COMPUTERIZED BRAILLE PRODUCTION

by

Jørgen Vinding

Let me start with a question: "Why do we not have very much Braille or why is the amount of Braille-reading so low?"

There are of course many answers to that, but some answers can be given.

1) The number of Braille-readers is small in comparison to the total number of blind and partially sighted persons.

2) The production of Braille is complicated and takes quite a long time.

3) The number of Braille transcribers is small.

4) The number of different titles on books is small compared to what is available to the sighted person.

5) The Braille system itself, taking about Grade 2 Braille, is difficult to learn and use.

6) The speed of reading Braille is slow compared to speed for normal inkprint.

7) Maybe the Braille cell, the dots, should be changed to another tactile cell?

Let us now look a little into what kind of material is produced in Braille. Generally there are two different types of material:

1. The mass production like newspapers, magazines, where the number of copies is big, more than 20. We use here the stereograph, writing on zinc, and the press for the embossing or we could use the English system, the solid dot. Because of the nature of the machinery this kind of production is normally centralized to big printing houses.

2. The single copy production like literature for the students and school children where the number of copies is one to five. The characteristics of this production are many different titles and few copies. We use here the perkins, writing the original copy on paper, and the following copies are obtained by using the thermoform machine for copying on plastic. This production is normally done "on the spot" and of course more decentralized.

Now the Braille systems should be analyzed a little, but first our national languages should be examined. In this case I shall talk about Danish, but I think there are some relationships to other national languages.

Let us have a look on a normal book. Such a book will consist at about 50,000 words equal to 300,000 characters. Among the 50,000 words you will find that the nine most common used words in the book will cover 20 percent of the book. Taking the 100 most common used words in the book will cover 50 percent. This of course gives you an idea what kind of full-word contractions should be used in a Braille system.

Now, if we look on the different characters in the book, we shall find that the space between the words takes up at least 17 percent of the total book. The letter E is used 12 percent of the time.

One sometimes get a crazy idea in the preparation of new Braille systems. Let us for a very short minute think of a Braille system where the nine most common used words were removed and the spaces between the words were removed, too. Then we have a "very nice" Braille system with a saving-value equal to 40 percent, and some easy rules for the use of the system. - Allright, the minute is over!

With respect to the preparation and changing of Braille systems I shall here enclose an article, written by V. Påske and myself with the title "The Perfect Braille System".

There are three main reasons for using contractions in Braille:

1) It is easier to read short words than long words in Braille

2) It is faster to write contracted Braille than non-contracted Braille

3) Braille books will take up less space and be easier to read

Number 3 cannot be discussed since it is obvious from every point of view.

Numbers 1 and 2 cannot be accepted just like that. Contractions have to be learned. If the Braille system is too complicated a long time is needed to learn it, giving a lot of difficulties for many Braille readers. Considering the job done by people, who have worked with Braille systems, this job has for a great deal concerned the development of complicated and advanced systems. In doing that people often insert a lot of contractions selected from the feeling that these contractions are of some value for the final system. In some cases already existing contractions show to be of almost no practical value.

The answer to the question "how to make the perfect Braille system" might be: make a system with as great a character savings as possible and with a very high degree of simplicity.

The following pages contain a report on the methods and the results of the revision of the Danish Braille system. The revised Danish Braille system has been used since its acceptance in September 1971.

The Material

The following material has been used:

1) The 4,089 most common used Danish words

2) "Midt i en jazztid" (a Danish Book with 46,000 words equal to 260,000 characters).

The 4,089 words are taken from all kinds of literature with a total of 1,000,000 words. The occurence of each of the 4,089 words are counted and the number of times the word occurred is said to be the word's value. Each one of the 4,089 words is then punched onto cards together with the word's value and finally codes for the Braille characters that the word will be translated to in Braille.

The book is punched into cards character by character, including spaces and punctuation.

Now the material is ready to be processed by a computer.

If we look at the material, we find that it is a collection of letters, spaces, and punctuation, which we will call characters. Putting together these characters we create different theoretical character combinations. If the combination of characters in the word we get a part-word.

A list of words is made by listing each word on a separate line. A text is made by setting the words after each other on one or more lines in a way that the words set together have a meaning.

Characters, words and part-words written with the normal alphabetic letters are called an inkprint. If instead we use the 63 Braille characters, we get Braille.

When we translate a word in inkprint to Braille, without translating character by character, but using one or more Braille characters we then have a full-word contraction in Braille. Translating a part of the word in inkprint to one or more Braille characters we now have a part-word contraction in Braille. Grade 1 Braille is when we translate a text to Braille character by character, while a Braille text containing full-word as well as part-words contractions is Grade 2 Braille.

The use of full-words and part-words in Braille gives an abbreviation of the inkprint text. The characters saved is the number of characters by which the text is reduced when going from inkprint to Braille. A six character word in inkprint that by translating to Braille only contains four characters, has a character saving of two characters. If we multiply the character saving by the value of the word we get a product that shows how important the abbreviation is. We say that the product is the savings-value.

