Computerised Braille Production

Today and Tomorrow

Edited by
D. W. Croisdale, H. Kamp, and H. Werner

Springer-Verlag Berlin Heidelberg New York

Computerised Braille Production

Today and Tomorrow

Edited by
D. W. Croisdale H. Kamp H. Werner

With 50 Figures and 10 Tables

Springer-Verlag
Berlin Heidelberg New York 1983

Papers Presented at the Conference Held in London
May 30, 1979 - June 1, 1979, under the Auspices of
The World Council for the Welfare of the Blind,
Committee on Cultural Affairs
(Sub-committee on Computerised Braille Production)

Derrick W. Croisdale
Royal National Institute for the Blind
224-228 Great Portland Street
London, W1N 6AA, Great Britain

Dr. Hermann Kamp
Universität Münster, Rechenzentrum
Einsteinstrasse 60, 4400 Münster, FRG

Professor Dr. Helmut Werner
Rheinische Friedrich Wilhelms-Universität
Institut für Angewandte Mathematik
Wegeler Strasse 6, 5300 Bonn, FRG

ISBN 3-540-12057-2 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-12057-2 Springer-Verlag New York Heidelberg Berlin

Library of Congress Cataloging in Publication Data
Main entry under title: Computerised braille production, today and tomorrow. Papers presented at the International Conference on "Computerised Braille Production, Today and Tomorrow," London, May 30 - June 1, 1979, under the auspices of the World Council for the Welfare of the Blind, Committee on Cultural Affairs, Subcommittee on Computerised Braille Production. Bibliography: p. 1. Blind - Printing and writing systems - Congresses. I. Croisdale, D.W. (Derrick W.), 1925 -. II. Kamp, H. (Hermann), 1941 -. III. Helmut, H. (Helmut) IV. International Conference on "Computerised Braille Production, Today and Tomorrow" (1979: London, England) V. World Council for the Welfare of the Blind. Sub-committee on Computerised Braille Production. HV1669.C583 1983 686.2'82 82-19414

Printed in Germany

Printing and binding: fotokop, Wilhelm Weihert KG, Darmstadt
2145/3140 - 543210

FOREWORD

The aims of the Sub-Committee on Computerised Braille Production are to seek international co-operation for the development of computerised braille production facilities and to encourage the cost-effective use of technology for braille production.

To fulfil these aims the Sub-Committee is endeavouring to:

- maintain an international directory of people and organisations involved or interested in computerised braille production;
- collect and disseminate information about current and future computerised braille production systems;
- hold international meetings.

The Sub-Committee comprises:

1) Mr. D. W. Croisdale (Chairman)
c/o Royal National Institute for the Blind,
224-228 Great Portland Street,
London, W1N 6AA,
United Kingdom.

2) Dr. H. Werner,
Rheinische Friedrich-Wilhelms-Universität
Institut für Angewandte Mathematik
Wegelerstraße 6
D-5300 Bonn
German Federal Republic.

3) Mr. R.A.J. Gildea,
Mitre Corporation,
Bedford,
Massachusetts 01730,
U.S.A.

The international directory is maintained by Mr. Gildea and the collection and dissemination of operational information is being planned by Dr. Werner.

Previous workshops have been held at Münster (1973) and Kopenhagen (1974) on private initiative, whereas this conference held in London 30 May - 1 June 1979 was the first of its kind held under the auspices of the Sub-committee. It is hoped to mount other conferences as circumstances demand.

The Sub-Committee would like to record their gratitude for the enormous amount of management, technical and secretatial assistance they received from the Royal National Institute for the Blind which helped to ensure the success of the conference. This was in addition to the RNIB's generosity in financing the conference, exhibition and associated social activities.

The Sub-Committee is heartened by the many letters of support and suggestions received from conference participants which will help to determine the Sub-committee's future work.

Finally, to end on a challenging note, it was remarkable that no conference material was produced in braille. It is easy to advance reasons or, more likely, excuses for this state of affairs. Nevertheless, it is a sad reflection on the state of the art of braille production that blind people attending a conference on computerised braille production were deprived of the raw material on which the conference was based. Let us hope that the application of computer technology will remedy that state of affairs in the near future.

