounds reminded me of R2D2” (*Birdfish*). The latter is a fairly reasonable account of how a sound as generic as a brief synthesized tone can be connotative of a robot. However, due to the listener’s tendency to establish a semantic coherence across various sound elements, unintended references to diegetic sounds of science fiction might influence the entire story, as evidenced in some of the examples provided earlier in this chapter.

A Semantic Paradigm for Electronic Music

Just as we can evaluate the extent to which the listener is inside or outside the musical material, we can also question how much of the material is internal or external to the listener. How much of the content of an electronic music piece is objectively out there in the concert hall? From a poietic standpoint (i.e. looking from within the artwork) the listeners can be outside of the diegesis. But the listeners, when exposed to the artwork, construct semantic universes around themselves in the physical domain of the concert hall and they *observe*.

When Meyer describes collective image processes, he refers to a common ground shared by every listener, which would inevitably be appended by the private, or individual image processes, and ultimately amount to an affective appraisal. The emotional assessment of sounds will naturally be attached to our individual experiences. However, the layer of meaning attribution, in which the diegesis emerges, generates overlapping universes amongst different individuals – listeners and composers alike – owing to a shared mental catalogue of auditory experiences. The semantic and physical aspects of electronic music discussed throughout this chapter cogently indicate that a composer's orchestration of cognitive cues can play a significant role in shaping the experience of a piece. We can hear both abstract and representational sounds in electronic music. The possibilities in between yield an experiential depth for the listener as evidenced by the diversity of categorical distributions for each piece.

The diegetic approach discussed in this chapter suggests new perspectives towards understanding and communicating the experience of electronic music. I have outlined a coalescence of representational modes informed by the discipline-specific interpretations of diegesis. This allowed me to situate electronic music in a broader context of artistic forms. Using this framework as a semantic paradigm for electronic music, I have delineated various relationships between the analyses of the listening experiments and the domains of experience in electronic music. For instance, meta-descriptors were used to explain the perceived presence of the composer in a piece. The source and perceptual descriptors were highlighted to articulate the boundaries of a diegesis. Other examples revealed how affective appraisal and quality judgements were conditioned by the cognitive disposition of a piece. Furthermore, the general impressions were used to reveal the contextual relationships between concurrent gestures. These examples demonstrated how the cognitive continuum mediates the interplay between the semantic and the physical domains of experience in electronic music.