Interactivity and multimodality in language learning: the untapped potential of audiobooks.

Emanuela Marchetti\textsuperscript{1}, Andrea Valente\textsuperscript{2}

\textsuperscript{1}Institute for the Study of Culture, University of Southern Denmark, Odense, Denmark
ORCID 0000-0002-7949-9953
emanuela@sdu.dk

\textsuperscript{2}Maersk Mc-Kinney Moller Institute, University of Southern Denmark (SDU), Odense, Denmark
ORCID 0000-0002-6295-9511
anva@mmmi.sdu.dk
Abstract. In this work we present three case studies, involving classes in primary and secondary schools, in Denmark. The studies, conducted in the past 2 years, show how audio contents can be generated and shared among teachers and learners, how audio materials can be made more interactive to offer fruition similar to that of digital games, and how language learning can benefit from adding a social dimension to audiobooks. All case studies were conducted in a user centered fashion and build on social semiotics, in which interactive audiobooks are seen as providing new ways to receive, interpret and share literary texts. Local primary and secondary schools were involved in ethnographic user studies and qualitative evaluations with semi-functioning prototypes. In the main case study presented, social interaction was chosen as key feature to allow high-school students and teachers to annotate audiobooks, then share and comment on the annotations; the social context in this case is a digitally-augmented English teaching class. To better investigate the potential of sharable audiobook annotations we also created a mockup supporting the workflow of the main case study, using standard YouTube annotations and freely available audiobooks. The findings and technical solutions explored in the three studies are the basis for design guidelines aiming at making audiobooks interactive and better integrated in learning contexts.

Keywords: e-learning · multimodal interaction · information presentation · knowledge management

1. Introduction

In recent years, audiobooks have become increasingly more accessible, both in the market and in free, often publicly supported, streaming sites and apps. Denmark follows this trend and many local schools (primary and secondary) consider free audiobook repositories as part of their resources. Our data suggest however that integration of audiobooks in the class poses many challenges. Our studies of primary and secondary school classes suggest that audio resources, despite their growing availability, are often used in a rather traditional way. Learners are typically required to listen to an audiobook perhaps on a portable device, then take notes (on paper or making use of a different application), and finally present and discuss their notes in class or in group discussions. This traditional fruition of audiobooks presents problems:

• It takes more time than reading, even when the text is in English and the reader is learning English as her second language
• Audiobooks cannot fully substitute a text, since they often offer a shorter version of the original text
• Audiobooks are perceived as more passive than regular books as it is often not possible to take notes and mark specific passages as with books and e-readers

We also found that in other occasions, such as during group read-aloud tasks in primary schools, speech (and therefore audio) seems to be the preferred modality. In this context learners and teachers could benefit from recording oral performances, to store as documentation and for further feedback later in the courses. However, audio-content creation is typically disregarded, even when current mobile technology has both the hardware and software needed for inexpensive, good quality recording and playback of audio contents, as well as powerful and fast audio compression, which allows in turn for large amounts of data to be stored on board a phone or tablet.

The general attitude towards audio contents is characterized by lack of active uses, for example in relation to generation, editing and sharing of content, and we could find virtually no audio-based interactivity. In the past two years we have conducted three case studies to explore the untapped potential of audio in the context of learning of English as a foreign language in primary and secondary schools. Our studies suggest that textual and video modalities are predominant in classes practice, even when most of the orchestration of learning activities is oral. We consider the lack of interactive audio as a missed opportunity to promote richer use of multimodal resources in class. Moreover, audio contents might support learners with special needs, such as: visually impaired people or pupils affected by dyslexia or when learning a foreign language ordering commutes for occasional, non-formal learning.

In the following sections we present related work (section 2), our three case studies (sections 3 and 4), discussions (section 5), our mockup for the main case study (section 6), and conclusions (7).

2. Related work

Several studies have already explored use of audiobooks as tools for creative engagement with literary stories and tools for learning foreign languages. For instance Furini [7] and Huber et al [10] challenged the typical use of audiobooks, which was found to be passive with respect to books. Both studies argue in fact that readers of audiobooks might gain a more passive experience, being constrained to listen to the story, therefore, they have explored possibilities to enable users to interact with non-linear narratives creating their own stories. The study conducted by Furini aims at turning the passive reader of audiobooks into the “director of the story” [7]. In developing his system he looks at the use of audiobooks through a cinematic metaphor, imagining the experience of editing new stories as if the reader was editing video sequences, referring to movies like Sliding Doors and Pulp Fiction as displaying non-linear stories. The system targets three main use cases: entertainment, education, and game applications. The design focuses on two main principles: transparency, as the book file should be standard and easily handled by the system, and security, as only the owner of the audiobook file should be able to play it and she should not be able to alter the original media files from which the audiobook was created. The article discussing Furini’s studies, however, focuses on the technical aspects of the design of the system and does not discuss in details the expected user experience or results from testing.

Huber et al. [10] take a similar approach and discuss the evolution of audiobooks into interactive media, suitable for editing non-linear stories. The authors propose to combine elements from computer games with the experience of listening to oral presentations, which are defined by Huber et al as immersive and entertaining. Moreover, sonification was used as a resource for interaction, in order to enable the users to interact with the system mainly through sound, but this interaction style was found difficult by the users.

A relevant study has been conducted by Alcantud-Díaz and Gregori [2], who propose an extensive review of the use of audiobooks in foreign language learning and two projects named Tales of the World and The Power of Tales: Building a Fairer World. The authors claim that even though in their country, Spain, audiobooks are not commonly used, they can see great potentials in supporting English learning in relation to the five skills listed in the Common European Framework of Reference for language learning: listening, reading, spoken interaction, and writing. The two projects discussed by Alcantud-Díaz and Gregori aim at spreading awareness of languages as scaffolding for intercultural values and respect for human rights in the educative community. The outcome from both projects were collections of tales, for the Tales of the world project 40 tales were gathered from underprivileged countries, for The Power of Tales 15 tales against violence were collected. All the tales were edited into free downloadable audiobooks. The format of audiobooks was chosen in order to give access to pupils with learning and visual difficulties; moreover, audiobooks were seen as a mean to improve learners’ English pronunciation. The other studies discussed in the review in [2] focus on the use of audiobooks for primary school pupils dealing with language learning, such as Wilde and Larson [22] who argue that audiobooks enabled children 8 to 12 years of age to find more time to read, hence reading more books. Moreover, Baskin and Harris [3] found that use of audiobooks supported students with learning difficulties, who find it challenging to interpret written text, in making sense of written texts and improving their reading fluency in English as first language.

