The Tertulia Project at Mirogoj Cemetery:
A constellation of voices, names and spectral transformations in space

Nicolás Varchausky
CCME, Universidad Nacional de Quilmes, Argentina
nvarchausky@unq.edu.ar
http://www.varchausky.com.ar

Scott Carver
Unaffiliated, USA
scott@artificia.org
http://www.scottcarver.org

Michael McCrea
DXARTS, University of Washington, USA
mccrea.michael@gmail.com
http://www.michael-mccrea.com

Proceedings of Korean Electro-Acoustic Music Society's 2011 Annual Conference (KEAMSAC2011)
Seoul, Korea, 22-23 October 2011

Abstract. This paper briefly comprises the research, ideas and techniques for creating the sound component of the 'Tertulia' project, an art installation composed in collaboration with visual artist Eduardo Molinari and produced by the 23rd Eurokaz Festival, realized at Mirogoj Cemetery in Zagreb during the early summer of 2009. Spectral Modeling techniques are discussed in relation to the micro- and macro-structure of the project, as well as to their poetic potency. The relations between sound and space within the context of 'Tertulia' are also revisited through the perspective of spectral transformations over time.

Tertulia is a large-scale sound and visual installation that took place at the Mirogoj Cemetery in Zagreb[1], Croatia. It involves real-time computer generated audio diffused through an array of 40 speakers and 14 points of visual interventions using everyday objects, videos and performance actions. It proposes a metaphorical conversation with our ancestors, a reflection on remembering and forgetting through the creation of a shared aesthetic experience within this intimate landscape. Visitors to the installation were allowed to walk through the Cemetery during the night, encountering a series of visual compositions along the way while being immersed in an ever-changing territory of sound.

Overview of the project

Can a cemetery be something else than scenery of pain?

The idea of a conversation with our ancestors is an old human desire that every culture approaches in a different way: it raises questions on the ephemeral and the eternal, on what's public and what's private, on the way we relate to past generations and on the interplay of individuals and their community [1].

We understand the space of the cemetery as a geography that articulates and organizes these possible relations, particularly those between individual and collective memory. As a kind of monumental space, it offers each member of a society an image of that membership, constituting a collective mirror. In terms of Henri Lefebvre, such sites do not have a fixed set of relations, but "a *horizon of meaning*: a specific or indefinite multiplicity of meanings, a shifting hierarchy in which now one, now another meaning comes momentarily to the fore, by means of – and for the sake of – a particular action" [2].

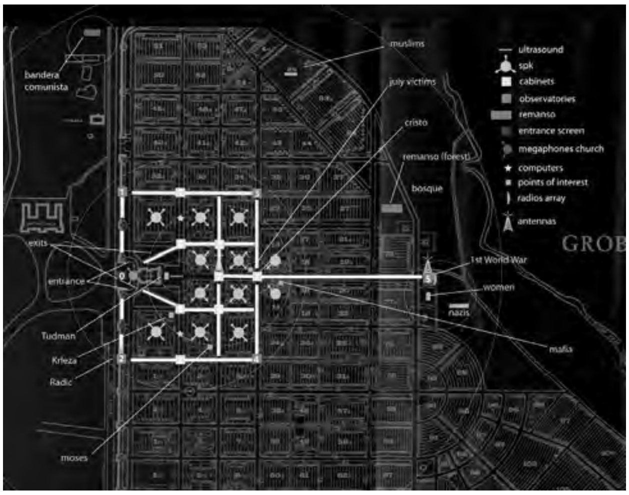

Figure 1. A map of the cemetery depicting spots of intervention, array of speakers, circulation paths and relevant historic figures.

History as a point of departure

Through the use of sounds, objects and images, we tried to craft an experience that would shift the way we relate normally to that space. To achieve this, we researched for two years on the history of Mirogoj and Croatia, with frequent trips to the site, to define its set of spatial *formants* – acoustic, geographic, architectonic, historic, symbolic, social and political. Finally we defined a smaller area in the cemetery (of approximately 200 by 200 meters) where to produce our intervention, and designed the possible paths that the audience would be able to take within it. By conceiving the act of walking as an artistic practice and a way to read and write on the territory, we proposed a nocturnal drift through the cemetery to generate a complex dialogue with the space and our place in it, that took into account both real and imaginary facts. For a cemetery is also a symbolic space suggesting a particular form of communication.

