Toward a Dramaturgy for Theatre Sound-Scoring Design

Kevin Purcell
Box Hill Institute
Centre for Performing Arts
Box Hill, 3128
Australia
k.purcell@bhtafe.edu.au

Abstract

This paper describes the foundation investigation into new ideas for the dramaturgical underpinning of sound design in drama within the context of stage productions. The ideas expressed are, in large part, transferable to other forms of dramatic narrative but have not been tested outside of the context of stage plays. The design paradigm of the author's Max/MSP patch (The MaxStage) as described, encompasses discourses in theories of auditory listening, semiotics of dramatic interaction and the semantics of human–computer interface design. Explained through a derived system model, The MaxStage is a collaborative testing environment to evaluate both the function and design of auditory aura as suggested by the system model. The system taxonomy for The MaxStage is theorised as part of a physical, interactive computer music system (ICMS). By virtue of being a prototype, however, The MaxStage is several steps removed from the physical technologies required to re-create intended auditory outcomes in a live theatrical context.

Introduction

My primary research focus in the application of computer music technologies is in the dramaturgy of sound design. As an investigative tool, I have authored a prototype application in Max/MSP called *The MaxStage* v.1.4 (Fig. 1.). Constructed via a system taxonomy after Laurel (1993) and incorporating Schaeffer's revised indexical modes of listening described by Smalley (1996) and Chion (1994), *The MaxStage* allows sound designers to explore the suitability of designed auditory aura within in a 2D environment. As a tool for non-specialist computer users, *The MaxStage* is designed for collaborative use with directors, choreographers and designers. In this sense, the application is like a sound designer's toolbox similar in philosophy to 'The Performance Animation Toolbox' project of Callesen, Kajo and Nilsen (2002) for the Narrativity Studio at the Interactive Institute in Malmö, Sweden.

Principles of Design in The MaxStage

As a simulation, *The MaxStage* adequately contributes to the following principles in the design of a multimedia interactive model for dramaturgical sound design

1. The constructed virtual environment represents complex knowledge bases as well as individual constructions of unique knowledge associations.
2. The nature and perceived value of these knowledge bases may be mediated not only by the designer's intent, but also by the unique associations generated by other individuals.
3. The presentation stimuli (visual, audio, text, etc.) are semantically and contextually compatible units. Segmentation of the units is meaningful and not simply display.
4. The ability to retrace and reflect through operating the system model, enacted in the interface, allows probable construction of auditory meanings neither possible nor likely outside of the environment.

The interface as a working presentation of the derived taxonomic system, assists in understanding dramaturgical relationships between intention and action. This would, hypothetically, deepen with increased usage; planning processes being enhanced by continuing manipulation of the system.

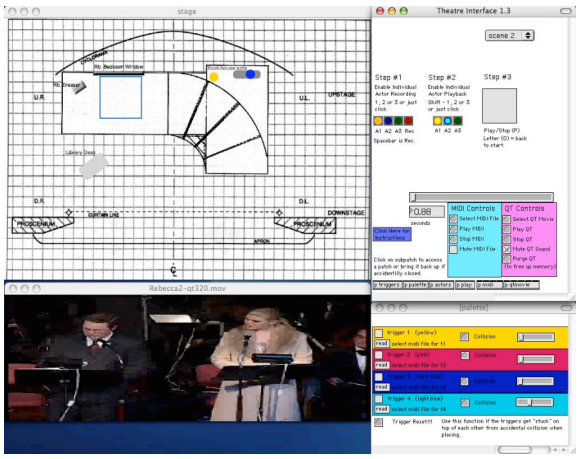

Figure 1. *A scene plot from The MaxStage. Circular sprites represent stage actors. The square shown upstage opposite prompt (o.p.) is a scaleable virtual sensor frame. Actor 'collisions' with the sensor frame trigger either audio or MIDI files. Multiple sensor frames and up to three actors can be accommodated. The system allows for storing and recall of up to ten separate scenes.*

The MaxStage environment is comprised of five modules that can be positioned in any configuration on the screen by the user. The five modules are

- The Stage

- Theatre Interface
- The Trigger Palette
- The QuickTime™ (QT) Movie Window
- The Instructions

