

12-2015

Synchronizing Oral History Text and Speech: A Tools Overview

Robin Camille Davis
CUNY John Jay College

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: http://academicworks.cuny.edu/jj_pubs

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Davis, R. "Synchronizing Oral History Text and Speech: A Tools Overview." *Behavioral & Social Sciences Librarian* 34.4 (2015): 234-238. Web.

This Article is brought to you for free and open access by the John Jay College of Criminal Justice at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.

This is an electronic version (postprint) of an article published in *Behavioral and Social Sciences Librarian*. The article is available online at <http://www.tandfonline.com/doi/full/10.1080/01639269.2015.1098790> (subscription required).

Full citation:

Davis, R. "Synchronizing Oral History Text and Speech: A Tools Overview." *Behavioral & Social Sciences Librarian* 34.4 (2015): 234-238. Web.

Internet Connection
Synchronizing Oral History Text and Speech: A Tools Overview
Shortened title: Synchronizing Oral History Text and Speech

Robin Camille Davis

Abstract:

This article explores three tools that synchronize sound and text for online oral history collections: the Oral History Metadata Synchronizer (OHMS), the Text Encoding Initiative (TEI), and YouTube. A detailed description and examples are given for each. Integrating audio/video recordings and transcripts enables searching and browsing, making oral histories more accessible and approachable.

While the choices of digital collections software are numerous, few are friendly to oral histories. Often, oral histories are presented as a set of files to download: an audio or video file and a PDF of the transcript. To produce these digital objects is a resource-heavy undertaking, even prior to dissemination (Stevens). Making both file types available online to researchers is a common practice: audio/video recordings provide linguistic information such as intonation, and text enables skimming and searching. But these functions remain separate. Integrating speech and text through dual displays and synchronized timestamps opens oral histories to greater searching and browsing capabilities, with the effect of increased accessibility and approachability. This article highlights three tools that offer synchronized audio and transcript displays for oral histories.

Oral History Metadata Synchronizer

OHMS enables transcript sections to display in sync with the audio or video recording. The transcript becomes interactive, allowing the user to see at a glance the topics covered in the interview and target a segment immediately. The audio therefore becomes more accessible, rather than remaining an undivided media file that may be hours long. By closely intertwining the transcript and recording, OHMS makes browsing the oral history much more user-friendly. OHMS is flexible and can be utilized in a number of content management systems.

OHMS software is two-part: there is a web-based app that requires a sign-in and a software package that must be installed on a server. Note that OHMS is not a repository or a CMS: the audio or video recording must be hosted elsewhere, such as Vimeo. Once the recording is online, the archivist working with the oral history feeds the recording's URL into the web-based OHMS app and inputs item-level metadata and time-tagged transcripts. These transcripts can be done on the fly or pasted in, and the OHMS app includes an audio player so that time tags can be easily synced. The transcriber or archivist decides how to split the transcript into manageable segments, usually topic-based (e.g., "Plantation Life in the 1930s"). The OHMS app exports an XML file containing the metadata, time-stamped transcripts, and display instructions; this file must be uploaded to the server where the OHMS software resides. At this point, the XML file is linked from an object's record in a repository or CMS. The public view of the object page will then display the recording and searchable transcript in the OHMS listening/reading interface.

OHMS requires a fair amount of technical "elbow grease," and the decisions to split the transcript into topic-based sections are non-trivial. But the pre-built software is free, open-source, and at the time of this writing remains under active development. OHMS can work with many CMSes, allowing the synced oral histories to be included with other digital collections.

OHMS is a project from the Louie B. Nunn Center for Oral History at University of Kentucky Libraries. Launched in 2014 after years of work, OHMS is open-source, free to use, and growing in popularity. Oral history collections that currently use OHMS partially or fully include the Brooklyn Historical Society's "Crossing Borders, Bridging Generations" project, which uses a Drupal CMS and audio hosted on Soundcloud; "United States of Aids" from Out History at The New School for Public Engagement, a project that uses a WordPress CMS and video hosted on YouTube; and the Richard Nixon–Frank Gannon Interviews from the Walter J. Brown Media Archives & Peabody Awards Collection at University of Georgia Libraries, a collection that uses a custom interface and videos hosted on YouTube.

The Text Encoding Initiative

The TEI is a standard, scholarly XML schema used across a multitude of publication projects, from classroom transcriptions of medieval manuscripts to a scholarly edition of Van Gogh's letters. The TEI is a descriptive markup language that can encode oral history transcripts with as much detail as desired. Item metadata, speaker names, timestamps, sections, and sentences can all be encoded, as well as linguistic phenomena such as pauses and coughs ("Transcriptions of Speech"). The end result of encoding a transcript using the TEI is a long XML file that is machine-readable. Using this XML file to sync recordings and transcripts requires a custom-built web publication system wherein speaker names and other elements can be automatically formatted, and timestamp data can be harnessed to sync playback. Building such a custom system that uses TEI-encoded transcripts to display the oral history may require many technical resources.

