

Structured Audio: Using Document Structure to Navigate Audio Information

Mark Hakkinen

The Productivity Works, Inc.

George Kerscher

The Daisy Consortium.

Structured Audio

- What is Structured Audio?
- Structured Audio and Digital Talking Books
- Design of Structured Audio
- Using Mark-Up Languages and SMIL
- Structured Audio Browsers
- Directions

What is Structured Audio?

Traditional Spoken Word Audio

- Linear Presentation
- Navigation limited
- Forward and Backward Movement
(time/tracks)

What is Structured Audio?

Structured Audio

- Structure of Source Document Linked to audio
- Audio is searchable via text
- Navigation by structural elements
- Structure can be described by playback system

Structured Audio and Digital Talking Books

Talking Book Producers have begun efforts to move toward Digital Formats

- **The Daisy Consortium**
- **Recording for the Blind and Dyslexic**
- **Library of Congress National Library Service**

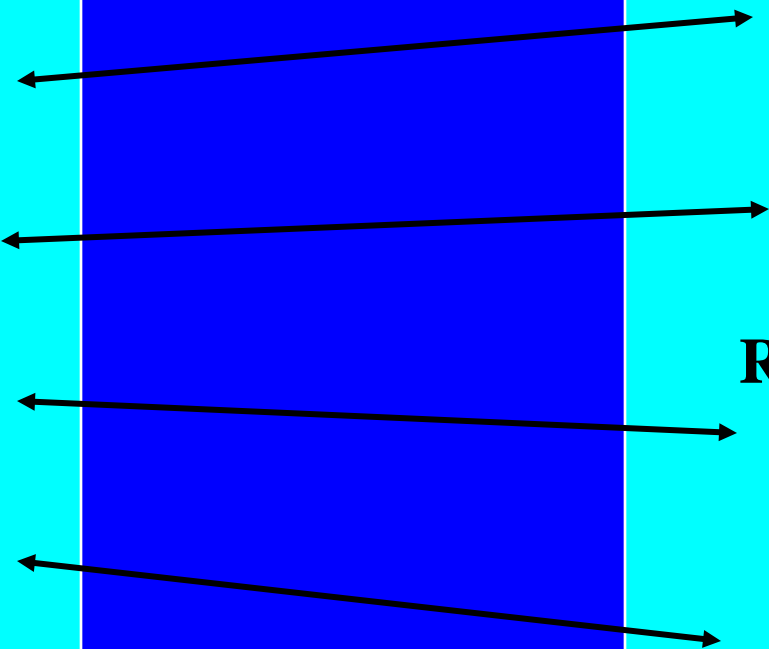
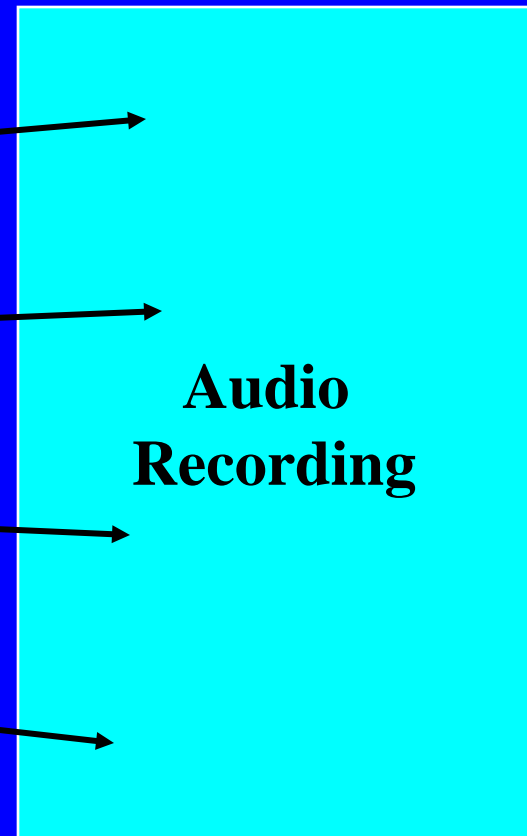
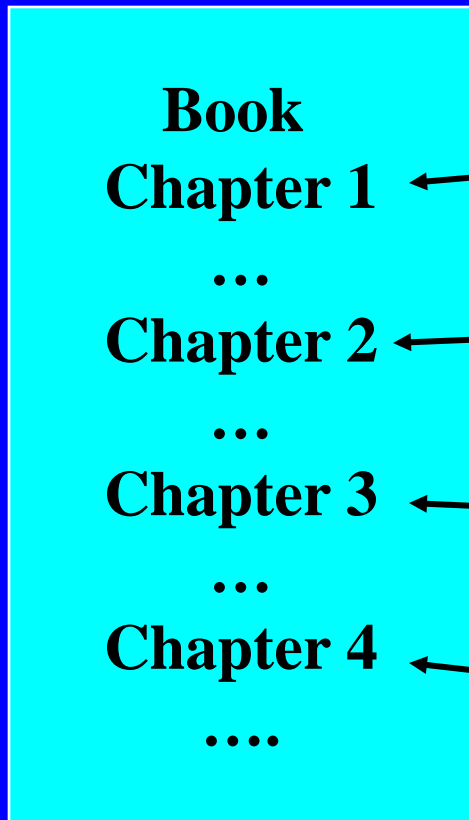
Goal: Next Generation Digital Talking Books

Elements of a Structured Audio Book

- Text Document with Structural Mark Up (E-Text)
- Digital Audio Recording
- Linkage Mechanism
- Playback Software/Device

Text

Audio



Representing Structure (And the Linkage)

Open Standards vs. Proprietary Formats

First Generation:

Daisy Audio (Phrase Recognition)

Structure based upon acoustic properties and table of contents

RFB&D Audio Plus

HTML and Real Audio

Second Generation:

Identify and Adopt Open Standards for Audio and Text

- HTML 4.0 for Text
- Audio Codec Independence
- W3C Synchronized Multimedia Integration Language (SMIL)

SMIL

- General Purpose Language for Creating Synchronized Multimedia Presentations
- Handles Linkage and Synchronization
- Text and Media Objects Independent of SMIL Definition

SMIL

- Structured Audio Books are a basic application of SMIL
- SMIL offers significant capability for creation of general, accessible multimedia content

Mark Up Languages

- HTML itself does not provide Sufficient Structure to represent books
- HTML and CSS offers one approach
- XML identified as ideal approach

XML and HTML/CSS

```
<canto id="co01">
```

```
<h1 class="canto" id="co01">
```

SMIL

```
<text src="CantoII.html#co01" begin="00:00:20.12"  
      id="co01"/>
```

Browser and Playback Systems

- Handheld Devices

Plextalk from Plextor

Victor from VisuAide

- Playback Software

Sigtuna and Daisy Software

- Web Browsers

SMIL Plug-ins

Directions

Sigtuna Project

Implementing using SMIL and HTML

Daisy Consortium

Implementing SMIL and HTML/XML

Library of Congress NLS/NISO Project

SMIL and XML under review for Digital Talking Book Standard

Directions

Tools and Playback Systems will soon be available to allow creation of rich, structured audio content

The content will be “playable” across platforms with a variety of user interfaces

It will change the way we ALL listen to and use recorded audio