For the preparation of this report we are indebted to Mrs. M. Möllenkamp and Mr. H. Mecke at the Computer Centre of Münster University for their dedicated help in rewriting parts of the manuscripts and redrawing several figures.

D. Croisdale H. Werner

TABLE OF CONTENTS

World Council for the Welfare of the Blind

Committee on Cultural Affairs

International Conference on

"Computerised Braille Production - Today and Tomorrow"

London, 30 May - 1 June 1979

KEYNOTE ADDRESS

by

Derrick Croisdale

Before we break for refreshment I would like to say a few words by way of setting the scene on the first two days of the conference - concerned with today's computerised braille production systems. Bob Gildea, USA, whom many of you may know should have been here to do what I am now having to do - and I must say again how sorry I am he is not able to be present.

In an article written by Bob Gildea in 1978 he distinguished four types of computerised braille:

1. Uncontracted braille which follows all the rules and conventions of what is sometimes referred to as grade I braille in many languages.
2. Contracted braille which follows all the rules and conventions of what is sometimes referred to as grade 2 braille.
3. What may be termed transliterated braille in which there is a one-to-one relationship between an ink-print character and a braille character; there may also be a one-to-one relationship between a non-printing control character and a braille character. This kind of braille sometimes referred to as 'one-cell' braille or grade 0. It is often used for computer output required by computer programmers or computer terminal operators.
4. Finally, there is a kind of braille used for the translation of technical computer manuals - and the like - which comprises many ad hoc technical conventions.

So far as the first 2 days of this conference are concerned, without doubt it is the second category that we are almost wholly concerned with i.e. computerised production of contracted braille - commonly called Grade 2 - because this accounts for the bulk of literary braille production.

In this area, it seems to me that there are no computerised braille production systems which are clearly cost-effective compared with the best manual systems. Computers are used not because they are cheaper, but because of the inability to recruit, train and retain manual transcribers. If I am wrong in so thinking - delighted to hear from the speakers and participants.

In common with the main trends in data processing in the 1960s and early 1970s, the bulk of the computerised braille production comes from centralised systems - but there are signs that the developments in micro-electronics may lead to a growth in distributed systems of great variety. We may even see paper displaced as the main medium for braille. Developments in solid-state electronics (silicon chip industry) and other allied fields (telecommunications, laser memories, etc.) may have a radical impact on braille.

The first generation of computerised braille systems used the computer almost exclusively for translation into Grade 2 braille. The current generation - such as the RNIB's - uses the computer not only for translation but for assisting 'manual' braille transcribers; input-output media are diversifying. New input devices or media are optical character readers for scanning printed text; direct input of printers' digitised tapes. Output may be directly onto paper embossing devices (high speed) or in digitised form as magnetic tape cassettes, floppy discs, etc

But in the next decade we may see a growth in independent, free-standing, braille production devices incorporating computers on a 'chip'. We may see a decline in the use of paper for carrying braille and an increase in magnetic tape - or floppy discs, or video-discs, or microfiche - coupled with refreshable tactile displays. We may see viewdata systems (TV & telephone) used to distribute braille or large print publications.

One thing is quite clear. Technology offers many exciting options for the production and distribution of braille. What is not clear is which option - or combination of options - will best satisfy the consumer. The consumer's interest is, or ought to be, paramount. But the consumer's real needs are the ones most lacking in research. Some fundamental questions need answering. As Grade 2 braille systems are designed primarily to save bulk (about 25%), is the expense of translation justified when saving of bulk is no longer important (because the braille is on magnetic tape or microfiche, or whatever)? The advent of highly compact braille storage media means that the design of a Grade 2 system can be looked at afresh with different criteria in mind e.g. ease of learning and speed of reading. If significant improvements compared with uncontracted braille cannot be achieved then one is bound to question whether Grade 2 braille systems are justified.