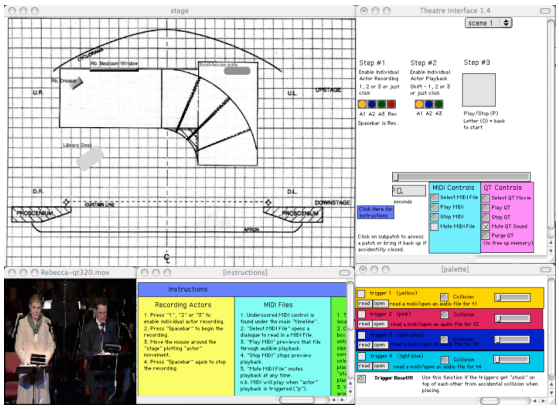

Figure 2. *Overview of The MaxStage Environment*

The Stage (Top Left)

In Figure 2., the background graphic is birds-eye view of a hypothetical stage plan. In this scenario, the set design includes two trucks connected at right angles: the higher-level truck positioned stage left. Access to the stage right truck is via a descending four-level stair. A cyclorama upstage of the trucks is used for back projections.

In the current version of *The MaxStage*, the only way to initiate changes of set, or properties, on the background graphic is by re-drawing objects onto the stage plan and re-saving the image in the Max file path preferences. In subsequent versions, calling up different scenes will also call up the appropriate stage graphic as intended.

Theatre Interface (Top Right)

A maximum of ten scenes can be recorded for subsequent recall. Scenes are selected from the drop-down menu in the top right corner of the Theatre Interface window. New scenes may be selected in Playback mode as well as in Stop mode. Only actor movements are stored within scenes in the current version of *The MaxStage*. Sensor-zone placements, MIDI files and QT movies have to be set manually for each scene as needed. Whereas this may appear to be a limitation, greater flexibility in auditioning different trigger placements and alternate MIDI or audio files can be accomplished without having to re-map predisposed actor movements.

Individual actors are enabled by clicking on 1, 2 or 3; or alternatively, by selecting buttons A1, A2 or A3. Each actor is indicated by a coloured circle: A1 (Yellow), A2 (Blue) and A3 (Green). Up to three actors can be represented simultaneously in this version of the application.

To permanently remove an actor off the stage, the corresponding sprite is re-selected and the mouse moved in any direction.

The Trigger Palette (Bottom Right)

There is a choice of four sensor-zones available from the palette window. Each sensor-zone is a virtual representation of physical sensor-based system designed to identify movement within a fixed boundary.

Sensor-zones in *The MaxStage* are placed on the stage by clicking and dragging from the square boxes. The size of any sensor-zone is scaleable by increasing the respective slider value on the right-hand side in the palette module. Each trigger is separately colour coded. When an actor 'collides' with a sensor zone, a separate 'collision' detected LED will be activated. If an actor moves quickly into and then out of a sensor zone, there may be no visual confirmation, dependent upon the processor speed of the user's computer. The Max/MSP software places a higher priority on audio events than it does graphic events. Irrespective of this, the pre-selected audio or MIDI file will be triggered.

A separate audio or MIDI file can be selected to be triggered when an actor crosses into a zone. Audio and MIDI files for each trigger (sensor-zone) are selected via the respective 'read' buttons. The designated auditory aura is designed to continue irrespective of whether the actor sprite moves outside of the predetermined boundary of the sensor zone. For clarification: this is important, as part of the collaborative process, to determine the respective physical area required for physical sensor coverage when the intended auditory aura is designed to take precedence within the semiotic hierarchy. Multiple Audio or MIDI files can be triggered sequentially irrespective of whether other sound files are still playing.

In the current version of *The MaxStage*, trigger location and relative sizing information is not captured in the scene information. Triggers must be set into the correct position at the beginning of each session. This is a shortcoming in the prototype version.

To delete a trigger on the stage, click on it and press delete. The 'Trigger Reset' button is used if two triggers are accidentally placed on top of one another. Triggers, however, can be placed within each other to create multi-zone frames.

