

# Shared listening experience for hyperaudio textbooks

Niels Seidel <sup>1</sup>, Robin Dürhager<sup>2</sup>, Marcel Goldammer<sup>2</sup>, Alexander Henze<sup>2</sup>, Frank Langenbrink<sup>2</sup>, Joachim Otto<sup>2</sup>, Veronika Stirling<sup>2</sup>

**Abstract:** The integration of multimedia elements into educational materials has the potential to improve students' engagement and understanding. In this paper, we present the development of an hyperaudio player that represents textbook information in auditory format, providing a joint playback experience for groups of users. Additionally, the Moodle-based hyperaudio player includes a voice chat feature, allowing for real-time discussions among users when the audio is paused. We aim to investigate the appropriateness of hyperaudio as a new online learning format.

**Keywords:** Hyperaudio, Multimedia Learning, Collaborative Learning

## 1 Einleitung

Informal knowledge transfer through podcasts, audiobooks, and radio documentaries enjoys great popularity. However, in formal learning settings at schools or universities, auditory learning resources are rarely used. Compared to text, audio productions are more expensive and have a less distinct academic tradition. Furthermore, audio-based learning is perceived as an individual learning activity. In this paper, we are contrasting common online textbooks with a well-designed interactive auditory complement summarized as hyperaudio. Furthermore, we propose hyperaudio as a format for collaborative learning. The combination of hyperaudio and collaborative learning shall provide a mobile learning experience comparable with a telephone conference among peers where the textbook is read aloud on request.

As a follow-up to our previous research [Se22] about the conversion pipeline and mapping from course text to a single hyperaudio document in Moodle, we present the design and development of an hyperaudio player that represents a textbook alongside its corresponding auditory representation, providing a joint playback experience for collaborative learning for groups of users.

---

<sup>1</sup> Forschungszentrum CATALPA, FernUniversität in Hagen, Universitätsstrasse 27, 58097 Hagen, niels.seidel@fernuni-hagen.de,  <https://orcid.org/0000-0003-1209-5038>

This research was supported by the Center of Advanced Technology for Assisted Learning and Predictive Analytics (CATALPA) of the FernUniversität in Hagen, Germany.

<sup>2</sup> Fakultät für Mathematik und Informatik, Lehrgebiet Kooperative Systeme, Universitätsstrasse 27, 58097 Hagen, {vorname.nachname}@studium.fernuni-hagen.de

## 2 Related Works

This section summarizes related work regarding hyperaudio learning environments as well as technical approaches to create audio from text-based resources. In contrast to other instructional formats such as hypertext, there is only little research on learning with hyperaudio yet [ZS14,ZM20,DB07]. Donker and Blenn laid a foundation for the use of audio in hypermedia applications with the Hyperaudio Encyclopedia [DB07]. In that project, unlike the use of screen readers, articles in the encyclopedia were not just recorded and read to the listener. The users were able to interact with the audio document. For instance, links, headings, and other salient text passages were highlighted by auditory markers and recognizably presented to the listener. In terms of learning, [Re09] encouraged receptive processes for listening rather than equating them with passivity. Receptive processes, such as listening to podcasts, must be contrasted with productive learning processes. Reinmann suggested improving the quality of storytelling and listening through modern technologies. In her doctoral dissertation [Ha18] revealed that the implementation of books that are read aloud using a human voice (audiobooks) can positively impact the reading experience, at least for elementary students. In a study, [ZS14] compared textually and auditorily represented information in linear and non-linear forms. The non-linear representation increased cognitive load compared to linear representations. In more recent research, [ZM20] investigated the design of mobile hyperaudio learning environments. Referring to these results, hyperaudio should be designed in a linear rather than non-linearity manner. Non-linearity increases cognitive load and is less conducive to learning.

