Diegetic Affordances and Affect in Electronic Music

Anıl Çamcı
University of Illinois at Chicago
anilcamci@gmail.com

Vincent Meelberg
Radboud University Nijmegen
v.meelberg@let.ru.nl

ABSTRACT

In this paper, we investigate the role affect plays in electronic music listening. By referring to a listening experiment conducted over the course of three years, we explore the relation between affect and diegetic affordances (i.e. those of the spatiotemporal universes created by electronic music). We will compare existing perspectives on affect with the psychologist James Gibson's model of affordances in the context of an electronic music practice. We will conclude that both the sounds themselves and the diegetic affordances of these sounds may elicit affective reactions, and that further study into the relation between diegesis, affordance, and affect may contribute to a better understanding of what we hear in electronic music.

1. INTRODUCTION

In contemporary music studies, *affect* seems to play an increasingly important role. This concept enables the articulation of the way music has an impact on listeners and artists alike. In this paper, we explore how affect "works" in electronic music, and how it is intrinsically related to affordances of an electronic music experience. More specifically, we will discuss the manners in which so-called diegetic affordances may evoke affects with listeners.

Listeners of electronic music may derive diegeses (i.e. spatiotemporal universes referred to by narratives) 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.

First, we will outline a listening experiment, the results of which will be used to further our discussion. We will then define affects and affordances, and how these can be applied in the articulation of musical experience. We will explore the similarities between these two concept to construct a framework that can be used to articulate artistic experiences. In doing so, we will suggest that the diegetic affordances of electronic music may evoke affective responses with listeners. Finally, based on the listener feedback gained from the experiment, we will demonstrate how this framework can be useful when discussing features of electronic music that are corporeally relevant for the listener.

2. OVERVIEW OF THE EXPERIMENT

Between May 2012 and July 2014, 60 participants from 13 different nationalities took part in a listening experiment that investigates the cognition of electronic music. 23 participants were female while 37 were male. The average age of the participants was 28.78. Ages ranged from 21 to 61. 22 participants identified themselves as having no musical background. Amongst the remaining 38 participants were musicians, music hobbyists, composers, and students of sound engineering and sonic arts.

The experiment aimed to explore how fixed works of electronic music operate on perceptual, cognitive and affective levels. The design of the experiment was aimed at extracting both contextual and in-the-moment impressions while offering a natural listening experience. The design involved: 1) an initial listening section, where the participants were asked to listen to a complete work of electronic music without any instructions pertaining to the experiment, 2) a general-impressions task, where the participants were allowed to reflect upon their experience in writing without any form or time constraints, 3) a real-time input exercise, where the participants were acquainted with a browser-based system in which they could submit descriptors in real time while hearing an audio material, and 4) a real-time free association task, where the participants listened to the same piece they heard earlier while at the same time submitting descriptors as to anything they might feel, imagine or think as they listen to the piece.

Five complete pieces of electronic music, in 44.1 kHz, 16-bit WAV format, were used in the experiments. Four of these pieces, namely *Birdfish*, *Element Yon*, *Christmas 2013*, and *Digese*, were composed by the first author of this paper. The fifth piece was Curtis Roads' 2009 piece *Touche pas*. Said pieces utilize a wide range of forms, techniques (e.g. live performance, micromontaging, algorithmic generation), tools (e.g. audio programming environments, DAWs, physical instruments) and material (i.e. synthesized and recorded sounds).

The results of this experiment, including a categorical analysis of the real-time descriptors and a discourse analysis of the general impressions, have been offered in previous literature [1, 2]. The current paper relies primarily on a semantic analysis of the general impressions and the real time descriptors. The general-impressions were provided in one or a combination of various forms, including list of words, list of sentences, prose and drawings. The vast majority of the descriptors submitted in the real-time free association task were single words or two-word noun phrases. A participant's prior experience with electronic music did not significantly impact the semantic qualities (e.g. representationality versus abstractness) or the number of the descriptors submitted by that participant. Technological listening, where a listener recognizes the technique behind a work [3], was infrequently apparent in the responses by sonic arts students.