The QuickTime™ Movie Window (Bottom Left)

The QT Movie Window is incorporated into *The MaxStage* environment to show an audience view of the stage and associated actor movement within the stage space. Its function is to act as visual reminder of established stage blocking taken from production rehearsals. Alternatively, it could be integrated to show back-projections; or any other pertinent visual information, as part of a production, installation or physical theatre event.

Instructions (Bottom Middle)

A brief set of instructions may be opened to remind users of how to record, playback, stop and store scene information.

Constructing a Framework for Enriched Auditory Experiences

The design of an enriched auditory experience for any drama setting: whether it be stage play, musical, or other hybrid music-drama setting, necessitates the design of a hierarchical system to codify and explicate the use of auditory stimuli that are separate from any pre-existing musical or audio source material. Treated in isolation, without reference to an overarching structural paradigm, any sound design risks becoming a concatenation of unrelated auditory events.

Theories on aesthetics of auditory information are able to describe the way in which space around a sound connotes a relationship with that sound: heard and felt on a narrative level, as a complex set of interconnected attributes. Similarly, it is inadequate to develop abstract notions of sound as auditory objects, designed to append meaning to character or action in drama, without reference to the associative fields of dramatic semiotic enquiry. It is through an appreciation of dramatic agencies of interaction that we can substantiate the need for additional auditory aura.

The process for determining the placement of suitable audio aura in drama is described by the following steps

1. Why is the sound required?
2. What is the function of the sound?
3. What is the nature of the sound to be designed?
4. How does the sound function in context with the narrative action?
5. Where, in the dramatic structure, is the sound to be placed (both in terms of timeline and spatial dimension)?

These questions outline the basic design requirements for a taxonomical system, the results of which may be tested on *The MaxStage*.

The other essential criterion for the system taxonomy is that it has a direct relationship with the fundamental devices in the creation of theatre. Aristotle defined the poetics of mimesis (artistic representation) into six discrete functions: action, character, thought, language, melody (or pattern) and spectacle [enactment] (Laurel 1993). Although Aristotle's poetics of mimesis have been extensively re-classified and extended by theatre semioticians in the twentieth-century, the classification of all sign systems remain derivative of Aristotle's definition (Aston & Savona 1991, pp. 5-9 and 105-08).

The auditory corollary for each of these six functions must be identified so that sonic representations may be mapped onto any aspect of the theatrical design as required.

Drama Semiotics

The performance of a theatrical work, whatever its derivation, is by nature polysemic. It draws upon a set of sign systems, which, according to Aston and Savona (1991, p.99) do not operate in a linear mode but in a complex and simultaneously operating network unfolding in time and space. Correspondingly, auditory stimuli as part of a sound design may also unfold in a non-linear temporal fashion. In order to do so, auditory 'signs' need to be hierarchised in such a way to help 'fix' meaning (Aston & Savona, p.101).

Codification of signs according to Keir Elam (1980, p.50-2) is divided into *theatrical* codes and *dramatic* codes, where the word code is understood to mean an ensemble of *correlational* rules governing the formation of sign-relationships. Two of these codes are pertinent to audio signs: *proxemics* (codes of governing the use of space) and *kinesics* (codes governing movement). Hence, the manner in which auditory signs interact with the performer within a predetermined space, have a bearing on the overall meaning attributed to the conveyance of all interconnected and simultaneous signs. Equally, kinesics: the ability to analyse and codify the performer's gestures and movements has an impact on the way auditory signs are designed. Changes in gesture or in movement by the performer within an auditory space should reflect changes in the audio outcome of the sound design.

This correlation between movement and auditory device is difficult to articulate, but in modern theatre is a function of actor 'motivation' - a dramaturgical philosophy initially espoused by Henrik Ibsen and Konstantin Stanislavsky (Cima 1983, pp.5-22).

This type of actor delivery has a significant impact on sound design, when a designed auditory outcome is reliant upon a perceived change of actor-based gesture, or movement. Correspondingly, the auditory sign should be an expression of the emotion motivated by the actor 'motivation' as evidenced by the delineated actor movement or gesture.

