Listening and Meaning:
How A Model of Mental Layers Informs Electroacoustic Analysis

DRAFT

Gary S. Kendall
Stockholm, Sweden
garyskendall@me.com

1.0 Introduction

1.1 Concerning the Analysis of Electroacoustic Music

A key question to be answered by the analysis of electroacoustic music is: How does the listener's experience of listening become *meaningful*? Amidst the novel timbres and spatial soundscapes, the raw electronics and clashing juxtapositions, the electroacoustic listener is instantly making sense of auditory sensations and experiencing something meaningful. For the more experienced listener, this may be a well accustomed encounter that immediately gives rise to a sense of clear and nuanced meanings. This listener likely has a history of similar encounters, possibly spanning multiple genres of works and composers, and perhaps this listener is particularly familiar with some electroacoustic works that have become exemplars of the idiom. For the novice listener, this could be a disorienting encounter with an aural landscape that juxtaposes the recognizable with the unknown, whose unfamiliar topography is best grasped in terms of eccentric objects and fragmented narratives. What is so easily lost in hese dissimilar experiences is how similarly both the expert and the novice are making meaning.

That meaning varies widely with individual listeners is clear. Meaning is a product of the individual listener's mental processes: whatever meaning the listener makes of the listening experience, that is its meaning to the listener. But as listeners become more familiar with a particular work of electroacoustic music, we should expect that these meanings become less divergent and more finely attuned to the specific implications of the work, essentially because these listeners' mental processes are increasingly guided along convergent lines, even allowing for their differences in personal history. Therefore, the analyst should be mindful not to ascribe meaning to the physical acoustic attributes of a work or acoustic representations in which in the listening experience is absent. For analysis to be ultimately tenable, it must be envisioned fundamentally in terms of human perception and cognition. This places a rigorous responsibility on our analytical techniques to justify their relevance to the listener's mental activity and to avoid unexamined, implicit assumptions about what is meaningful to an actual listener.

1.2 Cognitive Processes

Part of the analyst's dilemma is the lack of a framework for discussing the subjective experience of art, for putting its component parts up on the board for study. Clearly listeners have experiences that are meaningful and varied. But how does this come about and what are the mechanisms by which this happens? We understand that the listener's mental processes are at root neurological processes and that neuroscience provides an important framework for talking about listening, but we can also approach listening in terms of *cognitive processes*.

> 'A major objective driving what might be termed the cognitive deconstruction of artistic experience is to analyze the functional architecture of its underlying component operations.' (Donald 2006:10)

This statement is representative of an effort within cognitive sciences to address humanity's distinctive aesthetic capacity and to illuminate the mental basis of this capacity. We posit cognitive processes in order to facilitate a discussion of such mental activity in terms of functional components like working memory,

attention, schemas, etc. It is a goal of this essay to elucidate ways in which listening processes can give rise to meaning.

2.0 Mental Processes and the Five Mental Layers

Listening is a mental activity that is constantly unfolding. It unfolds moment to moment in time, and importantly it also unfolds in multiple, simultaneous layers. In order to address these issues, a *model of five mental layers* will be introduced, a model that is sufficiently complex to clarify how the notion of layers contributes to our understanding of the listening experience.

2.1 The Five-Layer Model

This model of the five layers starts at the bottom with the simplest sensory processes and proceeds up layer by layer to the most synoptic and abstract level. Numerous authors within cognitive science have postulated similar ideas for a variety of purposes (Donald 2001). This model owes a large debt in particular to Per Aage Brandt (2004:180-5; 2006: 173-6) who says:

> '. . . our organization of what we will call *meaning* is a process that occurs on many levels simultaneously. On a scale extending from the most dense, massive, and "concrete" sensations to the most transparent and "abstract" notional conceptualizations, we may distinguish a series of strata interrelated by processes of integration.' (Brandt 2006: 173)

Our purpose here is to account for listening processes, and for that reason our discussion of this model may omit many conciderations that might otherwise be important to cognitive scientists. At the same time, a focus on listening enables us to greatly increase our level of specificity within the bounds of Brandt's model. The five layers can be summarized as follows:

Layer 1. **Sensations** - *Perceptual organization and constancy of immediate sensation.*
Layer 2. **Gist** - *Framework of things and space extended over several seconds enabling sustained awareness in the short-term.*
Layer 3. **Locus** – *Self-governing actions in response to situations in the 'perceptual present' and slightly beyond.*
Layer 4. **Contexts** - *Framework for enlisting and assessing medium- and long-term, event-oriented schemas and expectations over an extended time frame.*
Layer 5. **Domains** - *Frameworks of background knowledge providing long-term constancy.*

While Brandt was in part influenced by considerations of neural physiology, this model is not about physiology *per se*. Rather, the model is being adapted from cognitive semantics, here acknowledging the multilayered, parallel nature of meaning making. In this sense, cognitive semantics is serves as a highly relevant and informative inspiration for understanding meaning in electroacoustic music. Brandt (2006) himself employed this model in a discussion of the aesthetic experience of specific works of visual art and makes the explicit connection to aesthetic experience. But a major difference between the contexts of visual art and music is the temporal flow and the 'in the moment' nature of music listening that finds no parallel in Brandt's discussion of visual art. The short descriptions of the layers above and a discussions ahead specifically attempt to address the issue of time. As witnessed in another departure, Brandt (2004:182) originally labeled the model's layers correspondingly as: Qualia, Things, Situations, Notions and Feelings. The relabeling of the layers for the specific context addressed here is generally in line with Brandt's original thesis with the exception of the last layer which Brandt later re-labeled as 'Emotions' (2006:174). This exception was primarily motivated by a different approach to feelings and emotions (to be addressed in a forthcoming article), which differentiates feelings and emotions by layer.

2.2 Productions

Each layer of the model can be described in terms of what it produces, that is, the kind of mental productions it constructs and manages. Each layer is constructing potential meanings from within its own

'vantage point', percolating meanings from which dominant meanings will ultimately emerge. An overview of the productions relevant to electroacoustic music listening is shown in Table 1.

Table 1. Productions of the five layers.

Layer:	Productions:
1. Sensations	• Perceptual binding and grouping: auditory streams, perceptual events, 'gestalted' forms (low-level attributes of the auditory scene) • Attributes of perceptual units, that is, perceptual features • Immediate trajectories of change • Influence of cross-modal integration
2. Gist	• Instanced schemas for things (taken broadly and including people), short-term events such as image schemas and space (high-level attributes of the auditory scene) • The attributes associated with instanced schema (properties of objects, forces, events, etc.), filling their empty slots with defaults, thereby reducing ambiguity • Organization and grouping: foreground/background, strongly related things and events (percussion, footsteps, etc.)
3. Locus	• Instances of action schemas for medium-term situations matched to situational characteristics • Attributes of such schemas including o priorities for action, o appropriate responses in focusing and ignoring, etc. , o filling empty slots with defaults, thereby reducing ambiguity of action. • Executing action schemas and managing transitions when updates occur. • Networks of associations among schemas providing context and deeper meaning
4. Contexts	• Instances of medium- to long-term event-oriented schemas that extend the scope of consideration beyond locus. These include such temporal patterns associated with musical/auditory patterns, scenarios, narratives, etc. • Attributes of such schemas, filling in their slots possibly with defaults and reducing ambiguity. Predictions and expectations are a direct consequence. • Managing/evaluating event schemas in the flow of experience
5. Domains	• Schemas of long-term, domain knowledge, typically 'abstracted' knowledge, meta-knowledge • Parameters of schemas, filling in their slots possibly with defaults and reducing ambiguity.

