FOUNDATIONS OF INTERACTIVE SOUND DESIGN FOR TRADITIONAL STORYTELLING

Lonce Wyse
Communications and New Media Department
National University of Singapore
lonce.wyse@nus.edu.sg

Srikumar Subramainan
Communications and New Media Department
National University of Singapore
srikumarks@gmail.com

ABSTRACT

The traditional practice of oral storytelling has particular characteristics that make it amenable to extending with interactive electroacoustic sound. Recent developments in mobile device and sound generation technologies are also lend themselves to the particular practices of the traditional art form. This paper establishes a context for interactive sound design in a domain that has been little explored in order to create an agenda for future research. The goal is to identify the opportunities and constraints for sound unique to live storytelling, and criteria for evaluating interaction designs. The storytelling domain addressed includes not only particular instances of telling, but also the variability of stories between tellings and tellers, as well as the mechanisms by which stories are passed between tellers.

1. INTRODUCTION

There is a tremendous variety of storytelling traditions across every culture that extend back to (and frequently even explain) the very beginning of time. For the purposes of this paper, we consider "traditional storytelling" to be characterized by a single teller addressing an audience using speech, physical gestures, possibly props, and non-speech sounds. Sounds might be generated vocally, with physical gesture, or with a sound-making device which could be anything from a staff to a percussion or musical instrument. Stories are in general fluid and involve improvisational components, and as part of an oral tradition, are fluently passed from one person to another. The objective of the research agenda outlined here is to provide wider access to sound for storytellers than the vocal and instrumental techniques they currently use provide, and to do so without disrupting their long-established modes of practice. This "design research" formulation implies some stringent constraints which will be studied by working with storytellers themselves.

Hand-held, sensor-rich, touch screen devices have proven their effectiveness for sonic interaction. A hand-held device does not constrain the kinds of physical gestures storytellers typically make any more than other hand-held props or instruments, and sound generation and transformation capabilities are no longer critically limited by computational power. For these reasons, this platform seems an ideal candidate to provide the extension of the sound palette for storytelling. However, there remain three main several outstanding challenges:

1) How to design an interface that works for the possibly very many sounds a story requires,
    a. The physical interaction must be appropriate for all the sounds,
    b. It cannot require physical gestures that detract from the story,
    c. It must be easily learnable,
    d. It cannot require cognitive bandwidth that detracts from the storytelling.
2) How to represent the sounds to accommodate the flexibility and improvisation that characterizes the stories they will accompany,
3) How to represent the sounds in such a way that they can be customized by different tellers,
4) How to represent the story with sounds in such a way that the whole package can be passed as fluently between tellers as they are in the classical oral tradition.

The term "rig" will be used for collection and organization of sounds and interfaces that are prepared in advance for a given story.

2. BACKGROUND

In most oral storytelling traditions, stories are generally not told verbatim from memory, nor are they entirely made up on the spot as they are told. They are instead typically passed from teller to teller, over generations, and sometimes over thousands of years (e.g. The Ramayana) undergoing constant mutation. For example, one can find dozens of variations of Little Red Riding Hood (many of which have been written down) – told for different purposes, and frozen at different times and locations in the course of their mutations. Red Riding Hood tales differ in every detail from the age group of the targeted audience, to who gets eaten in the end. Even stories that are read aloud are delivered differently each time they are presented. Thus one fundamental characteristic of storytelling is its combination of prepared material and structures with variability and improvisation.

A teller is armed with a variety of prepared elements to support an otherwise improvised tale. In addition to plot elements, characters, voices, specific memorized lines, and gestural and theatrical elements, there might also be a variety of props, media, puppets, and musical instruments ready to support a tale. The guitar is one of the most common instruments for self-accompaniment in western traditions due to low physical demands on anything but the hands. The pipa has been similarly used as a storytelling accompaniment in China (Werle-Burger, 1999).

Figure 1. Storytelling using (presumably) vocal sounds and gesture (Photo: McLuhan, M. & Fiore, Q. (1967) The Medium is the Massage.)

Figure 2. Fiddler Don Woodcock with Adirondack storyteller Bill Smith who combines elements of American and French-Canadian folklore. (Photo: Old Songs Festival press images.)