3. AFFECT IN MUSIC

3.1 Interpretations of Affect

The affective appraisal of music comprises successive stages that utilize different but interconnected perceptual resources. A particular component of this spectrum is the experience of affect, which has been studied within a variety of domains ranging from virtual reality [4] and painting [5] to politics [6] and sports [7]. This concept is not only adopted by a large array of disciplines but also subjected to a variety of interpretations. On the far end of the spectrum, Lim et al. [8] and Shouse [9] point to uses of affect as a synonym for emotion. While this approach begs the question of why affect would need to be demarcated as a separate concept, it nevertheless provides an insight regarding the context within which the concept is situated.

The use of affect in philosophy dates back to Spinoza's *Ethics*. Spinoza identifies affect as an affection of the body by which "the body's power of acting is increased or diminished" [10]. In his introduction to Deleuze and Guattari's *A Thousand Plateaus*, the philosopher Brian Massumi offers a related description of affect as a "prepersonal intensity corresponding to the passage from one experiential state of the body to another" [11]. Emotion on the other hand is personal according to Massumi: "Emotion is qualified intensity, the conventional, consensual point of insertion of intensity into semantically and semiotically formed progressions, into narrativizable action-reaction circuits, into function and meaning" [12].

Based on Massumi's interpretation, we have previously proposed the concept of a sonic stroke [13]. A sonic stroke is an acoustic phenomenon that induces musical affect upon impacting the listener's body. A consequence of this impact is emotion, which emerges once the affect is reflected upon (i.e. a sonic stroke is registered as a musical gesture).

3.2 Affect and Mechanisms of Music Perception

Music, despite lacking immediate survival value, activates brain mechanisms associated with pleasure and reward. The combined sensory and cognitive experience of a musical piece influences the listener's affective state [14]. Accordingly, existing research points to a mixture of cultural and physiological determinants of music appreciation [15, 16]. Brown et al. delineate musical universals, such as loudness, acceleration and high-registered sound patterns, which incite affective experience independent of cultural origin [17].

Juslin and Västfjäll emphasize a need to investigate the mechanisms underlying the affective appraisal of music [18]. They argue that the evocation of emotions in music is based on processes that are not exclusive to music. They enumerate several neural mechanisms that contribute to this phenomenon. Out of these, the *brain stem reflex* deals with the low-level structural and cross-cultural characteristics of the musical experience. Brain stem reflexes are hard-wired and are connected with the early stages of auditory processing. Sounds that are sudden, loud, dissonant, or those that feature fast temporal patterns signal the brain stem about potentially important events and induce arousal. This arousal reflects the impact of auditory sensations in the form of "music as sound in the most basic sense" [18].

Due to its attachment to the early stages of auditory processing, brain stem reflex is highly correlated with human physiology and the so-called universals (i.e. the low-level structural properties) of musical experience. A functional coherence between affect and the brain stem reflex is highlighted by their intrinsic reliance on the spectrotemporal and dynamic properties of musical sound. While affect represents the corporeal segment of the affective appraisal of music, it cannot be dissociated from an ensuing emotion. This is mainly due to the aforementioned interplay between the mechanisms underlying music cognition. The musicologist Marc Leman points to seminal neuroscientific studies, such as those by Antonio Damasio, Marc Jeannerod and Wolf Singer, that motivate a departure from the Cartesian view of "mind and matter" as separate entities; it is understood that the so-called subjective world of mental representations stems from our embodied interactions with the physical environment [19].

4. AFFORDANCES

An approach to perception that is commonly facilitated in musical research [20, 21, 22] is the model of affordances developed by the psychologist James Gibson. Gibson's studies on ecological perception stemmed from his experiments in aviation during the World War II. Focusing mainly on an active observer's perception of its environment, Gibson postulated that the invariant features of visual space represent pivotal information for perception. *I*nvariants are features of an object that persist as the point of observation changes [23]. While most items in Gibson's taxonomy of invariants pertain to the visual domain, his concept of affordances has been applied to other modalities of perception including hearing.

