

Research and Development Report

MA Scheme Media and Arts

London South Bank University

SPEAKERSON.NET

Sonic-inspired Language of the World Wide Web

Supervisor: Katrina Sluis

Student Name: Mariola Gajewska

Student number: 2514899

Introduction

My research project 'SpeakersOn.net' deals with the issue of the innovative use of audio in web interfaces. I believe that Web audio has great potential to enrich our aesthetic experience whilst browsing Internet sites, although to date it remains a largely unexplored area. I have thus aimed to investigate how sound is currently used on the Internet and have started to develop new ways of using sound in Web Interfaces, thus making Internet browsing a more interesting and memorable experience for the user.

In this report I will cover how my project has progressed since the Proposal Stage, how creating my prototypes has made me think differently about the project and how identifying my target audience has focused the development of my ideas. I explain how my project concept has gradually evolved as I have deepened my understanding of the three main research areas, starting from the notion of web interface, through understanding sound and how we listen, to materializing the concept of a virtual web-based soundscape.

Project Progression

Since the completion of my proposal in spring 2006, I have continued to research and develop my expertise in those areas which I consider relevant to my project, focusing specifically on web-related technologies, theoretical background and existing practices relating to the subject of Web Interface and Sound.

At the beginning of the Research and Development Unit I set up a web presence for SpeakersOn.Net and installed a blog using WordPress technology. My blog at SpeakersOn.Net is intended as a space for documenting project progression and exhibiting the sound prototypes. Additionally, it serves as an area where I can save useful links that I come across and also acts as a forum for discussion with visitors to the site.

My initial set of aims - to create new sound experiences on the Web by isolating the audio from the visual - has been extended to the idea of creating a virtual soundscape which could be navigable in an Internet browser.

To further my technical skills I have been attending Action Script workshops where I have improved my Flash and Action Script 2.0 skills. I have also taken part in Sonic Media workshops in order to acquire new skills of professional sound editing and field recording. My newly acquired knowledge about sound

editing and recording facilitated the process of making and then editing field recordings taken in Battersea Park.

Target Audience

Deciding on the target audience for the project has been the hardest task, as well as the crucial one since it has further developed and narrowed down the focus of the project. Researching the subject of interface and web audio instantly led me to the issue of Web Accessibility. Although I am very aware of Web Accessibility rules in my professional work, I have found it hard to empathise with the subject matter enough to want to research it further. I have realised that my MA is about my artistic advancement, and about creating something new and innovative which would inspire and challenge not only a visually impaired or blind user but an ordinary Internet user as well.

I have thus settled on my project character, which is to be exploratory, aesthetic, art-based and created for experimentation's sake. I have set out to create a non-linear sound space on the Web which can be navigated through a Web interface. I wish to discover what happens if the Internet user is provided with limited visual feedback and instead, navigation through the site is accomplished with audio cues.

This decision made me narrow my research and focus on the main three areas of my project – *interface*, *sound*, and *soundscape*. In the following chapters I explain how studying each of these areas inspired me and made me think differently about my final project.

Buzzing Fly Experiment

In the Research and Development Phase I have built two sound prototypes in order to explore what is possible with Flash 8 and Action Script 2.0 technology and audio.

The first prototype I have created is *Buzzing Fly*, which illustrates the impressive audio capabilities of Flash technology. It demonstrates dynamic panning of an object's sound in relation to the object's position on the screen. The fly is 'attached' to the cursor and when moved around the screen, the sound effect mirrors the visual by panning the sound from the left to the right channel. It also shows how the volume of the sound can be controlled along with the fly's apparent distance from the user which is varied by UP and DOWN arrow keys or the mouse wheel; the bigger (or "closer") the fly is the louder the sound becomes; the smaller it is the softer (or "more distant") it sounds. This experiment can be viewed at <http://www.speakerson.net/experiments/buzzingfly.swf>.

Creating the *Buzzing Fly* prototype has made me realise the new possibilities that can be achieved with Flash when using Action Script. It has also reassured me that it is the right tool for my project.