If we look at a five character word in inkprint and this word in Braille is three characters and the word's value is 100

then this particular word has a saving-value of 200. Taking another five character word, that in Braille is reduced by four characters and having a word's value of ten you get a saving-value equal to 40. We now can see that the first word is more valuable in Braille even though this word is only reduced by two and the other word by four characters. Of course, the reason for this is that the first word will be in a text more often than the second.

The use of contractions seems to be the case. But sometimes this usage brings up new problems since some misunderstanding of a word or a text can result from an abbreviation. In these cases we have to establish special rules for the use of certain contractions and sometimes even exceptions.

The Computer Programs

The technical research done on the material uses different computer programs. The result of these runs on the computer (IBM 360/75) is discussed later, but here a description of the programs will follow.

Program A. This program works on the 4,089 words. The word's values of these words are sorted in ascending sequence. At the same time the saving value for each word is computed and the words sorted are in ascending sequence by saving-value.

Program B. This program works on the 4,089 words as well as the book. Since the 4,089 words on punched cards contain information on how the words are translated into Braille, it is possible to compute how many characters are saved by the Braille abbreviation. There after the program reads through the book and counts how often each word appears and totals the number of characters in the inkprint text. Now the 4,089 words are sorted into ascending order by the number of occurence of each word in the book. It is now possible to print how many times the word occurred, how many characters the abbreviation of this word in the whole book saves, and finally the percentage this saving is of the total number of characters in the book. At the same time, the accumulated savings-value is printed.

Program C. This program works with the 4,089 words. For each word the contents of contractions is investigated. For each part-word contractions in a word, we compute the saving-value of this particular part-word contraction. Then the part-word contraction is sorted in ascending sequence. We also sort the part-word contractions according to the accumulated saving-value for each part-word contraction.

Program D. This program works with the 4,089 words. From the Danish alphabet we now make the possible theoretical character combinations of two characters (841). Then we go through the 4,089 words looking for the occurrence of any combinations set up. When we have an accumulated value for each theoretical combination and we sort according to this value.

It is possible to control all the mentioned programs so that subtotals can be computed at certain specified places in the program. It is, of course, possible to specify that only certain words, part-words or groups of words and part-words are to be taken into consideration while the programs are running. For all program it is possible to specify that the result are to be plotted as a curve or diagram.

During the work, it was necessary to make certain special programs to investigate special details.

The Results

As mentioned the aim of the revision was to create a Braille system which at the same time is simple and effective and where these two goals not at the same time was acceptable to reach a point where a natural selection was taken in relation to the two main goals.

This gives us the following:

1) We only use full-word contractions if they have an acceptable saving-value and only as many as we find reasonable from a pedagogical point of view.

2) We use part-word contractions only if they have an acceptable saving-value and only as many as we find reasonable from a pedagogical point of view.

3) The system is designed to have as few rules and exceptions as possible.

4) Exceptions and special rules for the use of contractions are avoided if possible. We chose simplicity instead of saving-value if the use of a given contraction forces us to set up a special rule for that particular contraction.

Comments on Point 1

In the research of the saving-value of full-word contractions we find that 100 is a reasonable maximum number of full-word contractions. Using Program B and the "old" Danish Braille system it is possible to compute for each of the full-word contractions and part-word contractions the saving-value. Accumulating the saving-values gives us the total saving-value for this Braille system. Since from different programs have the total saving of characters in Braille, in percent 26, we then can define the absolute value of saving, meaning that we know how big a value of saving is needed for 1 percent saving of characters. We now selected the 500 most common used words and allowed them to be contracted to one or two Braille characters. We then computed the saving-value for each of these words and sorted them in ascending sequence according to this value. This showed us that if we used more than 100 full-word contractions the percent of saving for the words above 100 would be very little. Using 100 full-word contractions gives a percent of saving equal to 18 while 500 full-word contractions gives us 22.5 percent. The additional 400 words only is worth 4.5 percent and this requires an additional 400 full-word contractions.

Comments on Point 2

We now look on the part-word contractions to get a better saving of characters than the extension of full-word contractions beyond 100.

In the research we found that the use of more than 30 part-word contractions is not reasonable. In the "old" Danish Braille system there were 58 part-word contractions. Using Program C we found that there was a great difference in the saving-value of each of the 58 part-word contractions. We got two groups, 29 good ones and 29 bad ones. Taken the accumulated value of saving in account we could see that the 29 good part-word contractions represented 95 percent of the total character of saving for all the 58 contractions. It was also obvious that the 29 bad part-word contractions were contractions which often used the so called contraction sign. Of course about 30 part-word contractions seemed to be the goal. But we had to make sure if there was some combination of letters which by their value of saving should be contracted and which in the "old" system was not and then to see if we had any more unused Braille characters. Using Program D and excluding all combinations which because of overlapping were worth nothing, we found 27 possible good ones of which 21 already exist. Of the 27 we could only have 22 plus five, three or four characters part-word contractions, which at the end gives us 27 part-word contractions. Ten of these are also full-word contractions and are the only ones, which can be used as full-word contractions also.

