

# Tobi

***“an extensible software tool to author next-generation accessible publications”***

**Daniel Weck, Software Developer, DAISY Consortium**



*The Tobi “DAISY Multimedia” logo: a light bulb wearing black shades, and 3 cubes (yellow, red, green) symbolizing components assembled together.*

## **Presentation Abstract**

- **Multimedia user experiences based on text, images, audio and video are now the norm. This rich ecosystem is mostly unstructured and inaccessible to persons with disabilities.**
- **The revision of the DAISY standard is incorporating new features to fulfill the scalability, interoperability, and accessibility requirements of the publishing industry.**
- **This presentation will demonstrate the Tobi open source software in a typical production workflow involving the synchronization of text and recorded audio. We will introduce the application's technical architecture, to showcase the plugin extensibility mechanism and its potential for innovative third party features.**

## About Daniel Weck

- **Writes software applications for the DAISY Consortium**
- **“Tobi” development team leader**
- **Architect and programmer of the “Urakawa SDK” (Software Development Kit for synchronized multimedia authoring)**
- **Contributing developer in the “AMIS” player (self-voicing user interface)**
- **Involved in Zed-Next standardization, video-in-DAISY working-group**



*The DAISY Consortium  
Logo*

# **Slides and Demos**

## ***Easy DAISY Production***

- **Digital Talking Book: synchronized text and audio**
- **Demonstration of the Tobi authoring tool**

## ***Workflow Overview***

1. **Mainstream editors: MS-Word, OpenOffice, etc.**
2. **“Save as DAISY”: DAISY 3 XML (DTBook, NIMAS)**
3. **“DAISY Pipeline Light”: automated production (TTS – Text To Speech)**
4. **“Tobi”: interactive authoring (record audio, edit waveform)**
5. **Press “play” and enjoy !**

## **Key Topics**

### ***End-Users / Content Makers: Production Workflow***

- **Using both automated and interactive production**
- **Document authoring (text, structure), mainstream tools**
- **Audio recording and editing**

### ***Solution Implementors: Open and Extensible Software***

- **Reducing costs: existing development resources**
- **Future-proofing a software investment: flexible architecture**
- **Mixing open-source software and closed-source proprietary extensions**
- **Adding features to already-installed application: plugins**

## **Tobi, In A Nutshell**

- **Free, open-source software developed by the DAISY Consortium**
- **Authoring tool for DAISY 3 full-text full-audio Digital Talking Books**
- **Extensible framework for editing next-generation accessible multimedia**
- **GNU LGPL license, business-friendly policy (allows commercial derivatives / close-source extensions)**
- **Windows-only application (XP SP3, .NET 3.5 SP1)**

# Screenshot

The screenshot shows the Tobi software interface. The top window title is "Tobi [Z:\Assets\DAISY\_books\valentin\_haüy\06-speechgen.opf.xuk]". The menu bar includes "File", "Edit", "Tools", "View", and "Help". The toolbar contains various icons for file operations and editing. On the left, a "Headings" pane lists a hierarchical structure of document sections, with "1. Research questions" selected. The main document area displays the heading "1. Research questions" in a large, bold font, highlighted in yellow. Below the heading, the text "My interest in Valentin Haüy, the father of education for the blind in Europe, was aroused while working at TPP on a book" is visible. At the bottom, an audio waveform is shown with a green signal. The waveform is divided into segments with time markers: 600ms, 1s 200ms, 1s 800ms, 2s 400ms, and 3s. A selection box highlights a portion of the waveform, with a label "1s 320ms Research questions" and a time range of "1s 627ms - 2s 680ms (1s 53ms)". The status bar at the bottom indicates "Waveform loaded.", "16 bits 22.05 KHz Mono", and "Time: 924ms / 3s 297ms | Selection: 1s 627ms - 2s 680ms (1s 53ms)..."

*Tobi Screenshot, text and audio waveform*

# Supported Formats

## ***Import***

- **DAISY3 (Z39.86–2005–3)**
  - **DTBOOK-only**
  - **Full-text / full-audio**
- **EPUB (experimental support)**
  - **DTBOOK**
  - **XHTML**

## ***Export***

- **DAISY 3.0 full-text full-audio**
- **DAISY 2.02 (under development)**

## ***Audio***

- **PCM 16 bits recording and editing (DirectX / DirectSound)**
- **Built-in MP3 decoding and WAV resampling**
- **MP3 encoding via Lame (under development)**
- **Planned support for real-time processing plugins (VST)**

# Features Highlights

## ***Fully Accessible***

- Screen-readers (JAWS, Windows Eyes, NVDA, Narrator)
- System colors (high-contrast theme, etc.)
- Scalable vector graphics user-interface (crisp icons, no pixel artifacts)
- Configurable keyboard shortcuts (under development)

## ***Easy Authoring Workflow***

- Integrated text and audio edition views, no popups
- Simple user-interface navigation and keyboard shortcuts
- Intuitive waveform display

## ***Extensibility, Plugins***

- Contributed by the community
- Available from commercial vendors

## Technical Highlights

### ***“Real” Authoring Data Model***

- Urakawa SDK (Software Development Kit)
- No need to maintain the DAISY fileset: SMIL, NCC/NCX, HTML/DTBOOK, OPF, etc. only “visible” at import and export time
- Extensible programmatically (not locked into a specific document type)
- Portable, interoperable XML markup (XUK format)
- Native undo-redo support (slim risk of corrupting data)
- Transparent management of WAV audio files (non-destructive authoring)

### ***Modular Architecture***

- Separate, well-scoped development sub-projects
- Plugin framework, extension points: deploy new features without re-compiling the main application
- Lightweight, incremental online updates
  - Download only what is needed (small files)
  - Opportunity for more frequent releases (bug fixes, etc.)

**Phew, Demo Time !**



## **First Release, Limitations**

- **No structure / text editing (main focus is full text/audio synchronization)**
- **No true WYSIWYG (CSS rendered only at export time)**
- **Minimalistic audio waveform editor (no advanced filters, fade in/out, normalize, amplify, etc.)**
- **No built-in TTS features (Text To Speech)**

# Development Schedule Overview

## ***October/November/December 2009***

- Finish ongoing feature implementation
- Finalize plugin framework
- Package application (installer, online updates, etc.)
- Testing

## ***January/February 2010***

- Support forums, user documentation
- First stable release, production-ready

## ***September 2010 (tentative, depends on funding/resources)***

- Support for emerging DAISY-next standard (video, etc.)
- Structure/text editing
- TTS features
- Localization (support for right-to-left languages, Japanese IME, etc.)
- New plugin contributions
  - VST audio processing
  - etc.

## Get Involved !

Want to contribute to the future of accessible multimedia ?

- Tobi community forum: to be opened shortly
- “Early access” for end-users (tool is not production-ready)
- “Developer preview” for programmers (design / code not finalized)

| <b>End-Users, Content Producers</b>  | <b>Software Developers</b>  |
|--|---|
| <ul style="list-style-type: none"><li>• Find crawling bugs</li><li>• Report performance issues</li><li>• Review usability, accessibility</li><li>• Request new features</li><li>• Write cookbook, tutorial</li></ul> | <ul style="list-style-type: none"><li>• Improve or customize the baseline application</li><li>• Design your own software extensions</li><li>• Write programming tutorials, sample plugins</li></ul> |

<http://www.daisy.org/projects/tobi/>