These questions are being posed to bring home to us that we are living in interesting times - many changes in society and technology are likely in the latter part of this century. It is your job to decide how technology should be used for the blind. The purpose of this conference is to help disseminate information to enable decisions to be taken in the full knowledge of their likely impact and cost.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

THE CLOVERNOOK SYSTEM

by
Gerald W. Mundy, Ed.D.,
Executive Director
and
Martin F. Droege,
Workshop Director

Clovernook Home & School for the Blind
7000 Hamilton Avenue
Cicinnati, Ohio 45231

The Clovernook System

Clovernook Home and School for the Blind was founded in 1903 to provide a home and place of employment for indigent blind women in the greater Cincinnati area. Initially, blind persons were trained to operate hand looms for the production of rugs, blankets, tablecloths, and other woven items. In 1910 the gift of a printing press converted to emboss braille gave birth to the Clovernook Printing House. From that moment Clovernook has consciously promoted the marriage between technology and one of its major purposes, that of providing employment to multi-handicapped blind adults.

I begin this presentation with the emphasis on the relationship between our sheltered workshop and technology because we, in the field of rehabilitation of the blind, must be aware of important technological advances that can be applied to our task of providing effective rehabilitation training and meaningful employment for blind persons who benefit from sheltered employment.

Clovernook has continued to progress with the times by incorporating new technology such as closed-circuit television, speech compression, electronic reading devices and electronic mobility devices within its programs. Our purpose in establishing the system of computerized braille production that is in operation today furthers our desire to utilize technology to its fullest.

The most important goal of the Clovernook system, in our view, is that it must assure employment for the people we serve. Examination of our sheltered workshop operation pointed out that we needed to create a cost effective method of producing braille that would serve the dual purpose of streamlining our production and assure steady employment for the people we serve. The conventional method of braille production still being used in the printing house has some drawbacks. The stereotypist must handle the zinc plates necessary to produce each page of braille, proof copies must be produced and proofread and then the plates must be corrected as the final step in this time-consuming process. While these factors have been present within the conventional method of producing braille, we have been able to make remarkable progress in increasing our production from the early days of Clovernook's printing house to the present. To illustrate, in fiscal year 1920-21, Clovernook's stereotyping department produced 1,250 plates. These plates were used to produce 80,700 pages of braille. In fiscal year 1975-76, after the addition of two modernized stereograph machines and more personnel, 55,000 plates were produced from which 66 million pages of braille were embossed. Today, with the computerized system in operation, it is possible to produce about 2,000 plates a month using two typists and a corrector. This is approximately double what a comparable number of stereotypists, proofreaders, and correctors can do within the same time.

From an operation employing less than ten people in the production of fewer than ten braille publications, we now have according to our latest annual report to the National Industries for the Blind, 59 blind employees in our sheltered workshop. Our printing house is now embossing approximately forty different periodicals. They include popular monthly magazines such as Better Homes and Gardens, Psychology Today, Playboy Magazine, New Horizon, Seventeen Magazine, and Popular Mechanics among others. In addition, periodicals, books, pamphlets, brochures, calendars, and season schedules for all major professional sports are embossed by the printing house. A major portion of the braille embossed by the Clovernook Printing House is completed under contract with the U. S. National Library Services (Library of Congress) and other governmental agencies. The remaining portion of braille embossed consists of magazines and journals for non-profit organizations, and manuals for profit making corporations marketing equipment for use by visually handicapped people.

Keeping with our desire to combine steady employment for the people we serve with today's technology, plans are underway to incorporate the expertise of Clovernook's blind braille transcribers into the automated process. An enhancement to our system of braille production that is contemplated is the utilization of electronic keyboards which would allow blind operators, working from print matter read onto magnetic tape, to provide direct input to the computer from which a proof copy and final press-ready plate would be produced. This aspect of Clovernook's operation will be touched on again later in this presentation.

The demand for braille production has been increasing steadily over the last few years. We anticipate a continuing demand and have therefore committed ourselves to a computerized system.

Clovernook System

When it became apparent that our blind stereotypers could not keep up with the demand for braille, we started shopping for an automatic translation system.

