**Producing Ebraille Books**

James Bowden

RNIB Braille Technical Officer

UKAAF Chair of Braille Coding Group

# Abstract

The RNIB has an extensive braille library available for customers to borrow in hard copy. Each book consists of several weighty volumes which are sent out and returned by free post.

With the advent of more affordable braille displays, more customers are asking if they can borrow the electronic files to read in digital form. This saves time, energy and resources.

In this paper we discuss steps to convert multi-volume braille files, which are suitable for embossing, to single volume files, more suited for use with braille displays.

We discuss the differences between embossable braille files and those for braille displays, as well as processes which can automate converting large numbers of existing titles and the changes to transcription processes that might be needed for producing new books in the future.

We touch on some of the main advantages and limitations of using electronic braille on current technology.

With no physical pages, we discuss the techniques customers can use to efficiently navigate the text using a braille display, and consider if new braille file formats might further improve navigational possibilities.

We concentrate on the braille layout used in the UK and by RNIB, but briefly consider any potential differences with the main features of other systems of braille layout and how these differences might affect navigating using current technologies.

# Electronic Braille Files

At their simplest level, electronic braille (or Ebraille) files are exactly the same braille as would normally be sent to a braille embosser to produce a hard copy book. The main difference is that Ebraille files are mostly read using refreshable braille displays, thus saving trees. You could say it is equivalent to the difference between using a paperback book or an eBook reader such as a Kindle.

The de facto standard for electronic braille is the BRF file, which essentially captures the image of a series of braille pages in a simple text format. Layout is given by using ordinary space and new line characters, as many as needed, with new pages indicated by the form feed character. Such a simple format is compact and easy to use, from a machine point of view, to recreate the embossed pages.

Most braille books consist of several braille volumes. There is often one BRF file per volume. So, a single book could consist of several files, several dozen in the case of longer novels or reference works.

It was felt, at the most basic level that having a single file rather than multiple files would considerably ease reading of Ebraille, so one file would correspond to one book, rather as it does in print.

## Creating Single Volume Ebraille

The simplest thing to do would be to concatenate all the original single volume files into a single unit, but this causes some curious effects: most notably the inclusion of rogue title pages and other preliminary information at various points throughout the work, where the original volumes were joined.

Additionally, a problem when reading even a single volume is that the flow of the main text is periodically interrupted by the page header and/or footer information, where the original braille pages changed, see Sample 1.

Clearly, some form of additional processing is required to improve the user experience of reading Ebraille. Call this process, "devolumising".

We came to the following basic steps, nearly all of which can be automated with sufficient programmer skill and taking advantage of heuristics of the braille layout.

### Basic Page Processing

The basic page processing can all be applied before considering the higher level features of the document.

* Delete the first line of every page if page headers are used.
* Similarly delete the last line if page footers are used.
* Delete all blank lines at the top of the first page of the first volume.
* Compress all instances of multiple blank lines to a single blank line.
* Similarly, remove all blank lines at the bottom of each page.
* Remove any trailing white space at the end of each line.
* If the book is double line spaced, then remove the double line spacing by removing every other new line. The rest of the blank line processing still applies.
* Remove form feed characters, they were used to mark braille pages.

Note: in the UK, braille page numbers are shown top right, on the header line. Page footer lines are not used. Removal of header and footer lines avoids the problem shown in Sample 1.

We suggest the removal of excess blank lines as above so there will not be multiple blanks when navigating using a braille display.

The removal of all blank lines at the top of the very first page means that when a user opens the file on a single line braille display, something should immediately be displayed (hopefully the title).

### Top and Tail the Volumes

Next, "top and tail" each braille volume. Essentially, keep the preliminary information of the first volume and the trailing information of the last, but delete the rest.

The preliminary information generally consists of a title page, perhaps a copyright page or additional title material, possibly a print and/or braille contents page, note to braille edition, dedication and so on. Eventually, the main text of the volume (at least in the UK layout) begins by repeating the book title.

So, given the repetition of the title, the extent of the preliminary information can be determined: it should match the first lines of the entire book.

Within the preliminary information, there are a few adjustments which should also be made:

* On the title page, any designation of how many braille volumes there are and which volume this is, should be replaced by an indication that this is an electronic edition, see Sample 2.
* In a print contents page, there may be information (not in the original print version) about which print pages are in which braille volumes. This information can be removed as there are no braille volumes any more. Note that a similar removal may occur in indexes.

The main body of the original print contents page can be kept, because the referenced print page numbers will still be relevant. However, a braille volume contents should be removed entirely, because there are no braille pages any more.

The trailing information of a volume, at its simplest, just consists of the indication of the end of the volume, such as "End of v3". This is often the last line of the volume containing printable characters, so is easily removed. We decided to keep the indicator from the last volume, "The end", because this is still relevant.