Other studies show that the use of audiobooks can enable pupils with limited views or blindness in accessing literary content, for learning as well as leisure. In some cases interesting ventures have emerged between commercial and non-profit organisations, in the production of new audiobooks targeted at individuals affected by visual disabilities. Adkins and Bushman [1] have conducted a survey to investigate which services were provided by public libraries to children affected by disabilities. The two authors start from data provided by The Census Bureau stating that 5.2 percent of children in schools in the USA are affected by a disability, which could deal with vision, hearing, cognition. The survey was conducted in a form of a questionnaire and was sent to 185 public libraries in the USA, of which thirty-nine sent back a response. According to the survey, different resources were provided to children with limited vision such as: audiobooks, large print books, and Braille books. In several cases it was found that libraries cooperate
with schools in providing material for children with special needs. Moreover, it seemed that audiobooks were found to be the preferred resource by children affected by limited vision.

Moving away from the learning context, we can find another form of interactive audio: audio walks or audio tours [20]. Similar to the audio material offered by museum, audio walks are usually implemented as mobile apps where users can follow predefined audio commentary while moving around a city or a building. An interesting commercial product of this kind is yapQ’s “Worldwide city guides”², a mobile app that offers audio walks in multiple languages and for many cities; the application uses geolocation and text-to-speech to generate interactive audio guides. The contents in this case are not user-generated. SoundCloud³ instead is an example of user-generated and socially shared resources: “SoundCloud is [...] social sound platform where anyone can create sounds and share them everywhere”. Among other sound collections, SoundCloud offers a selection of audio walks.

2.1 Theory - Social semiotics and multimodality

This study builds on the theory of social semiotics and multimodality, in particular on the works of Kress [12] and Kress and Van Leeuwen [13]. Our study aims at turning audiobooks into more creative and interactive tools to support literacy, intended as learning of foreign or native languages, but also self-expression through authorship of literary audio texts.

The discipline of social semiotics is concerned with the study of social communication through signs and the relation between the senders of the sign and the receivers in the process of making and remaking of meaning. According to Kress [12], the production of meaning is composed by three main features: the semiotic feature that deals with the form of the content, the conceptual feature that deals with the concepts represented by the content, and the affective feature, which deals with self-expression, the personal interest and investment of the maker of the message. Makers of meaning or rhetors, make use of these three features at the same time while creating a message. At the same time creation of meaning leverages on rules, which provides logic to the integration of the different features and of the signs used in the process of meaning making [13]. However, the process of meaning making continues when a newly created message is received by someone, the receiver(s), who will engage in decoding and interpreting the new message. This means that the receivers are not simply given a message, but as they engage in decoding and understanding the message, they are recreating the meaning embodied in the message eventually in new unexpected ways. It is in this respect that Kress argues in [12] that senders and receivers are both actively participating in a negotiation of meaning. The interpretation process in which receivers engage to understand new messages, implies nowadays different forms or authorship, through which receivers can be empowered and fluidly shift into creators of new messages. More specifically through the act of decoding a message, receivers are in fact creating a new meaning and new knowledge, which might be distant from the one intended by the original creator. This means, according to Kress that “knowledge is always produced rather acquired” [12], in the sense that acquisition is a “non-agentive” interpretation of the relationship between senders and receivers, neglecting the active role required by the receiver in decoding a given message.

The process of the production of meaning is affected by the social context and the power relation between sender and receiver, in this sense officialdom provides a central factor. In his book Kress provides several examples of public authorities sending official messages to individuals, such as limited circulation on specific streets in Salzburg because of an athletic event [12]. In the case the sender of a message is a political authority, as in the example presented by Kress about street signs, it is assumed by the authority that the receivers should be able to decode the meaning of the message as intended by the sender, so that the receivers are expected to conform to the behavior prescribed in the message and that forms of enforcement and eventually penalties could be applied, in case of the receivers’ misconduct.

Kress also reflects on the role of social semiotics as a theory that deals with meaning in all its appearances, in all social occasions and cultural sites [12-13]. In this respect he introduces the notion of

² Freely available on GooglePlay.
³ List of popular audio walk on SoundCloud: https://soundcloud.com/tags/audio%20walk (last seen on the 23rd of March 2017).
multimodality, intended as the normal condition of human communication. The notion of multimodality deals specifically with the richness of human communication, which takes place involving different modes of communication leveraging on different sensorial stimulation. For instance, verbal communication involves primarily auditory modes as people talk and listen to each other, where words together with modulation of the voice participate in conveying meaning. However, facial expressions and gestures enrich communication with visual modes, which also contribute to the exchange of meaning among individuals. Media like movies or videos make rich use of the visual and auditory modes of expression, in conveying meaning and emotions to the viewers. Written communication, nonetheless, represents a complex and rich form of visual communication [13], for example communicating through the visual format of a text, in graphic design the discipline of typography provides perceptual and sociocultural theories about how texts should be formatted in relation to the message being sent, the goal of the message, and the intended target group. Typography is based on the principle the designers of written messages should work creatively on typographic variables such as: the shape of the letters, which is commonly referred to as “font”, size, color, texture, orientation and more [14]. Moreover, written communication can become more effective making use of images. Examples are provided by Kress in [12], where he presents the case of two shops communicating with street signs, making use of text and of a simple diagram, how drivers can reach the rear parking lot of the shops.

Reflecting on the creation and fruition of literary texts in school, from the perspective of multimodality and social semiotics, physical books are still the main resource for literary fruition and learning. Physical books are considered the norm when it comes to how schools, as educational institutions, relate to schooling practice and the pupils’ learning process. Written communication and reading are seen as highly important skills for the pupils to gain, in order to be able to move forward in their studies and function as citizens in their society. Education is in fact determined by sociocultural values, in relation to the knowledge and the skills that are valued in a particular society, in contemporary western societies literacy happens to be among the most valued skill and this explains why so much effort is put by educational institutions in enabling children to learning how to write and read [19]. Moreover, writing and reading are acknowledged in different cultures as particularly hard skills to achieve, which demand for an institutional framework, and not as skills that can be achieved on an independent basis by autodidacts. This has implications for the production and fruition of literary texts, as institutionalized education requires a constant production of books for schools, which provide newly created reading materials as well as collections of existing literary texts.

2.2 Social semiotics and audiobooks

In recent years, we have experienced the appearance and penetration of digital book formats, like ebooks and audiobooks. Ebooks still leverage on the visual modes and reading, just in a digital form, so that readers will read their books on a computer or mobile device and not on paper. Audiobooks leverage instead on the auditory mode, so that readers will listen to a chosen book from a computer or mobile device, hence the reading experience will be more like music and radio drama fruition than reading as traditionally intended.