Let us illustrate the organization of these productions and their layers with a hypothetical example from the everyday, here demonstrating how the mind forms a single mental thread linking related productions across the layers. Our example starts with a series of clicking events that *Sensations* would group as an auditory stream and whose localization could be affected by multimodal influences such as visual motion. The connection of this stream to *Gist* is the start of a thread that *Gist* would understand as footsteps in the spatial context of a room, a thread that would likely be the foreground against background noise. Given this current situation in *Gist*, *Locus* decides where attention should be focused: on the thread of the footsteps that are approaching instead of the thread of the voice that is speaking far off in the distance. Motion approaching personal space takes priority and mental resources are invested there. *Contexts* holds the medium-term pattern that situates this foreground thread, which was learned when first one, then several, people walked past, each one pleasantly saying hello as they passed. *Contexts* anticipates this happening again now. *Domains* holds our background knowledge of personal greeting behaviors, maybe connected to an office setting, and the personal significance that such greetings hold for us. And while this example is taken from the everyday, it could just as easily be part of an electroacoustic work of art.

2.3 Harmonization

In order for moment-to-moment experience to have a sense of consistency, the productions of these autonomous layers must constantly be harmonized with one another. This is a dynamic process in which productions are constantly shifting in response to external and internal changes. As illustrated in figure 1,

we could say that the path from *Sensations* to *Domains* is driven by the onrush of sensory stimulation and that sensory input is compressed and/or contextualized at each step. The path from *Domains* to *Sensations* is driven by internal representations that impose constancy and structure on the dynamic input. Brandt (2004:182) labels the first 'dynamic processing' and the second 'figurative processing'. There are numerous tentative productions being birthed all the time, but the listener will become conscious of only a few. For something to rise to the level of conscious awareness, it must have achieved a significant degree of salience. The operative inhibition of weak, tentative productions eliminates irregularity and perceptual noise in the system that ought to be ignored. Importantly, the proliferation of productions is counterbalanced by the constant pull toward consolidation both within and across threads as the mind is attempting to weave a coherent tapestry, sometimes more successfully than at others.

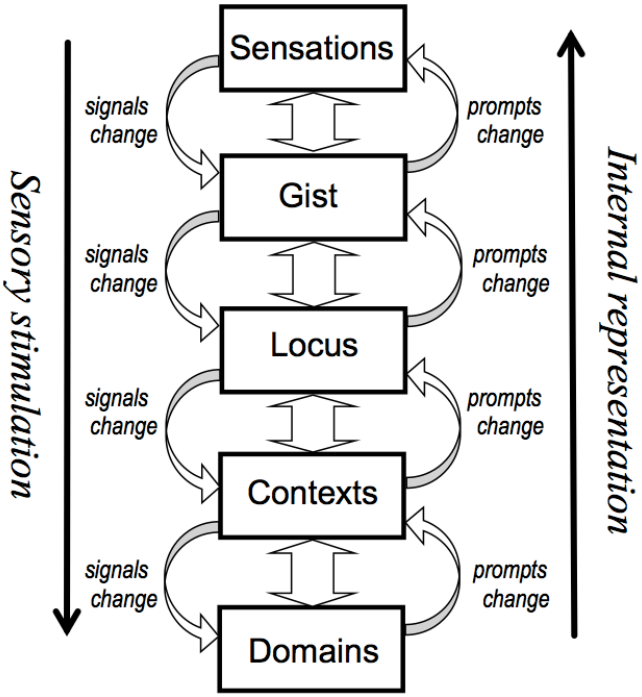

Figure 1. Harmonization within the Five-Layer Model.

To summarize how harmonization works, lets trace the path of a single thread up and down the layers. *Sensations* produces perceptual bindings and groupings while *Gist* connects these to instances of schemas for things in space thereby compressing multifaceted sensory information into something more manageable. *Locus* takes these things and connects them with schemas for the governance of medium-term tasks like deciding what to attend to. *Contexts* is trying to match these on-going situations to longer-term patterns that also generate expectations about the future. Finally, *Domains* sustains the background information that defines and maintains the long-term context. Tracing the thread's path down through the layers: *Domains* delimits the schemas that are appropriate to *Contexts*. *Contexts*' schemas determine what medium-term governance is prioritized in *Locus*. *Locus* causes the *Gist*'s organization to shift in response and thus trigger *Sensations* to retune to different auditory attributes.

2.4 Time Frames

The functional differentiation of the layers extends to their time frames. There are two independent considerations (which are often confounded in common discourse). The first we might call the *functional* time frame of the layer's focus and the second the *operational* time frame in which it responds to change. As a simple example, a listener might think about the form of Stockhausen's *Kontakte* that is 34 minutes long (functional time frame), and after three seconds (operational time frame) switch to the form of Varèse's *Poème électronique*. These two considerations reflect the relationship of the layers to memory: the sensory hold, working memory, and long-term memory, represented in figure 2.

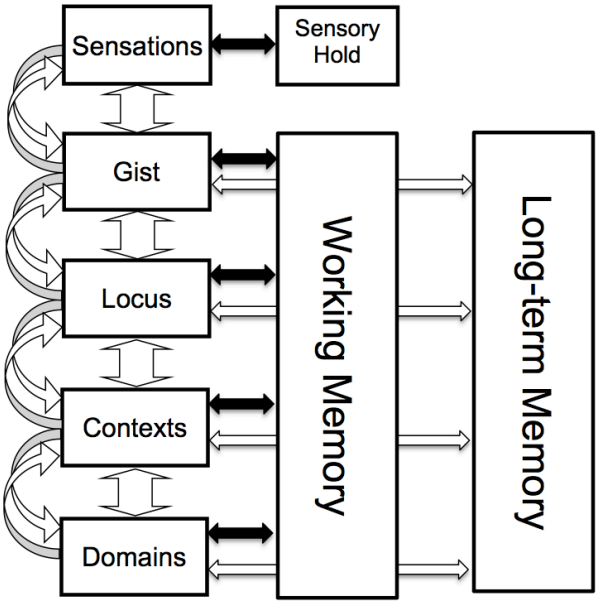

Figure 2. Five layers in relation to memory systems.

The knowledge associated with each layer is accessed and stored in long-term memory, while the current productions are held in working memory. The 'perceptual present' is captured in the inner workings of the layers with working memory and varies between 2 and 8 seconds, depending on the processing load (Pöppel 1997, 1998). Mental activity is always in the perceptual present, but what we think about can have a wide range of time frames. This is reflected in the way that each mental layer specializes in productions of varying functional duration. *Sensations* is the simplest because it registers immediate perceptual processes in sensory hold. The 'auditory present' has a duration of around 250 msec, less than the perceptual present during which auditory attributes can be changing. *Gist* extends the auditory present over several seconds giving short-term constancy to the auditory scene. *Locus* manages real-time tasks in the short- to medium-term, the time frame in which the body executes gestures, that is, the perceptual present and slightly beyond. It accesses action schemas (including listening strategies) from long-term memory while its temporary productions are held in working memory. *Locus* seemingly has a clock, while *Contexts* operates in abstracted time, generally medium- to long-term schema extending well beyond the perceptual moment. *Contexts* manages temporally organized schemas that create projections about the future. *Domains* we can think of as holding the permanent, background information for the present situation.