According to Gibson, objects in an environment, by virtue of their invariant features, afford action possibilities relative to the perceiving organism. For instance, a terrestrial surface, given that it is flat, rigid and sufficiently extended, affords for a human-being the possibility to walk on it [23]. His main motivation to propose this seemingly straightforward idea is to refute the prevailing models of perception, which assume that ecological stimuli are chaotic, and therefore the perceiver must extract a meaning out of sensory stimuli by imposing mental structures upon disorganized information. Gibson suggests that there are certain kinds of structured information available prior to perception in the form of invariants. The nature of these invariants is relative to the complexity of the perceiving animal [24]. In other words, an object will have different affordances for different perceivers: a stone, on account of its physical characteristics, affords the action possibility of throwing for a human-being, while at the same time affording the action possibility of climbing for an ant.

Gibson suggests that "perceptual seeing is an awareness of persisting structure" [23] and that knowledge exists in the environment, for a viewer to pick up. When viewed in the light of modern experimental studies on perception, we can consider Gibson's proposal of perceptual knowing as an addition to, rather than a replacement for, the existing models of learning that are based on memory processes. This sentiment is clearly materialized in Gibson's writing as well when he states: "To perceive the environment and to conceive it are different in degree but not in kind. One is continuous with the other" [23]. The ecological approach addresses certain stages of our perceptual experience and complements higher-level mental processes. In that respect, Gibson's model of invariants aligns with various models of experience, such as perceptual symbols and schemas [25].

5. DIEGETIC AFFORDANCES AND AFFECT

As the above discussion indicates, the concept of affect and the model of affordances have significantly convergent characteristics, even though they emerge from two separate fields of study, namely philosophy and psychology. Recalling our previous discussions of these concepts, a correspondence chart between these two concepts, as seen in Table 1, 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

Table 1. A comparison of the definitions of affordance and affect

This table shows that how 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 underplays 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" [12]. Affects, anchored in physical reality, are therefore both pre- and post-personal. This dual 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 [26].

In their article *Percept, Affect, and Concept*, Deleuze and Guattari elegantly describe "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" [27]. Affect, as we 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 an 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 evoked by this representation will have its own dimensions, landscapes, surfaces and objects.

However, such landscapes and surfaces will only afford so-called diegetic action possibilities to the listener. The narratologist Gérard Genette defines diegesis as "the spatiotemporal universe" referred to by a narrative [28]. The concept of diegesis can be traced back to Plato's dichotomization of narrative modes into imitation and narration [29]. 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 [30]. 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 [31]). And it is precisely such an imagined spatiotemporal universe, in this case created as a result of listening to electronic music, that we describe as consisting of diegetic affordances that are evocative of affects [30].

Gibson describes a behavior for surrogate objects in the visual domain, such as a photograph or a motion picture, that is similar to the diegetic action possibilities introduced above [23]. While these objects also specify invariants, they instigate indirect awareness and provide "information about" [24]. The electronic music listener can also make out acoustic invariants characteristic of a certain object. While a representation in electronic music will be "a structured object in its own right" [22], 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" [23].

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" [6]. Listeners of electronic music concoct dieges from the poietic trace left by the composer. In semantic consistency with these dieges, 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. As Gibson explains:

> 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 kinesthesis and visual self-awareness, but it is passive, not active. [23]

Accordingly, the listener of electronic music experiences passive aural kinesthesis. An inexperienced participant, who listened to *Digese*, 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.

6. THE AFFORDANCES OF IMAGINED SOUND SOURCES IN ELECTRONIC MUSIC

The concept of diegetic affordances can be useful when discussing features of electronic music that are corporeally relevant for the listener. Reverberation, for instance, affords a relative sense of space while low frequency gestures afford an awareness of large entities. Even purely synthesized sounds can afford the instigation of a mental association to a sound source. 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. In literature, this is referred to as the "principle of minimal departure", which describes the readers' tendency to relate a story to their everyday lives in order to resolve inconsistencies or fill the holes in the story. In electronic music, this tendency is informed by our mental catalogue of auditory events we have been exposed to thus far: we possess a sophisticated understanding of how a certain object in action will sound in a certain environment. As the design researcher William 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" [32].