Remembering and Forgetting

Questions about what means to remember and what means to forget are entwined with reflections about the various ways in which we build historical narrations. When we remember, we create a new instance of the past, embedding it in a new temporality. In Agamben's words, "memory cannot give us back what was as such: that would be hell. Instead, memory restores possibility to the past" [3]. Thus, memory is a territory to be rediscovered and newly inhabited every time. Following these ideas, we've attempted to transform the space into a living organism, by creating a constantly changing sound flux, empowered by the names of the people present at the cemetery.

Figure 2. View of Cabinet 7. (Photo: Azul Blaseotto).

Walking Seeing Hearing

Articulated in terms of two main metaphors, *polyphonies* and *constellations*, the sonic and visual materials of the piece were designed to generate non-linear and multi-threaded types of relationships and organizations.

Visually, the piece created small size installations conceived as a variety of *Cabinets* and *Observatories*: discrete interventions along the cemetery's paths made out of objects, furniture, images, video and performance actions[2], that functioned as temporary rests throughout our drift. A multiplicity of intervened visual records from the history of Croatia and the World were classified and exhibited within these stations creating complexes of mythological, political and fictional material. While Cabinets recreated every day life situations through the use of desks, lamps, typewriters, ashtrays, chairs, pocket calculators, washing machines, radios and actual cabinets, Observatories proposed more open connections through the use of video animations[3] also generated using intervened visual records. Some of these animations were played on TV screens installed in custom made furniture, while others used ATM cash machines to run the video on their screens.

Sonically, the composition set up a *polyphonic labyrinth* of transformations: a computer system took Mirogoj's database of names – Zagreb's nominal mass – to generate and transform a multiplicity of sound *Fluxes* in real-time, articulated by *Synchronies*, or moments were a single musical gesture would unfold throughout the space. The sounds were diffused and spatialized through an array of 40 speakers that created a grid on the space with the ability to evenly distribute sound throughout the site.

The voice of everything

Inspired by the basic acoustic components of the human voice, the sounds in Tertulia were organized into a series of:

- *Resonances*: choir like sounds.
- *Turbulences*: abstract versions of the natural elements (earth, wind, water and fire).
- *Explosions*: the machines and technologies of war.

The human voice was used not only as a reference for the macro organization of timbres but also as a specific material to be exploited on its own throughout the piece. Research in the Croatian Radio National Archive (HNR)[4] helped us gathered a myriad of voices that were incorporated both raw and transformed.

A final sonic layer was introduced using a low power FM radio transmitter. It broadcast a computer-generated voice – heard through portable radios placed at every Cabinet – delivering a series of algorithmic transformations of a fragment of Borges' poem "El Golem":

"*Gradualmente se vio (como nosotros)
aprisionado en esta red sonora
de Antes, Después, Ayer, Mientras, Ahora,
Derecha, Izquierda, Yo, Tú, Aquellos, Otros*."[5]

The listening perspective

The sonic set up was designed to create a situation with multiple listening points, where *each listener becomes the center of the listening*. All the decisions were made to open the possibility of listening from any position in the space in a unique yet equally fulfilling way. Wherever you were at any given moment of the piece, you were experiencing a particular perspective of a whole that could only be experienced incompletely. The scale of the site was such that allowed the audience to wander within a large and dynamic mass of sound, listening both to their immediate sonic surroundings as well as to sounds occurring farther away in the cemetery. Its geography and architecture naturally contributed to this, since it is a flat open space.

The design of the paths on the territory loosely resembled the shape of a radio. One of these paths (its *antenna*) would take us far away from the main area illuminated by the speakers into Observatory 5, which had a privileged view of Mirogoj's memorial to German soldiers who died on Croatian soil during World War II. This gave us a unique chance to listen at a distance to the full array of 40 speakers, adding a very interesting sonic perspective to the overall sound component of the installation.

Figure 3. View of a speaker-cluster at the center of a polje.

Synchronies and Fluxes

The form in Tertulia emerged from the alternation of a series of *Fluxes* articulated by *Synchronies*. The Fluxes were real-time generated sounds that used the database of names to create electronic abstractions of wind, water, fire and earth. The Synchronies were moments when the 40 speakers became a single *voice*, creating a common and recognizable gesture throughout the whole space, some of which were pre-composed.