2.5 Attention vs. Focus

As this model reveals, each layer has its own 'viewpoint' on heard events. The model then enables us to clarify an important distinction between *attention* and *focus*. (Wright and Ward (2008) make a similar distinction between overt vs. covert attention.) We will define attention as a question of *what* is listened to (mental thread), while focus is a question of *how* it is listened to (mental layer). If I should hear Xenakis' *Gendy 3* with my focus in *Gist*, I might find it simply noisy and rough. But if my focus is in *Domains*, which connects my listening to my background understanding of Xenakis and his works, then my experience is entirely different: Xenakis' choices have a nuanced and moving artistic significance. Active focus on a single layer also brings with it the expanded resources to strengthen and expand mental connections. If focused in *Gist*, then that aspect of listening will be enriched and most likely produce an enduring memory. Most listeners shift focus unconsciously and effortlessly, never noticing the resulting leaps of perspective. These effortless jumps occur so easily in part because all layers are continually active and accessible.

The conceptual importance of focus is collaborated by Brandt (2004:183-4) who describes two 'styles of consciousness', practical and communicative, and how they are differentiated by which layer is their 'center of gravity'. Practical consciousness is characterized by a focus on the layer he labels 'Situations' (what is referred to as Locus here), and communicative consciousness is characterized by the combination of Qualia (Sensory) and Feelings. We can relate Brandt's practical consciousness to everyday listening and his communicative consciousness to aestheticized listening (Oliveros 2005), a distinction carried by focus.

One of the most important constraints on moment-to-moment listening is that mental resources are limited. These are real limitations that can be measured in terms of available working memory, inherent neural delays, restricted neural representation, etc. The influence of processing load on limited resources is especially obvious when observing real-time tasks like listening where one often has to glide over the details in order to keep up. *Locus*'s job of prioritizing one sound while ignoring others is the most effective way of managing limited resources.

Attention and focus in active listening will be responsive to a composition's content (and as a corollary, we might say that the work itself will cue the listener to the *what* and *how)*. Then, in repeated hearings the listener tends to repeat the same patterns of attention and focus; in fact, one essentially replays listening patterns. On the other hand, any number of things might induce the listener to deviate from their usual habits, for example, to shift focus to a different layer. Well-practiced listening patterns require fewer mental resources, thus freeing resources for finding new meanings, especially in connection with alternative layers, say shifting from the immediate (*Gist*) to the long-term (*Domains*). The powerful influence of focus, attention and learned listening patterns yields additional insight into how a novice listener's hearing of a work of electroacoustic music can be so different from an expert's.

3.0 Mental Layers and the Emergence of Meaning

In the moment-to-moment flow of listening, just as in everyday life, the mind is constantly engaged in determining the meaning of things. A practical question that we are often implicitly posing is: What is going on? That question is most easily answered in terms of a limited number of options. The more that the details can be related to known models and categories, the easier it is to manage experience. Meaning emerges from the web of relationships we construct, essentially relationships between the onrush of sensory stimulation and the systemization of internal representation. Of course, we expect different listeners to experience differing meanings. Gist, which is closely coupled with *Sensations,* is the most likely layer to produce similar meanings across all listeners. Contexts and Domains, which are the layers most dependent on the listener's internal representations, are the most likely to produce differing meanings, especially when comparing novices and experts. But, while experiences of meaning may differ, the goal of the five-layer model is to capture the essential mental activities of individual listeners, possibly clarifying the differences between novices and experts in specific terms. The five-layer model enables us to separate each layer's contribution to meaning making. In this section, we will explore some repercussions of the five-layer model and some ways in which the layers specifically participate in the emergence of meaning.

3.1 Schemas

The electroacoustic listener's knowledge is both the direct knowledge of things and also the background knowledge that shapes a sense of context. Many components of such knowledge can be modeled as *schemas*, most simply understood as repeated mental activities that are codified as a pattern. For example, we all have repeated patterns of thinking for things like trumpets and loudspeakers, for people like DJs and electroacoustic composers, and for events like raves and electroacoustic concerts. Whether the listener is a novice or an expert, listeners make sense of experience according to the knowledge that they already hold, much of which is schematic. Schemas as such are essential to simplifying the complexity of direct experience by compressing information. They are also essential for thinking at more abstract levels of representation, for anticipating outcomes and for understanding contexts. We will use 'schema' as an umbrella term for what are the closely related mental phenomena of frames and scripts (Minsky 1975; Schank and Abelson 1977). And while the concept is quite broad, we expect schemas to

have some common structural features. For instance, we expect schemas to have slots (like key holes) in which information of a particular type will be matched and fitted (quite possibly with other schemas). Such slots often have default values that fill in missing gaps in specific ways, enabling us to make sense of situations in which our information is incomplete. Schemas are also often connected to exemplars that represent idealizations, for instance, not just a bell, but an idealized bell. Especially important for the arts, schemas are often coupled with feelings and emotions that could arise from our direct experience or from cultural associations. So, if we imagine the schema for the sound of a baby crying, we can envision a slot for filling in the cause of the crying (perhaps with a default value like hunger for a bottle). We can envision prototypical baby cries in our ears, and we immediately know our feelings when hearing this sound.

We made repeated references to schemas in our discussion of the five mental layers without fully clarifying how schemas are differentiated with respect to layer. When discussing Gist, we invoked schemas that capture our recurrent mental processes for objects, motion, space, force dynamics and events, inclusive of the *image schemas* that capture our most essential notions of objects and forces (Johnson 1987; Lakoff 1987; Lakoff and Johnson 1999). When discussing Locus, we largely invoked schemas that were procedural, which like scripts captured patterns of action, in this case, immediate listening strategies invoked in response to the current perceived situation. These schemas are potentially rich with signification---what should be important within a context and how important is it. At the level of Contexts we emphasized schemas that enable the listener to recognize medium- to long-term patterns, especially important for anticipating outcomes, whether in physical reality or in musical works. And at the level of Domains, we discussed schemas that connected what normally belongs together, background knowledge and meta-knowledge, all of which are essential to providing a continuing sense of context. Clearly, all of these differing kinds of schemas are important, and the five mental layers help us to separate and understand their complementary interactions.

The way that sensory stimulation activates schemas and schematic connections is an initial, essential level of meaning making. As was illustrated in figure 2, active mental processes within the layers access long-term memory and its vast interconnected network of schematic knowledge. However this process happens neurologically, triggering this access activates a functional categorization of schematic knowledge both within and across layers. Thus, a bright, high-pitched sound (Sensations) is made meaningful by being categorized as a musical sound, an instrument, a bell, a church bell, a small church bell, etc. (Gist). One of these, probably church bell, represents the most practical level for situational thinking (Locus), the basic-level category (Rosch 1978). The church bell schema has connections to other schemas relevant to Gist: the mechanics of a church bell, the player and the gesture of playing, for the feeling dynamics of the sound, etc. These connections in particular differentiate church bells from doorbells, and they effectively clarify the frame of meaning that we hold around church bells. At the same time, we can imagine many other connections being made to schemas at other mental layers: for the role of bells as signals for attention, for musical patterns with bells, for the role of bells in church services and other important social events. Thus, the activation of schemas at one mental layer may also activate schemas across the layers. Conversely, the schemas already present at other layers may influence which schemas are most salient and activated. Narrowing the range of potential connections is a particularly important role for domain knowledge. The network of associations we hold within a particular domain guides and limits new connections and projections. For example, if Domains is holding a schematic network for weddings, it will influence the schemas that may be activated by a bell sound. This is a consequence of the harmonization process across the layers that we described in section 2.3, a process that assists in managing limited mental resources by eliminating the less salient productions.

