A Rapidly Growing Electronic Publishing Trend: Audiobooks for Leisure and Education

Jan. J. Engelen

Kath. Univ. Leuven – DocArch group – ESAT-Dept. of Electrical Engineering
Kasteelpark Arenberg 10, B-3001 Heverlee – Leuven (Belgium)
jan.engelen@esat.kuleuven.be

Abstract

This contribution focuses on the relatively new phenomenon of the purely commercial availability of audiobooks, sometimes also called “spoken books”, “talking books” or “narrated books”. Having the text of a book read aloud and recorded has been for a very long time the favourite solution to make books and other texts accessible for persons with a serious reading impairment such as blindness or low vision. Specialised production centres do exist in most countries of the world for producing these talking books. But now a growing number of commercial groups have found out that there is a booming market for these products as people slowly get used to leisure listening to books instead of reading them. Some companies claim already having over 40,000 titles in spoken format in their catalogue. Major differences and possible synergies between the two worlds are discussed.

Keywords: audiobooks; talking books; spoken information; commercialization

1. Introduction

Electronic equivalents of printed books (e-books) have been around for a long time now; also multimedia documents have become more and more popular. Especially the spoken variants of books are continuously gaining popularity. Up to a few years ago, producing talking books was seen uniquely as a service to support reading-impaired persons but nowadays commercial interest is growing at a high pace.

We start by comparing the traditional specialised production processes with their equivalents in the commercial circuit. This will anyhow involve also some technical aspects. Digital Rights Management and copyright challenges are handled too. Finally we discuss a few implications of this phenomenon on the organisation of libraries and related cataloguing issues.

2. Specialised audiobook production centres

Most audiobook production centres in Western countries that are focussing mainly on consumers with a reading impairment, have now abandoned cassette distribution in favour of CD-based solutions. Cassettes had been around since the beginning of the sixties but recording, erasing and checking returned cassettes remained very time consuming activities for these production centres. Furthermore most books had to be put on a series of cassettes (due to their limited storage capacity) and clear indications on the cassettes and their boxes, preferably in Braille, were needed to keep some order in such a collection. But even at that time several measures for protecting copyright (nowadays called Digital Rights Management, DRM) were taken: special cassette formats or non-standard tape speeds were used to have some copying protection.

In the middle of the nineties internet technology and especially web documents with hyperlinks to other
documents became widespread. Within the European Digibook project several hybrid books were developed, containing both the text and the linked audio files of the same book. The linking was done on sentence level [1]. Similar initiatives were developed at the Swedish production centre TPB.

In 1996 a large group of specialised production centres on a global scale has created the Daisy consortium [2] in order to study and to standardise the future audio recording of talking books and, very importantly, how a navigation structure could be added to the books in question.

This lead to the Daisy 2.02 and 3.0 standards which have been turned into US standards by NISO but are accepted worldwide by all specialised production centres in order to permit the exchange of this new generation of audiobooks.

Most centres distribute their productions nowadays on a data CD [3], or to a minor extent via the internet [4]. Data CD’s permit a trade off between quality and recording speed that is not possible for audio CD’s (e.g. with music). As the human voice can be recorded with a much lower sampling frequency than high quality music, data CD’s can contain easily 50 to 70 hours of speech.

Technically the Daisy format describes the content of the book (in XHTML or XML type files) while the audio is recorded as a collection of mp3 files (.wav is rarely used). The Daisy CD can also contain the text of the document and a whole series of timing links (in SMIL format) between the two. That way one can have a computer or reading device searching the text content but the user still can listen to the corresponding audio output.

Furthermore a Daisy book permits easy and rapid navigation through complex documents as up to six levels of table of contents are possible.

Daisy books are read with computer programmes (AMIS, Easereader, TPB reader…) or special players. These players are actually CD-ROM readers with Daisy reading software. Some look even like a CD walkman.

Since last year mini Daisy players have reached the market. These smaller devices with PDA or mobile phone dimensions use SD memory card readers instead of CD’s.