In terms of hyperaudio technology, existing frameworks such as popcorn.js, waivsurfer.js, or timesheet.js [CQR11] can be used for implementation. Also, several commercial services like Soundcloud or YouTube offer the possibility to markup continuous audio documents similar to hyperaudio players. Capable text-to-speech systems for the German language are mainly provided by big players like Amazon Web Services (Polly), Google (text-to-speech), and Microsoft Azure. Open source systems like Thorsten Voice<sup>3</sup> are promising but cannot create customized time markers and features like pitch and pauses. For manual audiobook production, among others, the open source tool Obi<sup>4</sup> supports standards like DAISY 3<sup>5</sup>, W3C Audio Books and Accessible EPUB 3.

To the best of our knowledge, no hyperaudio player is yet available that is capable to convert text to audio, presenting rich audible content, and facilitating collaborative learning.

---

<sup>3</sup> See <https://www.thorsten-voice.de/> (accessed 2023/04/02).

<sup>4</sup> See <https://daisy.org/activities/> (accessed 2023/04/02).

<sup>5</sup> See <https://daisy.org/activities/standards/daisy/daisy-3/> (accessed 2023/04/02).

### 3 Design and Implementation

Since hyperaudio is an interplay of hypermedia and audio, neither ordinary audio players nor ordinary hypermedia systems can accurately represent such content. For this reason, a new hyperaudio player was conceived as a blend of audio player and hypermedia application dedicated to collaborative learning. The hyperaudio player was developed as an activity plugin for the learning management system Moodle. The source code is provided under GPLv3 license at [https://github.com/nise/mod\\_hypercast](https://github.com/nise/mod_hypercast).

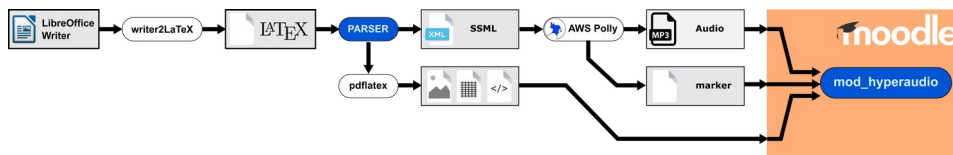


Fig. 1: Workflow of hyperaudio creation form course texts as proposed in Seidel [Se22].

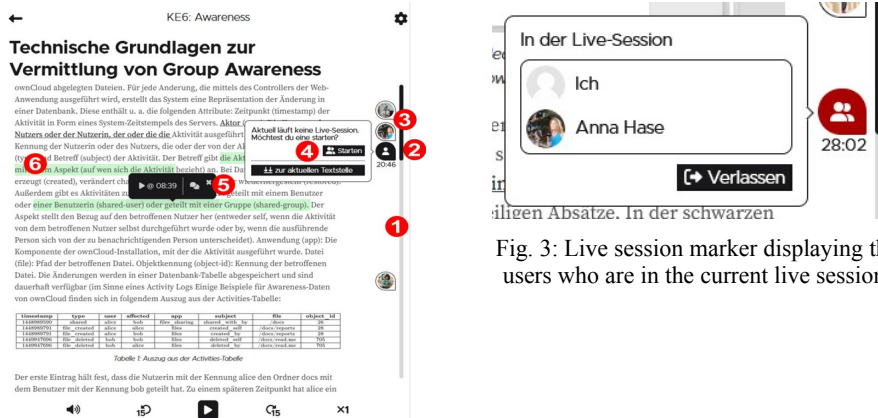


Fig. 3: Live session marker displaying the users who are in the current live session.

Fig. 2: Hyperaudio player design with the (1) vertical progress bar, (2) personal marker, (3) group members' markers, (4) live session start option, (5) comment pop-over and a (6) comment highlighted in the text.