The global form

Approximately every 25 to 35 minutes, a *Synchrony* would take place. Because of the scale of the site, this was likely to be the least amount of time you would decide to walk around the space, thus ensuring the occurrence of at least one of these formal relationships. Each *Synchrony* was always introduced by the sounds of machines used in early electro-magnetic experiments by Nikola Tesla, recorded at Zagreb's Technical Museum. These sounds – refer to as *Raps* in the score – often lingered on, along with Synchronies, into a dissolved version of them or *Echo,* as a way of transitioning into the following section. As the piece unfolded over time, the human voice as a raw material was slowly revealed, becoming fully exposed by *Flux 4*, which was entirely made of recordings of voices.

What follows is a brief description of each Synchrony for a better understanding of the form and for documentation purposes:

1. Bells: this was the opening of the piece. Each speaker played a different kind of bell sound, either real or synthesized. During its 12 minutes of duration, bells would transition from having each one its own rhythm to being grouped by clusters. At that point, each cluster had its own rhythmic structure and their resonances would liger longer until becoming a huge mass of slowly decaying resonances.
2. Geysers: voices that started as a rumble from one speaker-cluster, emerged as short syllables and spilled outward into adjacent clusters as water droplets. The gestured tried to emulate geysers spitting out fragmented voices that emerged at irregular rhythms.
3. Hallo-Hallo: the word "hallo" (taken from the first broadcast done by Radio Zagreb), stopped and stretched at the "l" sound, played at slightly different rates and organized by cluster at irregular attack intervals.
4. Cannon: the sound of Zagreb's famous cannon that fires out daily at noon, recorded at 12 different positions in a circle of 300 meter radius[6]. Each recording was played from a different speaker cluster, first in unison, progressively becoming asynchronous in a 4-minute lapse.
5. Breaths: as in the Geyser Synchrony, breath sounds of different lengths would start at a cluster and expand at different rhythmic rate and spatial widths into adjacent clusters.
6. The Golem: a recorded fragment of Borges' poem in his own voice, played back in unison on all speakers.
7. Coda: the real sound of wind, water, earth and fire.

On the other hand, the sequence of Fluxes was defined as follows:

1. Wind.
2. Earth.
3. Water + Wind.
4. Voices: during this Flux, only voices were used, mostly raw or unedited, generating stronger connections local to each Cabinet or Observatory.

5. Wind + Earth + Water + Fire.

Finally, the space was divided into three layers defined by their particular treatment:

1. Global: using the complete array of speakers to create trajectories over the entire site.
2. Local: specific kinds of sounds (like the real voices of Flux 4) heard only in a particular Cabinet or Observatory, creating clear connections to the images and objects displayed in them.
3. Ethereal: the broadcast of the computer-generated voice.

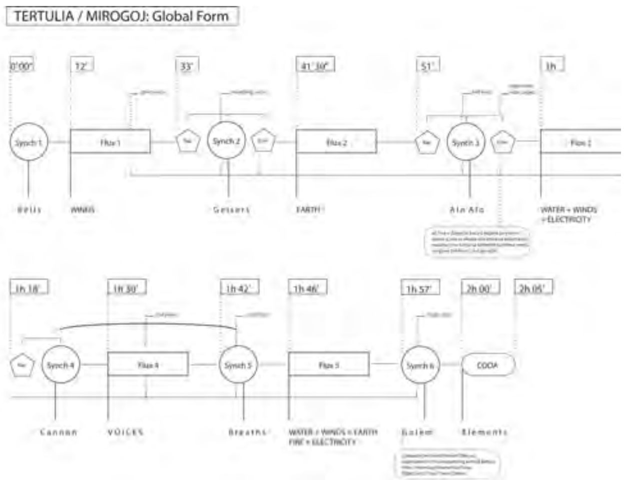

Figure 4. Schematic of Tertulia's global form.

The cycling structure of Synchronies and Fluxes – recognizable by the return of the Tesla machines and the global gesture – coexisted with more linearly evolving structures. As briefly described early on, the sound of the human voice would develop a long curve as a material, undergoing different treatments that would slowly reveal it as such. It would first appear as ghostly choirs by the end of Flux 1, become fragmented syllables during the Geysers Synchrony, appear recognizable as a fragmented voice repeating a single syllable over and over during the Hallo-Hallo Synchrony, to being openly revealed during Flux 4 (roughly 3/4 into the piece), and finally be dissolved into the air during the Breaths Synchrony. At the same time, each Flux would either introduce a new Element sound or combine a previously heard Element with a new one, progressively adding new sounds or combinations to each Flux. The superposition of structures at different rates, with different lengths, sounds and logics was a way to reconcile the different ways the audience could experience the piece in time: either staying at the installation from the beginning to the end or unpredictably entering or exiting it at any point in time, but hopefully still grasping its formal structures at a smaller scale.