At that time, two translation programs were available. The American System and the Duxbury System offered similar programs. Since the American System was only available as a complete unit including hardware, Clovernook selected the Duxbury Translation Program while looking for suitable hardware. The Duxbury System Translator Program is a direct descendent of the DOTSYS III Program which was designed in 1970. DOTSYS III was written in COBOL programming language in order to render it usable on a wide variety of computing systems. Because of this adaptability and the quality of its translation it was found most desirable.

The Duxbury System translator is an adaption especially for Data General and similar mini computer systems and offers considerable improvements over its predecessors. Probably the most significant improvement is the option of a simpler and more natural form of input. For example, text is mixed case and visually formatted thus avoiding many of the explicit controls required by DOTSYS III.

The translator is table driven, which means that a file separate from the program itself serves to specify many details of the Algorithm. This feature makes it relatively easy to correct deficiencies in a translation.

Operation

To operate the system, input text is prepared in the form of a file that can be processed by the operating system containing lines of text in full ASCII code. The translator is then run with this file as input and produces as output a file containing the braille in Snipas code, a set of sixty-four symbols that represent the sixty-four possible braille configurations on a one-for-one basis. (This code is named after Mr. Dick Snipas, the developer, who passed away a few years ago). This file may be edited, if desired, to remove errors. The final step is to use the editors code file as input to a simple transfer process that drives the LED 120 embosser or PED 30 automatic stereotype machine to produce braille.

Conformation to Braille Rules

The translation effected by the Duxbury System Translator is in general conformance with the rules of literary braille, for example, English Braille, American Edition. Because some of the rules are ambiguous or require adjustments that cannot be made by a practical automatic process, perfect conformity may require editing of output. It has been our experience that all errors are preventable by suitable annotation of the input.

Formatting Capabilities

The following formatting features are provided by the system:

1. Properly formatted inkprint page number inserted upon input command.
2. Automatic braille page numbering.
3. Automatic recognition of paragraph breaks and line breaks.
4. Running page titles.
5. Automatic centering of titles and headings.
6. Tabulation controls to allow left-aligned, right-aligned or mixed columnar materials.
7. Resetting of braille page numbers.
8. Adjusting the braille page margins.

Output Editing Capabilities

Additionally, the system offers output editing capabilities since the output file from the translator can be edited as a collection of braille pages providing ready access to any individual page. This capability then makes it possible to emboss the entire output file, or only selected pages on the LED 120.

Translation Speed

The translator, when operating on the required computer system, and when operating without interference of other tasks, is capable of processing at least one hundred input characters of general text per second.

Hardware

At the heart of the "Clovernook Braille Plate Production System" is a PDP 11/34 computer. This computer system serves as the data collection, translation, and transfer device in the production cycle.

Currently configured with 48K words of main memory, this system supports two 2.5 megabyte removable disk packs (RK05's) and an 8-line asynchronous multiplexor (DZ11). Data is entered into the system via the 4 CRT terminals (VT52) which are connected to the system via local current loop lines.

Data is stored on disk until a document is completed. Output from the disk is routed through the multiplexor to one of two embossing devices (LED120 or PED-30).

One of the more important aspects of this computer-based production system is the flexibility of the computer. As our production demands have expanded, we have added peripheral devices to the system to support this demand. We have added a TS03 magnetic tape unit which now serves as the primary back-up device for all files, and is an economical method for acquiring other literary data already in machine readable form. We are also adding additional disk space (an RK05F) of 5 megabytes which will double the current capacity for storing data. Finally, we are adding additional memory as we expand the number of concurrent data entry users to the system.

Other media also supported by the system are industry standard cassette tape drives and floppy disks both of which are potential media sources for additional literary data.

Our current configuration can be further expanded to accommodate 5 additional disk drives (25MB), 3 additional tape drives, and 8-10 more data entry stations.

Triformation Systems, Inc., with a grant from the U. S. National Library Services (The Library of Congress), has developed the first PED 30 Automatic Stereotype Machine. This machine has been placed with Clovernook for evaluation.

