DAISY BOOK PRODUCTION AT THE NATIONAL ORGANIZATION OF THE BLIND (ONCE)

Rafael Pérez Arnaez
Servicio Bibliográfico ONCE
C / La Coruña, 18
28020 Madrid
SPAIN

ABSTRACT

DAISY book production at the ONCE, with the digitization and subsequent reformatting of our stacks to the DAISY standard, has pursued two key objectives: to digitize our collection as quickly and efficiently as possible and to lower the production costs of digital talking books in DAISY format. This discussion describes the highlights of that process.

1 The ONCE Bibliographic Service

The ONCE began to produce talking books in the 1960s. That means that the organization has a huge collection, covering a wide range of topics and recorded using any number of methods. Nearly fifty years later, conservation and recovery, as appropriate, of our recordings and adaptation to the possibilities afforded by new technologies had become absolute imperatives.

The ONCE is a full member of the DAISY Consortium since 1996, the year when the first trials were conducted to create books in the DAISY 1.0 standard, which has since all but disappeared.

The ONCE has vested its Bibliographic Service (Spanish initials: SBO) with responsibility for furnishing its members with all (primarily written) materials that blind people and print-disabled people need to access education, employment, culture, leisure... in a word, information.

The human and material resources directly involved in producing audio, Braille and relief materials for the aforementioned purposes are grouped, then, under the SBO umbrella. It has a head office (SBO) with two locations, in Madrid and Barcelona, five services under its aegis specializing in the bibliographic adaptation of educational materials, one at each of the organization’s educational resource centres, and 33 documentary adaptation units, one in each ONCE regional department and Administrative Centre. Each node on the network has a specific purpose and works in conjunction with the other centres and services.

Our priorities in bibliographic adaptation obviously depend on the needs of our members, be they students, employees or job applicants. That does not mean, however, that we neglect the adaptation of books and other bibliographic materials (primarily periodicals) that provide general access to culture and entertainment.

1.1 Digital book production. The beginnings

Initially, while we waited for the first beta versions of LpStudio software to become available, we tested and analyzed our needs and the tools we would have to use. We knew very well that unless we digitized our audio collection, we risked losing it. The hardware and networks needed were permanently installed in the year 2000. To prevent noise from interfering with our recording studios, we opted to install a 1-Gb fibre optic network for our audio network communications. That phase of the process was characterized by computer installation as well as staff recycling and retraining to learn digital technologies. The underlying premise was that staff should be able to perform their tasks as independently as possible. We wanted our narra-
tors, for instance, to be able to make and fully edit their recordings, delivering a final master for control and subsequent compression with no need for help from a sound engineer. This was a gradual process that culminated around 2003, when we undertook the digitization of our collection in earnest.

As we reported on an earlier occasion, the ONCE has over 200,000 hours of recorded material in its audio collection. And much to our misfortune, both the media involved and their state of conservation vary widely. Consequently, the use of a direct system, specific hardware for instance, to digitize the entire collection was unthinkable. The process necessarily had to be conducted title by title, with a specific solution for each problem that arose: in several four-track cassettes, for instance, we found substantial differences in signal quality from one track to another.

1.2 Production process characteristics

In light of the foregoing, the ONCE took a three-prong approach: conservation of the analogue collection, use of the full potential of new technologies to create talking books in these new formats, and continuation of analogue production during a transition period.

1.2.1 THE DIGITIZATION PROCESS

Since neither the quality, noise nor signal of our recordings was standardized, the obvious solution was to digitize with software.

Essentially, the system involved first digitizing all the material semi-automatically and then converting the output to DAISY.

By simultaneously digitizing eight tracks at double speed we were able to convert 12 hours of recording in approximately 45 minutes. The software designed in-house enabled us to retrieve the tracks as 22,050-Hz audio files, directly.

The next step was to review the files to analyze quality and minimize a series of previously established parameters such as background noise, signal differences in one and the same file and so on.

Once the recordings were approved, they were filed on duplicate CDs for subsequent conversion to DAISY.

One of the things that had to be borne in mind was that these same files had also to be usable for producing cassette tapes. For this reason, both the format and the number of tracks per CD had to be as required to ensure that books could be produced by a computer connected to an external cassette recorder using four tracks and half the normal reproduction speed.

1.2.2 CONVERSION TO DAISY

Converting the digital books to the DAISY standard posed other types of problems. Initially, the intention was to produce books with a structure as complete as possible. We could have converted our collection to DAISY based on a very simple structure, for instance, at a single level, or we might have created a new structure from time to time. The more comprehensive approach naturally entailed investing much more time, particularly considering that the conversion was not automatic, but performed on a book-by-book basis. In the first few years, we found that we would be better advised to do the conversion ourselves, for a number of reasons:

We knew better than anyone else how the recordings had been made.

We could devote more time to audio aspects when we deemed it necessary and in our own interest to do so.

Because of the nature of our organization, we assumed that we would deliver more detailed structuring, more in keeping with our own philosophy, than any other company. This means that we were in a better position to more readily ascertain the type of structure and detail needed for each book. Moreover, we had specialized staff able to decide the degree of detail required in any given structure.
This was the procedure used in 2002, 2003 and 2004, when we were not yet delivering DAISY books to our users. The number of works digitized in those years grew steadily.