3.2 Mental Spaces

While schemas are accessed from long-term memory, they are instantiated and made active in working memory, organized within *mental spaces*, the temporary assemblies that are constructed, interconnected and modified in moment-to-moment thinking (Fauconnier 1994, Fauconnier & Turner 2002). The concept of mental spaces enables us to model the schematic connections that are made on the fly and moment-to-moment. For example, we expect that any heard event instantaneously activates a mental space in working memory. Depending on context, this space will contain the schemas that frame the

event (say, source category, type of agency, etc.) along with their connections to other potentially relevant schemas. We will call these *schematic associations,* adjunct connections, including connections with other mental layers, that may be reinforced depending upon context or the particular associations that are strong for the user. Thus, we can visualize that sensory stimulation at the level of Sensations evokes a mental space for a perceptual event recognized as striking a bell at the level of Gist. Due to limitations of resources, the schemas activated with this mental space can only represent a fraction of all relevant schematic associations. Depending upon the listener's predispositions and the context, especially the context of what else is in working memory, this space is potentially connected with other schemas and mental spaces: school bells, being startled, standing in the school yard, feeling imposed upon by school, etc., some of which relate most closely to Locus, Contexts and Domains. A new event or change of context will likely evoke changes in the mental spaces and their interconnections. How mental spaces affect one another, how new connections are made, how significance is recognized, etc.---this is the work of the mental layers that will dynamically reconfigure the mental spaces of working memory.

So, at the beginning of Jonathan Harvey's *Mortuos Plango, Vivos Voco* (Harvey 1981, 1990) when we hear the sound of church bells, a mental space is instantly created which frames church bells and which activates schematic associations. When considering typical listeners, we can expect that some of the listener's schemas are cultural, shared and easy to anticipate (how church bells are rung, the typical occasions for ringing church bells), while others are more individual and personal (personal memories of hearing church bells). It is important not to lose touch with the fact that the interconnection of mental spaces is a creative act of meaning making that can produce a variety of results. And, importantly for the arts, we can anticipate that this cultural/personal mix also encompasses the feelings associated with church bells: the cultural solemnity of church bells, personal feelings about churches, rituals, etc.

In *Mortuos Plango, Vivos Voco*, when soon after the church bells we hear the entrance of a boy soprano singing in Latin, another mental space and its schematic associations are evoked in working memory. Much of its extended network of associations will overlap the first. The overlap immediately clarifies the conceptual background of the piece, especially in Domains, because areas without overlap, may be less important to the piece's sense of meaning---associations of boys with childhood mischief, for example. These may never become sufficiently salient to participate significantly in the listener's sense of meaning. Other associations may be quite salient and participate in meaning because their incongruences are highlighted. For example, the sonic disparities between the bells and the boy at the level of Gist turn out to be very important to the continuation of the piece.

3.3 Blending

We have said that meaning in electroacoustic music can emerge directly from schematic associations, that is to say that meaning is a product of the schema's connections within the vast network of schemas. In the example of Harvey's *Mortuos Plango, Vivos Voco* above, we discussed the schematic associations of church bells and boy sopranos, and we also proposed that the broad context of the piece was clarified by the overlap of two schematic networks. Significantly, meaning can also emerge when the listener forms combinations of mental spaces in the process called *blending*, sometimes called conceptual blending or conceptual integration (Fauconnier and Turner, 2002). Fauconnier and Turner have developed models for the most common forms of blending called *conceptual integration networks*. The prototype integration network shown in figure 3 has four mental spaces: two represent the spaces for the input elements that are being blended; one represents the space of the common elements that relate the two inputs, the *generic* space; and one represents a new space, the *blend*. In blending, selective content from the input spaces is projected into the blend, which now contains relationships that cannot be represented by either input. Blending produces emergent structures not contained within the input spaces and most importantly emergent meanings that are unequivocally a product of the blend.

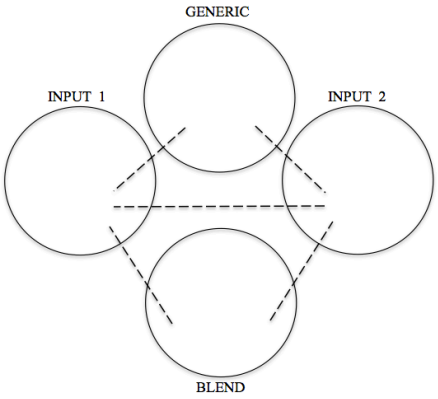

Figure 3. A connection of mental spaces representing a prototype integration network, after Fauconnier and Turner (2002).

In *Mortuos Plango, Vivos Voco*, we already mentioned the sonic disparities between the bells and the boy soprano. Shortly after the opening section, new sound material is introduced that is a product of cross-synthesis between the sounds of the bell and the boy soprano. This is art: a sound that is physically impossibility is realized through the magic of digital signal processing. How does one understand the deeply evocative sound of a hybrid bell-boy? It requires an imaginative act of meaning making. The blend of the two is an emergent entity with many hybrid properties (as illustrated in figure 4), one that is rich with schematic associations and not necessarily harmonious ones. For example, the bell is a passive object that must be acted upon by an agent in order to produce sound, while the boy soprano is the agent of its own sound production. Especially in the context of contemporary arts, blends often contain the incompatible or the oppositional, the novel or the profoundly obvious. Here the disparity is greatest at the level of Gist, where there are also discrepancies between the sound sources (metal and flesh) and between acoustic envelopes (impulsive and sustained). As the individual sections of *Mortuos Plango, Vivos Voco* unfold one after the other, the flexibility of the cross synthesis enables the potentials of the bell-boy blend to be explored first one way and then another, creating novel events that mix bell, boy and hybrid attributes. Passive vs. active agency is negotiated, as well as metal vs. flesh source material. There are bells seemingly singing and clusters of boys sounding struck tones.

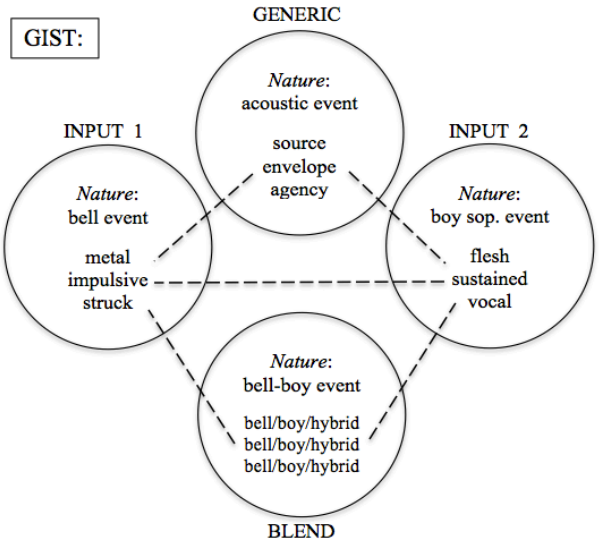

Figure 4. Prototype of the blending of church bells and the boy soprano in Jonathan Harvey's *Mortuos Plango, Vivos Voco*.