The appearance of new digital book formats has several implications from the perspective of fruition and dissemination of newly created and classic literary works. In his studies Kress reflects on how digital technologies have provided new resources, formats, and affordances for the design and implementation of meaning. When using the term design, Kress intends the process through which “the meanings of a designer (...) become messages”, where a design could be a public speaker or a teacher or any participant in everyday interaction [12]. At the same time design is defined by Kress as “a theory of communication and meaning”, which acknowledges the work of individuals in their social lives. In this respect, design moves away from traditional conventions shifting towards an understanding of communication as equitable participation of individuals in the shaping of the social and the semiotic world [12]. Furthermore, Kress argues that the availability of the new affordances provided by digital technologies within the global framework of contemporary information society is contributing to a redefinition of the notion of power and authorship. For instance new understandings of authorships have emerged in our society since design resources for the creation and dissemination of messages are provided to a larger numbers of individuals than in the past. As a result, Kress argues that recent technological developments are causing
rearrangements in power, mainly as shifts from a vertical to a horizontal structure of power, allowing for more open and participatory relations among the individuals involved in contemporary communication.

In relation to our study, we find that a similar trend towards rearrangements of power is emerging also with the introduction of new book formats. The availability of book formats that leverage on different sensorial modes has, for instance, provided a basis for increased access to the fruition and creation of literary texts [2]. Moreover, we argue that within these new formats the term “text” acquires a new broader meaning, indicating both a written text but also an audio recording of a literary text being read aloud by a professional actor. This broader meaning of text is accompanied by new possibilities of access to literary fruition, for individuals with special needs, such as limited vision abilities, blindness [5], and dyslexia [2]. In the case of individuals affected by limited vision abilities through the use of audiobooks the very fruition of literary texts becomes a possibility, while in the second case it becomes easier and can actually prepare the path towards reading skills for pupils experiencing challenges [2]. Moreover, the combination of different modes and formats can enrich literary fruition, so that individuals can enjoy the same texts through different experiences and in more flexible ways, shifting from reading on a traditional book to a mobile device, and listen to the same book read by professional actors on a mobile device while driving or being occupied with other activities. Furthermore, a professional actor might be able to convey the emotions embodied in a literary piece in compelling ways, so that reading can become like the experience of music and theatrical pieces. These different formats can also provide richer experiences for pupils learning a different language as in the already discussed case of Alcantud-Díaz and Gregori in [2], where the combination of the written and the auditory modes, can provide support in future readings, knowing the correct pronunciation of written words, and also contribute to the students’ understanding of the texts through the emotional interpretation of the text provided by actors. Ease of access to digital books is increased also by initiatives like The Guttenberg Foundation, which are engaged in the diffusion of classical literary works in digital form, hence contributing to the dissemination of literary culture among young people and people who might not afford access to books for different sociocultural and economic reasons.

In agreement with [12], we find that the new digital formats of literary texts have provided new opportunities for authorship, so that it became easier for anybody who is interested in writing new literary works, which can be easily distributed for free through websites that enable writers to gain Open Source licenses on their work, so that other people can access these literary works for free but are bound to give credit to the author when referring or reutilizing excerpts from their works. The publishing industry has attempted to stop this flow making enforcement of the right of the authors and of the publishers in gaining credit and economic compensation for their work. In this respect we find that the current situation is complex and these attempts in controlling the distribution of digital book formats is de facto attempting to limit people’s access to literature and other forms of artistic production. We refer to Kress in our analysis of the current situation as he argues that authorship “is in urgent need of theorizing” [12], in order to sensibly take advantage from the new affordances provided by digital formats, which are contributing to a reconfiguration of the power relations between senders and receivers, democratizing access to the fruition and creation of messages. Kress points out how the notion of authorship itself, as it was intended for traditional formats, might have become obsolete. For instance the very notion of plagiarism has changed since new forms of content creation have appeared, which include mesh up, cutting and pasting material from existing media content. In this way excerpts from other content, which could be written texts or audio-video material, are being reused in the creation of new messages, hence acquiring new meanings from those intended by the original authors. In this respect these materials are not simply being stolen, but are actually reinterpreted in new messages by the new authors and tools, like Open Source licenses were created to secure that credit is given to the original authors.

We see our study as contributing to this discourse of democratizing access to fruition and creation of literary content, providing new scenarios and new tools promoting authorship in the fruition of audiobooks. We start from the increasing presence of audiobooks in Danish society, in particular we find an interesting exchange between schools and libraries, who are providing access to audiobooks to citizens since primary school. Our study suggests that audiobooks or audio resources for learning are used in many Danish schools at different levels, as part of the pupils’ training in foreign languages and literacy. Hence we see traditional books, audiobooks and ebooks as media complementing each other in promoting literary fruition in our society. However, when looking at audiobooks we identified several limitations in relation to
authorship and appropriation of the content from the readers. For instance readers can easily annotate a traditional book or an ebook, visual reading also enables readers to choose themselves their own speed, if they want to read all in details or just quickly scan through a text searching for a specific passage. Currently these operations are not adequately supported in the fruition of audiobooks, in this respect we see audiobooks as offering a more passive reading experience, relegating the readers to the role of receivers and hindering their opportunities for authorship. Taking this into account, our study aims at exploring how we can enrich the fruition of audiobooks and reconfigure the role of the audiobook user, so to:
1. Promote social inclusion in learning for individuals with special needs
2. Provide a complementary and alternative fruition of texts in situation where traditional reading might not be comfortable
3. Enrich young people creative palette to support self-expression in manipulating audio files
4. Enriching of current modalities of assessment in schools through more meaningful use of the audio mode

These points are central when adopting a universal design perspective, in which special attention is given to individuals with special needs. At the same time, as discussed in the previous section we find these points central also when accessing the needs of learners as individuals who leverage on different sensorial modalities and ways of learning as pointed out by Gardner’s theory of the multiple intelligences [8]. In this perspective we see multimodality and universal design as complementing each other, providing respectively a theoretical and methodological framework to support our study, in relation to redefine audiobooks to improve access to literary texts as well as enriching the fruition of audiobooks; in the next section we discuss our understanding of universal design within the context of our study, defining what universal design is and then clarifying how we build on it in our study.

2.3 Methodological framework

From an overall methodological perspective our research takes inspiration from the principles of universal and inclusive design or design for all, especially on the work discussed in [4, 16, 21].