From the percipient's perspective, proxemic and kinesic codes (as auditory signs) are perceived in one of the following three ways

1. Conventional appearance within the dramatic narrative timeline (sound effects).
2. To cultivate an illusion of time and space being altered, thereby appearing to break up a strict linear narrative (invariably non-diegetic sound).
3. To reinforce, or distance (distort), the validity of the apparent meaning as suggested by the immediate narrative text and actor movement.

Listening Perception Theory Reviewed

Denis Smalley (1996, p. 78 and 1986, pp. 61-93) and Michel Chion (1994, p.25) have separately re-evaluated Pierre Schaeffer's indexical modes of listening taking into account the advancements in computer technology since the 1960s. In summary, Schaeffer's four modes of listening are

1. Information-gathering
2. Passive reception
3. Appreciating and responding to attributes of sound
4. Acquiring abstraction (Chion refers to this as 'semantic listening')

Smalley has distilled Schaeffer's modes into three relationships. Partially established with reference to Ernest Schnactel's (1984) theories on subject-centred and object-centred perception (termed *autocentric* and *allo-*

centric perception respectively) the three relationships are defined as

1. The indicative relationship - corresponds to Schaeffer's mode 1
2. The reflexive relationship - corresponds to Schaeffer's mode 2
3. The interactive relationship - combines Schaeffer's modes 3 and 4 as a form of 'reduced listening'

Given that neither Schaeffer nor Smalley's listening modes can be separated or received as isolated events, it is important - from the perspective of sound design - not to design hierarchical classifications of auditory stimuli that subjugate causal listening, but rather to enrich causal listening with an ever-evolving development of sound as an influence on perception. Chion emphasizes this criterion in observing that

> Perception is not a purely individual phenomenon, since it partakes in a particular kind of objectivity, that of shared perceptions. And it is in this objectivity-born-of-intersubjectivity that reduced listening, as Schaeffer defined it, should be situated. (Chion 1994:29)

Surrogacy as Sound Gesture

> 'Surrogacy', by Smalley's own definition is the existence of new types of sounds which are more remote from physical, gestural origins than was previously possible…and can be recorded as they occurred in their original, cultural context and be incorporated in a musical work. (Smalley 1996:85)

He defines three types of surrogacy as

1. *First order surrogate* - an identifiable instrumental sound-source, irrespective of whether it normally has a musical connotation or not. It also includes synthesized sounds where they are representing a known instrument.
2. *Second order surrogate - a sound with created spectro-morphologies that has no apparent link to an identifiable sound. The human gestural activity does not give a precise explanation for what is heard.*
3. *Remote surrogacy - physical origins of the sound are masked to a point where neither gesture-type, nor source of the sound can be identified. The gesture field operates within the psychological domain relying on the listener to exercise considerable gestural imagination.*

The principles of field relationships and surrogacy become an important constituent in a composer's toolbox of sound design when extrapolations of these principles can be handled empirically in an audio-visual medium, such as *The MaxStage.*

It is a requisite of all such components within a hierarchical system model as proposed, therefore, that contextual applicability of components - whether used singularly or as layers - are determined by a consistent approach to the laws governing perception; and through perception, to arrive at an implied meaning.

In the same manner that Schaeffer's theory of 'reduced listening' can be broken down into a set of subsidiary constructs, a similar reductive process can be applied to the constituent parts in designing auditory stimuli. This process is regulated by an awareness of the semiotic considerations of the physical (actor gesture / movement, object as sign, etc.) combined with appropriate sonic figurations mediated through precisely determined and imposed layers of listening.

Listening in a Sounding Space

Smalley proposes that the connectedness of space and the relational structure of content within that space cannot be treated separately. He proposes that there are, conceptually at least, three categories of relational structure. These are

1. The relationship and behaviour among sounds within the composed musical space.
2. The movement between successive musical spaces, or the transformation of space in the work.
3. The interaction between the musical space and the listening space which for the listener is the sum of the three categories.