Flash As a Tool for Creating an Innovative Web Interface

Lev Manovich (2001) defines interface as a language for communication between two entities. In the case of the Web interface the two entities are a Web browser and the website content. Since the Web interface replaced the cinema, library, gallery or book interface, its role has shifted from being just a technology to being the “filter for all culture”. Manovich (2001) also argues that: “Content and interface merge into one entity, and can no longer be taken apart”. The interface and its form will influence the way the user interacts with its content.

Certainly, if this is true, a web designer could easily break away from a standard web interface and create new rules of web browsing. This could take a form of a web interface that remains intuitive but is more enjoyable to use and incorporates the unexplored area of sound. Thus, my project aims to make the user temporarily forget about usability by rewriting the rules of web interface navigation.

The most suitable software, which gives a web designer desired flexibility and powerful tools to build innovative interfaces is Macromedia Flash. Nowadays, it often serves as a platform for net artists to create a web-based aesthetic experience which can be enjoyed by using a standard Internet connection:

“Flash has held out the promise of full multimedia environments and an aesthetic experience of navigation for the bandwidth impoverished regions of netspace.” (Munster 2003)

Not only is Macromedia Flash used for multimedia but also for creating the whole Web interface, often challenging Internet users' browsing and navigating habits. An experimental website such as www.dontclick.it can serve as an example here. The site can be fully navigated without a single mouse click, thus challenging arguably the most common of website design rules. Moreover, with advancing computer technology, faster Internet connection speeds and better computer performance, combined with improved web tools, such as new versions of Flash 8 and Action Script 2.0/3.0, there is much room for exploration and experimentation for web professionals.

The latest version of Flash comes with reasonably good quality codec for compressing digital sound. In words of Dr Anna Munster:

“This has allowed artists to produce sophisticated sonic dimensions for their web interfaces without major sacrifices to download speed. As a scripting application, Flash also allows for sound to be triggered nonsynchronously from the image, often as a result of user-activated events such as a mouse click or rollover.” (Munster 2003)

Despite some limitations concerning the quality of compressed sound, it seems we are now ready to open our ears, or enter what McLuhan (1989) calls an “acoustic space” when browsing the Web. Although web audio is still used very sparingly, there are several examples of effective and interesting way of using sound in the web interface. An experimental website soulbath.com, created by London web design group “hi, Res!”, presents an intriguing layered sonic environment. The site uses multiple short event sounds at different areas of the interface which are triggered both synchronously and asynchronously to the onscreen visuals.

Net artists are not the only people who are exploring sound in web interface design. [Opodo.com](http://opodo.com), a renowned travel site, has recently used the idea of a soundscape effectively; newsletter subscribers were challenged to identify the city where the streamed sound came from to encourage them to enter a competition to win a free flight. Additionally, Opodo’s promotions section on the website (<http://promos.opodo.co.uk/city-breaks>) lists ten popular city-breaks;

each city has a dedicated page featuring a short description and photography accompanied by sounds of the city.

Invisible Maze Concept and Related Academic Practice

My initial concept for the final project involved navigating an invisible maze, using arrow keys or the mouse to physically move around the maze, and controlled volume and panning of the sound to aid localisation. However I felt that it lacked a purpose or a context. I have gradually fine-tuned and defined my idea after researching various related academic and art projects which have changed my way of thinking and inspired me.

While investigating the concept of a maze, I have come across a project created by renowned sound artists from Greyworld (www.greyworld.org). A fascinating sound installation has been produced in the historic maze at Hampton Court Palace, called *Trace*, which is designed to illustrate the intrigue and mystery of the maze's rich heritage. When travelling through the maze, visitors trigger a random series of sounds – a fragment of music, a whisper of an illicit conversation, the rustle of fine silks etc. – which enhance their journey and help them to work out a strategy to find the way through the maze. There are over a thousand different sounds; consequently the visitor encounters different series of audio stimuli every time that they traverse the maze.

The *Trace* installation has made me question how I could recreate a similar experience, i.e. taking a journey through a maze enriched by sound, but in the virtual world, using a Web browser and a pair of speakers.