3.4 Context and Domain Knowledge

As becomes obvious from much we have said so far, meaning arises in a context. Speaking in regard to cognitive semiotics, Per Aage Brandt says:

> 'Anything meaningful is meaningful in a 'context'; contexts supply relevant frames for the contents of our consciousness, and they thereby allow us to draw inferences from these contents. . . . contexts are structured within distinct semantic domains, . . .'
>
> (Brandt 2004:33)

And while there is no universally accepted model for the organization of schematic knowledge, the notion of distinct domains of knowledge is useful to understanding how listeners relate content and context, especially how the network of schematic associations is negotiated by the layers of Contexts and Domains while listening. Mental spaces are partial activations of knowledge within a particular domain. Drawing again from cognitive semantics and following the lead of Brandt, let us postulate at least three largely independent 'base' domains for the listener's knowledge that we can call *nature*, *culture* and *spirit*. Brandt in fact describes these as 'world types':

> '1. NATURE: the macro-physical, material and gravitational, geo-, bio-, and zoological environment;
> 2. CULTURE: the collective horizon formed by groups of fellow human beings densely informed by intentional and mimetic behaviors of all kinds, practical or symbolic; . . .
> 3. SPIRIT: the sphere of direct interaction with other minds by expressive contact, allowing for the sharing of thoughts and feelings with individuals in a face-to-face relationship.'
>
> (Brandt 2004:23)

Music might also be considered an independent domain, especially when considering the area of seemingly autonomous musical knowledge (harmony, grammar, etc.). Of course, interconnections between aspects of music and other domains can be demonstrated; some music theorists have related musical intelligence to linguistics (Jackendoff and Lerdahl 1983) and others to ecological knowledge (Clarke 2005).

Within the field of electroacoustical music, soundscape composition, for example, clearly focuses on the context of the *natural* domain. Barry Truax (2000) says that the first two characteristics of soundscape composition are the recognizability of source material and the listener's knowledge of the environmental and psychological context (relating to Gist and Domains respectively). And ironically, *nature* is also the base domain for Denis Smalley's seemingly antithetical concept of 'remote surrogacy', in which a sound event is detached from its natural, causal agent (Smalley 1986, 1997). This fact points up what a central concern *nature* and *natural* events have been for electroacoustic composers. Natural ecology has played a central role in electroacoustic theory, whether it is the attempt to create 'sound objects' devoid of ecological meaning (Schaeffer 1966) or to hear electroacoustic compositions in the manner of natural soundscapes (Truax 2000). A particularly interesting twist on ecological listening is David Wessel's *Anthony* (1977), which is a kind of perceptual Rorschach test in which objects and spaces are evoked by the most totally un-*natural* aggregates of sine waves. Purely electronic sounds like these lie in what we might call the *natural* subdomain of *technology*. All of this extended attention to *nature* reflects a shared style of meaning making (analogous to Brandt's (2004:183-4) styles of consciousness) that concentrates mental activity within Gist while repeatedly evoking the domain of *nature*.

We have discussed how the cross synthesis of the bell and the boy soprano in *Mortuos Plango* produces a bell-boy blend which reconciles discrepancies in source material and agency, properties that lie in the domain of *nature* and at the level of Gist. But there are also discrepancies at the level of domain knowledge. What is a conceptual bell-boy, not just a sonic one? In this situation, forming the sonic blend compels the listener to form the conceptual blend. The conceptual blend is formed at the level of Contexts (and is eventually integrated into Domain's knowledge of electroacoustic music). This is represented in figure 5. The bell's schematic network interconnects with schemas in *nature* and *culture* that we have already discussed. The boy's network interconnects with those and also with *spirit*---we identify with the boy's internal thoughts and feelings (part of what Dan Dennett (1987) calls the

'intentional stance'). And here we come to the compelling artistic core, the emergent meaning of these blends, which depends on an intersection of base domains. In alignment with the piece's background of church and spirituality, the listener is led to blending the animate and the inanimate. How that blend is understood by the listener must vary with personal history and receptivity, and it may lead the listener to the realm of the spiritual that was a primary concern for Jonathan Harvey. That was the association of both Michael Clarke (2006) and Patricia Lynn Dirks (2012) in their analyses of the piece. Dirks came to "interpret the bell as representing the dead and the boy's voice, the living."

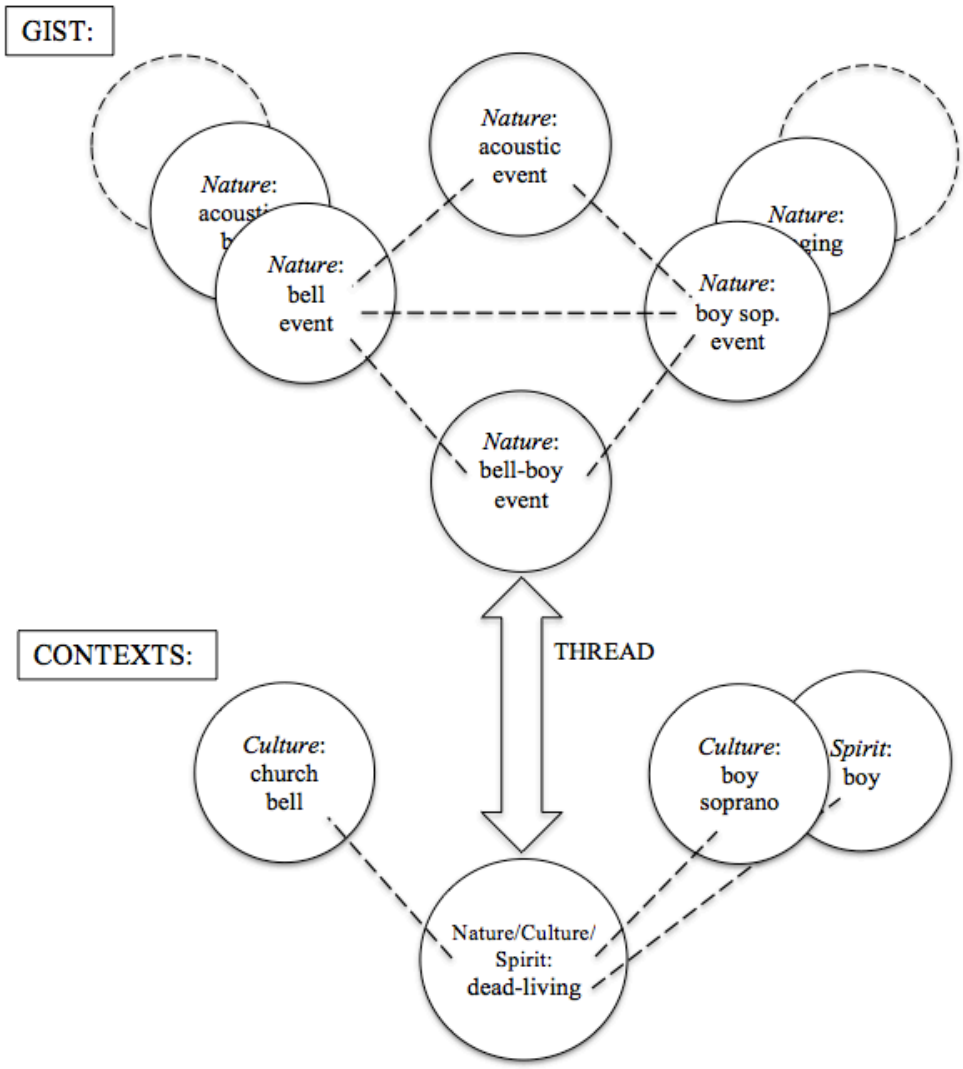

Figure 5. The multi-layer blend of church bells and the boy soprano at the level of *Gist* also compels blend at the level of *Contexts*, a blend of intersecting base domains.

