Archetypal Sound Effect Icons for Improved Multimedia Accessibility

Karen Collins[1] and Peter Taillon[2]

[1]Canadian Centre of Arts and Technology, University of Waterloo, Waterloo, ON
[2]School of Computer Science, Carleton University, Ottawa, ON
Canada
voice: [+001](519)888-4567 x38326
e-mail: Karen Collins: collinsk@uwaterloo.ca, Peter Taillon: ptaillon@scs.carleton.ca

Abstract

The paper introduces "SoundSign", a prototype symbolic representation of sound effects for multimedia, using an innovative icon and compass that indicates direction, sound cue and proximity. In this way, users who have disabled the sound in a multimedia device, who are hearing impaired or who otherwise are unable to hear a particular sound may be able to still have important information relayed to them. A description of SoundSign and the results of a usability test are presented. Potential uses outside the video game market for which the technology can be adapted (such as in PDAs, film/television, educational software, industrial or military uses).

Keywords: User interfaces, accessibility, usability, multimedia, games, audio

1 Introduction

Accessibility in user interfaces has been the subject of much work in interaction design, but research and technology in video game accessibility has lagged behind, in part due to the view of academics that games are not as "important" as other software applications in terms of accessibility, and in part because the game industry has not seen the financial value in incorporating accessible options into games [MOA+08]. As others have pointed out [Za06], however, in addition to serious games and educational games which serve a direct function (learning, rehabilitation, and so on), games play an important role in social play, and the creation of accessible games can ensure social inclusivity and reduce feelings of isolation or difference amongst those with impairments. Game accessibility research has focused primarily on creating games that are accessible for the blind, cognitive or mobility-disabled. [SBH+03], and [FG04] for instance discuss games for the blind (a popular area of research), and more recently this has advanced to games that attempt to be accessible to all [OMA08]. There has been very little research into making games more accessible to the hearing impaired, although other areas of hearing impaired accessibility permeate the human computer interaction literature [LFU07], [MFM05], [ABA+08].

When it comes to hearing-impaired gaming, the focus of accessibility has been on dialogue at the expense of music and sound effects. At best, in most media music is represented in captions by a title and/or musical note to indicate that music is playing, but in many games music is used to indicate rhetorical, emotional and affective information to the player. One particularly ignored accessibility issue with video games is sound effects. Sound effects add information to the narrative, foreshadow events and actions, hint at off-screen action, serve as symbols or leitmotifs for characters and locations, create a sense of time and place, provide proximity and location cues of adversaries, and most importantly, can warn the player to take a particular action [Co08]. In many cases, playing without the sound effects in a game leads the player into peril. For example, in many stealth games (a sub-genre of first-person shooters), sound effects inform the players of nearby enemies, and also cue enemies to the player's whereabouts. The more noise the player makes, the more likely they are to be tracked by an enemy.

2 Previous Approaches

While many games are still released without captioning, there are an increasing number of games that include some form of closed or open-captions (text and/or emoticons) for some auditory information. For example, colour coding of captioned text has been used sometimes to distinguish between speakers, and some games show a speaker portrait of the speaking character next to the caption. Animated speaker portraits or avatars can also increase readability by showing facial expressions to add context (e.g. *Freedom Force vs. The Third Reich*, VU 2005). Action captions, originating from comic books, have been used very rarely to represent sound effects on screen using onomatopoetic text ("bang!" "pow!"). *Speech balloons*, similar to those found in comic books, for example, help to tell the user which character is speaking, or where a sound source is emanating from (e.g. *Mario and Luigi: Partners in Time*, Nintendo 2006).[vT06]

There are several difficulties with text, however, most notably the speed with which text must be read. Games occur in real-time in high speed and the player needs to be able to read very quickly in order to take in all the information required, as well as react to that information. Language barriers are common, in that even where games are localized (translated for another market), there are generally only a limited number of languages that games are translated into-and some players play in a second language, which means reading speeds vary. Third, there are, as intimated above, problems of interpreting speaker, location and emotion in text captions. Finally, of course, the use of text captions assumes that all users can read, whereas many children play video games before they can read (or at least, read at the level/speed required).

Alternatives to text do exist. It is possible, for instance, to use the sub-woofer to convey some information through the vibrations that occur with loud bass sounds, although other than a message that there is some kind of sound occurring, the communication is limited. Videos of animated sign language could be implemented in games, although they have yet to be. However, these have the difficulty of distracting the player for long periods from on-screen action, along with localization difficulties (there are many regional variations of sign language), and are more expensive to implement. Other visual information used to indicate what is traditionally in games communicated through sound share programming and cost limitations, and have only been implemented to provide very limited information (e.g. a threat meter).