The PED 30 accumulates data in its buffer which has a capacity of 2048 characters. As the buffer receives information it relays it to the mechanical device which embosses the double folded metal plate at a rate of thirty (30) braille characters per second. At the end of one page, the page is manually turned and the process is repeated. When the second side is finished, a new blank plate is installed and the process is repeated for subsequent pages. The completed plates are then ready for embossing paper. Operational experience has shown that approximately thirty (30) double plates can be produced per hour.

Braille Production System (as shown on schematic)

To illustrate further, a finished braille plate is produced through the following sequence of events.

Step I

Print material is fed into the computer by a typist operating a CRT. A few special annotations provided by the CRT operator are required to format the material for the translation process. After finishing a file, the computer is directed to translate the material and the output in braille is completed by the LED 120 at a rate of one page in eight seconds. These pages go to the proofreading department where mistakes are marked. Ninety percent of the mistakes found are due to imperfect input, the rest are translation errors.

Step II

The material is returned to the Computer Department, and is called back to the CRT to be corrected. Another copy is made on the LED 120 and proofread for the second time.

Step III

The process of correcting is repeated after which an embossed plate is produced on the PED 30, Automatic Stereotype Machine.

Although formatting of tables must still be done by an expert blind transcriber, nearly ninety percent of the material embossed is translated and formatted by the Duxbury Translator Program.

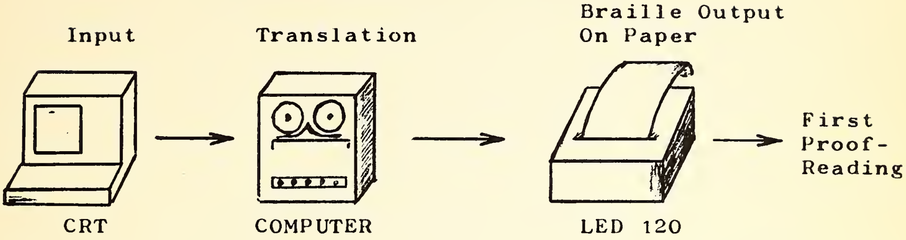
Planned Future Research

To include the trained blind braille transcriber in this automated process, we are preparing to test the possibility of the use of the PD 80, in conjunction with an electronic braille keyboard developed by Mr. K. P. Schoenherr, of Stuttgart, Germany. This will function as an input device into the computer. From this stage on, the process is identical to the one mentioned above.

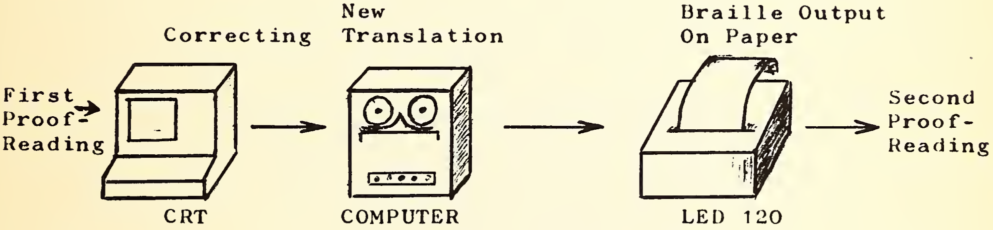
The greatest advantage of the electronic keyboard for a blind person, is its ease of use since plates would no longer be handled and the correcting of zinc plates thus omitted. Additionally, the electronic keyboard would replace the old-fashioned and costly stereotype machines.

SCHEMATIC DIAGRAM
OF AUTOMATIC BRAILLE PRODUCTION SYSTEM

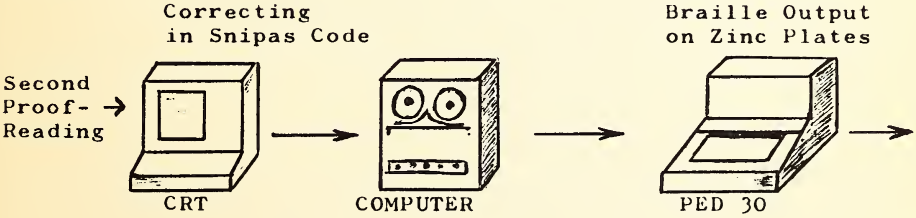
STEP I



STEP II



STEP III



World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

THE LINGUISTIC PART OF AUTOMATIC BRAILLE TRANSLATION AT THE COMPUTING CENTER OF MÜNSTER UNIVERSITY

by

Jochen Splett

Germanistisches Institut der Universität Münster
Johannisstr. 1-4
D-4400 Münster

The Linguistic Part of Automatic Braille Translation at the Computing Center of Münster University