The Database

For the generation of each Flux, the database of names from the Mirogoj Cemetery was used as a source of data. It holds about 400.000 names, and for each one it details date of birth, age at the time of death and the location in the cemetery (as a polje number). The system took each name and read it letter by letter, assigning each one a value according to its frequency of occurrence (as a percentage) in the Croatian language. These values were mapped in a variety of ways, most commonly to the frequency and amplitude parameters of resonant filters (within various predefined frequency ranges) that shaped the output of different noise generators and granular engines. Age data was typically mapped to durations; while location in the space was used to decide which series of names were picked at a given moment in time. For the creation of each Flux though, this information was organized differently. Sometimes names were ordered alphabetically; others sorted by age, creating textures that evolved from shorter to longer sounds.

The information extracted from the database would be used alternatively to define the different stages of sound production:

- Generation, mapping data to noise generators, granular synthesis, analysis and re-synthesis techniques.
- Transformation, through the use of resonant filters, granular processes, spectral modeling techniques.
- Articulation, helping to create gestures, dynamic envelopes and larger timbre and time structures.

Using the database was a way of involving everyone at Mirogoj in the creation of the piece, even those away from the sector of the cemetery where the installation actually took place. Since arguably names are what a cemetery actually preserves, it became also a strategy to make the piece *portable*, opening the possibility for it to happen at any other cemetery, as there's virtually not one of them without nominal records.

Sound equipment overview

The sounds were played through 40 speakers that received 40 discrete channels of audio ran through 4 Motu 828. Two computers MacBookPro running OS X 10.5.8 were synched together using their Ethernet ports, each one using two of the Motu's connected through Firewire. A custom computer framework coded in SuperCollider 3 was developed for handling, generating, playing back and spatializing the sound. All the manipulation, mixing and mastering of the sound was also done in the computers, eliminating the need for a mixing board or any extra hardware, thus considerably simplifying the set up. Signals were routed directly from the sound cards to the speakers. Unfortunately, it was very hard to get 40 speakers of the same brand, model, size and specifications, but most of the speakers used were EON15. The radio broadcast was done with a Ramsey FM stereo transmitter (model FM25B) and a dipole antenna with an approximate range of 300 feet. Most of the speakers were gathered in clusters of 4, placed at the center of each parcel or *polje*, pointing outward at 90º form each other towards paths' intersections (see Figures 1 and 3). These clusters were approximately 30 meters away from each other, keeping always the audience that walked through the paths not closer than 15 meters away from any given speaker, thus attenuating considerably the precedence effect (or rather making it relevant only at a larger scale).

The Virtual Speaker System

A variety of software tools were developed for the piece, mostly in SuperCollider, in order to achieve the different sonic transformations, spatial trajectories and structures, real-time processing and control.

One of these tools was a virtual speaker simulation. This tool was essential for composing spatially for such an unconventional site and array of speakers. It proved to be crucial in order to have an idea, as early as possible, of how things would sound before getting to the site. It allowed us to place speakers on a map of the cemetery and preview the sound at different locations. All composition could be done within this environment using only headphones or a modest ambisonic setup, giving us the chance to devote on-site time to final adjustments.

It had three core objects:

- VirtualSpeaker.
- VirtualSpeakerArray.
- VirtualSpeakerArrayUI.

The Virtual Speaker

This object reads audio from a 1- channel audio Bus, listener position/rotation from a 3-channel control Bus, and adjusts the audio based on the position of the listener.

Adjustments include:

- Delay.
- Amplitude scaling.
- Filtering based on distance (ISO 9613-1:1993).
- Filtering based on speaker angle (i.e. dry if speaker is facing listener, more filtered if speaker is facing at a right angle or directly away - rough equalization curves based on gathered speaker impulse responses).

Audio is then turned into a panned 4-channel B-Format signal, based on the relative angle to the listener. B-Format signals from all VirtualSpeaker's are summed to a common B-Format Bus, which can then be converted to any speaker setup for working. During actual performance, VirtualSpeakers skip all this and wrote their audio directly to hardware outputs, and the real speakers (so no distance attenuation, panning, obviously - this happens in the world!).