In summary, there are many limitations in the available approaches, which has meant that there has been very little adoption of any attempt to make games more accessible to the hearing impaired. Moreover, there are as yet no standards for these visualizations, they are used sporadically in very few games, and require considerable extra programming effort (and therefore cost).

3 An Approach to Improve Iconic Sound Effects: SoundSign

Based on the difficulties encountered with existing systems, we created a new sound imaging system called SoundSign that:

- is symbolic (not reliant on language), is easy to learn, and intuitive to use;
- is hardware and software independent and integrates easily into existing audio software or game engines and web technology;
- is icon or font-based, scalable and able to be rendered in black and white as well as colour;
- adapts to future change or game-specific needs (potential options include: motional cues using colour, blinking for immediate actions required, and user-generated fonts to suit styles of game-kids/cute, aggressive/militant, regular, sleek/modern, old, etc.);
- could use sound metadata or sound information retrieval techniques to indicate when symbols are necessary (either text-based metadata or actual reading of sound file);

- displays in a succinct, accurate manner positional information of sound events and proximity information.

The first step was to decide which sound effects required a sound icon, since some sound effects are more important than others, and having too many sound icons would distract from the purpose of the system (to indicate sound cue information). We used quantitative and qualitative methods to study game sound from hundreds of games by playing games with the sound on and off and recording which sounds were most missed. We also compiled sound asset lists from sound designers and audio programmers working in the game industry. We settled on a list of 26 of the most important sound effects commonly used in a variety of genres, using 26 as a number to map the icons to a letter of the alphabet so that the icons could be created as a font for easy implementation. Each icon has a corresponding alphanumeric key, so that implementation by a sound designer can be made with minimal effort. These symbols sit on a layer above the background compass, a rotational pointer which indicates direction of the sound. The SoundSign compass is turned on or off by the player at the game's option screen, and when turned on is stationary, positioned in the lower right hand of the screen. This way, players can become accustomed to looking in one spot for their audio information, and the particular position in this case would not occlude game display information. In games that already use this position for game display information, the SoundSign may have to take another position.

For 2D games and media (typically platform or casual games that take place in a 2-dimensional space), the compass can act like a clock face, with 12 rotational settings based on the sound's proximity and direction to the player. The compass graphic sits in one position which gets rotated according to (x,y)-coordinates (and, typically, stereo positioning of sound). Proximity is calculated from (x,y)-coordinate of the player and the sound, with four proximity settings that alter the opacity of the compass (50%, 75% and 100%, with 0% representing no important sound present). A demo can be downloaded here: http://www.gamessound.com/ss.htm

4 Experimental Framework

4.1 Prototype Basis

A prototype of SoundSign was developed based on the game *Assault Cube*, an open-source first person shooter. A recent usability study of user interfaces in first person shooters ranked that visual information regarding ammunition, radar and health were all highly important, and in particular that many players prefer having a visual rather than audio warning if critical levels are being reached (for instance, only 19% of players preferring an audio warning over a flashing indicator for health meters) [LF08].

4.2 Prototype Implementation

The prototype for this experiment required embedding code directly into the target platform source code. The original source code was written in C++ and linked to the OpenGL graphics libraries. The prototype was developed and tested under Linux (Ubuntu 8.10) using the GNU Project gcc compiler version 4.3.3.

During execution, the game program maintains in real-time a 2-dimensional grid-map data structure, locating all the players (entities) in the environment. Relative distance information is readily computed as the entities navigate through the environment, and this is crucial given proximity is an important factor in game-play. The map provides both (x,y)-coordinate information for the sound sources, and a relative sound level, depending on the distance from the player. Code that interpreted the sound direction and distance was inserted into the main game engine and the audio driver, and when action on the screen was updated, the appropriate icon could be displayed along with direction and level information. A small queue was implemented to facilitate the simultaneous display of several icons in the event that more than one sound was heard. Each icon had a timer instantiated with it, so that after three seconds, a given icon would disappear. A maximum of three simultaneous sound events was displayed (three icons) at any one time. In this case, when there were more than 3 simultaneous sounds, the three closest to the player were shown. Typically, sounds are assigned priority codes in games and these priority codes can be used to select which sounds should be displayed.