by J. Splett

The method for automatic translation from inkprint into contracted Braille which is going to be presented here has been in production at the computing center of the university of Münster since the summer 1978. It is an improvement of the translation procedure introduced by WERNER and DOST which translate a given character string character for character from left to right. Contrary to this the new method works by help of a modified Markoff-algorithm and yields better and correcter results. About 40 per cent of the text is translated word to word by applying a data set which contains a series of letters for the most frequent corresponding contractions of German grade two Braille. This is a data set, named A, of nearly 650 entries. All words which are not recorded in this data set cannot be translated in a single phase; they are scanned successively whether they contain a series of letters which correspond to a Braille contraction. This is done by matching them with a data set, named B, in which all character strings are stored for which German Braille grade two applies contractions. This data set starts with the longest character entry GESELLSCHAFT, and in the last we have the entry EX. The data set contains altogether about 300 of such entries.

If in the word to be translated no character string of the data set B is found, it is translated letter by letter. Most words, however, contain a series of letters for which a valid Braille contraction exists. If this is the case, the word is run through a block of translation rules. This is a data set of rules which is linked to one entry of the data set B in that way that for each entry of the data set B there exists a corresponding block of rules. These blocks of rules are compiled in different ways, according to the entries to which they belong, and they vary in size. So, for instance, the block of rules for the entry ALL consists of more than 250 translation statements. Form and sequence of the translation statements within the blocks of rules control the translation of a given word in a manner that an adequate Braille translation is the result. After a matching with the entries of the data set B and searching the corresponding blocks of rules, every word is translated so much that the possibly remaining word segments can be translated letter by letter. This may be explained further by means of the example SCHALLTRICHTERN.

At first the word is matched with the entries of data set A in which no character string identical to the given word is found. Then it is compaired with the entries in the data set B, beginning with the entry GESELLSCHAFT. The result is that RICHT is found as the first entry which is a substring of the given word SCHALLTRICHTERN; therefore it is linked with the block of rules for RICHT and processed with its translation statements. The block of rules consists of the following translation statements:

1 PATENTRICHTER → #=PA=T<EN>T=<,RICHT><ER>#
2 STADTRICHTER → #=<ST>ADT=<,RICHT><ER>#
3 HAFTRICHTER → #=<HAFT>=<,RICHT><ER>#
4 HəOCHSTRICHTER → #=H<əO><CH><ST>=<,RICHT><ER>#
5 KUNSTRICHTER → #=K<UN><ST>=<,RICHT><ER>#
6 BLUTRICHTER → #=BLUT=<,RICHT><ER>#
7 RICHTUNGS → #=<,RICHT><UNG>S=#
8 RICHTUNG → #=<,RICHT><UNG>#
9 TRICHTER → #=TR<ICH>=T<ER>#
10 MOSTRICHT → #=MO<ST>=R<ICH>#T
11 ESTRICHT → E#<ST>R<ICH>#T
12 _STRICHT → _#<ST>R<ICH>#T
13 KEHRICHT → #=K<EH>=R<ICH>#T
14 SPRICHT → #=SPR<ICH>T#
15 ABRICHT → #AB=<,RICHT>#
16 BRICHT → #=BR<ICH>T#
17 TəORICHT → #=T<əO>=R<ICH>#T
18 RICHT → #<,RICHT>#

On the left you find the character sequences in which RICHT is contained, on the right of the arrow the corresponding translations into German contracted Braille. The number signs indicate that the part of the word included by them, has already been translated, and consists therefore of a different alphabet. The peaked brackets enclose groups of letters for which a corresponding contraction exists. The equal symbols indicate that if necessary the word can be hyphenated in these places. The commercial AT-sign preceding a vowel denotes the umlaut.