All these approaches are seen by researchers as different names to refer to approaches that strive for the development of products that can be accessed by a wide range of people possibly by the entire population, in spite of differences in age and ability or other special needs. Referring to the definition of inclusive design provided by the Design Council [6], in [4] Clarkson and Coleman argue that universal and inclusive design are not a separate, new genre of design, but rather a “general approach to design” in which designers “ensure that their products and services address the needs of the widest possible audience, irrespective of age or ability” [6]. Universal and inclusive design are defined by the Design Council as two major trends, where the term inclusive design is mainly used in Europe and Universal Design or Design for All are mainly used in the USA [4, 16]. In general these two approaches emerged to meet the needs of individuals with different ages, abilities, and sociocultural background [21].

Despite this general agreement, there seems to be a challenge in finding a uniform definition of the approaches of universal and inclusive design. In [16] it also argued that inclusive design is aimed at creating main stream products that could be used by as many people “as reasonably possible” without any need for special adaptations. In this respect [16] identifies a subtle difference between design for all and inclusive design, since the word “reasonably” can be interpreted as if the application of the principle of universal inclusion could be limited by costs or other constraints. Nevertheless the authors also point out that inclusive design is not a fixed system of criteria but it is a “constantly evolving philosophy” [16]. For instance the application of inclusive design in the field of education refers to the need of creating learning applications and environments that could be used by anyone [16]. According to [16] this challenge is caused by the lack of a uniform definition of the concept of “accessibility”, whose meaning might change in relation to design approach and the designers’ goals. Interestingly, paper [16] points out that accessibility has become a central concern in the law of many countries, with the goal of reducing discrimination in relation to different level of accessibility, also determined by cultural and financial constraints.

Nevertheless, Clarkson and Coleman [4] argue that even though we might find different definitions and perspectives on accessibility and inclusive or universal design, the emergence of these approaches have contributed to eliciting awareness about how design can enable or disable people, in connection with features embodied by a product, the contextualization of the product, and the design process itself. This in
turn implies that designers should actively strive for enabling all the population to make use of new products and services.

In our study we specifically refer to the application of inclusive design or universal design within education, which refers to the need of creating learning applications and environments that could be used by different learners in spite of their different needs [16]. Typically the focus of universal design within education focuses on learners with special needs and abilities, in this respect we find that by combining multimodality and universal design we can gain a more comprehensive understanding on how the design of new learning tools can contribute to enrich the fruition of literary texts combining different modalities so to answer to learners’ different needs. Moreover, Gardner’s perspective on the different intelligences [8] enables us to see that every learners has different needs, in spite of not being diagnosed special medical conditions. Starting from this framework, our aim is to create interactive audiobooks that could facilitate the fruition of literary text in audio forms, in order to support individuals who might have challenges in learning another language. At the same time, we aim at expanding the experience of reading literary texts in contexts that would normally hinder reading activities: for instance, while doing physically engaging activities (Audook project section 4.1) or while travelling by car avoiding sickness and nausea (Carbooks section 3.2). Moreover, we find that the fruition of audiobooks poses more limitations to the freedom of the readers than normal paper books, as the voice of the actor might impose a specific rhythm and mood to the reader, who might not able to "read" at a desired speed. Audiobooks do not afford for activities such as: annotation, bookmarking and interactive game books. However, combining the principles of universal design with multimodality, we see that in making the use of audiobooks could open up towards richer and more flexible user experiences enabling readers to enjoy literary texts in more different ways.

In conclusion, combining universal design principles with multimodality and the theory of the different intelligences, we see the creation of interactive audiobooks as a way to support individual needs, in relation to exploring different contexts for reading, freeing interactivity with audiobooks, and supporting different modes in relation to the theory of difference intelligences as well as challenges in learning foreign languages or in reading in general.

In the following sections we will show more possible ways in which audio can be made interactive and we will explore the possibilities offered by social creation and sharing of audio data, based on the our studies conducted in cooperation with Danish schools and our own students.

3. Two supporting case studies

The main case study described in this paper is supported by unpublished data from 2 other case studies conducted in the past two years, which provided insights on the advantages of interactive audiobooks. All the three case studies adopted the User Centered Design method supported by qualitative methods. Our students had to conduct a full design iteration consisting of: a field study investigating the practice, in which users participate in; a phase of analysis in which design requirements are formulated; a phase of conceptualization through brainstorming and prototyping techniques; testing in which a semi-functioning prototype is evaluated with the users. The testing was conducted as a play test session with focus groups, involving users in demonstration of the prototypes. Qualitative methods were chosen for several reasons: first of all our students engaged with a limited number of users, weather high-school, primary school classes, or focus groups. Second, the students’ goal was to closely explore current user experience and opportunities for improvement, also enabling the users to propose possible ideas. Specifically, for the main use case, our students adopted visual ethnography in situ, semi-structured interviews for which they were requested to prepare a minimum set of pre-defined questions for the users [23]. The students were therefore required to analyze the gathered video recordings scrutinizing how users interacted during class activities and how they talked about their practice (verbal and non-verbal language) during the interviews, with the goal of identifying aspects that needed improvements or support. Semi-structured interviews and observations were also adopted in the two supporting studies. Given that our students were still learning about User Centered Design and qualitative methods, we took part in many of the phases of the 3 studies, complementing their field work with our notes and reflections. The findings discussed in the following sections are the result of this process.
3.1 Audio Deliverables

The audio deliverables application originated from the supervision of 4 groups of students attending the Software Engineering and IT bachelor at the University of Southern Denmark (SDU); the semester long project, run in fall 2014, was about developing user centered software solutions to better support English teachers in 2 Danish primary schools. The field study started with observations of 2 classes of 4th graders learning English, one in each school. After a preliminary visit and meeting with the 2 teachers who agreed to participate in this study, the groups of SDU students visited the school repeatedly and proceeded by defining requirements and producing a few prototypes, from low-fidelity ones to partially working horizontal prototypes (created using MIT’s AppInventor⁴).

