SPEECH: creating a virtual audio-visual artwork

Roger Alsop

University of Melbourne
ralsop@unimelb.edu.au
rogeralsop@gmail.com

ABSTRACT

This paper accompanies and discusses the audio-visual work SPEECH, a virtual and evanescent audio-visual work that considers the ephemeral and immediate nature of current media art through the equally ephemeral and immediate processes of speech.

Speaking has many inherent gestures, such as the physical gestures of propelling air and shaping the mouth, and the resulting temporal sonic gestures of amplitude, and frequency shifts. These gestures coalesce to reify and communicate mental gestures. These can be represented through notation systems that allow an abstracted representation of speaking that can be stored, reorganized, and published silently.

SPEECH interprets phonemic notation of the poem Ambit [8] and movement in its environment as sound and image. It may be presented in concert or installation settings, as a part of the ambience or environment, or for the individual listener via a located computer and headphones, this undefined approach to presentation is essential to the work as it reflects the un-located transience of current media.

The version presented here is a bespoke version SPEECH for this presentation. It is part of a trilogy of similar works that include MOTION, and VISION [1] that explore similar themes. These works require complex interactions between their audience and the works themselves, and are designed for prolonged and intimate interaction, in which the audience directly influences the creation and re-creation of the work.

1. INTRODUCTION

The digital world is essentially ephemeral and evanescent; it is dependent on a continually developing, mutating, inventing and reinventing palette, with consequent traditions, paradigms and processes being developed, discarded, uncovered and understood.

While much art is made through and with digital technologies, it is often made to create a permanent and immutable example or indicator of the author's intention at the time of conception or creation. This can be seen in the works that result from the various image and video editors and digital audio workstations; these computer-based systems replicate or are heavily influenced by analogue systems. Use of these systems can be seen in the work of Bill Viola, Gillian Wearing, and Gina Czarnecki.

There are works of less permanency but equally deterministic approach to outcomes, such as those of Ryoji Ikeda, and companies such as Troika Ranch, Company in Space and Chunky Move. These works are often built on collaborations in art forms other than music.

Audience interactive works, such as those of Arik Levy, Kurt Ralske, and Robin Fox, show a more reactive and ephemeral approach to art making and its outcome. Here the author takes the role of representing and/or influencing an environment and interactions within it, and interpreting them through the artistic outcome. It should be notes that, like many media oriented artist, those listed often work across these areas.

SPEECH is in fact a process in which its environment is immanent in the created artwork. It forms a Möbius-strip like link between the author (sender), audience (receiver), and the work itself, (intended message) and environment, or context, as seen in Figure 1. This creates an ephemeral outcome that interprets the interactions of its environment filtered through the structure outlined in Figure 2 below and that the audience and environment potentially adjust. All adjustable parameters are shown in Figures 2, 3 and 4.

This approach re-considers the often unidirectional, linear understandings of 'artwork (to) audience', 'author – artwork – audience' or 'sender – intention – receiver' (to paraphrase Peirce [6], Jakobson [2], Nattiez [5] and many others). Instead it consider the interaction as being simultaneously and constantly bidirectional in the authorship, creation and apprehension of the work, and within the environments distilled from the ideas of the authors above, outlined in Figure 1.
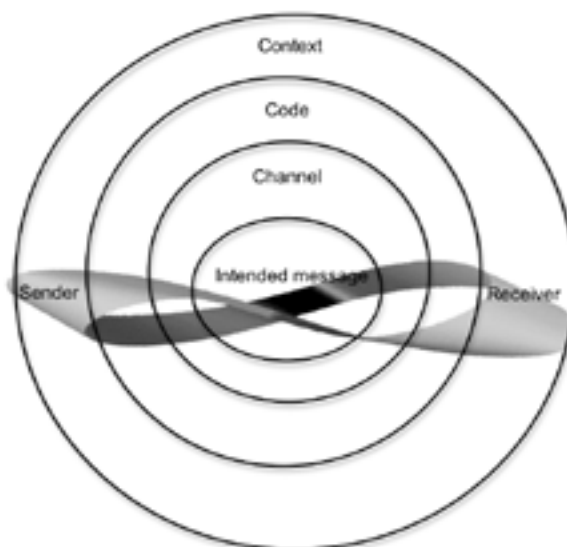

Figure 1: Environments in which communication occurs

The Möbius strip in Figure 1, which paradoxically provides only one continuous side, represents the idea that both sender and receiver are simultaneously engaged in the creation and apprehension of an art work within an environment.

These environments are explored in SPEECH, MOTION and VISION, in which the sender and receiver are simultaneously engaged in the making, perception, interpretation and understanding of the resulting work.

2. PROCESS

The version of SPEECH presented here is a does not require the depth of audience creative and interpretive interaction of the works listed above. Instead it presents the authors interpretation of the presentation environment.

Room acoustics, light variations, motion of the audience and so on are aspects of the environment that impact on and are reflected in the version of SPEECH. These aspects are responded to in the presentation of SPEECH[1].