Now the word to be translated is scanned whether it contains one of these sequences of characters. As the last entry in each block of rules is identical with the entry of data set B, to which it is linked, one of the translation statements can always be applied. In our case it is the nineth entry: TRICHTER - #=TR<ICH>=T<ER>#. According to the rules of German contracted Braille the word contraction for RICHT must not be used here, which is avoided by the given translation statement. Compound words with the root RICHTER and a preceding partword ending on T would also be covered by this translation statement which in their case would yield an incorrect result. To avoid this there are the first six translation statements. The word PUNKTRICHTER is superfluous here because it is already covered by a translation statement in the block of rules for PUNKT. This is because entries of same length in data set B are arranged in alphabetical order.

The translation statements seven and eight are necessary to avoid incorrect translations of ST in words like EINRICHTUNGSTEIL. This cannot be done in the block of rules for ST because the necessary context to control the contraction in such words would already have been translated. Words like STEIL in which ST has to be contracted could no more separated from those in which it has not.

The following translation statements have a similar function as those mentioned before. They are arranged in that way that in words with the sequence RICHT, but without the genuine meaning of RICHT, the corresponding contraction is not applied. Compound words like MOSTRICHTOPF and ESTRICHTÖNUNG, word forms like (ER) SPRICHT and (ER) BRICHT, the word TÖRICHT with its different word forms and derivations are translated correctly because of the translation statements in the block of rules for RICHT. The underscore preceding _STRICHT means that this translation statement must be used only for those words in which this sequence of characters occurs in the beginning.

After applying a translation statement from one block of rules, the word to be translated is again matched with the entries of data set B. The comparison starts again with the entry of which block of rules has just been processed because it might be the case that the same sequence of letters occurs a second time in the word to be translated. In our case the matching would start again with the entry RICHT. In the further processing of this data set the entry ALL is found to be the next which is contained in the already partly translated word SCHALL#=TR<ICH>=T<ER>#N. The block of rules which is linked to ALL is now treated in the same way as before the block of rules for RICHT. It is scanned whether one of the sequences of characters is contained in that part of the word which has not yet been translated. In our case it is the sequence SCHALL with the corresponding translation statement #<SCH>A<LL>#. Because of this statement our word is completely translated except for the letter N. In this case it is not necessary to go again through the block of rules which is linked with SCH. Here it is even so that a matching with all remaining entries of data set B is superfluous and can therefore be omitted. The remaining N is then translated letter by letter. With this the translation process is finished, and the next word of the text can be treated in the same way until the last word of the text has been translated.

The scheme of the translation mechanism is briefly this: Move a word into the workspace. Matching with the entries of data set A. If one of the entries is identical with the word, complete translation and move the next word into the workspace. If no identity was found, matching of the word with the entries of data set B. If an entry of data set B is contained in the given word, the corresponding block of rules is searched. From the translation statements select that one which is a substring of the given word. If the translation is finished, move the next word into the workspace; otherwise match again with the entries of data set B, starting with that entry which was used immediately before. Continue the matching and apply the appropriate translation statements of the corresponding blocks of rules until the word is completely or nearly completely translated. The remaining characters translate letter by letter.