4.3 Measuring Usability

We undertook a usability test of our prototype SoundSign with the following questions in mind:

- Do users utilize the tool (i.e. respond to events indicated by tool) when the sound is turned off or kept on?
- How much do users utilize the tool when sound is off, and when sound is on?
- Did users find the tool useful?
- How well did the users understand the icons, having been given no instructions?
- Was the tool too much of a distraction for the user?

The effectiveness of the user interface for such a system is best measured in both subjective and objective ways, by recording reaction of the participants. Only attempting raw statistical measurement, for example monitoring if game scores improve over time, is difficult to formulate and the results would lack relevance. As such, we chose to use think-aloud protocols as well as structured interviews to measure subjective response, and eye tracking to measure objectively. Most participants, however, had difficulty with think-aloud while playing what for all of them was a new game (and a game that had much simultaneous action) that required considerable attention, and abandoned think-aloud during their play despite encouragement. As such, post-game interviews enabled us to identify areas of difficulty as well as success while the participants reflected on gameplay.

The objective testing (eye tracking) enabled us to record data on the use of the SoundSign during game play. As indicated by [PCV05], eye tracking can be useful for usability testing when conventional usability methods (click analysis, questionnaires, interviews, think-aloud) cannot provide enough data. It is particularly useful in determining where the user focuses visual attention. An additional benefit, of course, is that it can expose response bias.

The test participants consisted of 12 hearing-abled students at the University of Waterloo, recruited through on-campus signs announcing a study of video game sound. Approximately half of the participants were regular gamers, and half considered themselves to be casual gamers. With the exception of one participant who classified himself as an expert gamer, however, there was no significant difference in response to SoundSign between casual and regular gamers. The expert gamer reported that he did use SoundSign during gameplay, and his eye tracking results supported this statement.

Participants were tested by a research assistant, and were given no instructions with regard to what was being tested, and so were not focused on any area of the game. This lack of instruction also enabled us to test how intuitive the SoundSign system would be to new users. As one of our primary goals was to create a simple system that would be easy to learn, it was important that the system not require any instructions. Participants were brought into the lab and played *Assault Cube* with both the sound and SoundSign on for ten minutes. Sound was then turned off, and players continued to play for an additional ten minutes.

5 Usability Experiment Results

5.1 Subjective Feedback

Participants were asked if they found SoundSign useful when sound turned off, asked if they relied on it for sonic information during play, and whether they found it difficult to follow the icons while playing the game at the same time. They were also given a chance to raise questions or ideas about SoundSign. No participants reported being too distracted by the system, although one student thought that their visual attention was somewhat divided. A few participants reported some minor difficulty having not been given any instructions, although most found it very intuitive. The most positive responses to the system occurred when the sound was turned off, with most participants saying that they came to rely on the iconic system for information previously obtained through sound.

5.2 Objective Feedback

The participants were all eye-tracked during game play, using a monocular system by Applied Science Laboratories (ASL), in which the participants had to sit with their head in a chin rest mounted in front of them on a desk in front of the computer monitor playing the game. The sample rate was 60 Hz. The SoundSign system was fixed in the bottom right-hand corner of the screen, which became our area of interest (AOI). The number of fixations within the AOI was recorded, along with the gaze % (proportion of overall time) spent in the area of interest.

User	With sound on, gaze %	With sound off, gaze %	Difference
1	10.65	13.91	+ 3.26
2	11.67	16.97	+ 5.3
3	3.93	10.86	+ 6.93
4	1.82	0.045	-1.775
5	7.78	3.95	-3.83
6	4.67	15.67	+ 11
7	1.67	0.67	-1
8	.96	1.12	+.16
11*	2.41	10.59	+ 8.18
12	2.6	4.68	+ 2.08

Table 1: Table 1. Gaze % (proportion of time) on AOI (*data tracking failed on participants 9 and 10.)

As Table 1 shows, the gaze % mean was an increase of 3% of the total time spent playing looking in the direction of SoundSign when sound was turned off, suggesting that participants used the system significantly more when sound was off. Out of the recorded participants, 7 out of 10 had longer fixations on that area of interest with the sound off, and an average gaze duration (cumulative duration playing time as a percentage of the overall time playing) of 7.85% was spent looking in the area of interest.

Table 2 shows the number of fixations within the AOI. As can be seen, while some participants did not use SoundSign to a significant extent (indeed, some very little, and some less when the sound was turned off), those who did use it to any significant extent all show a marked increase in use when the sound was turned off.

User	Sound on	Sound off	Difference
1	122	176	+ 54
2	136	203	+ 67
3	46	120	+ 74
4	26	5	-21
5	83	43	-40
6	46	178	+ 132
7	24	10	-14
8	11	14	+ 3
11	36	161	+ 125
12	32	62	+30

Table 2: Table 2. Number of fixations within AOI (accounting for saccades) (*Average total data points per participant was 1105.)