2.1. Structuring SPEECH

SPEECH uses the phoneme-notated sonic core of a text, to form a core melody represented as a string of integers relating to frequency, amplitude, duration and the temporal relationships. This then forms the structural 'background' (in Schenker's sense of a 'foreground', 'middleground', and 'background'). The raw output of the string can be adjusted and then fed through an algorithmic process to create a number of co-incident strings, which in turn form a counterpoint to the initial string, resulting the heard foreground.

The tools for structuring the work have three sections, shown in figures 2, 3, and 4 below. The parameters of these tools can be adjusted as is appropriate to the presentation environment.

Figure 2: Creating the structure for SPEECH

Here the frequencies and temporal relationships of the structural 'background' are generated from the raw integer string of the FORM section.

The 'foreground' is generated from the 'background' through adjusting parameters in the SPACE section, As well creating the sonic outcome of this 'foreground' SPACE also generates the placement of images representing the local environment on screen by providing Cartesian x,y coordinates that specify the position and size of images placed on the screen.

Figure 3: Effecting the image input and output

Here images of the environment gathered via a camera are mediated both as they are being gathered and displayed. These mediations may result in the screen image not being easily recognisable as the presentation environment.

Figure 4: Creating the audio output

Here the sounds heard are synthesised. In this iteration additive sine wave synthesis in used, with each part having its own envelope and frequency range.

The choice of this sound source references early electroacoustic works, such as Stockhausen's *Elektronische STUDIE II* [9] and Young's Dream House [10] audio component.

The controls shown in Figures 1, 2 and 3 are designed to be initially arcane until explored; the controls in Figure 4 may be less so to those who have developed skills in sound synthesis techniques. This represents the process of exploration inherent in art making/perceiving, in which the various knowledge bases used by the author and perceiver are of varying depth.

2.2. Algorithmic and indeterminate approach

This is an inherently algorithmic/arithmetic approach to developing a structure ('background') and outcome ('foreground') of a work. The algorithms themselves are quite simple and reflect approaches taken in later serial/pitch class oriented composition [3, 4, 8]. By adding the indeterminate qualities of an unknown and continually changing environment the inherent intransigence and predictability of such a system is subverted.

This allows for an interpreted representation of the immediate environment to be mediated by the structure of the system, and alternatively the system to be mediated by the maker (sender) and audience (receiver) within the vagaries of the environment; this creates the Möbius like reciprocal interaction within a mediating and indeterminate environment.

3. OUTCOMES

It is inherently impossible to predict the exact outcome of the SPEECH, this is also the case with tape or score based electronic or acoustic instrument-based music, which are reliant on their presentation environment. As said earlier, this predictability is not a preferred outcome of SPEECH. However in this situation the outcome has greater predictability than was intended in the original versions of SPEECH, MUSIC, and VISION.

In these works the outcome is defined by the viewer both in the act of viewing, and through their ability to directly interpret the outcomes of the work by adjusting any and all of the parameters available. The unique parameters for SPEECH are shown in Figures 2, 3, and 4; MOTION and VISION have similar parameter sets, which cater to the direct needs of each work.

This iteration of SPEECH has closer resonances to traditional composition and installation processes, in which the author defines the most likely outcomes of their work through setting the available parameters of their work. Here these parameters are set in direct response to the environment in which it is being presented. Therefore this version of SPEECH is initially representing the author's response to that environment, through the setting of parameters, which then represent the presentation environment, which includes the audience's, response to the work.

4. REFERENCES

[1] Alsop, R. "*Mapping Gestures in the Creation of Intangible Art Objects*." PhD by Project, Royal Melbourne Institute of Technology, 2011.

[2] Jakobson, Roman. 1960. "*Closing Statment: Linguistics and Poetics*." In Style in Language edited by Thomas Sebeok, 350-77. New York: Wiley.

[3] Forte, A. 1973. *The Structure of Atonal Music*. Yale University Press, New Haven.

[4] Morris, R. 1987. *Composition with pitch-classes: a theory of compositional design*. Yale University Press, New Haven.

[5] Nattiez, Jean-Jaques. 1990. *Music and Discourse. Toward a Semiology of Music*. Translated by Carolyn Abbate. New Jersey: Princeton University Press.

[6] Peirce, Charles Sanders. 1931. *The Collected Papers of Charles Sanders Peirce*. Edited by Paul Weiss Charles Hartshone. Cambridge: arvard University Press.

[7] Schenker, H. 1977. *Harmony*. MIT Press, Cambridge.

[8] Skovron, Alex. *Sleeve Notes*. Sydney: Hale & Iremonger, 1992.

[9] Stockhausen, Karlheinz. "Studie II." In Karlheinz Stockhausen – Studie I / Studie II / Gesang Der Jünglinge / Zyklus Für Zwei Schlagzeugern. Europe: Doxy Music, 2011.

[10] Young, La Monte, "Dream House: Sound and Light Environment", MELA Foundation http://melafoundation.org/lmyresum.htm#works (accessed January 18 2011).

[11] Wourinen, C. 1979. *Simple Composition*. C.F. Peters, New York.