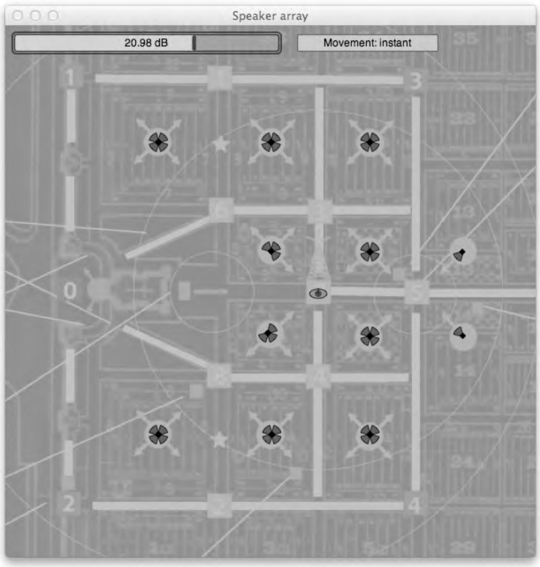

Figure 5. The virtual speaker simulation GUI used for composing before getting to the real site.

The VirtualSpeakerArray

Inside it, we have:

- All VirtualSpeaker's.
- Listener Bus (a 4-channel B-Format audio bus).
- Listener position Bus (a 3-channel control bus: x, y and rotation of the listener).

Among its features we have:

- Amplitude adjustments across all speakers.
- Control of the position and movement of the listener through a function that takes a series of x-y coordinates and walks the listener along the path, to audition what a person would hear.
- It can run speaker tests (to check levels, connections, etc).
- Speakers can be accessed by numeric index.

- Speaker groups (clusters) could be accessed by name (i.e. *north-middle*).
- VirtualSpeakerArray could be serialized to disk, so every time the project was reloaded the speaker positions would be recalled.

From here, a series of simple player synths would send the B-Format listener signal to headphones, 4 channels, 6 channels, etc. for auditioning.

The VirtualSpeakerArrayUI

This was developed so we could display all speaker positions, with rotation, at scale on a map of Mirogoj. Individual speakers could be clicked on to play test tones and noise. Also, the listener could be positioned by clicking on the map, and could be moved by either using keys or dragging paths for it to follow.

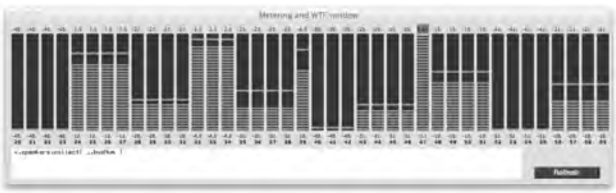

Figure 6. Virtual level meter displaying all 40 channels.

The Spatialization tools

During the initial stages of the project, a series of poetic spatial gestures were designed, conforming a catalog of 10 to 12 separate and very unique acoustic-spatial environments. Some of these gestures attempted to create sonic metaphors like *geysers*, *arcs*, *whorls*, ghostly *echoes,* and *tides*. The challenge was to create a spatialization system that could smoothly transition between all of these types of spatialization, without making compromises to the aesthetics of any of them.

We looked at several general spatialization models and evaluated their pros and cons considering our setup:

- Ambisonics would only work with adjacent speakers, and would only sound good at the midpoint between them.
- Vector Based Amplitude Panning could use all speakers, but it's a very precise panning model, so there's little adjustment to be done if things don't sound right with a particular spatialization model.

We finally settled on a hybrid approach, breaking up the space into two main areas:

- A perimeter semi-circle of speakers, for more *distant* sounds (panned using adjacent pairs + spread)
- An interior rectangle (panned using adjacent pairs + spread factor)

In addition, we introduced the *Geyser panning* for speaker clusters (sounds emerging first from one cluster, but then spreading in various directions) plus several other panning methods that would be pre-recorded and played back from individual speakers.

The Behavior System

Any given sound could be placed in the perimeter space, or the interior rectangle. Its position could be controlled with one or two location parameters (one in the case of the perimeter, two for the rectangle), and a spread parameter (0 being only adjacent pairs, 1 distributed to all speakers). Location of a sound was specified by a "behavior". When a behavior of a sound was changed, it would smoothly cross-fade from one behavior to the next, over a specified period of time. Behaviors could be as simple as a fixed location and spread in one of the two speaker arrays. However, any arbitrary function or UGen could be provided for any of the parameters, and saved as a behavior. These complex, composite behaviors could also be cross-faded. What follows a three behavior examples, as we would define them in the code.