The two teachers, here called Anders and Britta for anonymity, were also interviewed; they showed very different approaches of using technology in their teaching. Anders can be considered a designer of content. He states openly that he has limited IT skills but he is very creative in the design and generation of new content. In the first visit he showed us how he wrote a short dialogue with 4 roles, for his students to read aloud. In fact, spoken interaction and comprehension are the main goals for the 4th grade English curriculum. The dialogue was about 3 friends who interact with the waiter (the other role) in a British restaurant, and have to order, confirm their orders, eat and pay the waiter, who in turn has asks typical questions about their choice of food, beverages and how they want to settle their check. It was clear that Anders compensates the lack of interactivity in his material (which was not given to the pupils in digital format, but written at his computer and then printed) with role play and social interaction. Britta is much more in touch with IT and in particular likes to use what is available online, but she re-contextualizes it according to her pupils’ needs. She has a toolbox approach and often uses tools that are not originally pedagogical, like video editing, comics authoring tools and online audiobooks in English. In our first visit Britta brought her class to the IT lab for the English lecture; the pupils kept switching from audiobooks to cartoon editing, to chats with the teacher and each other.

We found these two approaches very intriguing and believe they should be further studied. However, in this paper we are mostly interested in user-generated audio contents, therefore, we will focus on the group of SDU students working with Anders’ class. They noted the various problems he had orchestrating the class with his printed material: the pupils were divided in groups of 3 to 4, and had to read the text a few times, waiting for Anders to drop by, listen to them and provide feedback. The result was the audio deliverable application, a mock-up mobile app that allowed pupils to read an English text aloud in groups, and deliver it to the teacher as an audio recording; a walkthrough of the mobile app is visible in figure 1. These audio deliverables afford good peer interaction and make the communication with the teacher more asynchronous. Moreover, they represent a form of audio content generation that is natural and very easy to master for 4th grade pupils who are typically proficient in the use of smart phones; the focus was mainly on reading skills.

The audio deliverable application was tested iteratively during the semester project. In particular, the final version of the prototype was tested and assessed by the teacher, Anders, and his class. In the interview that followed the testing session, Anders explained how recordings enable more asynchronous teacher/pupils interaction, since he does not have to be physically present at each English practice session; he also liked the idea that recordings can be preserved to serve as a learning diary to make Anders and his students more aware of their progress. We also observe that using audio recordings as deliverables opens the possibility of peer reflection. We observed pupils recording and submitting their read-aloud English exercises, and we noticed that audio content can be easier to generate than written English, at least in the context of Danish 4th graders.

⁴ AppInventor’s official page: http://appinventor.mit.edu/explore/ (last seen on the 23th of March 2017).
Fig. 1 Walkthrough of the "Audio Deliverables" mobile app. After loggin-in the children users can play various games, including read-aloud and recording of a text in English. The recording can be sent by email to a teacher. The teacher can also log-in and access her "inbox" to hear the recordings and provide feedback to the children.

3.2 Carbooks

This study demonstrates the versatility of audio as a communication modality, by mapping gamebooks into mobile-friendly, interactive audiobooks. The goal of this project was to offer an entertaining and relaxing experience to kids who often get car-sick in long car trips, and have problems reading or watching videos while traveling. In this case playing videogames using mobile devices is not an option; audiobooks instead can offer relief and help passing the time in a fun or perhaps educational way. However, audiobooks provide a passive experience and can become boring in long trips, so we wanted to investigate how non-linear narrative can be used in audiobooks, to create interactive and enjoyable experience for kids and young adults. A focus group was created to play-test the interactive audiobooks, composed of 10 young adults (age 19 to 25) and 2 kids (10 and 12); the family of the 2 kids was among the other stakeholders involved in the project. The Carbook bachelor project tested various ideas, running in the fall 2015 semester and through 3 iterations, with the central focus to develop an audio-only interactive application for android platform. The main tools were Unity5 and Google Text-To-Speech.

Removing the graphical user interface while retaining the interactivity typical of digital games proved one of the major challenges; the project also explored possibly mappings between input modalities and choice in the non-linear narrative. A mobile phone offers gestures, microphone and orientation/motion detection. Typical gestures we considered are touch, hold and swipe. As for microphone input, voice recognition was too complex to work in practice and it would have been mostly limited to English language, so volume level was used instead; microphone input was used in the second iteration of the interactive audiobook prototype, but turned out to be unreliable and difficult to use by the players, who got frustrated by the experience. In the third (and final) prototype microphone was replaced by orientation (basically reading the state of the phone’s gyroscopes). These input modalities were to be used in steering the narrative of the interactive audiobook, mostly without the player looking at the screen, and that required some analysis too; background audio clues were also used (in version 2 of the prototype) to help players orient themselves while exploring the locations in the story. In printed gamebooks the player is often faced with 3 to 6 options to select from, but in Carbooks we had to break down the player options in sequences of

5 Unity3d website: https://unity3d.com/ (last seen on the 23th of March 2017).
binary choices. The users commented positively on this restructuring of the choices, and told us how they preferred few binary alternatives instead of a single choice among multiple options. However, we would argue that the use of binary alternatives has limited the non-linearity of the narrative, reducing de facto the branching factor of the multi-linear plot.

The Carbooks project shows that interactivity can work in audio-only (or audio-first) applications, and that the user experience is similar to that of slow-paced exploration/adventure video games, such as classic text-based games of the 1980s. Smart phones, with their current computing power, audio support and their wide range of input modalities, were commented by users a reasonable choice of platform for audio-only interactive applications. The main limitation of the project however, was that it did not focus on content creation, so while we have evidence that interactivity and audio work for simple, fun non-linear stories, we have to progress further with our studies before we can directly link interactive audiobooks to language learning.

4. The main case study: social audiobooks

The last and main case study was conducted in relation to an elective course in Media Sociology, the course lasted for five weeks in the fall semester of 2015, and focused on e-learning with students from the Multimedia Design program (MMD for short) at the Lillebaelt Academy in Odense, Denmark. The course involved 21 students who had to work on a mini-project in groups of three or four, in cooperation with Nyborg gymnasium, a high-school located in Nyborg a small town on the island of Funen, Denmark. From the point of view of the Lillebaelt Academy, the learning goal of the mini-projects was to create conditions for the MMD students to conduct a rigorous user centered design process, actively involving users, to adopt a contextual perspective on the design of learning technologies, and to critically reflect on how their new solution contributes to teaching and learning practices in the gymnasium. On the other hand the gymnasium in Nyborg was eager to explore and test together with MMD students new interactive solutions, which could enrich the current learning and teaching practices.