Electronic and computer instruments have also been used by, for example, Laurie Anderson in her "Stories from the Nerve Bible", although for a variety of reasons including the complexity of the instruments, these stories are not retold by others. The aim of this research is to design a computer-based platform providing storytellers with the ability to create auditory scenes, sonic elements, and vocal transformations that are controllable in real time in order to support the telling, retelling, and sharing of stories.

3. ASPECTS OF SOUND AND STORY

3.1. Auditory Scenes

Sound has many unique qualities that make it an important part of storytelling (no matter what the form). Sound can create atmosphere in ways that neither words nor images can, evoking a strong sense of space and place (Schafer, 1994). For example, the sound of fog horns, seagulls, water lapping, and an irregular pattern of metallic "dings" of ropes hitting masts can locate us at the waterfront. We can be positioned in historical time by, for example, the kind of ring a telephone makes, or the particular quality of engine sounds.

Sounds can inform us about things we can't see, and nondiegetic sound (Chion, 1990) can set mood. Environmental scenes are difficult for a story teller to create by traditional means, though stylized versions are sometimes created with the help of instruments.

Specific sound effects are also frequently called upon by the story teller. They might be made vocally (the "thumb-scrrrch" of the hobbled man who creeps up on unsuspecting campers) or with the help of the hands and feet. Instruments are also sometimes used for extraneous sound effects, such as a guitar for door knocking. In the Tholu Bommalattam theater of Tamil Nadu, the sometimes sole performer of puppets, music, and sound effects rigs up a pair of wooden planks that can be clapped with his foot when needed (Verle-Burger, 1997). Enriching the palette of sounds available to the story teller for creating scenes and specific effects with technology would be a natural extension to the vocabulary and traditional practice of the story teller.

A vast expansion of the sound palette could be supplied by synthetic sound. A further advantage that synthetic sound would have over physical instruments is that they could be designed to be flexible for real-time manipulation by the performer. A storm, for example, may need to wax and wane through rage calm as the story unfolds. A range of techniques from physical modelling (Cook, 1997; Smith, 1992) acoustic modelling (Arfib , D., 1979; Horner, Beauchamp, & Haken, 1993; Serra, 1997; Wyse, 2004), and sample based techniques can be used to provide flexible, interactive, and when appropriate, realistic sounds under the real-time control of a story teller.

3.2. Text and sound

One of the most important elements of storytelling is the creation of voices for different characters. The sonic rig would offer the storyteller vast possibilities for extending their voice for characters as well as for other sounds.

Of course, manipulating the voice has long been an important part of synthetic and electroacoustic sound arts. Karlheinz Stockhausen in *Gesange der Junglinge* (1955-56), for example, explored the relationships between voice and synthetic sound with the use of electronic and tape techniques for manipulation. Live electronic processing of voice has also been used in text sound and sound poetry works such as the Swedish artists Bengt Johnson and Lars-Gunnar Boden, although much of this work involved utterances and manipulations to create forms understood more as music than as linguistically meaningful text (Wendt, 1985).

Storytelling presents a stringent demand for live processing because the control must be intimately coordinated on short time scales with the vocal work of the storyteller. Microphones and amplification already commonly support storytellers, so that a well-designed hand-held interface is all that is needed for the storyteller to have intimate control of how their voice sounds to the audience.

3.3. Interaction

One of the most challenging aspects of integrating a sound rig into the story telling performance is the interaction design. Constraints and challenges come from several aspects of the specific performance culture.

The roots of storytelling are grounded in personal, family, and community contexts that do not always bear a resemblance to the concert auditorium. Storytelling has also always been accessible and practiced by communities of every economic stratum. It would be self-defeating to design a system necessitating outrageously expensive electronic and computer equipment or a controlled reverberation hall with 64 Genelec speakers in order to work.

One approach to addressing this issue is to use inexpensive unobtrusive wearable sensors (Kapur et al., 2005) which have become so widely available. Another is to consider mobile devices such as phones and tablets as a platform. In large parts of the world, mobile phones are far more pervasive and affordable than personal computers. New devices are rich in sensors, and have processors with ample power for complex sound synthesis (Essl & Müller, 2010). An effective strategy for amplifying sound from mobile devices is an open question, though one interesting wearable solution has been demonstrated (Oh et al. 2010).