5.3 Discussion

Both objective and subjective data suggests that a visual iconic system that is intuitive and provides information regarding proximity, sound effects and direction are useful in video games when the sound is disabled. Some players spent significantly longer times gazing at our SoundSign system when audio cues were removed from gameplay. Without audio cues in regards to enemy direction and proximity, for instance, players can be at a significant disadvantage without sound. Of particular note is that the findings suggest that sound effects play a significant role in game play (at least in this type of game). Sound is often an afterthought when it comes to game design, and yet the reliance on sound is clearly indicated by the usefulness of SoundSign.

There were several limitations to our study, including the realism of play: the fact that players rarely play a game without moving their heads, and a slight lag in the game's graphics due to the eye tracker's recording (a limitation of the eye tracking technology used). The limited number of tests provided useful anecdotal data but will need further testing for reliable statistical data. Another test, where users have no SoundSign icons during the first part of play with sound on would also be useful, as would a test on hearing impaired gamers. Of great interest would be the impact of compensation, where it might manifest that hearing-impaired players react differently in an experimental scenario when using SoundSign. Follow-up tests will be conducted in a strict framework, using hearing-impaired participants.

6 Conclusions and Future Work

This paper illustrates the value of a visual alternative to sound effects for video games, and highlights the need for more work in this area. Participants found that an alternative to sound effects was necessary in cases where important game information is relayed through sound effects. While this particular game is *playable* without sound effects, many users found the game more enjoyable and easier with the SoundSign system on when sound was turned off.

A second prototype with a variety of potential revisions is in development. We will also allow the user to select placement of SoundSign on the screen and test a variety of types of games. This eliminates the usefulness of eye tracking, but other testing methods will be employed. A long-range milestone is to design a general-purpose module that would integrate into existing game audio engines (e.g. Audiokinetic's Wwise). This would maintain the SoundSign functionality, without the requirement for specific game-dependent coding. A finals stage would be to integrate into and test the technology in non-game applications that might make use of alternate representations of sound, including PDAs and other handheld devices.

Acknowledgments

Funding for the project was provided by C4 Consortium and the Canadian Foundation for Innovation.

7 References

[ABA+08] O Aran, T Burger, L Akarun, A Caplier, Gestural Interfaces for Hearing-Impaired Communication. In *Multimodal User Interfaces*, D. Tzovaras (ed.) Springer Berlin, 2008, pp. 219-250.

[Co08] K Collins, *Game Sound* The MIT Press, Cambridge, 2008.

[FG04] J Friberg, D Gärdenfors, Audio Games: New perspectives on game audio. In: *ACE '04*, 2004, pp. 148 - 154.

[LF08] K Lenz, D Fox, Examining the Critical User Interface Components of First-Person Shooter (FPS) *Games Usability News* Vol. 10 Issue 2.

[LFU07] DG Lee, DI Fels, JP Udo, Emotive Captioning. In: *CIE 5/2* 2007.

[MFM05] T Matthews, J Fong, J Mankoff, Visualizing non-speech sounds for the deaf. *ASSETS'05* 2005, pp 52-59.

[MOA+08] K Miesenberger, R Ossmann, D Archambault, G Searle, A Holzinger, More Than Just a Game: Accessibility in Computer Games, LNCS *USAB 2008*, Graz, Austria, 2008, pp. 247-260.

[OMA08] R Ossmann, K Miesenberger, D Archambault, A Computer Game Designed for All. K. Miesenberger et al. (eds.): *ICCHP* 2008, LNCS 5105, pp 585-592.

[PCV05] M Pretorius, A Calitz, D Van Greunen. The Added Value of Eye Tracking in the Usability Evaluation of a Network Management Tool. *Proceedings of SAICSIT* 2005, pp. 1 -10.

[SBH+03] J Sánchez, N Baloian, T Hassler, U Hoppe, AudioBattleship: Blind Learners Collaboration through Sound. *CHI 2003: New Horizons*, pp 798-799.

[vT06] RA van Tol, The Sound Alternative. *Game Accessibility*, 2006. http://www.accessibility.nl/games/ accessed 1 Aug 2009.

[Za06] B Zahand, Making Video Games Accessible: Business Justifications and Design Considerations. *Microsoft.com* 2006. http://msdn.microsoft.com/en-us/library/bb172230(VS.85).aspx accessed 1 Aug 2009.