Comments on Points 3 and 4

In all the research we have found that any expansion of the number of rules and exceptions trying to get a bigger saving of characters is of no good at all.

Of course it is necessary to have some to exclude misunderstandings.

The final result of the research was a proposal for a revision of the "old" Danish Braille system. The characteristics of the new system are as follows:

1) 100 full-word contractions (plus 30 extra because of certain compound words, which have pedagogical values)

2) 27 part-word contractions

3) 28.6 percent character saving

4) Three rules

5) Five exceptions for the use of full-word and part-word contractions

6) Punctuation rules.

The following table shows the comparison between some Braille systems.

	The "Old" Danish System	The Danish System	English American Usage
Number of full-word contractions	49	130	140[1]
Full-word contractions usable as part-words	36	10	practically all of them
Other part-word contractions	22	17	48[1]
Character saving-value	26 percent	28.5 percent	26.5 percent

Conclusion

The "perfect" Braille system is obtained by the following rules:

1) Use about 100 full-word contractions

2) Use about 30 part-word contractions

3) Use as few rules and exceptions as possible

4) Do not use the so called contraction sign

5) The first paragraph of "how to write Grade 2 Braille" must be "it is never wrong to write Grade 1 Braille!"

The new (revised) Danish Braille system has a high degree of simplicity and is pedagogically very simple. This is also true when talking about computer-translated Braille.

I now shall give you the plans for the computerized Braille production in Denmark.

This system is composed of three parts:

1) The preparation of input and punching the inkprint text on cards

2) The translation of inkprint to Grade 2 Braille using a computer

3) The production of Braille copies using a minicomputer and a lineprinter

I. The Input

The inkprint book is punched character by character on cards, including spaces and punctuation. In the punching operation the keypunch operator will insert special symbols into the text to indicate format controls like start paragraph, new line, new page, tabulation, inkprint page numbers, etc. As you will see the keypunch operator does not need to know Braille at all. Of course some kind of planning has to be done. But in this case there is no difference from what is done already today. The planning group inserts handwritten remarks in inkprint text and the keypunch operator will now know to insert the special symbols.

On a later date the system will accept input from other media like paper tape and OCR (Optical Character Reader). I think that OCR is the future, since the already available OCR reading devices are able to read many types of written characters and even some handwritten text.

For a standard book the punching time will be less than a week. The punching can be done locally or in a key-punching service bureau. If your local capacity for a short period of time is too little, you just send the material to the

service bureau.

2. The Translation

The translation is done on a computer. We load into the computer the necessary translation program and the dictionary containing the rules for the translation from inkprint to Braille. This dictionary can be set up to contain information about the rules for the Danish Braille, the English Braille, the German Braille, etc. We still use the same translation program and during the translation we have the opportunity to switch from Danish to English. This is of course very flexible.

On a later day it might be possible to insert into the dictionary the rules for translating mathematics and music.

The translation program reads the cards and translates word by word the inkprint to Braille. The Braille is written to a magnetic tape with a code for each Braille cell.

The translation speed is very fast, about 4,000 words per minute, depending on the computer. One standard book (50,000 words) will be translated in less than 15 minutes.

The translation program is programmed in a high level language, FORTRAN. Therefore the program can be executed on most of the computers available today. This again means that you do not need a big computer yourself. You simply buy the computer time you need in a service bureau.

3. The Output

As mentioned, the output from the translation is a magnetic tape. This is the original copy and the only one you keep in the library. Every time a copy is ordered, you simply mount the tape on a tape reader and while the tape is read, output in form of Braille will be printed on a line printer. The system is set up by three parts:

1) A tape reader

2) A minicomputer (controlling electronics) and

3) A Braille line printer

All the mentioned devices are commercially available from big companies (IBM, HONEYWELL), which is rather important with respect to maintenance and repairment. The system produces 420 Braille lines per minute equal to about 20 Braille pages per minute. Therefore the time for printing a standard Braille book will be less than 15 minutes. The book will be printed on one side of the paper (there is no interprint). I suggest that the copy is given permanently to the person who has ordered the book. By this, the library cost will be reduced very much since you only have to take care of the magnetic tapes.

About the cost for producing a book I cannot give you exact prices, but the prices seem to be at about the same from the thermoform techniques (plastic).

The Braille line printer connected to a minicomputer and a tape reader is the new idea.

The flexibility for the system is rather high, since interface controls can be developed for the purpose of connection of other devices to the minicomputer. If you want to control an automatic stereographing machine, this is possible, too. By this, the system can be used both in the mass production and in single copy production.

The total investment for the above described system will be about the amount of 150,000 dollars.

For further information, do not hesitate to contact Jørgen Vinding.

References

1) A Frequency Count of Symbology of English Braille, Gr. 2, American Usage by Kederis, Siems, Haynes in The Education of the Blind, December 1965.