The next generation of these devices will connect automatically (through WiFi) to the internet and will then automatically download books (or newspapers, cf. below). Currently only a few UTP-cable connected devices (Webbox, Adela) do exist but their WiFi versions are under development.

3. Commercial audiobook production

3.1 Booming commercial audiobook popularity

Over the last years we have witnessed an enormous increase in audiobook popularity outside the “traditional” user group of persons with a visual impairment. Several commercial groups, some linked to traditional publishers, some completely new ones have popped up. Audible.com [5] is the leading online provider of digital spoken word audio content in English, specialising in digital audio editions of books, newspapers and magazines, television and radio programmes and original programming. Through its web sites in the US and UK and alliances in Germany and France, Audible.com offers over 40,000 programmes, including audiobooks from well-known authors such as Stephen King, Thomas Friedman, and Jane Austen, and spoken word audio content from newspapers including The New York Times and The New Yorker. However these newspapers are only made available in excerpted form.
Meanwhile in Belgium and in the Netherlands (two small countries, 15 million Dutch speaking inhabitants) about a dozen specialised publishers have popped up in a short time. Curiously enough customers will seldom buy audiobooks in bookshops: they seem to be used to downloading music and therefore expect also audiobooks to be downloadable. On the other hand many public libraries have reacted to an enormous interest in audiobooks by adding them to their collections.

There is also a growing interest in spoken versions of education material and course material [6].

3.2 Technical formats and standards for audiobooks; copy protection

A very important issue is the type of audiobook standard that is used. As stated above, within the sector of audiobook production for reading-impaired persons the Daisy standard is very common (and in fact globally accepted). Commercial publishers on the other hand do NOT use the Daisy standard but rely on several alternatives for distributing their audiobooks:

- Some companies provide documents on standard [7] audio CD’s (e.g. Dan Brown’s “Da Vinci Code” spans 13 audio CD’s). The main reason for this choice is the universal usability on any audio CD-player developed since 1980.
- Others use data CD’s with audiofiles in mp3-format and for reasons explained above. Up to 40 hours narration on one CD is not uncommon.
- A few, including the largest one (audible.com, cf. above) provide their audiobooks in a DRM-protected format. Some of their more expensive books however can be burned onto (a pile of) audio CD’s by a legal buyer. A special version of NERO CD writer is needed to do this.
- Audible.com has developed the proprietary “.aa” format and provides free software for playing (legally acquired) .aa files on 290 platforms. This format also caters for different quality levels.
- Apple i-Tunes used mainly the proprietary MP4 format (a container format, including the media and DRM info) which made it impossible for some time to use the files on non-iPod players.
- Since the beginning of 2007 more and more music on the internet became available without DRM, although generally at a somewhat higher price. Many see DRM now as a thing of the past (cf. below).
But the most striking difference of all these solutions with the Daisy format is the lack of any sensible navigation system through the audio files. The available solution, the Daisy standard is not used in the commercial audiobook world!

A very important aspect of audiobook (and music) distribution on CD’s (or via the internet) is copyright protection, often seen as copying protection. Digital rights management was once seen by the music industry as the method to prohibit illegal copying. In practice however DRM lead to quite a lot of customer frustration as it hindered copying in general or made it sometimes impossible to play the legally acquired files on a whole series of devices.

In practice, all widely-used DRM systems have been defeated or circumvented when deployed to enough customers. Protection of audio and visual material is especially difficult due to the existence of the “analogue hole” [8], and there are even suggestions that effective DRM is logically impossible for this reason. A much more complex situation for illegal copiers arises when books become interactive and the sequential nature of the narration is abandoned.

3.3 Business models

A special audiobook issue is the business model used by the publishers: some companies, including again the largest one, prefer a subscription model with monthly instalments – worth approximately one audiobook. Audible.com’s business model is closely mimicking the well established marketing system of “book clubs”, i.e. one gets the possibility to download a number of books by paying a monthly membership fee. Buying individual books is possible too, but at much higher prices. Its main competitor LeisureAudiobooks, on the contrary, stresses that no subscription is required (cf. figure 1)

Others charge different prices for the different audio qualities available. E.g. at audiobooksforfree.com, the lowest quality is for free but users are charged for better quality files. In fact the company stores high quality audiobooks but degrades them for those who want to pay less.