In their Media Sociology course, the MMD students were introduced to five research articles applying a specific learning theory to learning contexts and to the design of a digital solution. One particular group of three MMD students explored the design of an application to support interactivity with audiobooks, these students chose to work with the studies conducted in visible learning by Hattie and Gan [9] and in the sociocultural theory by Marchetti and Petersson Brooks [15]. Hattie and Gan explain how visible learning can affect learning practice, discussing the role of teachers in enabling the students in formulating learning goals and success criteria, in providing descriptive feedback, which enables students to improve their skills, and formative assessment, aimed at collecting evidence of the student’s achievement. Paper [15] instead adopts the sociocultural theory in the design of a digital exhibit, aimed at enriching the social interaction between guides and visitors during guided tours. The study aims at enriching the interaction between guides and visitors, looking into guided tours as a sociocultural activity, which is influenced by the traditions and practices of museum contexts. The project of our students aimed at designing an interactive solution to enrich learning practice and social interaction in English language class of the Nyborg gymnasium.

4.1 Audook: social experience of audiobooks

The Audook mini-project by one group of three MMD students explored how interactive fruition of audiobooks could enrich learning practice in classes of English literature and language, with the cooperation of a gymnasium teacher (here called Sanne) and her class, 15 students of approximately 15-16 years of age. The outcome of the Audook mini-project represents an attempt of transduction of reading assignments from the visual to audio mode. Transduction is defined in social semiotics as a translation, in which meaning-material is moved from one mode to another, for instance “from speech to image, from writing to film” [12]. Since each mode has specific material qualities and entities to be manipulated, for instance speech has words and images have colors, each mode has also a different history of social use. This in turn has implications on how the same meaning-material is formulated and transmitted by the
sender, and on how the message is received and interpreted by the audience, so that the same message might be slightly altered in its meaning through the transduction process. Audiobooks represent for instance a case of transduction from the visual book format into an auditive one. As showed by related studies like [2], the fruition of the same story both through reading and in audio form affects significantly how leaners experience reading, in some cases even enabling them to improve their skills.

Through their field study the three MMD students found that English classes in Nyborg, involved mostly reading and analyzing texts. The English teacher Sanne was concerned with choosing samples of English literature that the students could find interesting to “motivate her pupils to read and analyze the texts”. For this reason she said: “I am trying to look for novels that can be interesting, handling topics about social relations and adventures”. Her strategy involves “books that have become popular in recent years, often because they were adapted into movies, so that they have heard about them”. During our study for instance the class was reading “The Beach” by Alex Garland, which is also the subject of a popular Hollywood movie starring Leonardo Di Caprio. In this way the teacher was already encouraging a multimodal fruition and analysis of the assigned novel.

Fig. 2 Overview of the interface of Audook. The top row shows the log-in screen and the initial access to the audiobooks library. The second row shows how the text can be visualized by the reader, and the sharing and annotation features.

We found that the Nyborg students are typically assigned a set of pages or entire chapters to read for a certain date. While in class they are asked to discuss in groups the read chapters and to fill a form with questions or aspects to reflect upon, such as the maturation of a character, the social conflicts, or narrative techniques adopted by the writer; afterwards, a group discussion is conducted in class. The students also watch the movie based on the novel they are reading, together with the teacher. This is supposed to keep them motivated to read and reflect on how the novel could be interpreted, and Sanne added with satisfaction “they often prefer the novel to the movie!” as the students notice that in the movie many elements were omitted or the actors representing specific characters do not match their imagination.

The gymnasium students complained, however, that reading requires a “total” involvement; several of them said that they can read mostly while on the bus or at home, but unfortunately they cannot read while running or walking in town. Reading is also perceived as isolating, so that for sharing impressions on specific passages they have to either meet or write through social media.

The design process that led to the creation of Audook, an application aimed at providing an alternative fruition of literacy texts. The central idea was to operate a transduction of novels into audio, and create a gesture-based app for mobile phones. The requirements involved being able to use a hand gesture to add a bookmark on a specific passage, while listening to an audiobook; users should also be able to add...
comments in spoken and in written forms by opening a visual interface, and share their comments and bookmarks through social media.

The resulting prototype mobile application (visible in figure 2) offers a richer, multimodal experience than just reading and showcases the extension of annotating and sharing comments from a book to an audiobook. A summative evaluation provided criticisms and positive feedback. For example, several students from Nyborg were critical towards audiobooks, five different students argued that they cannot imagine to “replace ordinary books with audiobooks, because audiobooks take longer to ‘read’” In general the students said that they were uncomfortable in turning visual reading entirely in an auditory experience, saying that “reading books gives you more control than listening”. Both we and the group developing Audook agree with these statements that it is not a good idea to substitute visual reading entirely with audiobooks, because of the importance of reading as a cognitive skill and of seeing the text, especially in language learning. In that respect, we explained to the students of Nyborg that it was not our goal to replace visual books with audiobooks, but that we envisioned interactive audiobooks as expanding the available resources for reading with the goal of enriching the experience of reading. The students also expressed their doubts on the quality of the voices obtained via text-to-speech, with respect to those of actors and native speakers reading the texts. Concerns were raised by the teacher in relation to how she could fetch audiobooks for her students; the fully developed application should be able to connect with the collection of audiobooks of the school or of the local library, which is already available online, enabling the teacher and her students to easily get the novels they need.

On the positive side, the students from Nyborg acknowledged that audiobooks can be “read” also while doing sports or other physical activities, someone said: “I could continue learning about the book while walking or running”. Another student argued: “I often get sick when I try to read on the bus, but I still want to use that time to study”, similar statements were expressed also by other students, who acknowledged that audiobooks can be easier to access than books (and e-books) while travelling on public transportation with less chances of motion sickness. It was also asked by a few students if it was possible to listen to an audiobook while watching the e-book version (a scenario similar to existing karaoke applications): in this way, users could learn more effectively how to pronounce new words. Generally, the students commented positively on the social interaction that Audook should support, as when fully developed Audook should enable users to share electronically books critique and commentaries in preparation for group discussion in class. This social scenario was perceived by the students in Nyborg as a natural enhancement of their regular learning activities in class.

Finally, the Audook app was positively evaluated as an interactive alternative to normal reading, expanding opportunities for multimodal fruition of novels and for sharing personal reflections on texts. In general the social aspect of the application and the possibility to listen to the story while engaging in outdoor activities were particularly appreciated as if they were making the experience of reading less isolating.

5. Discussion

The main case study and the 2 supporting studies show the wide spectrum of opportunities offered by audiobooks in language learning, from content generation to social and game-like interactivity. The main contributions of this paper are design insights to make audiobooks interactive and better integrated in the social interaction emerging in learning contexts, between learners and teachers but also among peer learners. At the same we aim at exploring how the transduction of literary texts could foster different experiences, when moving from the visual and tangible modes associated to the experience of physical books and e-readers, to the auditory modality enhanced by interactivity.