Thereafter, I encountered another inspiring academic practice conducted by Canadian college lecturer Andra McCartney, who, as a part of research for her PhD dissertation, documented a soundwalk on-line, which took place in August 1997 with Hildegard Westerkamp through the landmark Queen Elizabeth Park in Vancouver, British Columbia. The visitor can experience the virtual sonic environment of the park by clicking on different navigable areas of the park's map. Andra McCartney used standard HTML web technology and a web browser to present her research, which can be examined in more detail at: <http://cec.concordia.ca/econtact/soundwalk/indexandra.html>. The idea of a soundwalk and using a park as a space appealed to me, and I have realised that this idea can be made more powerful by using Flash instead of HTML technology.

Another interesting academic research I have come across and found very relevant to my project was conducted by Prescilla Chueng from University of Huddersfield. Her study circles around the role of discrimination and expectation when designing auditory interfaces (Chueng 2002a, 2002b, 2002c) as well as using minimal audio content to create a sense of presence in a digital virtual space (Chueng 2002d). Priscilla Chueng (2002c) suggests that it is feasible to

create a “minimal auditory design that holds minimal information (but enough to provide the sense of presence and characteristics of a virtual space)”, instead of designing an advanced realistic auditory virtual space.

According to Prescilla’s study, when we operate an interface, typically we will easily match the sound with its representation. This is mostly possible due to human's ability to expect to hear a sound in a particular location and to discriminate the sounds which do not belong to a particular context. (Chueng 2002a). Whereas this is true, in my research I would like to see what happens if one ceases to rely on visual feedback and thus the element of expectation is removed.

Understanding Sound and Why Use It?

My theoretical research has led me to the theory of sound itself and the notion of ecological approach in sound design, introduced by Gaver (1997). The study has helped me to deepen my understanding of how we hear and provided my project with a suitable theoretical background.

In his studies, Gaver (1997) believes that vision and hearing are the two primary senses that help us perceive the world and obtain information about it without direct contact. Hearing is often regarded as secondary to vision; it may seem that

we use it only to tell our eyes where to look. Because vision appears to be richer, has more detail and is more precise than sound, it has been ruling the design world and graphical user interfaces (Gaver 1997).

However, Gaver (1997) argues that hearing is very different to vision and provides us with different kind of information:

“What we see about an object is not what we hear about it. What we see is patterned variations of light frequencies, usually reflected from the surface of objects in the environment around us. ... But what we hear is patterns of moving air, often emitting from objects as they vibrate due to some event.”

Vision hence provides us with information about the objects size, surface and shape. Sound, on the other hand, can carry information about the object's hollowness and consistency, timing and dynamics, interaction and interior. Sound differs from vision fundamentally when it comes to its temporal nature – visual objects tend to be more stable whereas sound objects are more transitory. In the words of William Gaver: “sound exists in time and over space, vision exists in space and over time” (1989).

To sum up, sound can be a powerful method of conveying information and a crucial complement to vision. Using sound in the Web interface can have many advantages, starting from providing information about distant objects and events, through feedback about user's action, to creating mood and peripheral awareness (Gaver 1997). In the natural world as well as man-made environments such as film or multimedia, vision and sound interact smoothly. The area that is still lacking such smooth interaction is using sound to complement the web interface.

Everyday Listening and Web Interface

Gaver (1997) introduces a relatively new approach to sound and hearing, i.e. *everyday listening* combined with ecological approach to auditory perception. He suggests that we do not hear sound waves or acoustic signals, we hear events instead. The majority of sounds can be characterised in terms of their source, and their attributes. Although the area of everyday listening is still new and unexplored, Gaver argues it has already been useful in interface design:

“Most importantly ... we can build auditory interfaces using this framework. Instead of mapping information to sounds, we can map information to events” (Gaver 1997)

A human being is able to hear the *localisation* of the sound source, both in terms of its *distance* and *direction*, and sometimes even its environment. As we are now able to recreate this environment and the sound location artificially, we have powerful tools to hand when designing auditory interfaces and virtual reality systems - "In sum, then, sound provides information about an *interaction of materials at a location in an environment*" (Gaver 1993)

Web audio is often considered to be unnecessary noise and distraction, but this view mostly relates to sound being used as background music without a real purpose, or to isolated foreground sounds used intermittently in the interface. How then can we recreate the everyday listening elements on a web page? How can we apply the ecological approach to web browsing? Lastly, how can we avoid web audio becoming an unwanted distracting noise?