The composed space, as suggested by Smalley, possesses an acoustic topology in which, theoretically, the listening space is enclosed. He concludes that perceived musical space is always a *superimposed space*. If it can be shown, as Smalley contends,

> that the superimposition of spaces can create 'consonant' or 'dissonant' relationships between composed and listening environments changing indicative interpretations to an extent often not envisaged, or even considered, by the composer (Smalley 1996:91)

then the need to devise a sound design process that incorporates how superimposed space is perceived becomes increasingly relevant. The importance of this proposition is supported by David Worral's theory which states: when background ambience is complex and variegated, the distinction between ambient and non-ambient sounds is made clearer, as it is the ambience itself that creates the space in which other sounds are heard. (Worral 1998, p.97)

Superimposed spaces further indicate that layers of auditory stimuli are present. Katharine Norman (1996, p.10) argues that the composer [or sound-designer] can offer superimposed layers of sonic transformation while appearing to preserve the temporal duration of real-world scenario. This implies that a degree of planning is necessary to coordinate the use of variegated ambient sound (whether diegetic or non-diegetic) that acts on multiple temporal planes simultaneously.

The philosophies outlined above indicate that the creation of a virtual system model incorporating superimposed composition and listening environments is justifiable; not only for testing the applicability of de-

signed auditory stimuli, but to explore different relational fields of visual and sonic experience.

In themselves, acoustic topologies of listening in a sounding space do not configure a sound world that will encompass the intended emotional spectrum of response. Smalley contends that to achieve a true indicative contribution depends primarily on how well the composer circumvents the confinement of the superimposed, listening space. In other words, is the listener transported to a real or imagined environment beyond the immediate walls, or is the listener in the midst of musical activity within the space? Is it a case of 'reaching out' or of 'closing in'? (Smalley 1996:92)

Domains in a Sounding Space

Theo van Leeuwen (1994, p.4) has identified six semiotic 'domains' to describe the function of sound. They are

1. Sound perspective and social distance
2. Sound time and rhythm
3. The interaction of voices
4. Melody
5. Voice quality and timbre
6. Modality

Van Leeuwen believes that, the semiotics of sound concerns itself with describing what you can 'say' with *sound*, and how you can interpret the things other people 'say with sound'. He prefers the term 'meaning potential' rather than the term 'code', as often used by semioticians. As he sees it, this distinction is significant because the term 'code' implies literal transference of meaning; whereas, 'meaning potential' allows additional information to be received before a level of event perception is conveyed in its entirety.

The first two domains are particularly relevant to the issues of sound design in a listening space as previously described. Domains 3 - 6 are less relevant in this context as they relate to verbal utterances and pitch-specific music.

Domain 1. Sound Perspective

Sound, according to Leeuwen, creates a relation between the subject it represents, and the receiver(s) it addresses. He argues that this is managed in two ways. The first way is by perspective, whereby sound can be placed either in the foreground, middle ground or in the background. Murray Schafer (1986, p.157) codifies the environmental distribution of sound perspective by defining 'Figure' as the actual sound signal; the focus of interest, 'Ground' as the immediate setting or context, and 'Field' as the place in which the 'Figure' is represented; the soundscape.

The second way is by means of *social distance* (one of Smalley's archetypal fields). The degree of formality that distances the signifier and the receiver (defined as intimacy, informality or formality) reflects a tangible meaning potential. Furthermore, *social distance* is normally conditional upon sound *perspective*. One further related aspect of sound perspective is the way in which sound acts dynamically. Leeuwen (1994, p.18) sees this as sound that is able to move us *towards* or *away from* a certain position [and] can *change* our relation to what we hear. This is equivalent to Chion's axiom that sound presupposes movement in relationship to the associated visual stimuli.

Domain 2. Sound Time and Rhythm

Leeuwen (1994, p.7) views unmeasured time as a particularly apt signifier for 'eternity' - it literally negates time and goes 'on and on'. It is also used a sonic metaphor consistent with Ambrose Field's (2000) theory of 'sonic objectification'.

Leeuwen (1994, p.54) proposes that there are two main kinds of unmeasured time

1. *continuous time* - which lacks any form of phrasing and either does not vary in pitch at all, or wavers in pitch in slight or irregular ways, and
2. *fluctuating time* - which also lacks phrasing, but does shift between different pitches, at more or less regular intervals which are, however, too long to produce a clear sense of regular pulse or periodicity.