Comparing the three case studies (as summarized in table 1), we can see that interactive audiobooks are preferred to non-interactive ones by potential users, who in our testing consistently described typical audiobooks as eliciting passive experiences. Interactivity with the text was evaluated positively both in relation to exploring non-linear stories, but also in contexts of language learning (Alcantud-Díaz and Gregori in [2]). As pointed out by Kress in [12] the transduction of literary text into an auditory format can significantly alter how readers relate to the text.
Case study | Participants/stakeholders | Main findings
--- | --- | ---
Audio deliverables | 2 Danish primary schools, 2 classes of 4th graders learning English and 2 teachers | Audio deliverables can ease teachers’ orchestration of class activities enabling asynchronous teacher/pupils interaction. Audio deliverables can serve as learning diary to enable students to keep track of their progress. Pupils were more at ease creating audio content than written. Direct way to evaluate and practice pronunciation.

Carbooks | A focus group composed of 10 young adults (age 19 to 25) and a family with 2 kids (10 and 12) | Users found difficult to interact without the graphical user interface and the use of the microphone. Users were positive on the use of binary options through the story, instead of having multiple options per choice as in normal game books. However, binary options limited the non-linearity of the narrative. Users found smart phones the ideal platforms for the fruition of audio stories.

Audook | Nyborg Gymnasium, English language class, 15 pupils and their teacher | Audiobooks can complement but not replace reading. Audiobooks make reading into a less isolating and flexible experience. Social sharing of critiques through the app can naturally enrich class activities.

| TABLE 1: A summary of the findings in our 3 case studies. Each study lasted for 1 semester and involved iterative development and testing of prototypes, in various schools in Denmark. |

The auditory modality can make the reading activity more flexible and accessible for learners, for instance the possibility to create audio deliverables can support adoption of pedagogical approaches like visible learning (Hattie and Gan in [9]), in which learners and their teachers can afford longitudinal monitoring of spoken language competences. The recordings created during language learning open the possibility to apply analysis techniques and data mining on audio content. This is valid also for learners who have a busy day and see in the auditory fruition of novels a better support for multi-tasking, enabling them to “read” also when traveling and reading might get them sick and when engaging in outdoor activities. Moreover, the audio modality can better support children who are still in the progress of developing writing skills in their own or in a foreign language, as well as learners with linguistic difficulties. Finally, the study in Nyborg provides new insights on how interactive audiobooks could contribute to turn reading into a social experience, as according to sociocultural theories in learning. Adopting a sociocultural perspective (like Rogoff in [19]), learning is seen as a social practice in which learners are facilitated by an expert adult, the teacher, but can also support each other, in a persistent and asynchronous way. Enabling learners to share their thoughts and bookmarks with each other, Audook can contribute to the emergence of a shared understanding of the text at hand enriching the process of textual analysis and reflection.

Building on these case studies, we propose insights on how audiobooks could be turned into an interactive medium:
• Support generation of audio as well as fruition. Audio just requires a bit of technical support, for example, Google docs can be extended to allow voice comments on texts, by using add-on like Kaizena\(^6\).

• Leverage on social and asynchronous communication between teachers and students, and provide support for peer-learning.

• Consider multiple storylines in audiobooks. Multiple storylines can allow for experiential learning (as discussed in \([7,2]\)) and support case-based reasoning. A major drawback of authoring non-linear narrative is the need to create multiple, potentially modular storylines; non-linear audiobooks in particular have always been human-intensive. Our Carbooks project, however, shows that text-to-speech technology is currently widely available (on laptops and even mobile devices) and good enough at least for English. All teachers in the schools we visited have at least basic IT skills, hence they have no problem in generating English texts and potentially create written non-linear narrative; our experience with Carbooks convinced us that by leveraging on text-to-speech and gesture-based non-visual interfaces, non-linear audiobooks in English can potential be created by the teachers themselves, in this way supporting language learning.

• Socially generated audio content as a kind of social media data. We suggest to consider the audio content generated by a group of students learning English as similar to the content produced in a social media. Since voice data-mining is still very complex and dependent upon pronunciation, often imprecise and typically works for English and very few other languages, we consider social media approaches like user-created tags as the best option to classify and search through audio contents.

• Audio as a complement to visual modality. Based on our studies we do not aim at replacing the visual modality of reading, but at providing complementary auditory alternatives that could enrich how people experience literary texts.

The exploration of interactive audiobooks is not new, as we can see in current research, however, we may argue that these studies have taken a limited perspective, mainly supporting the authoring of non-linear stories. On the other hand, when coming to learning these studies seem eager to argue that audiobooks can offer better support to learners in acquiring linguistic as well as intercultural competences (for instance in \([2]\)). In our studies we take instead a more cautious position, as results from our testing suggest that visual reading is perceived as more personal and active, as readers can decide for themselves how quickly they want to read, they can imagine for themselves the features of a character or a setting. At the same time audiobooks do not allow for that freedom, as they impose a specific timing and the voice of the reader, which could be found unpleasant or expressing feelings in an inappropriate way for the sensitive of the listener.

Audiobooks have many faces (or voices) and seem to us to possess untapped potential. The students from Nyborg gymnasium appeared eager to identify both the new possibilities offered by the Audook application, but were also aware of some intrinsic limitations of audiobooks.

6. Mockup

Based on the above insights on audiobooks, we created a mockup of the Audook application, using YouTube. We were interested in testing some use-cases with a simple semi-working prototype, so we started by finding a free audiobook and uploading it to the YouTube channel of one of the authors. We used a free, read-aloud version of the copyright-free book “Frankenstein; Or, The Modern Prometheus” by Mary Wollstonecraft Shelley\(^7\); the Project Gutenberg has many free audiobooks read-aloud by volunteers, as well as the corresponding copyright-free book, which makes it possible to follow the text while listening. The audiobook is provided in multiple compressed audio files, one per chapter. We decided to focus on chapter 5, where the protagonist Victor Frankenstein reanimates dead organs to create his monster. However, YouTube only allows free uploading, sharing and annotation of videos, so we generated a video

\(^6\) Kaizena’s webpage: https://kaizena.com/ (last seen on the 23th of March 2017).

\(^7\) A copy of Frankenstein is freely available at: http://www.gutenberg.org/ebooks/20038 (last seen on the 23th of March 2017).
version of the Frankenstein audiobook, adding just a single static frame for the entire duration of the video (a freely available cover image of an old edition of the same book).

The video for chapter 5 of Frankenstein’s book was uploaded on the YouTube channel in two copies, so that each of the authors could proceed to annotate independently. This was done to enact one of our use-cases, where multiple students attach annotations on multiple copies of the same audiobook on an online repository (here a YouTube channel) that they share, class- or group-wide.