Slow drift, slowly moving around perimeter speakers:

- array: \perimeter
- position: {SinOsc.ar(0.1, 0.5, 0.5)}
- spread: 0

Surrounded, spread between all perimeter speakers and moving rapidly:

- array: \perimeter
- position: {LFNoise2.kr(10).range(0, 1)}
- spread: 0.7

Centered and waiting, at a fixed location, slightly jittering (remember center takes an x and y location):

- array: \rectangle
- position:{[05 + LFNoise2.kr(10, 0.01), 0.75 + LFNoise2.kr(10, 0.01)]}
- spread: 0

Once these behaviors were defined, a score could be made:

- 0:00 synth.setBehavior('slow drift') – Define the type of behavior.
- 0:30 synth.setBehavior('surrounded', 60) – Fade to surrounded behavior over 60 seconds.
- 1:30 synth.setBehavior('centered and waiting', 1) – Quickly move external sound mass to a point inside.

In this way, very complex behaviors could be build and perfected on their own, and then scored in a simple and aesthetically transparent way.

The Event System

Each section of the piece was being developed separately. We ended up having code from three different people with three different backgrounds, coding styles and experience levels, written over a period of over a year. We needed a way to queue, test, and play each section without requiring the adoption of a heavy, overarching system that would require extensive modification of things already written. Instead of an *über-player*, or an *über-scoring* system, we opted for the most lightweight system imaginable: we created *MiroEnvironment*. This environment was a subclass of SC's Environment, which is a private namespace for keeping track of variables. That way, each section of the score could stand-alone.

Figure 7. The Environment player used to load, play and stop each section of the piece.

In the MiroEnvironment, any independent section of the score had to define four basic functions:

1. ~initialize: run ONCE, at startup, to ready the score. This included loading data and audio files, creating persistent objects, etc. We were very worried about resource management (memory issues, etc) over the course of the performance, so how much is loaded here was flexible.
2. ~prepare: run JUST BEFORE a section of the score starts playing. This should load audio buffers, load complex synths if needed, queue things so that playback can start as quickly as possible.
3. ~play: starts the playback. Hopefully, all the heavy lifting was done in ~prepare, so this is quick and lightweight, and happens instantly.
4. ~free: stop playback and free resources allocated in ~prepare, but NOT in ~initialize.

After initializing, the rest of the functions could be called over and over during testing, without any additional resources being built up.

This Environment also had an ability to signal that they were nearly done, so the ~prepare function of the NEXT section could be called ahead of time. Thanks to this development, it was trivial to adapt any sort of coding style to this paradigm. Plus, all the over-arching Player object had to do was call, in turn, each MiroEnvironment's ~prepare and ~play functions, then wait for them to end and call the next.

The Voice Instrument

The Voice instrument was designed to allow any sonic material (the *source*) to activate a process in which the material resonates with the harmonic structure of an altogether different sound (the *target*), while fluctuating with the internal dynamics of a singing voice. In the case of Tertulia at Mirogoj, we gathered a series of samples from Slavic folk *a capella* music to use as targets, while the source sounds were the Elements used throughout the composition (electronic sounds of wind, water, fire, earth). These vocal recordings were edited into short fragments (typically between 1 to 5 seconds), isolating moments of stable pitch that used different voiced sounds, intervals, timbres, rhythmic patterns and brief melismatic gestures that are recurrent in that particular kind of music. This gave us a kind of metonymic matrix of Slavic music vocal sounds; condensed and concentrated information of its structural elements.

The AtsSnapshot

Though the design of this instrument was informed by the artistic direction of the project, it was created to be flexible enough to employ the constituent processes in a number of ways. These processes include the extraction of partial frequencies from recorded material from a specified window of time, building a filter bank based on the extracted data, and being able to transform between these spectral “snapshots” intelligently and with a variety of parametric controls.

The construction of a Voice begins with gathering spectral snapshots of the target material to be reconstructed with a filter bank. To do this we made use of Juan Pampin’s Analysis-Transformation-Synthesis

(ATS) [4]. This is a powerful analysis method with the advantages of combining sinusoidal plus critical band analysis, accounting for perceptual information such as spectral and temporal masking. The user performs an ATS analysis of the target sound files and selects moments or regions of time within that sound source from which to take a "snapshot" of frequency and amplitude information of the constituent partials. The user chooses a threshold for the lowest acceptable partial amplitude as well as a window of time over which to average the partial information. The data may also be transposed by scale or shifted in frequency. Various methods are provided to allow a user to test, view and manipulate the results of the analysis to modify the analysis based on the goals or constraints of the application. For example, the number of partials used for re-synthesis was a limiting factor in the real-time application of the Voice for Tertulia, so snapshots needed to be modified to return a reasonable number of partials. From this analysis an AtsSnapShot is generated, with the amplitudes of gathered partials normalized, which is the blueprint for the filter bank constructed by the Voice instrument.

