Novel Audio Interface Design for an Interactive Mixed-Reality Museum Exhibit

Jon Friskics

Digital Media Department, University of Central Florida
P.O. Box 163121
Orlando, FL 32816
United States of America
jfriskic@mail.ucf.edu

I. Introduction

In the Fall 2009 semester, the graduate-level Applied Interactive Story course at the University of Central Florida designed and prototyped an interactive museum exhibit for a proposed NSF grant entitled “Dancing The Earth.” The general purpose of this exhibit is to teach museum visitors about some aspect of an ecosystem in the Everglades, and our specific scenario focused on feeding habits of the wood stork. Planning for such a large project required specialized research in the areas of interactive storytelling, level design, environmental design, character modeling, volunteer training, educational assessment techniques, video and sound design, and interface design. The class pooled together their respective strengths, and by the end of the semester had completed a process called “play testing” multiple times. I chose to focus on sound design and audio system design because of my prior experience in the field, and by the end of the semester I developed a novel touch screen interface for remote audio control of the exhibit.

II. Problem and Background

The core of the exhibit relies on remote control of the active scenario by a museum volunteer that is located away from the museum. Using a single volunteer to control the exhibit, and placing that volunteer so far away from the museum presents several large problems, the most significant of which are listed and described below:

1) **Bandwidth** - real-time audio streaming from the volunteer to the museum is infeasible due to the normal ratio of bandwidth needed to bandwidth available.

2) **Information overload** - since the volunteer will need to give attention to multiple screens/interfaces, anything more than simple control will be information overload, and is likely to cause fatigue.

3) **Feedback Loops** - since the volunteer needs to communicate with people in the exhibit, and the people in the exhibit need to communicate with the volunteer, the potential for feedback loops must not be ignored.

4) **Troubleshooting** - if something goes wrong on either end, the experience gets very quiet very quickly.

For the sake of scope and clarity, the rest of this paper assumes that audio and sound are the only elements under consideration, and that story-writing, video interface development, and volunteer training have been handled by external forces.

III. Interface Mockups + Background

In the course of pre-production, three distinct interfaces were planned for remote museum volunteer control of the exhibit audio systems. For each of these interfaces, it is assumed that the volunteer will be given all of the technology in a turnkey system that they can simply un-box, plug-in, and be ready to run with minimal setup and configuration. It also needs to be assumed that the systems will have the ability to reset all preference files and system configuration files on reboot, so as to minimize the chance of an unrecoverable error caused by a human mistake. The work of usability designer Stephen Few was...
consulted for inspiration into the interface
design. Few writes “The information must fit
on a single screen, entirely available within
the viewer’s eye span so it can all be seen at
once, at a glance,” [1] so all of the interfaces
discussed below are envisioned as fitting on a
single screen. Few also outlines 13 common
mistakes in dashboard design, including “3)
Displaying excessive detail or precision, 10)
Highlighting important data ineffectively or
not at all, and 11) Cluttering the display with
useless decoration,” [1] and these rules have
received much consideration in the design of
the interfaces.

Reeves cites Oviatt and Stanney and suggests
“cognitive science literature on intersensory
perception and intermodal coordination has
provided a foundation for determining
multimodal design principles… and to
optimize human performance in multimodal
systems, such principles can be used to direct
the design of information presented to users,
specifically regarding how to integrate
multiple modalities… such as voice and
gesture.” [3] This suggests that further advice
should be sought from experts trained in
cognitive psychology to fully explore all of
the issues surrounding this type of interface
design.

The minimum diagonal screen size of the
audio interface should be 20" with a
resolution of at least 1280x1024, and the
screen should support multi-touch control.
The three interfaces are presented and
discussed below in the order that they were
developed.

IIia. Interface 1 - “Scales"

Description
Scales features a 7x7 grid of angular boxes
(called scales because they looked like scales
on a reptile after I drew them on a large sheet
of paper) that display information about
currently loaded audio files. The scales can
be configured through software to show
information, and the idea was that they would
be available in four preset configurations:

1) Sound effect playback
2) Music playback
3) Preset dialogue playback
4) Push-to-talk

A single scale located at the origin of the grid
(row 1, column 1), would serve as a system
menu to select between exhibit scenarios, and
the bottom-right scale contains a push-to-talk
button that will activate volunteer-exhibit
microphone communication only when that
scale is held down. Loading a new scenario
would pre-fill the remaining 47 scales with a
combination of sound effects, music, and
dialogue presets. Each normal scale's display
would contain latch controls for play and stop,
a momentary switch to solo the audio
triggered by that scale, and a draggable
volume fader.