As a solution I propose creating a *virtual soundscape* which contains multiple sounds coming from different sources, as they exist in a 3D space. A soundscape encompasses all the factors mentioned above and is a great space for exploration as it contains navigational clues in the form of sounds, with controllable source, location and direction that the user can interact with.

Soundscape as a Sound Canvas of the Web

The term *soundscape* was first introduced by a Canadian composer R. Murray Schafer in the late 1960's. His soundscapes were constructed as musical composition and referred to as "a masterpiece of nature" or "what is perceptible as an aesthetic unit in sound milieu" (Augoyard and Torgue 2005).

In order to apply the concept of soundscape to the Web interface, a more detailed analysis or a map is required. Ferrington (1996) argues that a real life soundscape consists of three layers of acoustical information: foreground sounds, contextual sounds supporting the foreground sounds, and ambient sounds. Foreground sounds, isolated and random in nature, are the ones that attract our attention (for instance an auditory feedback on a click of a button). They are also the sounds typically used in the interface design.

Visual bias must be put aside by designers if they are to create effective interface using the soundscape concept. Macalalay *et al* believe that:

"...by taking a holistic view of the design and by considering the narrative presented by the soundscape, impressive interfaces to applications can be developed" (1998).

They also argue that, from the sound editing point of view, sounds should not be separated and cut from one another when compiling a soundscape. Instead each sound should smoothly fade in and out.

This being true for other media such as film and multimedia, we must remember that Web audio is still limited by the speed of the Internet connection. A Flash website would become very bandwidth-heavy if it was to incorporate long audio files. For a Web soundscape to work, I put forward a concoction of Prescilla Chueng's minimal ecological sound design (2002d) and sounds functioning on three acoustical layers of a soundscape proposed by Ferrington - foreground sounds, contextual sounds, and ambient sounds (1996).

A Sonic Map of Battersea Park Prototype

The second prototype that I have created can be viewed at:

www.speakerson.net/experiments/sonicmap.swf. Its purpose is to explore the idea of a *virtual soundscape* as used on the Web which is based on the open space of Battersea Park in South West London. Battersea Park gives me a lot of potential for experimentation since it has boundaries, can be navigated, and contains an abundance of interesting and varied sounds. It is also a space which carries a cultural significance and plays an integral role in relaxation of Londoners.

The prototype entitled *A Sonic Map of Battersea Park* is a simplified version of my final project, demonstrating movement of the “Walker” and his interaction with sound hotspots. Left and right arrow keys turn the Walker around, and backward and forward arrows move him ahead. A few sonic areas are specified (by use of x and y coordinates on the screen), which are activated when the Walker approaches them. The sounds which are attached to the hotspots fade in and out depending on the Walker's distance. Additionally, the stereo-pan of the sound is continually adjusted depending on the direction that the Walker is facing relative to the location of the hotspot.

In this prototype, I decided to abandon the graphical representation of the map, and instead use simple shapes representing the visited areas, which appear only when the Walker is within a specified distance. I am thus challenging the user's reliance on visual feedback aided by expectation. Also, all the sounds used in this prototype are real sounds which I have recorded in Battersea Park using professional recording equipment. I have then edited the sounds, cutting them into short audio loops, before compressing them and attaching them to the relevant sound hotspots.

My prototype aims to investigate the quantity of sounds that are required to create a sense of presence in a sonic virtual world. I would also like to examine users' response to removing visual feedback and replacing the element of

expectation with surprise. Will it aggravate the user or intrigue him or her to explore the interface further?

What's next?