Another stringent challenge for the development of the interface is the nature of story's delivery. The storyteller's eyes are busy with expression and communication, so an interface demanding visual attention is out of the question. It is clear that an ideal instrument would burden the already fully engaged body as little as possible and be controllable within and between the gestures that a storyteller otherwise makes. Wearables and hand-held mobile phones again seem well suited to the task.

Sounds need to be as flexible as necessary for the specific demands of the story, but not more so in order to minimize the cognitive load of dynamic control strategies for a heterogeneous collection of sounds. Mapping strategies (Hunt, et al. 2003; Wanderley & Depalle, 2004; Pendharkar, Gurevich, & Wyse, 2006) could play a key role for this design goal.

The spatial nature of sound is also important for creating a sense of place and realism. Sounds must exist in the proper environment (e.g. room size, resonance, and reverberation qualities), and perhaps most suggestive of the need for real-time control, sounds need to move depending on the storytellers needs.

3.4. Rigs as social media

Stories in oral traditions often spread from teller to teller and come down through the ages. They are passed between tellers with no dependency upon physical support mechanisms - if one hears a story, one has it. In this way too, stories can simultaneously be passed from one to many people since there is no single and authentic copy. This property of stories is mirrored in nature of digital media as well (much to the chagrin of commercial producers and distributors).

Another important implication of the oral mode of transfer itself is that it facilitates, indeed almost insures, variability. This kind of mutability also bears a striking kinship to what we now call social media (Cheliotis & Yew, 2009). CCMixter ("ccMixter," n.d.), for example, is website through which people share, transform, and mix audio samples to create new works that are again shared. Because of the interactive and computer coded foundations of the proposed sound rigs, if the oral traditions of sharing and mutability are to be respected then something like a social media network is called for.

The mutability requirements for interactive media are more complicated than for the fixed media currency of CCMixter. While many people with lay skill are capable of mixing and editing audio and video files, developing code is more specialized. Graphical coding and scripting may help address this issue, and Max/MSP patch sharing or the game "modding" cultures (Scacchi, 2004) are potentially useful models.

3.5. Prototype

Although the goal of this paper is to articulate a research agenda, the specific platform choices that must be made have important implications for the feasibility of meeting the design criteria. For that reason, several experimental prototypes have been developed. One promising version utilizes Javascript, HTML5, and the Web Audio API (a draft proposal to the W3C with implementations available for WebKit based browsers). A server running on a low-latency local network configures controllers and synthesizers via plain URLs and performs real-time message routing between them. Sensors and touch screen activity on mobile devices are accessed through the Javascript API provided by mobile browsers. Synthesis capability is provided by the WebKit browser running on computers on the same network. The advantages of this system with respect to the design criteria identified in this paper are that

1) It is accessible - there is no need to download any application specific code to the mobile device, and it just works on major mobile platforms.
2) Web-based authoring tools for this platform can be developed to minimize the programming skills needed to customize sound in ways consistent with the mutability of stories, and
3) The sound rig lives in the "cloud" which means that the platform for performing, customizing, and most importantly, for sharing the stories is identical.

4. CONCLUSION

We have identified the key opportunities and constraints for the design of an interactive sound system to accompany storytellers, as well to support a social networking-like infrastructure modelled on oral traditions for passing evolving versions of code-based media between practitioners.

In this paper we have attempted to identify the characteristics that define the genre, and interpret them as design constraints on a system the offers new sonic capabilities to storytelling. A prototype was developed on a system that embodies many of these first-iteration design goals, notably a platform for transmitting stories in an easy and flexible way between people.

The goal of the research is not to revolutionize storytelling with a "disruptive" technology, but rather to integrate the wondrous capabilities of interactive audio into the existing fabric of the storyteller's art. The motivation is that by providing these capabilities to the storyteller, the age-old genre might appeal to a new generation, and develop in new ways without sacrificing the live and improvisational qualities not possessed by its more popular "big media" rivals such as television, film. This approach is based in the belief that preserving a tradition does not mean freezing it, but rather enabling it adapt to a new context without its essence getting lost or obliterated.