The annotation process itself proceeded as follows:

- **Login in personal YouTube account, and access the channel**
- **Open the “Creator Studio” page (see figure 3)**
- **Create annotations, specifying the text to display and the start and end time for the text to be displayed (as visible in figure 4)**
- **Create a set of deep links (also known as chapter markers, in analogy with movie chapters in DVDs) to allow users to jump directly at the beginning of a specific annotation without having to listen sequentially to the audio book. Figure 5 shows how deep links are created and how they appear to the user.**
- **Manually write a webpage that embeds multiple fragments of the YouTube video (via iframes). By listening to each fragment in sequence one obtains an overview of the audiobook, a kind of audio summary of Frankenstein.**

The annotation process lasted a week, and then we shared our videos and commented on them using YouTube comments. This represents another of our use-cases for the Audook application, where students should be able to leverage on social media to exchange their annotations and work as a group to build a more deep understanding of the book they are listening to.

Using our mockup, we quickly come to the realization that annotations on audio contents work in a similar way to annotating user footage in qualitative research methods. We refer in particular to visual ethnography (Pink in [17]) and interaction analysis (Jordan and Henderson in [11]). These methods are widely adopted in research through design (Zimmerman in [24]) in which scholars engage in a design process with the goal of pursuing new knowledge about the design of a specific category of product, for instance e-learning applications targeted at schools, like our own. These inquiries can also aim at finding new knowledge about design methods, addressing issues like how to better conduct design processes, or in other cases a design process can be undertaken in order to discover more about human nature, psychological responses to certain conditions or cultural values. In all these kinds of studies, the researchers observe and shoot video footage on the activity taking place in context, afterwards they will analyze video footage from observations in the field to identify specific moments pointing at daily practices and issues.
that could be addressed in the design process. This particular way of using video material in design inquiry is found particularly relevant when gathering knowledge about tacit routines, that are actions or habits people engage in without rationalizing them, simply as part of their daily work, but which might be vital to the fulfillment of the activity itself.

![Fig. 4 YouTube provides a good editor for creating annotations, that allows for specifying start and end time for when the annotation will be displayed (during playback) and of course the text to be displayed](image)

As people are not fully aware of these routines or do not consider them worth of attention, they would not likely mention them during interviews. In order to identify potential issues and tacit routines, the researcher is supposed to scrutinize and edit the video material collected during field work, for instance: a researcher might need to add bookmarks to the video footage using specific software when interesting events are taking places, to annotate the video adding commentaries or simple keywords in order to remember was a particular event worth of attention, and to cut short clips or capture screenshots which might serve for reflection or documentation to attach to research articles. Bookmarks, annotations, clips and screenshots can be seen as providing support for reflections and creative discussions on the direction to follow in the design process.

![Fig. 5 Editing of deep links for the video, on the left. How deep links look when the user listens to the audiobook (on the right)](image)
In another use-case, we imagine a student (or a group of students) interested in marking various kinds of passages in the audiobook. This corresponds to marking in different colors on the pages of a physical book, and giving a specific meaning to each color. Color coding can be used also with YouTube annotations (see figure 6).

![Figure 6](image)

**Fig. 6** The text of the annotations is colored red. Here we decided that red annotations will be used to mark passages that describe Frankenstein’s monster. Other colors will be used to group different aspects of the story.

We also wrote a clickable index of all color-coded annotations for chapter 5 of Frankenstein, grouped by color (i.e. by meaning), so that during listening the YouTube player allows to jump directly to any of the colored annotations, hence offering a semantic chapter marking to the listener (figure 7). To further explore these semantic annotations, we manually created a webpage with an audio summary of Frankenstein, as visible in figure 8.

![Figure 7](image)

**Fig. 7** Here deep links (on the right) help the user navigating through the audiobook, by jumping directly to fragments that highlight specific aspects.

We decided to create our mockup using YouTube annotations, but audio files with accompanying annotations are another possibility. Playback of interactive audio might sound more complex than just listening to an audio file, which can be done on any device with a standard player. Interactive audiobooks would seem to require special apps or web-based services in order to understand the subdivisions inside the audio resource, and to be able to skip among fragments during playback. However, common file formats
might provide simpler ways to express chapters in audio files: for instance the MP3 ID3 tags\(^8\) have proposed already in 2005 a specific format for defining chapters in MP3 files. Unfortunately the ID3v2 standard is currently not supported in many MP3 players.

![Audio summary of Frankenstein](image)

**Fig. 8** This figure shows the webpage created to summarize chapter 5 of the Frankenstein audiobook. The audio fragments (represented here by multiple embedded YouTube videos with a specific start and end point) are grouped by color (i.e. by meaning). Listening to all fragments from the same color group provides a summary of a particular aspect of the audiobook: e.g. red annotations are all about the way the monster is described, while green annotations explain how Victor feels.

In conclusion, our mockup is based on the insights to turn audiobooks more interactive, and helped us to explore the feasibility of use-cases for the Audook application. The mockup also had clear limitations:

- In YouTube annotating is not possible while listening
- The mockup offered no support to keep the annotations coherent
- It relied on manual intervention to create deep links
- It required us to manually compile the audio summary using deep links, and group them by color
- Moreover, annotations and deep links do not work on mobile devices, which is the target platform we have in mind for Audook

We are currently building a prototype based on a client/server system, using Javascript and jQuery for the client part, and Node.js together with MongoDB as server.

### 7. Conclusion

The main contributions of this paper are insights on how to make audiobooks interactive and better integrated in learning contexts, in particular when learning English as a foreign language. The three case studies discussed show the large spectrum of opportunities offered by audiobooks in language learning, from content generation to social and game-like interactivity. The prototypes developed with our students provide evidence that audiobooks can help in documenting learning (thanks to audio deliverables), in supporting different learning experiences and styles, and in complementing visual information when exploring non-linear narrative. Moreover, we created a mockup using standard YouTube annotation tools, which provides a preliminary feasibility study in how our insights can be turned into use-cases and implemented as web-pages or mobile apps. We believe that the experience obtained in the three studies and the insights we gained can be used as design guidelines to develop more interactive audiobooks and audio-

\(^8\) The definition of the ID3v2 standard can be found at: [http://id3.org/id3v2-chapters-1.0](http://id3.org/id3v2-chapters-1.0) (last seen on the 23th of March 2017).
enabled applications. A fully functional mobile application is currently under development, based on the outcome of the main case study and the experience gained with our mockup.

References