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 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 a mechanical 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 *Digese*, 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", "champaign" (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 submitted 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 *Digese* 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*. 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-level semantic associations 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. The researcher Nancy VanDerveer draws attention to temporal coherence as "possibly the primary basis for the organization of the auditory world into perceptually distinct events" [33]. To examine the effects of temporal factors in the identification of environmental sounds, Gygi et al. used event-modulated noises (EMN) which exhibited extremely limited spectral information [34]. 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. Articulation of a so-called physical causality through the temporal configuration of sound elements is apparent in most of the pieces we used in our experiment, and particularly in gestures that bridge consecutive sections of a piece (e.g. 0‘33” to 0‘39” in *Christmas 2013* and 1‘27” to 1‘30” in *Digese*).

In *Birdfish*, short-tailed reverberation and low frequency rumbles were utilized 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 gestures with high frequency content 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.

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.

The linguist John Ohala points to the cross-species association of high pitch vocalizations with small creatures, and low pitch vocalizations with large ones [35]. 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. 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 an affordance of threat. Featural descriptors can therefore be viewed as indicative of affect.

Gliding pitch variations in intonation are expressive of not only meaning [36] but also personality and emotion [37]. Furthermore, this is true not only of humans but also of vocalizing animals in general [38]. 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”.

7. CONCLUSIONS

Music listening is a complex activity in which affect plays a crucial role. As our discussion of listening to electronic music has revealed, both the sounds themselves and the diegetic affordances of these sounds may elicit affective reactions. It is the latter kind of affective reaction, in particular, that has a decisive influence on the manner in which electronic music may be interpreted by listeners. In the experiment results, we observed that diegetic affordances guide the listeners to higher-level semantic associations, which inherently inform their affective interpretation of a piece. As a consequence, we believe that further study into the relation between diegesis, affordance, and affect may contribute to a better understanding of what we hear in electronic music.

8. REFERENCES

[1] A. Çamcı, “A cognitive approach to electronic music: theoretical and experiment-based perspectives,” in *Proceedings of the International Computer Music Conference*, 2012, pp. 1–4.

[2] ——, “The cognitive continuum of electronic music,” Ph.D. dissertation, Academy of Creative and Performing Arts (ACPA), Faculty of Humanities, Leiden University, Leiden, 2014.

[3] D. Smalley, “Spectromorphology: explaining sound-shapes,” *Organised sound*, vol. 2, no. 02, pp. 107–126, 1997.

[4] L. Bertelsen and A. Murphie, “Félix Guattari on Affect and the Refrain,” in *The affect theory reader*, M. Gregg and G. J. Sigworth, Eds. Durham: Duke University Press, 2010, p. 138.

[5] G. Deleuze and F. Bacon, *Francis Bacon: the logic of sensation*. Minneapolis: University of Minnesota Press, 2003.

[6] B. Massumi, “The political ontology of threat,” in *The affect theory reader*, M. Gregg and G. J. Sigworth, Eds. Durham: Duke University Press, 2010, pp. 52–70.

[7] P. Ekkekakis, *The measurement of affect, mood, and emotion: a guide for health-behavioral research*. Cambridge: Cambridge University Press, 2013.

[8] Y.-k. Lim, J. Donaldson, H. Jung, B. Kunz, D. Royer, S. Ramalingam, S. Thirumaran, and E. Stolterman, “Emotional experience and interaction design,” in *Affect and Emotion in Human-Computer Interaction*, C. Peter and R. Beale, Eds. Berlin: Springer, 2008, pp. 116–129.

[9] E. Shouse, “Feeling, emotion, affect,” *M/C Journal*, vol. 8, no. 6, p. 26, 2005.

[10] B. de Spinoza, *A Spinoza reader*, E. Curley, Ed. Princeton: Princeton University Press, 1994.

[11] G. Deleuze and F. Guattari, *A thousand plateaux*. Minneapolis: University of Minnesota Press, 1987.

[12] B. Massumi, *Parables for the virtual: movement, affect, sensation*. Durham: Duke University Press, 2002.

[13] V. Meelberg, “Sonic strokes and musical gestures: the difference between musical affect and musical emotion,” in *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009)*, 2009, pp. 324–327.