5. ACKNOWLEDGEMENTS

This work was supported by Singapore MOE grant FY2011-FRC3-003, "Folk Media: Interactive sonic rigs for traditional storytelling".

6. REFERENCES

[1] Arfib , D. (1979). Digital synthesis of complex spectra by means of multiplication of non-linear distorted sine waves. *Journal of the Audio Engineering Society*, *27*(10).

[2] ccMixter. (n.d.). Retrieved February 15, 2012, from http://ccmixter.org/

[3] Cheliotis, G., & Yew, J. (2009). An analysis of the social structure of remix culture. In *Proceedings of the fourth international conference on Communities and technologies* (pp. 165-174). University Park, PA, USA: ACM.

[4] Chion, M. (1990). *Audio-vision; sound on screen*. Columbia University Press.

[5] Cook, P. (1997). Physically Inspired Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, *21*, 38-49.

[6] Essl, G., & Müller, A. (2010). Designing Mobile Musical Instruments and Environments with urMus. In *New Instruments for Musical Expression*. Presented at the NIME2010, Sydney, Australia.

[7] Horner, A., Beauchamp, J., & Haken, L. (1993). Methods for Multiple Wavetable Synthesis of Musical Instrument Tones. *Journal of the Audio Engineering Society*, *41*(5), 336-356.

[8] Hunt, A., Wanderley, M. M., & Paradis, M. (2003). The importance of parameter mapping in electronic instrument design. *Journal of New Music Research*, *32*(4), 429–440.

[9] Kapur, A., Yang, E., Tindale, A., & Driessen, P. (2005). Wearable sensors for real-time musical signal processing. In *PACRIM. 2005 IEEE Pacific Rim Conference on Communications, Computers and signal Processing, 2005.* (pp. 424-427). Victoria, BC, Canada.

[10] Marshall, M., Malloch, J., & Wanderley, M. (2009). Gesture control of sound spatialization for live musical performance. In *Gesture-Based Human-Computer Interaction and Simulation*, Lecture Notes in Computer Science (pp. 227–238).

[11] Oh, J., Herrera, J., Bryan, N. J., Dahl, L., & Wang, G. (2010). Evolving The Mobile Phone Orchestra. In *Proceedings of the International Conference on New Instruments for Musical Expression*.

[12] Pendharkar, C., Gurevich, M., & Wyse, L. (2006). Parameterized morphing as a mapping technique for sound synthesis. In *Proc. of the Int. Conf. on Digital Audio Effects (DAFx-06), Montreal, Quebec, Canada* (pp. 45–48).

[13] Scacchi, W. (2004). Free and open source development practices in the game community. *IEEE Software*, *21*(1), 59-66.

[14] Schafer, R. M. (1994). *Our Sonic Environment and the Soundscape: The Tuning of the World*. Rochester, Vermont: Destiny Books.

[15] Serra, X. (1997). Musical Sound Modeling with Sinusoids plus Noise. In Roads, C., Picialli, G., & De Poli, G. (Eds.), *Musical Signal Processing*. Swets & Zeitlinger.

[16] Smith, J. O. (1992). Physical Modeling Using Digital Waveguides. *Computer Music Journal*, *16*(4), 74-91.

[17] Wanderley, M., & Depalle, P. (2004). Gestural Control of Sound Synthesis. *Proceedings of the IEEE*, *92*(4), 632-644. doi:10.1109/JPROC.2004.825882

[18] Wendt, L. (1985). Sound Poetry: I. History of Electro-Acoustic Approaches II. Connections to Advanced Electronic Technologies. *Leonardo*, *18*(1), 11-23.

[19] Werle-Burger, H. (1999). Interactions of the media Storytelling, Puppet Opera, Human Opera and Film. In V. Bordahl (Ed.), *The eternal Storyteller; Oral Literature in Modern China*. Curzon Press.

[20] Wyse, L. (2004). Hierarchies, Tied Parameters, Sound Models and Music (pp. 144-147). Presented at the International Computer Music Conference.