In the production stage I will build on the second prototype *A Sonic Map of Battersea Park*, while following Gaver's ecological approach (1997), which has become an essential reference for sound design in spatial-oriented environments. I am planning to do more field recordings in the park and having selected the best ones, I will incorporate them into the interface. I intend to use the three layers of sounds distinguished by Ferrington (1996): background, contextual, and foreground sounds. Background sounds will be associated with a hotspot by defining its coordinates on the screen, e.g. the sound of players on a football pitch. Contextual sounds will be moving sounds, such as a flying insect or a passing runner, which will appear occasionally with random positions and paths and will be audible over the background sounds. Foreground sounds will include the sound of the Walker's footsteps, which will vary depending on the surface he is on.

My project could also serve as a vehicle for visitors from all over the globe to relocate themselves and experience a London park sonically, an important element of the capital. Thus, local will become global.

Once the project is completed I intend to present it on speakerson.net website as well as submit it to some net art sites, such as rhizome.org or fibreculture.org. At the same time I hope speakerson.net will become a space for discussion about alternative ways of using web audio.

Bibliography

1. AUGOYARD, J. and TORGUE, H. (eds.) (2005) *Sonic Experience: A Guide to Everyday Sounds*, McGill-Queen's University Press
2. BHANGAL, S. (2004) *Flash Hacks: 100 Industrial-Strength Tips & Tools*, O'Reilly Media
3. CHUENG, P. (2002a) Designing sound canvas: The role of expectation and discrimination, in *extended proceedings of CHI'012 Conf. on Human Factors in Computing Systems*. 848-849. ACM Press
4. CHUENG, P. (2002b) Designing Auditory Spaces: The role of expectation, in *extended proceedings in extended proceedings of British HCI 2003*, London.
5. CHUENG, P. and MARSDEN, P. (2002c) "Designing Auditory Spaces to Support the Sense of Place: The Role of Expectation." Position paper for *The Role of Place in On-line Communities Workshop, CSCW2002*, New Orleans

6. CHUENG, P. (2002d) Minimal ecological sound design: the issues of presence in virtual environments, in *extended proceedings of British HCI2002*, London
7. FERRINGTON, G. (1993). 'Audio design: Creating multi-sensory images for the mind', World Forum for Acoustic Ecology (WFAE), [Online] Available from:
<http://interact.uoregon.edu/MediaLit/WFAEResearch/sndesign> (accessed 25 October 2006)
8. FERRINGTON, G. (1994) Keep Your Ear-Lids Open. *Journal of Visual Literacy*
9. GAVER, W. (1997) Auditory Interfaces. In M. Helander, T Landauer, & P. Prabhu (Eds.), *Handbook of human-computer interaction* (2nd ed., pp. 1003-1042). Amsterdam: Elsevier
10. GAVER, W. (1993). How do we hear in the world? Explorations of ecological acoustics. *Ecological Psychology*, 5(4): 285 - 313
11. GAVER, W. (1989) The Sonic Finder: An Interface that Uses Auditory Icons. *In Human-Computer Interaction*, 4 (1) p. 67-94

12. GAVER W. (1993) What In The World Do We Hear? An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, 5(1), 1-29
13. GROSVENOR, S. (2005) *The Flash Anthology: Cool Effects & Practical ActionScript*, SitePoint Pty Ltd
14. MANOVICH, L. (2001) *The Language of New Media*, MIT Press, 2001
15. MACAULAY, C., BENYON D., CRERAR, A. (1998) Voices in the Forrest: Sounds, Soundscapes and Interface Design, Dahlbäck, N. (ed.) *Exploring Navigation*. SICS Technical Report T98:01, SICS, Stockholm
16. MCLUHAN, L. and POWERS B.R. (1989) *The Global Village: Transformations in World Life and Media in the 21st Century* New York: Oxford University Press
17. MUNSTER, A. (2003), *Compression and the Intensification of Visual Information in Flash Aesthetics*, Melbourne DAC, [Online] Available from: <http://hypertext.rmit.edu.au/dac/papers/Munster.pdf> (accessed 5 April 2006)

18. TURNER, P., MCGREGOR, I., TURNER, S. and CARROL, F. (2003)
"Evaluating Soundscapes as a Means of Creating a Sense of Place",
International Conference on Auditory Display, Boston, MA