Encoding a text in TEI can be time-consuming, but the XML file can describe the structure of the transcript with a low or high level of detail. XML and TEI

conventions can vary by project, just as transcription practices do, but this textual format is reliable and extensible. The standards are readily accessible online, and the TEI is regularly taught in publishing and digital humanities workshops. The TEI is not shackled to certain software, and the encoded transcripts have value beyond the website. The data model makes archiving and reusing the encoded transcript possible.

One example of an oral history project that uses the TEI is Voices of the Holocaust, housed at the Illinois Institute of Technology. Using the TEI-powered data model, each oral history includes audio, transcripts, English transcript translations, geographic information, item-level metadata, and scholarly commentary (English 2). The timestamps in the transcript encoding are read by a Flash-based media player that syncs audio and text, with the transcript highlighting each speaker's sentence as it is uttered.

YouTube

Perhaps the most economical of the three tools highlighted here, YouTube offers built-in functions and easy distribution that make it a good choice for small oral history projects that use video. On YouTube, users can stream high-quality video at very fast rates, a feat that can be hard to accomplish on institutional servers without many resources and technical work. Videos can be linked or embedded from any webpage, so this solution can work on an institutional website. But most interestingly for oral history projects, the captioning and transcript function for YouTube is robust. A transcript without timestamps can be uploaded along with the video, and speech will be automatically aligned with the text. (YouTube discourages audio that is difficult to understand or videos that are over an hour long.) Transcripts are accessed on YouTube video pages from a dedicated button in the video description, which opens to display a timestamped transcript that highlights text as it is spoken. Each line of text, usually 5–10 seconds apart, links to the corresponding place in the video. Transcripts also auto-fill the closed captions, accessible from the playback controls that are visible even for videos embedded elsewhere. Both options visually synchronize speakers' utterances with their speech.

To make the video even more accessible to users, video descriptions can include links to timestamps next to segment names. This is most commonly accomplished in the description section of a YouTube video page, but it is also possible to create a link to a video elsewhere such that it begins playback at a certain timestamp.

The downsides of using YouTube for oral histories include that it is designed for video only. However, an audio recording could theoretically be played against a static image, and in fact many casual YouTube users share music in this manner. Another downside is that the best way for a user to view the video and its transcript is on YouTube.com, rather than a collection's page. Lastly, the oral history files and the interface are subject to YouTube's changeable policies and design.

Examples of YouTube videos with auto-synced transcripts include some oral history videos from the Richard B. Russell Library for Political Research and Studies

at the University of Georgia. The videos' YouTube pages are linked within the oral history collections' finding aids.

Use cases

Of the tools presented, when and why choose one over the others?

- OHMS: The oral history collection will be presented in a pre-built or custom-made CMS, such as CONTENTdm or Drupal, and the institution has the resources to install software on its own server and spend a moderate amount of time splitting transcripts into named, time-tagged segments.
- TEI: The oral history collection will be presented in a custom-made or heavily customized CMS, and the institution has the resources to create custom software and spend a significant amount of time encoding each transcript in detail.
- YouTube: The oral history collection will be presented as links or embedded videos on a webpage, and the institution has the resources to spend a small amount of time creating this webpage and uploading videos and transcripts to YouTube.

(Note that these are not the only tools available for this task. These three tools were selected because they are in recent use, offer high-quality results, and are all cost-free solutions.)

Presenting oral histories as integrated archival objects, rather than as a set of files, makes the collection much more accessible to web users. Tools like OHMS, the TEI, and YouTube produce oral histories that are listener- and reader-friendly, and may therefore increase use of the collection.

References

- "Add subtitles and closed captions." YouTube Help, YouTube, 2015. Web. 10 Sept. 2015.
- English, E. "Bringing It All Together: Integrating Text, Audio, Metadata, GIS, and Scholarly Criticism in a Holocaust Oral History Archive." *Journal of the Chicago Colloquium on Digital Humanities and Computer Science 1.2* (2010): n. pag. Web. 14 Aug. 2015.
- Stevens, K. Weatherford and B. Latham. "Giving Voice to the Past: Digitizing Oral History." *OCLC Systems & Services: International digital library perspectives 25.3* (2009): 212-220. Web. 4 Sept. 2015.
- "OHMS: Oral History Metadata Synchronizer." Louie B. Nunn Center for Oral History, University of Kentucky Libraries, 2013. Web. 14 Aug. 2015.
<<http://www.oralhistoryonline.org>>
- "Transcriptions of Speech." Text Encoding Initiative, 2015. Web. 10 Sept. 2015.
<<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/TS.html>>