Figure 1 – The “Scales” Interface Mockup

Pros and Cons
The biggest feature that Scales offers in terms
of interface design is the ability to have all
information displayed at all times AND be
accessible to the operator at all times. The
downside to this approach is, of course, great
potential for information overload. Another
question considered was "does the operator
need to have full control of dialogue, music,
and effects?" When we asked this question
and attempted to write down a comprehensive
list of all sounds that would need to be
manually triggered (not automatically
triggered), we realized that the operator will
most likely need to be able to switch scenarios
that contain sets of pre-scripted sounds driven
by exhibit triggers, and not need the level of
control that the Scales interface offers. If this
is assumed to be true, then Scales amounts to
a lot of wasted screen space and unnecessary clutter.

IIIb. Interface 2 - "The Flag"

Description
The Flag borrows from the iconic American flag metaphor with the hope that operators will be so used to the layout of the flag that they will have no trouble locating information once they are acclimated to the type of information that the stars and stripes sections contain. The stripes are each clickable bars that will cause a specific scenario to load, and the scenarios are ordered 1-13 from top to bottom. The currently loaded scenario will be highlighted in contrast to the other scenarios so that it is immediately obvious which scenario is active. The stars section will contain a screen that shows relevant information about the currently loaded scenario such as current position vs. duration of cues, overall volume, and the ubiquitous push-to-talk button.

---

![The Flag Interface Mockup](image)

**Figure 2 – “The Flag” Interface Mockup**

Pros and Cons
The idea behind The Flag is that it improves on a number of shortcomings of Scales. Since Scales displayed so much unnecessary information across multiple tiny panels, The Flag removes all of those tiny panels and presents the operator with one single display window at the top-left corner of the screen. It is assumed that the stars window will be anywhere from 1/3 to 1/2 the width/height of the display. Another thing that makes this interface unique is that the operator can instantly tell how far along in the sequence of scenarios the exhibit has progressed. A downside to this interface is that the stars window can only display information from one scenario at a time, and there is no way to "look ahead" to sounds in a future scenario (although arguably this feature is not necessary).

IIIc. Interface 3 - "Wheel of Sound"

Description
The Wheel of Sound is a series of concentric circles emanating from a single center circle. The spacing between each circle is large enough to display important contextual information about what that ring does, but is also small enough so as to not waste space. The idea is that the most used sound features would be accessible from the singular center circle, while less used features would be accessible from the outer rings of the "solar system" of circles. It was decided that the most used feature would likely be push-to-talk mode, and that this should occupy 100% of the space of the inner circle. The first ring would contain cross sections assigned to music cues, while the second ring would contain sound effects, and the third would contain dialogue. The beauty of this system is that ranges of the 360-degree pie would be confined to sounds from a specific scenario, and the selected wedge (scenario) of the pie would be highlighted in contrast to the other wedges. It is not expected that the operator will need to control individual volumes or playback of any of the specific sound effects for reasons discussed above in the descriptions of previous interface mockups, but it is desirable that the volunteer be able to see what sounds are possibly going to play in a specific scenario. It is this subtle interface design that allows for information about all scenarios to be visible at once, but through creative use of the "highlighted wedge", the information overload overhead associated with showing that much information at once will be greatly reduced. Other than hitting the push-to-talk button, the operator will only need to potentially select a new wedge to trigger the start of a new scenario early or
later than anticipated due to variations in audience participation/comprehension etc.

Figure 3 – The “Wheel of Sound” Interface Mockup

Pros and Cons
In addition to the unique interface tweaks discussed above, we anticipate that this interface will be the most intuitive for both non-trained operators and our key volunteer demographic of a potentially older generation that doesn't have intuitive experience reading touch screen interfaces. For the exhibit as defined this semester, we were hard pressed to find any disadvantages to this interface design short of wasting some corner screen space due to the circular nature of the interface paired with the rectangular physical construction of the touch screen monitor.

IV. Summary and Recommendations
Figure 4 identifies a number of key metrics that naturally evolved from the information above mapped across each of the three interfaces to create an "interface comparison" list. From this presentation, it appears that the Wheel of Sound is the most appropriate of the three interfaces to consider for implementation into a full remote volunteer control system. It would be beneficial for the interface programming team to work with an interface designer on issues such as typography, color palette, and resolution of each of the interface elements, but for the most part the Wheel of Sound mockup explains enough about its implementation to start developing it immediately.

![Figure 4 – Comparison of the key features of all three interfaces.](image)

Another issue that has not received enough attention is the process of usability testing. Interface developers should plan to engage in a complete user experience study, from generating personas and use-case scenarios, to conducting focus groups. When conducting tests with novel interfaces, it is important to follow some of the tenets that Spolsky set forth, like “When you sit somebody down in a typical usability test, you’re really testing how learnable your interface is, not how usable it is.” [3]

V. References