4.0 Example: Stockhausen's *Telemusik*, 'Structure 16'

As an illustration of how the concepts discussed above might be applied in the analysis of electroacoustic music, we will describe a hypothetical hearing of a short excerpt from Stockhausen's *Telemusik*, the section identified as 'Structure 16'. (What for lack of space cannot be covered in this one excerpt we hope to illustrate in future analyses.) This is an attempt to describe a reasonably informed, hearing, not necessarily an idealized one. 'Structure 16' is arguably the simplest of the 32 'Structures' that make up

Telemusik. It begins like all of the others with the sound of a Japanese temple instrument, here a small mental bowl called a *rin* (Kohl 2002). And like the other 'Structures', its length is a Fibonacci number of seconds, here actually the sum of large and small Fibonacci numbers, 55+2 = 57 seconds. Excluding the initial percussive sound that marks the beginning of the section, the content of this 'Structure' is exclusively made up of high-frequency ring-modulated Gagaku music that is completely unrecognizable as such. What the listener does hear is glittering clusters between 5 and 7 kHz that are changing more or less every couple of seconds. Such high-frequency clusters are emblematic to *Telemusik*. In Stockhausen's *Hymnen*, similar sounds are mimicking short-wave radio signals, and from that perspective, *Telemusik*, tunes into the 'vibrations' around the earth, in this case, picking up indigenous music from around the globe. In a less literal way, these clusters capture a continuous, dynamic flow of energy (as if the background sound of the world is not silence but a buzz). Similar sounds occur numerous times before 'Structure' 16, where the high-frequency ring modulation is interrupted three times by rapid, stepped changes in modulation frequencies that produce somewhat melodic sequences in a lower frequency range, 0.5-1.5 kHz. The score to 'Structure 16' is shown in figure 6 (Stockhausen 1969).

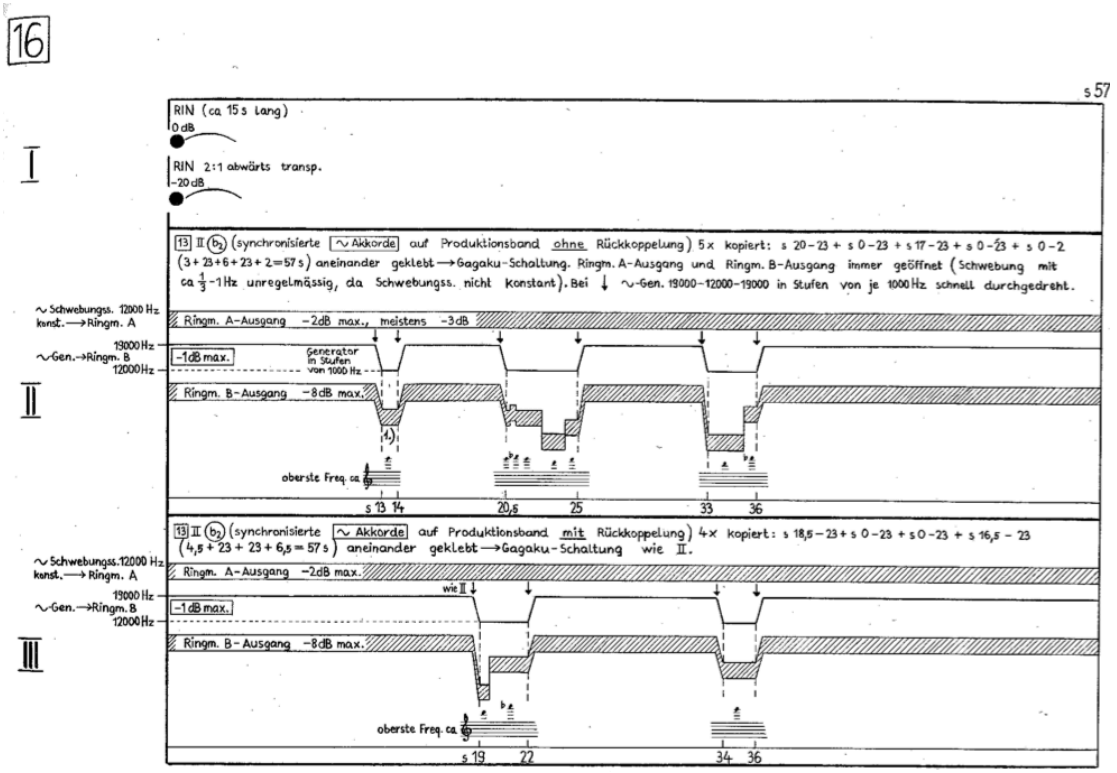

Figure 6. The score of Stockhausen's *Telemusik*, Structure 16 (Stockhausen 1969).

Most interestingly, the last 20 seconds of 'Structure 16' are occupied only by uninterrupted high-frequency clusters. From an acoustic perspective, almost nothing happens for 20 seconds. But that hardly means that the listener's mental processes are inactive: any analysis based solely on the acoustics would skip right past the significance of such a passage. A sonogram of 'Structure 16' with numbered markers for critical moments in time is shown in figure 7.

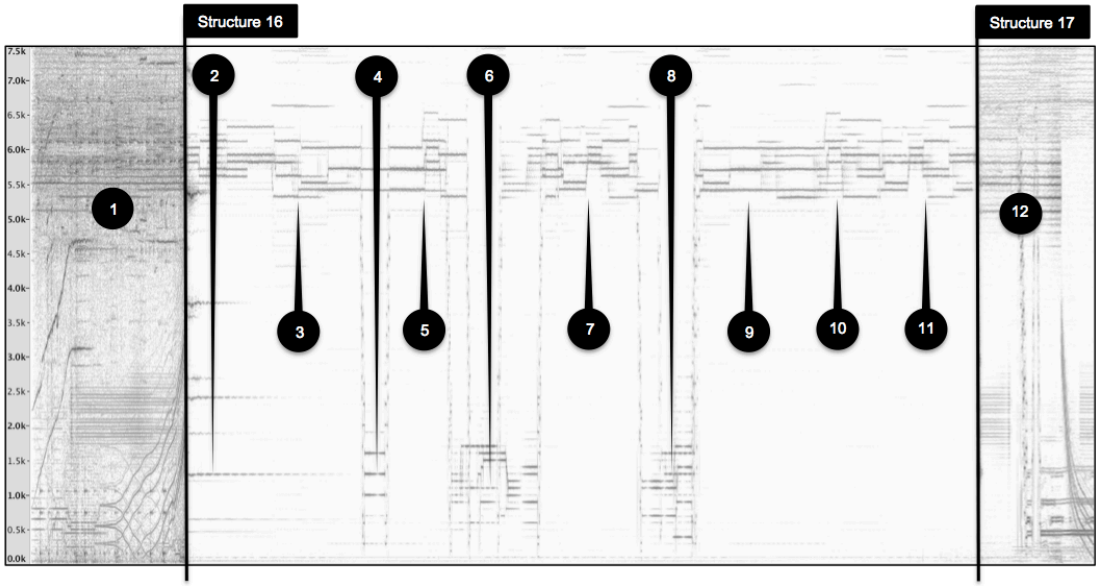

Figure 7. Sonogram of 'Structure 16' with time markers related to changes in the activity of the mental layers.