The hyperaudio player is based on a hypertext representation of a textbook chapter. For our intended use case a chapter may consist of 40 to 60 standard text pages including all kinds of content representations including images, tables, code, formulas as well interactive controls (e.g. links) and even videos or other audio. Referring to our previous work [Se22] we were able to map text-based design elements of typical course texts to an auditory hyperaudio representation using Amazon Polly as text-to-speech systems (TTS) as well as audio effects (e.g. panning for bulletin lists) and audio cues (e.g. acoustic indication of non-audio contents like images) during the playback. As shown in Fig. 1, we intended to provide teachers with a way to convert existing course text into

ready-to-use hyperaudio documents to be implemented in Moodle courses. Thus, the course text remains a nicely readable text rather than a text-only transcript of the audio. In contrast to audiobooks or podcasts effects and audio cues are included at runtime considering recent interactions with other users (e.g. comments).

Other than common audio players we've turned the timeline of the audio player from horizontal to vertical to be in line with the scrollbar used for navigating the text (Fig. 2). Concerning a mobile-first design, the playback controls are arranged at the bottom of the screen. The text scrolls synchronously to the audio playback progress. While the vertical timeline indicated the current playback position with a personal marker, the current text passage gets visually highlighted. Navigation within the text and audio can be achieved through a click on the timeline or a text passage. The graphical user interface is currently optimized for right-handed users. For instance, the timeline is placed on the right to be controlled using the right thumb. The same accounts for the playback controls. The users are encouraged to take their eyes off the screen while listening to the audio. Due to the audio cues, they get notified about non-audio content visible on the screen.

Collaboration is facilitated on a group level. Large student cohorts can self-organize collaborative activities by creating, managing, and joining one or multiple groups. These groups are persistent across all existing hyperaudio instances within a Moodle course. However, it is advisable to limit group size not to congest the collaboration space with the user's virtual representation (e.g. playback positions, avatars) and traces (e.g. comments). Group members share equal privileges except for the creator of a group is allowed to change group settings (e.g. maximal group size, group name). All group members are represented on their last playback position on the timeline. Their representation consists of a profile image, name, and current online status. This means the hyperaudio player provides support for workspace awareness.

The player supports asynchronous and synchronous communication facilities. Content-related communication is mainly supported through comments that are anchored to text passages that are linked to precise playback positions in the audio. Comments are visually highlighted in the text and audibly emphasized in the audio. A comment can be classified as a question, remark, or note to invoke direct replies from other group members. As a result, text and audio can be augmented with multiple threaded discussions.

Synchronous communication is achieved through so-called *live sessions*. In contrast to listening to the audio individually, the live session is shared by multiple group members who are currently online. Group members joining a live session synchronously listen to the audio. Group members can join and leave the live session whenever they want. Each participant in the live session is in control of the audio playback affecting all participants. As long as a group member stays in the live session the audio plays or pauses for all participants at the same temporal position. Concurrency control is ensured by a social protocol and group awareness instruments indicating the current status of the live session and a list of participants (Fig. 3). In case one is changing the playback status (e.g. from play to pause, changing playback speed, seeking on the timeline) other participants of the live session are notified by a toast notification. When the audio is paused the participants can talk to each other using voice chat if they enable their

microphones. The voice chat terminates as soon as the audio playback continues. Online group members who did not join the live session also see on the timeline who has joined the live session and where they are currently located in the text and audio.

## 4 Evaluation

Due to the novelty of the collaborative use of hyperaudio, usability and user experience was the main scope of the evaluation efforts presented in this paper. The evaluations aimed at getting insights on how to improve the design of the application and to increase the acceptance of individual and collaborative use in online learning. Ensuring high acceptance rates regarding the usability of the application and the audible representation of text learning is a requirement for future evaluations of learning success and implementation in teaching. As part of the design-based research process, expert reviews have been conducted at the end of each of the seven development iterations. To evaluate the usability of the final hyperaudio player we applied the think-aloud principle as follows.

*Participants and design:* Seven participants joined the think-aloud test remotely, 4 female and 3 male. The live session was performed in a wizard of oz mode.

*Material:* A course unit was prepared as a hyperaudio document and delivered through an online Moodle environment using the plugin presented in the previous section. The course unit entitled “Group Awareness” was part of a current course taught in a Computer Science Master's program at a German university. The participants were provided with a list of 15 instructions that covered the range of provided features. These instructions covered group management, audio player, live sessions, settings, and comments.