As such, aesthetic meaning can be invested in the action through the contradiction of auditory events portrayed in time measured, as compared to those perceived in unmeasured time.

Upon reflection, there are significant similarities expressed amongst theorists in the disciplines of theatre semiotics and the ontology of listening perception and reception. The requirement to maintain a causal connection between action and perceived outcome, whether the action is a result of visual or aural stimulation is consistently articulated. This does not, furthermore, preclude the use of sounds whose surrogacy levels involve a relationship remote to the perceived action (whether it be movement, gesture or narrative dialogue).

Intending meaning through additional auditory stimulus is not as important as eliciting an emotional response that directly supports the semiotic considerations of the drama. Without a fixed POV in the theatrical setting, cogency between visual, textual and sonic artefacts should be predicated upon a strict observance of the accepted classification of mimesis. In its most concise form, the poetics of Aristotle provides a manageable framework in the development of a taxonomic system for sound design as a dramaturgical device in its own right.

Evoking simultaneous layers of experience requires the development of an aural imagination whose elements can be codified, not only for evaluative purposes, but also as a means of systematically organizing the construction of a model in which that aural imagination may be mediated.

It is similarly apparent that the manner in which auditory stimuli behave in spatial terms, in relation to intended semiotic significations, also impacts on the system design model.

Reading the System Model

The system model (Fig. 3.) is, ostensibly, a map articulating not only the function of an overall conceptual sound design but also a process for the design sequence of auditory aura. It defines properties of auditory aura in terms of indexical and symbolic content.

The System Model (rotated 90º CCW as shown in Figure 3.) is read from top to bottom and from left to right. Laurel's iteration of Aristotle's six qualitative elements of structure in drama, and their contextual corollary in the design of *The MaxStage* environment, are the overarching constructs in this meta-representational diagram. Each element is the formal cause of all those below it; and each element, is the material cause of all those above it. This two-way organizational structure is consistent across all modules in the system design, irrespective of the direction from which they are viewed.

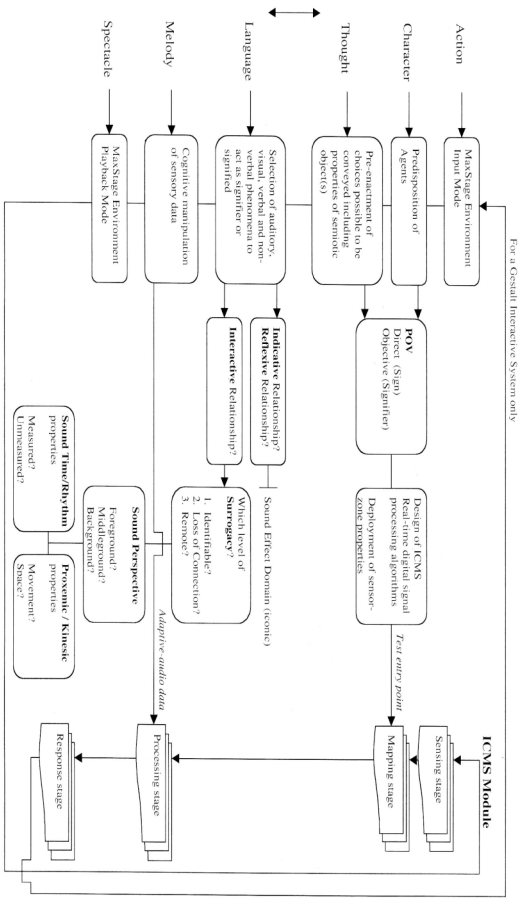

Figure 3. *Schematic of the System Model*

Implicit, also, within this two-way structure is the proposed theoretical machine learning protocols to be implemented in a physical interactive computer music system (ICMS). Entry points within the process chain for evaluating the efficacy of auditory objects as they are designed are clearly marked at both the mapping and processing stages.

The following steps should be taken to derive the constructs of any intended sound design object

1. Establish whether the POV for a character's predisposition toward action, or a pre-enactment of thought leading towards action (or non-action) is based on a direct sign or on an external signifying influence. This decision relates directly to the design of real-time signal processing algorithms in a physical system and the subsequent deployment of sensor-based technologies within a specified zone.