In more complex cases, trailing information may include a section of footnotes from the original volume. Ideally, these should all be collected together to one footnotes section and each footnote should ideally be unique.

### Comments

At this point it is worth noting some of the document features which have not been adjusted:

* Line length remains untouched.
* All document content, including layout remains untouched.
* Braille code used remains untouched.

In other words, the devolumising process keeps the main body of the text intact.

Note that to insert the indication "Ebraille edition" correctly, the system must be able to determine the correct position and braille code for the text to be added. The correct position can be established by centring according to the longest line length detected. The braille code – such as grade 1 or 2 braille (uncontracted or contracted), whether Unified English Braille or the older Standard English Braille (UEB or SEB), and in the case of SEB whether capitals are shown – can be determined by searching for various typical characteristic strings which naturally occur in the different braille codes, see Sample 3.

# Navigating Ebraille Documents

As mentioned previously, the BRF file format is a text file, containing no high level semantic information about the structure of the content. This means that braille displays do not themselves allow a user to jump with a special command to a next or previous heading, table, list, or follow a link from a contents page.

In effect, the most powerful command that a user has in order to navigate is the Find command. But this requires the user to know what they are looking for. Fortunately, though books use a variety of ways to denote parts, chapters and so on, the UK layout has various "end markers" and the "print page indicator" which will help, see Sample 4.

Once the user has identified the text style used for parts and chapters, they can jump straight to what they need, remembering to include some initial spaces to indicate that what they are looking for is not at the left margin – that is, likely centred. All text must be entered in exactly the same way as it appears in the book, e.g. contracted or uncontracted, and in the same braille code.

For example, a user can search for the word "Contents", look for the entry they require and note the print page it occurs on. Then, search for the print page indicator for the top of the page they want.

If there is no Contents, the user can search for the book title and then read forward to find out how chapters are marked, (e.g. with or without the word "Chapter", and then find out how the numbers are shown). If the first section is a prologue, introduction or other section before the first chapter, search for an end marker by searching for several colons or dots 2-5. The end marker typically marks the end of the section, so reading forward should locate the next.

Though not absolutely ideal, strategies such as the above can provide a viable method to navigate.

## Reference Works

In the case of reference works, where it is vital to uniquely and unambiguously identify all reference points, additional tweaks can be made to the text.

For example, in the case of a Bible, each chapter can be headed with the name of both the book as well as the chapter. Now a user can search for "John 3" or "1 Corinthians 13". Of course, some book names may be a little complex to spell!

In the case of dictionaries, each head word can be written in grade 1 (uncontracted) and preceded by a grade one word indicator (dots 5-6, 5-6).

Other methods can be devised as needed for other types of books. The point is to devise a unique string which can be easily thought of and found by the user. If necessary, explain the method used at the start of the work.

## Footnotes

The case of footnotes is worth mentioning: generally in the UK, a footnote reference consists of a number, preceded by some kind of indicator, most commonly an asterisk in SEB titles or a superscript indicator in UEB.

To locate the text for the footnote, it would be ideal if the reference marker would be repeated in full and be unique in any notes section. A user could then just set a bookmark where they found the reference point, search again for the reference mark, thus going directly to the text of the footnote, then return to the bookmark to resume reading at exactly the right place.

This ideal may require changes to transcription processes. Currently, several further stages of navigation may be needed to find the correct footnote, especially if notes restart numbering from 1 for each chapter and if the footnote text does not contain exactly the same reference mark.

## Indexes

To use an index, a user can search for the heading "Index" (probably centred, so precede with some blank spaces). Then, search for the desired keyword. Page references are to print page numbers, so once an entry is found, search from the beginning of the file for the required page number and then finally search for the keyword (which should appear on that page). In effect, this is the same process used when managing a paper book.

## Line Length and Wrapping

We should also consider the question of line length and line wrapping. If the line length of the braille file matches the length of the braille display, there is no problem and all lines can be shown exactly as originally designed, including indentation.

However, if the reading line is less than the line length in the file, some processing is required by the display to avoid the very tedious problem of one full length line alternating with a short line, which can prove very vexing for the reader, see Sample 5.

Fortunately, several modern displays can wrap lines, by treating a single new line in the file immediately followed by a printable character as if it were a space. Since nearly all new paragraphs are indicated in braille by either an indented line or a blank line, this simple strategy works well for ordinary text passages. But it doesn't work well where the braille layout has used hanging paragraphs or various different levels of indentation.

So another point for consideration for transcription is whether all paragraphs should have the overrun lines at the margin, or even if line breaks in ordinary text should be abolished. We do not attempt to answer such questions here.

## BANA Layout

Much of the foregoing can equally be applied or easily be adapted to books using the BANA layout conventions. Perhaps the most notable difference is that BANA generally tends not to use end markers, instead marking a major heading with a blank line above and below. Blank lines are more difficult to search for, unless one has an advanced Find function which can search for control characters. However, centred headings, such as "Contents" and the names of chapters, and print page numbers can be located in similar fashion as above.