We will examine a few of these moments in depth before scrutinizing those last 20 seconds. Marker #1 shows the activity at the end of 'Structure 15' that abruptly stops at Marker #2. From the perspective of *Sensations*, there is a new perceptual event that is pitched, impulsive and sustained. That perceptual event triggers a match to the schema for 'Bell' (from long-term memory) by which *Gist* makes sense of the perceptual event. The Bell schema also invokes a network of schematic associations among which are connections to world music and the ceremonial contexts evoked earlier by the piece. From the perspective of *Gist,* there is a Bell in the foreground, a new thread that is posted to working memory. *Locus* reacts to this simple scene with high certainty and a simple strategy (from long-term memory): attend to the thread of the foreground object, the bell. This is a more complex moment for *Contexts*, because everything that it was holding in working memory about the previous section is no longer in sync. But the meaning of the moment is confirmed by the bell strike which matches the schematic pattern learned earlier in the piece (and held in long-term memory) that such a sound marks the beginning of a new section. *Contexts'* productions for 'Structure 15' (its schemas, blends, threads, etc.) must be closed off and incorporated into long-term memory. Long-term memory also holds the listener's background knowledge of Stockhausen, his music and its language—what is typical and prototypical. *Domains* sustains the relevant knowledge in working memory and consolidates new information about *Telemusik*, creating domain knowledge of the piece, such as its typical textures, types of sounds, etc. The present moment with the bell sound seems entirely consistent with previous knowledge of *Telemusik*.

At marker #3, some changes occur. *Sensations* registers an on-going auditory stream that is high and rough, a second thread. With the cessation of the bell, *Gist*, *Locus* and *Contexts* must sort out relationships: is this bright, coarse cluster foreground? If so, *Locus* would switch attention to the clusters thread. But given the persistence of the Bell schema in the perceptual present (working memory), *Gist* takes the clusters for a background entity. *Locus'* strategy is less certain but simple: ignore the background thread and attend to the next foreground event. *Contexts* knows from earlier in the piece (long-term memory) that clusters do not likely change or create significant patterns; this influences *Locus*. So, all three layers are harmonized. *Contexts* must now make sense of this new section, so there is an opened-ended search anticipating the start of some new schematic pattern within the section.

Marker #4 indicates the first melodic event. *Sensations* registers a pitched, mid-range event in addition to the continuing high, rough stream. *Gist* understands this as new thread, a melodic foreground event against the continuing background thread. The event is similar to ones encountered earlier in the piece, and repeats a pattern of blending the natural and the technological domains to produce an emergent vocal-electronic event (top of figure 8). This kind of pattern at the level of *Gist* reoccurs and becomes part of

the domain knowledge about *Telemusik*. In this particular case, the event lies within a vocal range (domain of *Nature*) while it is also clearly electronic when jumping through stepped frequencies, like clicking through settings of an oscillator (domain of *Technology*). The life-like quality of the sound probably comes from the fluctuating quality of the original Gagaku music. This blend at the level of *Gist* compels the listener toward a conceptual blend at the level of *Contexts*, a blend that combines behaviors from within the domain of *Music* and the subdomain of *Electroacoustic Music* (bottom of figure 8). This conceptual blend of behaviors has a unique human-technological felt quality and the potential to make predictions about future behavior. *Locus* attends to the melodic thread with high certainty while ignoring the background. This gives *Contexts* the focus to take in the event as the beginning of a new melodic pattern (based on the conceptual blend and incorporating typical musical behaviors), after which it can narrow its search and anticipate continuations. For most listeners, attention at this point is on the melodic thread and mental focus is in *Contexts* because it is the most important layer for making sense of the on-going experience. That provides expanded mental resources with which *Contexts* is likely creating numerous productions (that will never become conscious and will eventually be discarded). By marker #5 when *Sensations* once again registers only the on-going stream, *Gist's* foreground/background relationships are stable as well of *Locus'* listening strategy.

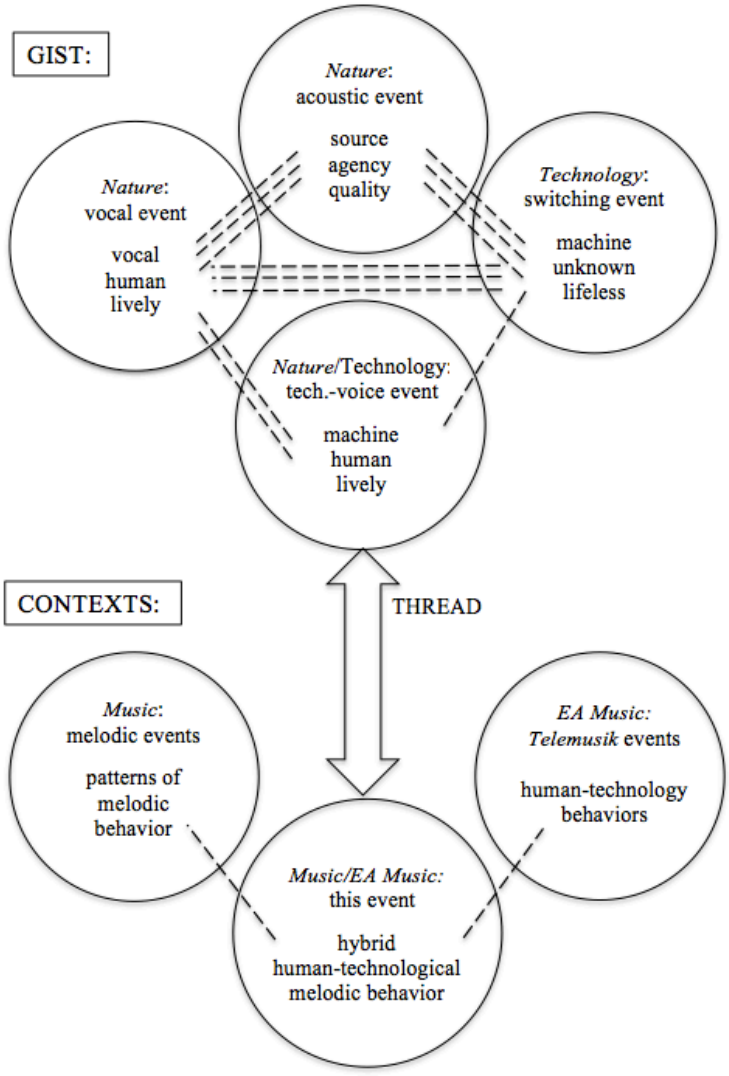

Figure 8. Multi-layer blend of the 'melodic events' of Structure 16.

The second melodic event at marker #6 is a continuation of the melodic thread still active in working memory, now a duet, which gives additional support to the vocal behavior of the sonic blend. This is also a confirming moment for *Contexts* because possible melodic schematic patterns are reduced to repetition-variation, which enables *Contexts* to more securely project continuations and to anticipate outcomes. Up to this moment, the sense across the mental layers has been one of certainty and confirmation. But after the second melodic event, the feeling arises in *Domains* that the absence of other events and other threads is inconsistent with its long-term knowledge of *Telemusik*. This is far too sparse!

The third melodic event at marker #8 confirms *Contexts'* projections and strengthens the anticipation of a consistent continuation to the thread. *Locus* is highly certain. But by marker #10, working memory is being stretched beyond the perceptual present, and nothing has happened. *Contexts* holds to a pattern that has not closed and now must search to find some new pattern or context that makes sense of this unanticipated lack of continuation within a typical time frame. *Locus* must refocus to find something in the foreground. *Gist* is signaled to reevaluate, but with only the cluster present, the foreground/background distinction is ambiguous. By marker #11, the cluster thread shifts to foreground because there is nothing else to attend to. This pushes *Sensations* to retune and register any potential changes in the clusters as perceptual events, although nothing quite qualifies. *Gist* experiences the clusters now as an amalgam of activity.