The Voice instrument uses the AtsSnapShot to construct a filter bank through which any sound source may be routed, expressing the spectral balance of the target sound articulated by the temporal dynamics of the source material. Because the target material selected for the project were chants and vocal songs, we chose to impose upon the resultant synthesized sound the dynamics of a singing voice. This included time-variant vibrato placed on the resonant frequency of the partial filter. Which allowed the effect of masking the vibrato when the filter's bandwidth was sufficiently wide, which in turn made the emergence of the target "voice" more pronounced once the filter closed in on the partial's center frequency.

One of the compositional goals for the instrument included the ability to transform between various internal states of the Voice. Some of these transformations were accomplished by transposing the target AtsSnapShot, and by opening or closing the bandwidth of the filter bank dynamically. Because the source material included the core Elements of the piece, the source material was sufficiently broadband to activate all of the filters (thereby activating the full spectrum of the target material) when desired. When the filter bandwidth was fully open, the Elements could pass through transparently, and the target voice could emerge to a variable degree by simply narrowing the filters over a specified duration and onset curve.

In addition to modulating the internal state of a Voice through its filter parameters, the spectral state of a Voice could also be transformed. This was accomplished by transitioning from one AtsSnapShot to another. For instance, the syllables "tza" and "zee" were isolated from a song, AtsSnaphots were taken of them and a timbre was created that would evolve from the open vowel sound "tza" to the close vowel sound "zee". Playing with the filter's bandwidth, we could make one of this vowel sounds emerge from the Element's broadband timbre, modulate into another vowel sound and smoothly return into the turbulence where it originally came from. A bubble sort algorithm was used to find common frequencies (within a specified deviation from the current state) from one snapshot to another. The partial parameters were then shifted to the new target state over a specified duration and rate of change. Partials that had no match to the new target Voice were faded out and new partials not yet present were faded in, all drawing from the common source routed to the Voice.

Figure 8. An ATM used to display video at Observatory 4. (Photo: Azul Blaseotto).

The Ethereal Voice

A computer-generated voice was broadcast through a low power FM transmitter placed at the center of the main area. It was heard through portable radios placed at each Cabinet, materialized out of the ether into these intimate spaces, creating a local sound stratum. It would also produce an interesting contrast to the multiplicity of human voices heard through the main speaker system, building an omnipresent character and incarnating an omniscient *acousmatic being*. This disembodied voice uttered in Croatian endless variations of Borges' selected fragment from The Golem, gradually shifting from a more exclusive way of addressing the audience to a more inclusive one. As you kept walking through the space and arriving at each distant Cabinet, you would encounter this voice once and again, attempting to find yet a new twist to the poem over and over as if engaged in a mechanical mantra. Though it remained unclear whose voice this was, its mediated and intimate presence would become unavoidable and eerie.

HMM speech synthesis

The voice was created using a trainable speech synthesis system designed by Ivo Ipšić, Sanda Martinčić-Ipšić and Miran Pobar at the Department of Informatics at the University of Rijeka and based on hidden Markov models (HMM). Inputs to the system were speech utterances and their phone level transcriptions. It used a method of context-dependent acoustic models and Croatian phonetic rules for speech synthesis. [5] Although the quality of generated speech is "vocoded" buzzy speech, it can be understood. [6] The mediation of the FM transmission and the little speakers on the radios would somewhat mask these issues or rather help us listen to them as acoustic features of this ubiquitous voice.

For generating the speech utterances that this system would use as inputs, renowned Croatian actor Zlatko Burić Kićo recorded over 900 *logatomes*: artificial words of one or more syllables, which obey all the phonotactic rules of a language but have no meaning. Examples of English logatomes would be the nonsense words snarp or bluck. [7] Some of the Croatian ones were, for i.e., "apta", "hatata" or "teba". For the recording, the speaker inserts the logatome in a simple sentence like "Reci *xxxx* opet" ("Say *xxxx* again"), substituting *xxxx* for the actual logatome. The speaker shouldn't have to stress the logatome, but the last word ('opet') and try to keep a monotone voice. Also, when a logatome contains a combination of phones that is hard to pronounce, say 'cb' in 'acbi' the speaker shouldn't try to pronounce the individual phones, but read the whole word naturally and without putting too much effort. In this way the phonetic changes that occur in natural speech are captured.