# A New File Format

To combat some of these limitations with current technology, is there a need for a new electronic braille file format? Such a format could contain all the higher level semantic information to allow a braille display to properly display the text, no matter what the line length, and also allow quick navigation, similar to the functions users have come to expect with screen readers in web browsers and even word processors.

While this sounds desirable, and several possible candidates already exist, there is a threefold problem to overcome: manufacturers of braille equipment, developers of braille software and production houses which create content, must all simultaneously be willing and able to agree and adopt a common standard.

To date, though approaches have been made to some likely organisations who could drive this forward, sadly, there appears to be little appetite to develop such a standard.

For what it's worth, the two preferred candidate file formats that contain much of the desired functionality would be either an enhanced and expanded Portable Embosser Format (PEF) file, or a hypertext based system such as EPUB, but containing braille characters instead of print, and perhaps a simplified lexicon of tags (simplified, only to allow more of the lower cost devices to implement the standard).

Such a format would allow a user to jump to the next or previous heading (of various levels), select a link in a contents page, footnote reference, index page etc, using commands on the braille display, and also negate the problems of line length, because a hard carriage return would only be indicated where necessary, not at the end of each line occurring simply as an accident of physical page size.

# Appendix

## Sample 1

Example of braille broken by irrelevant original page header:

"What is it, then?" I asked, for

his manner suggested that it was some strange

creature which he had caged up in my

room.

"It's a new patient," he

192 The Engineer's Thumb 57

whispered. "I thought I'd bring him

round myself; then he couldn't slip

away. there he is, all safe and

sound. I must go now, doctor, I have

my dooties, just the same as you," and

off he went, this trusty tout, without

even giving me time to thank him.

[Taken from The Adventures of Sherlock Holmes by Sir Arthur Conan Doyle.]

## Sample 2

Example of original and modified title page:

**Original:**

1

[blank]

The Call of the Wild

by

Jack London

[blank]

[blank]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

In Two Volumes

Volume One

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

[blank]

Produced and Published by …

**Modified:**

The Call of the Wild

by

Jack London

[blank]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ebraille edition

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

[blank]

Produced and Published by …

## Sample 3

Examples of characteristic strings to determine braille code, needed for insertion of any new text, such as "Ebraille edition":

The following strings typically occur in contracted (grade 2) Standard English Braille (SEB):

1. &a &! =a =! (a (! )a )!

2. 0a 0! 96a 96! 6a 6!

3. -e -+ a# /a# n,n /,n

4. re,y typic,y a4$ mi4le

5. 7in 7! ,7in ,7! was7 was40'

The following strings typically occur in contracted (grade 2) Unified English Braille (UEB):

Lines 1 and 2 now have spaces between the words.

1. & a & ! = a = ! ( a ( ! ) a ) !

2. by a by ! 9to a 9to ! to a to !

3. come com+ able /able na;n /a;n

4. r1lly typically add$ middle

5. "<9 "<! .<9 .<! 0"> was4,0

For capitalised braille, check for dot 6 preceding common words:

Capitalised: ,a ,! ,I ,he ,%e ,x

Uncapitalised: a ! I he %e x

## Sample 4

Some typical markers and page indicators:

The following are used as part of the typical UK braille layout and are not part of the UEB specification.

End marker (SEB): centred 12 × dot 2-5, e.g.

333333333333

End marker (UEB): centred dot 5, 2-5, + 10 × dots 2-5, e.g.

"33333333333

Colons end markers like these typically mark the end of major sections, e.g. parts and chapters.

Dot 2 end marker (SEB): centred 12 × dot 2, e.g.

111111111111

Dot 2 end marker (UEB): centred dot 5, 2-5, + 10 × dots 2, e.g.

"31111111111

Dot 2 end markers like these often mark the end of subsections and special items such as tables and extracts.

Print page indicator (SEB and UEB): centred dots 5, 2-5, immediately followed by the number, e.g.

"3#egb

Print page markers mark the start of the new print page. The number is as per print, in the example above, page 572.

In the BANA layout, end markers are generally not used and the print page indicator is a full line of dots 3-6 with the print page number right aligned.

## Sample 5

Example of long-line short-line:

This may occur if line lengths between source document and braille display are mismatched and there is no mitigation in the display.

It was almost ten o’clock before we

heard the

sound of wheels. We all got up

then; and my

mother said hurriedly that, as it was so

late, and Mr.

and Miss Murdstone approved of

early hours for

young people, perhaps I had better go to bed.

I

kissed her, and went upstairs with my

candle

directly, before they came in. It

appeared to

my childish fancy, as I ascended

to the

bedroom where I had been imprisoned,

that they

brought a cold blast of air into the

house which

blew away the old familiar feeling

like a

feather.

[Taken from David Copperfield by Charles Dickens.]