For most listeners, attention has shifted to the clusters thread and mental focus has shifted to *Gist* because that layer is most important to making sense of the on-going experience. *Locus* focuses on the clusters with a lack of certainty and *Contexts* attempts to find some pattern in this, but there is nothing to base a pattern on. There is no increase or decrease of any attribute; there is complete stasis. There is no way to make a projection about the future, and the listener's feeling may become somewhat apprehensive lacking direction. The sparseness and the lack of continuation intensify *Domain's* feeling of inconsistency with the rest of *Telemusik*. During these last 20 seconds, the thread that was previously background has become foreground. And, in the effort to make sense of things, focus has shifted from *Contexts* to *Gist.* Acoustically the clusters are just about the same as they were at the beginning of the Structure, but the listener's mental processes and the artistic meaning of the clusters has shifted significantly from what it was back at marker #6.

At marker #12 when 'Structure 17' begins with the sound of a wood block, the mental processes are very similar to the start of 'Structure 16'. *Contexts* recognizes the beginning of a new section and therefore closes its productions from the previous section. This means that the unusual sectional pattern for 'Structure 16' is incorporated into long-term memory. *Domains* too must integrate this unusual behavior in long-term memory so that the domain of *Telemusik* is expanded. Along the way, a trace of the listener's experience and the surviving productions of each layer have been written into episodic memory as a long-term record of the listening experience.

We have explored one hypothetical hearing. With repeated hearings, one would expect that many of the same mental activities would occur over and over again, especially activities that are driven by sensation: *Sensations, Gist, Locus*. The listener might also free up the mental resources to expand connections around the emergent meaning of its many blends or to marvel at the mystery of Stockhausen's odd decisions (domain of *spirit*). And repeated hearings could also deepen the experience of 'Structure 16', as the unique incongruity is embraced, intensifying the artistic nuance that the listener has learned to anticipate.

5.0 Conclusions

The key question posed at the beginning of this essay was: How does the listener's experience of listening become *meaningful*? We have adapted a model of five mental layers from Per Aage Brandt in an effort to elucidate the multi-layered nature of listening. And, we have integrated schema theory and blending theory into this model with examples from a well-known electroacoustic work by Jonathan Harvey. The potential of this expanded model has been illustrated by a very detailed analysis from another well-known work by

Stockhausen. It is not the purpose here to propose this detailed example as a general template for analysis, rather to illustrate the degree to which modeling the mental processes is essential to any understanding of how listening becomes meaningful. One lesson that this example clearly teaches, especially the last 20 seconds, is that electroacoustic analysis based solely on acoustic information, in either scores or sonograms, easily misses essential mental processes that are fundamental and basic to meaningful listening experiences. Therefore, future developments in analysis should take this essential lesson to heart.

In focusing on electroacoustic music, we have addressed not only meaning in its most general sense, but also meaning in the context of an auditory artistic experience. Electroacoustic music provides a distinctive domain for the discussion of cognition and meaning, especially in relation to blending theory where examples are typically drawn from non-auditory sources. The electroaoustic listener has meaningful experiences unfolding moment-to-moment without the need for conscious reasoning about sound; meaning is instantly grasped in sonic thinking especially at the level of Gist. Then too, a question that we have touched on here is the nature of the mental activity that distinguishes the expert listener from the novice. How does the experience of listening shift from a possibly disoriented encounter to nuanced meanings? To answer such a question, we must be able to describe the distinguishing mental activities. Then too, emotion and feeling are essential components of artistic experience. How do we integrate these in the context of listening's mental activity? It is hoped that this essay and the five-layer model helps to provide a foundation for the future exploration of such questions.

References

Brandt, P. A. 2004. *Spaces, Domains, and Meaning: Essays in Cognitive Semiotics*. European Semiotics: Language, Cognition, and Culture, vol. 4. Bern: Peter Long.

Brandt, P. A. 2006. Form and Meaning in Art. In M. Turner (ed.), *The Artful Mind: Cognitive Science and the Riddle of Human Creativity.* Oxford: Oxford University Press.

Clarke, E. F. 2005. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning.* Oxford: Oxford University Press.

Clarke, M. 2006. Jonathan Harvey's *Mortuos Plango, Vivos Voco*. In Mary Simoni (ed.), *Analytical Methods of Electroacoustic Music*. New York: Routledge.

Dennett, D. 1987. *The Intentional Stance.* Cambridge: MIT Press.

Dirks, P. L. 2012. An Analysis of Jonathan Harvey's "Mortuos Plango, Vivos Voco". Downloaded on January 3, 2013, from http://cec.sonus.ca/econtact/9_2/dirks.html.

Donald, M. 2001. *A Mind So Rare: The Evolution of Human Consciousness.* New York: Norton.

Donald, M. 2006. Music as Communication. In: M. Turner (ed.), *The Artful Mind: Cognitive Science and the Riddle of Human Creativity.* Oxford: Oxford University Press.

Fauconnier, G. 1994. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. New York: Cambridge University Press.

Fauconnier, G. and Turner, M. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Harvey, J. 1981. Mortuos Plango, Vivos Voco: A Realization at IRCAM. *Computer Music Journal*, 5(4).

Harvey, J. 1990. Mortuos Plango, Vivos Voco. *Computer Music Currents 5*. Wergo, WER 2025-2.

Kohl, J. 2002. Karlheinz Stockhausen: Telemusik (1966). In: T. Licata (ed.), *Electroacoustic Music: Analytical Perspectives.* Westport, CN: Greenwood Press.

Jackendoff, R. and Lerdahl, F. 1983. *A Generative Theory of Tonal Music*. Cambridge: MIT Press.

Johnson, M. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* Chicago: University of Chicago Press.

Lakoff, G. 1987. *Women, Fire and Dangerous Things.* Chicago: University of Chicago Press.

Lakoff, G. and Johnson, M. 1999. *Philosophy in the Flesh*. New York: Basic Books.

Minsky, M. 1975. Frame System Theory. In P. N. Johnson- Laird and P. C. Wason (eds.), *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press.

Oliveros, P. 2005. *Deep Listening: A Composer's Sound Practice*. New York: iUniverse, Inc.

Pöppel, E. 1997. A Hierarchical model of temporal perception. *Trends in Cognitive Science* 1(2).

Pöppel, E. 1998. *Mindworks: Time and Conscious Experience.* New York: Harcourt Brace Jovanovich.

Rosch, E. 1978. Principles of Categorization. In Rosch, E. & Lloyd, B.B. (eds), *Cognition and Categorization*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Schaeffer, P. 1966. *Traite´ des objets musicaux*. Paris: Seuil.

Schank, R. C. and Abelson, R. P. 1977. Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures. Hillside, NJ: Lawrence Erlbaum Associates.

Smalley, D. 1986. Spectromorphology and Structuring Processes. In S. Emmerson (ed.), *The Language of Electroacoustic Music.* Basingstoke: Macmillan Press.

Smalley, D. 1997. Spectromorphology: explaining sound-shapes. *Organised Sound*, 2(2).

Stockhausen, K. 1969. *Nr. 20 Telemusik* (score). Vienna: Universal Edition (UE 14807)

Truax, B. 2000. The aesthetics of computer music: a questionable concept reconsidered. *Organised Sound*, 5(3).

Wessel, D. 1990. Liner notes for *Antony.* In *Computer Music Currents 10*, Wergo CD# WE119.

Wright, R. D. and Ward, L. M. 2008. *Orienting of Attention.* Oxford: Oxford University Press.