   For example, if the auditory object functions to explicate the meaning of a sign (in semiotic terms), the processing algorithm may be conceived as one that increases frequency, variations in timbral spectrality and intensity (volume) as a character moves increasingly closer to the imagined source of sound. Conversely, if the auditory object functions to indicate a signifying influence, the ICMS may be disposed toward changes in sonic duration, alterations in the envelope of sounds and qualification of the directionality of those sounds.
2. Establish the relational mode of listening intended toward the sign or signifier. If the relationship is indicative, where direct causality of the sound is intended, then the sound is likely to be either a sonic metaphor, or sonically iconic and treated as a specific spot effect. If an interactive listening relationship is intended, the design of the sound must be treated in the context of where it is positioned in relation to any simultaneous dialogue or other events. This is particularly important in the application of sonic simile, where a rhetorical code suggests an association between two contrasting contexts (i.e., different time-space corridors, or simultaneity of events occurring in different places.)

   Foregrounding the auditory object will result in a distinct shift in dramatic focus, contingent upon its level of surrogacy. The level of surrogacy is inversely proportional to the potential shift in dramatic focus that may occur. With an auditory object of extended duration, it may be necessary to change both its perspective placement in the overall field and level of surrogacy over time. Neither of these two factors should be considered independently. Each factor influences the other. Therefore, both surrogacy and sound perspective are, to an extent, interchangeable in position in the System Model diagram.

   Establish whether the sound involves movement, either by an object, or by an actor. If the sound has kinesic properties, consideration should be given to including a measured, or rhythmic component to the auditory object. Conversely, if the sound is proxemic in nature, then consideration should be given to eliminating any sense of measured time associated with the auditory object.

Conclusion

As a prototype, *The MaxStage* provides only a point of departure for future discussion and research on the dramaturgy of theatrical sound design. One of the main limitations of the prototype is that it does not take into account the inherent and very real problems of acoustic diffusion, reflection and delay in audio reproduction. Similarly, it does not satisfactorily model auditory aura in any multi-channel speaker format.

In spite of these limitations, *The MaxStage* does tackle two problems of sound scoring dramaturgy head on. These are:

1. It defines a response system (the system model) that is consistent with basic drama semiotic function

And

2. It is an environment into which possible auditory aura can be both iterated and mediated.

References

Aston, E. and Savona, G. 1991. *Theatre as Sign-System*. London: Routledge.

Calleson, J. Kajo, M. and Nilsen, K. 2004. "The Performance Animation Toolbox: developing tools and methods through artistic research" *New Visions in Performance*. eds. Gavin Carver and Colin Beardon. The Netherlands: Swets and Zeitlinger.

Chion, M. 1994. *Audio-Vision:Sound on Screen*. New York: Columbia University Press.

Cima, G.G. 1983. "Discovery Signs: The Emergence of the Critical Actor in Ibsen" *Theatre Journal*. 35. 5-22.

Elam, K. 1980. *The Semiotics of Theatre and Drama*. London: Methuen.

Esslin, M. 1987. *The Field of Drama*. London: Methuen.

Field. A. 2000. "Simulation and Reality: The New Sonic Objects" *Music, Electronic Media and Culture*. eds, Simon Emmerson, Ashgate, Aldershot.

Laurel, B. 1993, *Computers as Theatre*. reprint edn. MA: Addison-Wesley.

Norman, K. 1996. "Real-World Music as Composed Listening" *Contemporary Music Review*. 15, 1-2. 1-27.

Schaeffer, P. 1966. *Traité des objets musicaux*. Paris: Editions du Seuil.

Schafer, R.M. 1986. *The Thinking Ear*. Toronto: Arcana Editions.

Smalley, D. 1996. "The Listening Imagination" *Computer Music in Context*. 13, 2. 77-107.

Smalley, D. 1986, "Spectro-morphology and Structuring Processes" *The Language of Electroacoustic Music*. ed. Simon Emmerson. London: Macmillan.

Van Leeuwen, T. 1999. *Speech, Music, Sound*. London: Macmillan.