Up until that time, DAISY books accounted for a mere token percentage of the total, but as our narrators gained experience with the system, that number also gradually climbed.

### 1.3 Books distribution

We began to distribute DAISY books in 2005. This placed an additional work load on our staff, for we not only produced, but also distributed our production and also had to handle possible incidents, ranging from distribution to recording or other errors detected in some of the books.

Under these circumstances, the SBO decided that book production would have to be outsourced to private companies to raise output and lower costs. We therefore favoured the creation of several such companies and asked them to produce books using PLEXTALK PRS software. The DAISY books we are presently recording feature audio and NCC. The PRS software proved to be extremely well suited to these books, because it called for scant user training. Moreover, it could be used to structure books in DAISY from the digitized format. One of the companies that undertook this task was FUCODA, a Fundación ONCE firm which, like TBS, another participant, is an equal opportunities employer that prioritizes the hire of people with all manner of disabilities.

This, as the graph shows, led to a perceptible increase in our output while lowering our costs by reducing the need for specialized staff. Indeed, presently our primary tasks consist in quality control and DAISY mastering.

Our current output is on the order of 250 books per month.

Note the steep decline between 2004 and 2005, which is when we restructured and outsourced our production. But that trend was soon reversed, with the added advantage of quality controls that were absent under the former arrangements. In 2005 the Barcelona SBO also became involved in the production of DAISY books, ushering in the decentralization of book production and distribution both. Approximately 39,000 hours of talking books are presently produced yearly, which translates into an annual output of nearly 3,000 books.

In 2008 we permanently discontinued cassette tape recording. Interrupting the production of cassette tapes entailed the interruption as well of initial digitization. In other words, the books produced on request today are recorded in the DAISY standard from the outset, cutting out one of the steps in the reformatting process.

### 2 Present production process

The present production process is based primarily on the requests received from ONCE members. But at the same time we are creating a DAISY library in all the ONCE regional departments and administrative centres to ensure the shortest possible turnaround for members, particularly those who live in remote areas or far away from such centres.

The digitization process per se is nearly complete: only a few works on tapes and of relatively scant interest have yet to be converted. At present, works are only digitized when explicitly requested by one of our members.

In any event, the production process is the same. Users have access to a book catalogue either online or in their regional departments and can request any of the books available.

When a request is received, the first thing we do is check to see whether it already exists and in what format. If it is in DAISY, a copy is made and sent to the member’s address. This is a very speedy process, for books...
actually ship on the day the requests are received. If a request lodged by a user with his or her regional department or administrative centre is available locally, a copy can be made then and there.

When the work is not on file in any of our production centres, an order is placed for the respective recording. The SBO channels production to the Braille or audio department and, in the event of the latter, to in-house or outsourced production. Once the book is recorded, it is checked for quality and the respective masters are generated: the original for filing and the compressed version for distribution directly to the user and for uploading to the ONCE’s digital library.

Requests for books can be made by e-mail, telephone or letter, at the member’s discretion. Students usually send the book they want to have recorded together with their request.

The tools we use are LpStudioPro, Sigtuna DAR3 and MyStudioPC. The recording companies, in turn, use PRS software.

Well-known software such as Validator and Regenerator is also routinely used for quality control.

We are presently running tests with synthetic speech to streamline book production. Trials are also underway with the DAISY 3 standard, which has not been adopted yet, among other reasons because of the impact it may have on our users, particularly the elderly whose ability to adapt to change is more limited.

In addition, we are analyzing the possibility of generating XML files from the text processing software used to produce books in Braille. As in the case of the American NIMAS standard, these files could then be used to generate either Braille copy or DAISY recordings.

3 Storage systems

Books have not only to be produced, but must be systematically stored. We are presently using two different storage systems: formerly, we recorded all the original DAISY masters on CD-ROMs, i.e., without compressing the audio, and the masters for distribution, with compressed audio, on hard disks.

To streamline order dispatching, each book was imaged for quicker CD recording and delivery. With this system, recording and printing a CD took around two to three minutes.

This procedure has been largely maintained, with certain changes. Firstly, the SBO has created a digital library containing all our DAISY books. This in itself is an important change in our users’ habits. We will have the opportunity to talk about the library in a separate presentation.

Moreover, recording, printing and storing CDs is a complex task, not to mention the space required. Storing our present collection of 250,000 CDs is costly and access to the information is slow. The CD containing the work requested has to be located, physically inserted into a computer and downloaded for processing each time it is needed. In light of this and of hard disk price and storage capacity trends, we have decided to store the masters directly on hard disks rather than on CDs. High capacity four- or six-TB units are presently available that can be set up in RAID 5 array with a redundant disk. If back-up copies are also made, the information can be regarded to be reasonably safe. We have only just begun this process and believe that the present stacks will take up around 30 TB, with expected growth of about 7 TB/year. That will mean purchasing ten 6-TB units, five for the main and five for the back-up files, plus the additional units needed to accommodate yearly growth. Nonetheless, this hardware cost is really small compared to the cost of recording all the information on CDs or DVDs.

In other words, we believe that hard disk storage is currently the most suitable system for handling and storing our collection. It will enable us to make changes where necessary, and above all, to speedily and automatically update books to any new standard, converting from DAISY 2.02 to DAISY 3, for instance.