3.4 Audio: human voice vs synthetic voice

A major distinction between audiobooks must be made according to the type of audio: is the narration done by a human person or by a computer (synthetic voice)? Everyone agrees that, even still nowadays, human voice is much more agreeable to listen at than synthetic voice although very good quality text-to-speech software (TTS) is available.

However, for some applications only electronic conversion is an option. E.g. during the production of the
spoken Flemish daily newspaper project – Audiokrant - with full text coverage of all articles, there are only some 30 minutes available after copy closure time to produce 12 to 15 hours of speech and to physically record the subscribers’ CD’s [9].

3.5 Growing synergy between commercial and not-for-profit audiobook publishers

Up to now, the worlds of the commercial and the not-for-profit publishers have been very segregated. Commercial publishers often state that their products also benefit reading-impaired persons but they show no interest in using the Daisy standard.

On the other hand, specialised production centres are clearly exploring the commercial possibilities of the large archive of spoken books most of them have created over the past years.

Sometimes, specialised and commercial productions go hand in hand. The Royal National Institute of the Blind (UK) recording of Terry Darlington’s ‘Narrow Dog To Carcassonne’ has won the APA ‘Audies’ award for 2007 in the category of best unabridged non-fiction. The book was produced by RNIB both as a DAISY Digital talking book for RNIB clients and also as a commercial audio book on CD (ISIS publishing).

In the Netherlands, the largest specialised audiobook production centre, “Dedicon” has created at the end of 2006 a commercial branch, named Lecticus [10]. Mainly linearly organised books are provided as a series of mp3 or wma files. Books can be downloaded but also can be delivered on a cheap audio mp3 reader (USB stick size).

4. Cataloguing Issues

The problem of how to find an audiobook in a library is clearly somewhat complicated by the fact that the number of productions centres is increasing rapidly.

Furthermore a comprehensive cataloguing process for audiobooks requires a whole new series of descriptive items including but certainly not limited to:

- flags for abridged [11] and unabridged versions; a field for total reading time;
- fields for technical recording specifications (e.g. audio quality/sampling frequency; file types, use of Daisy standards 2.02 or 3.0 etc.);
- a field to distinguish between recorded/synthesized speech;
- fields for the narrator’s details (Experienced narrators or books read by their author constitute selling arguments for commercially produced books!);
- flags for pronunciation details (UK English vs American, Austrian or Swiss German vs Standard German, Dutch vs Flemish intonation etc.). No standard for covering these subtle language differences is available;
- fields describing the audio-to-text linking mechanisms used in the audiobook (if the text is made available too): synchronisation between text and audio on a word, a paragraph or a page level.

Some of these requirements resemble the cataloguing needs for books in large print. These topics are actually under the remit of a special section within the International Federation of Library Associations (IFLA) that caters for the needs of reading-impaired users [12].

5. Conclusions
Due to the explosive booming of commercial audiobooks, a huge number of titles theoretically becomes available for reading-impaired users too. This process requires however new business models for the traditional specialised centres and probably also a completely new societal vision on who is willing to pay for what type of audiobook service in the future.

6. Notes and References


[4] Downloading is not yet a very common procedure due to most internet providers’ data download volume restrictions, although this is changing rapidly to permit more multimedia downloads.


After having been a minority shareholder for some years, Amazon.com fully acquired Audible.com on January 31, 2008.


[8] The “analogue hole” means simply that any audio or video signal has to be transformed into an analogue signal to be interpretable by human beings; but analogue signals can be re-digitised afterwards.

Internet music stores have more or less given up DRM protection. E.g. it was found that a new iTunes music track (with DRM) made available from Apple needed less than 3 minutes to become available elsewhere on the web in an unprotected audio format.


[10] Lectorius audiobook shop: http://www.lectorius.nl