[14] V. N. Salimpoor, I. van den Bosch, N. Kovacevic, A. R. McIntosh, A. Dagher, and R. J. Zatorre, “Interactions between the nucleus accumbens and auditory cortices predict music reward value,” *Science*, vol. 340, no. 6129, pp. 216–219, 2013.

[15] S. E. Trehub, “Human processing predispositions and musical universals,” in *The origins of music*, N. L. Wallin, B. Merker, and S. Brown, Eds., 2000, pp. 427–448.

[16] M. E. Curtis and J. J. Bharucha, “The minor third communicates sadness in speech, mirroring its use in music.” *Emotion*, vol. 10, no. 3, p. 335, 2010.

[17] S. Brown, B. Merker, and N. L. Wallin, *An introduction to evolutionary musicology*. Cambridge: MIT Press, 2000.

[18] P. N. Juslin and D. Västfjäll, “Emotional responses to music: the need to consider underlying mechanisms,” *Behavioral and brain sciences*, vol. 31, no. 05, pp. 559–575, 2008.

[19] M. Leman, *Embodied music cognition and mediation technology*. Cambridge: MIT Press, 2008.

[20] S. Östersjö, “Shut up ’n’play,” Ph.D. dissertation, Malmö Academy of Music, Malmö, 2008.

[21] W. L. Windsor, “A perceptual approach to the description and analysis of acousmatic music,” Ph.D. dissertation, City University, London, 1995.

[22] C. O. Nussbaum, *The musical representation: Meaning, ontology, and emotion*. Cambridge: MIT Press, 2007.

[23] J. Gibson, *The ecological approach to visual perception*, ser. Resources for ecological psychology. Mahwah: Lawrence Erlbaum Associates, 1986.

[24] J. J. Gibson, *The senses considered as perceptual systems*. Boston: Houghton Mifflin, 1966.

[25] D. A. Schwartz, M. Weaver, and S. Kaplan, “A little mechanism can go a long way,” *Behavioral and Brain Sciences*, vol. 22, no. 04, pp. 631–632, 1999.

[26] M. Gregg and G. J. Seigworth, *The affect theory reader*. Durham: Duke University Press, 2010.

[27] G. Deleuze and F. Guattari, *What is philosophy?* New York: Columbia university Press, 1994.

[28] G. Genette, *Figures I*. Paris: Seuil, 1969.

[29] Plato, *Plato, The Republic*, S. H. D. P. Lee, Ed. London: Penguin Books, 1955.

[30] A. Çamcı, “Diegesis as a semantic paradigm for electronic music,” *eContact*, vol. 15, no. 2, 2013, [Online], http://econtact.ca/15_2/camci_diegesis.html. Accessed May 12, 2016.

[31] R. Bunia, “Diegesis and representation: beyond the fictional world, on the margins of story and narrative,” *Poetics Today*, vol. 31, no. 4, pp. 679–720, 2010.

[32] W. W. Gaver, “What in the world do we hear?: an ecological approach to auditory event perception,” *Ecological psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[33] N. J. VanDerveer, “Ecological acoustics: human perception of environmental sounds.” Ph.D. dissertation, ProQuest Information & Learning, 1980.

[34] B. Gygi, G. R. Kidd, and C. S. Watson, “Spectral-temporal factors in the identification of environmental sounds,” *The Journal of the Acoustical Society of America*, vol. 115, no. 3, pp. 1252–1265, 2004.

[35] J. J. Ohala, “Cross-language use of pitch: an ethological view,” *Phonetica*, vol. 40, no. 1, pp. 1–18, 1983.

[36] C. Gussenhoven, “Intonation and interpretation: phonetics and phonology,” in *Proceedings of Speech Prosody 2002, International Conference*, 2002, pp. 47–57.

[37] P. N. Juslin, “Cue utilization in communication of emotion in music performance: relating performance to perception.” *Journal of Experimental Psychology: Human perception and performance*, vol. 26, no. 6, p. 1797, 2000.

[38] A. Amador and D. Margoliash, “A mechanism for frequency modulation in songbirds shared with humans,” *The Journal of Neuroscience*, vol. 33, no. 27, pp. 11 136–11 144, 2013.