*Procedure:* The participants have been invited to participate online using a video conferencing software. In the respective online meeting, they have been introduced to the think-aloud method and provided with a list of instructions including the credentials to access the Moodle course.

*Analysis and measures:* During the study with the individual participants the screen and audio conversation were captured. Furthermore, the interactions on screen have been observed by the examiner. Clickstream data was collected in the standard Moodle logstore. After the think-aloud test, the participants were asked to express their opinion about the application.

*Results:* The evaluation revealed a couple of usability problems and user experience pitfalls. Regarding the group awareness support, the participants struggled to recognize their representation on the timeline. For seeking the timeline the participants were expected to click on an arbitrary position on the timeline rather than dragging their marker to the desired position. Also, the menu provided to switch between textbook chapters was not recognized and thus was not used properly. Finally, Moodle controls distracted from activities within the player. All the mentioned usability problems have been addressed in the latest revision of the Moodle plugin.

## 5 Conclusions and Outlook

In this paper, we presented a hyperaudio player for collaborative learning. Considering related works we discussed design considerations regarding the hyperaudio player, group management, and synchronous and asynchronous communication. The resulting Moodle plugin was evaluated to improve the usability of the novel learning format. The following contributions to the research about technology-enhanced learning have been made: We presented a new online learning format as a blend of hypertext and hyperaudio to be used for individual and collaborative learning. Our implementation of a Moodle plugin enables a joint listing experience for audio-based learning resources incorporating synchronous live sessions with voice chat and anchored threaded discussions of text/audio passages. We achieved an integrated presentation of rich-media course text (e.g. textbook chapters) with its corresponding audio document by mapping text-based design elements to audible representations. The Moodle hyperaudio plugin lays the foundation to further research about the effectiveness of hyperaudio learning. To further improve the audio quality we aim at tweaking the pronunciation and speech variety using up-to-date text-to-speech systems. In the next winter semester, we are planning to implement the hyperaudio player in a university course with about 80 participants. Using the implemented logging capabilities we'll be able to track click behaviour, reading behaviour, and audio usage on a fine-grained level.

### Bibliography

- [CQR11] Cazenave, F.; Quint, V.; Roisin, C.: Timesheets.js: When SMIL Meets HTML5 and CSS3. In: Proceedings of the 11th ACM Symposium on Document Engineering. ACM, New York, NY, USA, pp. 43–52, 2011.
- [DB07] Donker, H.; Blenn, N.: Gestaltung von Hyperlinks in einer Hyperaudio- Enzyklopädie. In Gross, T. (ed.): Mensch & Computer 2007: Konferenz für interaktive und kooperative Medien. Vol. 7, Oldenbourg Verlag, München, pp. 139–148, 2007.
- [Ha18] Hartell, A. W.: Audio Books' Impact On Students' Reading Experiences, Doctoral dissertation, University of South Carolina, 2018.
- [Re09] Reinmann, G.: iTunes statt Hörsaal? Gedanken zur mündlichen Weitergabe von wissenschaftlichem Wissen Mündliche Weitergabe wissenschaftlichen Wissens. E-Learning – Lernen im digitalen Zeitalter 51, pp. 256–267, 2009.
- [Se22] Seidel, N.: Mapping course text to hyperaudio. In (Henning, P. A.; Striewe, M.; Wölfel, M., eds.): DELFI'22. Gesellschaft für Informatik e.V., Bonn, pp. 69–74, 2022.
- [ZM20] Zumbach, J.; Moser, S.: Examining the Effectiveness of Hyperaudio Learning Environments. In: Mobile Devices in Education: Breakthroughs in Research and Practice. IGI Global, pp. 421–437, 2020.
- [ZS14] Zumbach, J.; Schwartz, N.: Hyperaudio learning for non-linear auditory knowledge acquisition. Computers in Human Behavior 41/1, pp. 365–373, 2014.