Conclusions

We have presented the main ideas and software development for the creation of the sound component of Tertulia at Mirogoj Cemetery. Because of the scale of the project and space limitations for this paper, we have only focused on the sonic component of the piece, leaving an analysis of its visual elements and an integrated discussion of all of the project's strata for another occasion. An ad-hoc spatialization system for a non-traditional multichannel array was proposed along with esthetic strategies for occupying artistically an unconventional site. Speech synthesis and spectral modeling techniques used in the generation of sounds were described within the conceptual framework of the piece.

Acknowledgments. We wish to thank DXARTS (The Centre for the Digital Arts and Experimental Media – University of Washington, Seattle, WA, USA), Proyecto de Investigación I+D S.A.E.T.A.S. (Sistemas Algorítmicos de Espacio y Tiempo en el Arte Sonoro – Universidad Nacional de Quilmes, Buenos Aires, Argentina), Eduardo Molinari and Azul Blaseotto, Saša Božić and Simon Bogojević-Narath, Srđan Nogić and Tomislav Pletenec, Diana Meheik and Vedrana Klepica, Branko Brezovec, Gordana Vnuk and Kornelija Čović at Eurokaz Festival. And to all the members of the team that participated in the creation of Tertulia.

References

[1] Barley, N. (1995) *Bailando sobre la tumba. Encuentros con la muerte*. Trad. Federico Corriente. 2000th ed. Barcelona, España: Editorial Anagrama.

[2] Lefebvre, H. (2007) *The Production of Space*. Trans. Donald Nicholson-Smith. Paperback ed. USA: Blackwell Publishing.

[3] Agamben, G. (2002) *Difference and Repetition: on Guy Debord's Films*, in McDonough, T., et al. *Guy Debord and the Situationist International. Texts and Documents*. Ed. Tom McDonough. Cambridge, Massachusetts: MIT Press.

[4] Pampin, J. (1999). *ATS: a Lisp Environment for Spectral Modeling. Proceedings of the 1999 International Computer Music Conference.* Beijing: Computer Music Association.

[5] [6] Martinčić-Ipšić, S. & Ipšić, I. (2006) *Croatian HMM-based Speech Synthesis*. *Journal of Computing and Information Technology - CIT* 14, 2006, 4, 307–313.

[7] Retrieved from Wikipedia. http://en.wikipedia.org/wiki/Logatome

---

[1] A first iteration of the project took place at Recoleta Cemetery, Buenos Aires, and was part of the Proyecto Cruce (Crossroads Project), produced by the 5th Festival Internacional de Buenos Aires, in 2005.

[2] Saša Božić coordinated these performance actions and was in charge of the Dramaturgy of the project.

[3] These animations were done in collaboration with a team of artists directed by filmmaker Simon Bogojević-Narath at Bonobostudio in Zagreb.

[4] Diana Meheik and Vedrana Klepica were of great assistance during this research.

[5] *"Gradually, he found himself (like us)/imprisoned in this sonorous web/of Before, After, Yesterday, Meanwhile, Now,/Right, Left, I, You, Them, Others."* (Translated by Nicolás Varchausky).

[6] These recordings were courtesy of Srđan Nogić.

Further information:

http://archivocaminante.blogspot.com/

http://www.dxarts.washington.edu/

http://unq.edu.ar/

http://www.eurokaz.hr/

[Abstract in Korean | 국문 요약]

미로고즈 묘지의 테르툴리아 프로젝트: 음성과 이름, 공간에서의 스펙트랄 변환 배열

니콜라스 바르차우스키, 스콧 카버, 마이클 맥크리

이 글은 테르툴리아Tertulia 연구에서 소리를 만들어내는 아이디어와 기술 방법에 대한 논의로 대략 구성되는데, 테르툴리아 프로젝트는 시각 예술가 에두아르도 몰리나리Eduardo Molinari와 협동으로, 2009년 초여름 자그레브Zagreb의 미로고즈Mirogoj 묘지에서 열린 23번째 유로카즈 페스티발Eurokaz Festival에서 발표된 예술 설치 작품이다. 스펙트랄 모델링 방법이 이 프로젝트의 크고 작은 구조 체계와 이들의 시적 가치와 함께 논의된다. 테르툴리아 프로젝트 맥락 속에서의 소리와 공간의 관계는, 시간의 흐름에 따라 변형하는 스펙트럼의 관점을 통해 다시 논의되기도 한다.