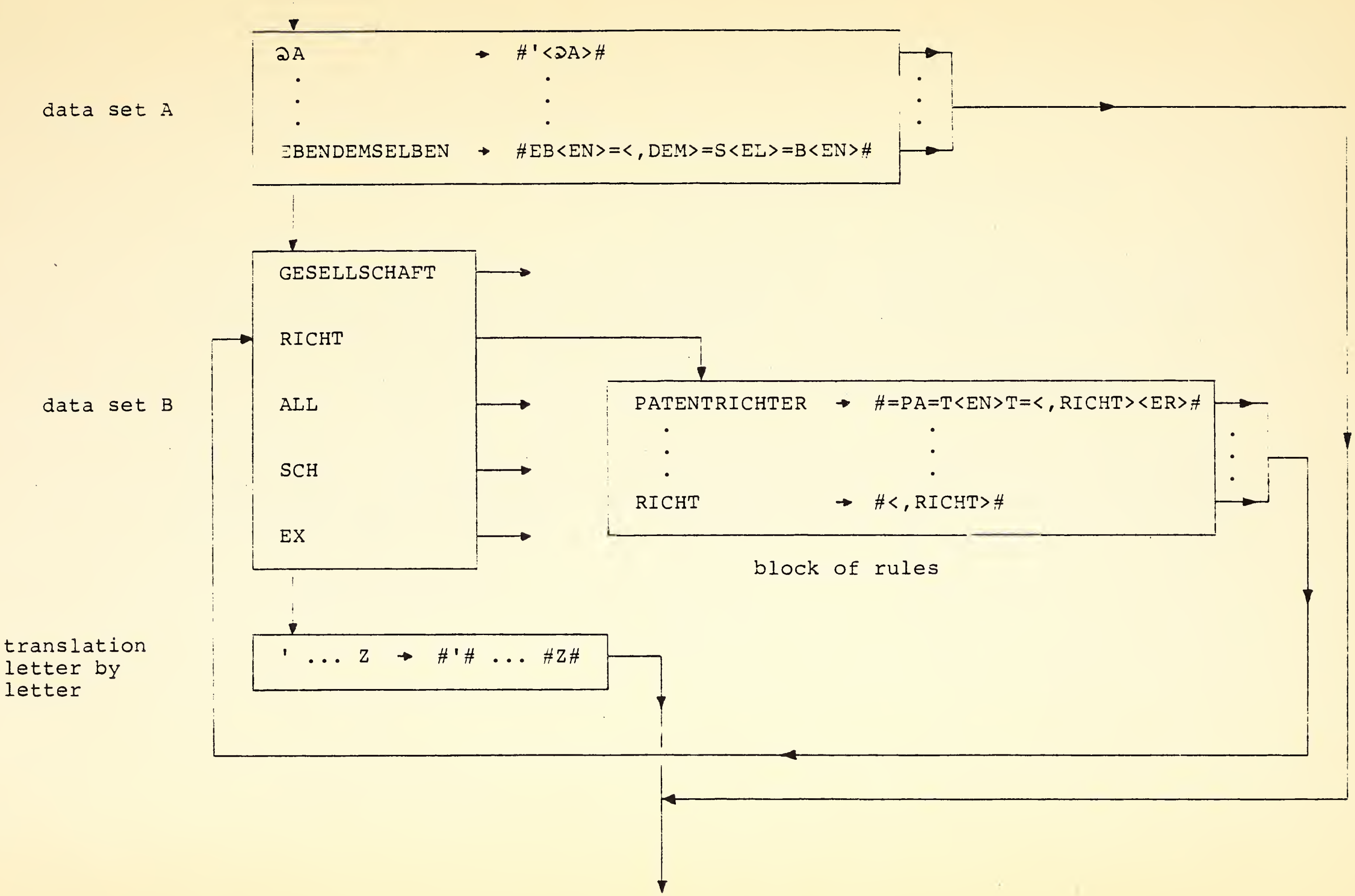
Behind this the real linguistic work was the compilation of the different blocks of rules. The basic scheme of this work may be briefly explained. It was started always with that translation statement which is now the last in a block of rules. The sequence of characters is always equal to the entry of data set B to which the block of rules is linked. Apart from the pseudocontractions like GER or BER this entry corresponds to a contraction of German contracted Braille. Those sequences of letters in which the character string in question occurs, but must not be contracted in the same way, must precede this entry together with their corresponding translation statements. If in such a preceding translation statement again a sequence of letters is contained which would yield an incorrect translation, another translation statement with its corresponding sequence of letters must precede this one, and so on. The word-context, that means the letters preceding or following a sequence of letters which correspond to a Braille-contraction, control the translation process; they govern the form and order of the translation statements within a block of rules.

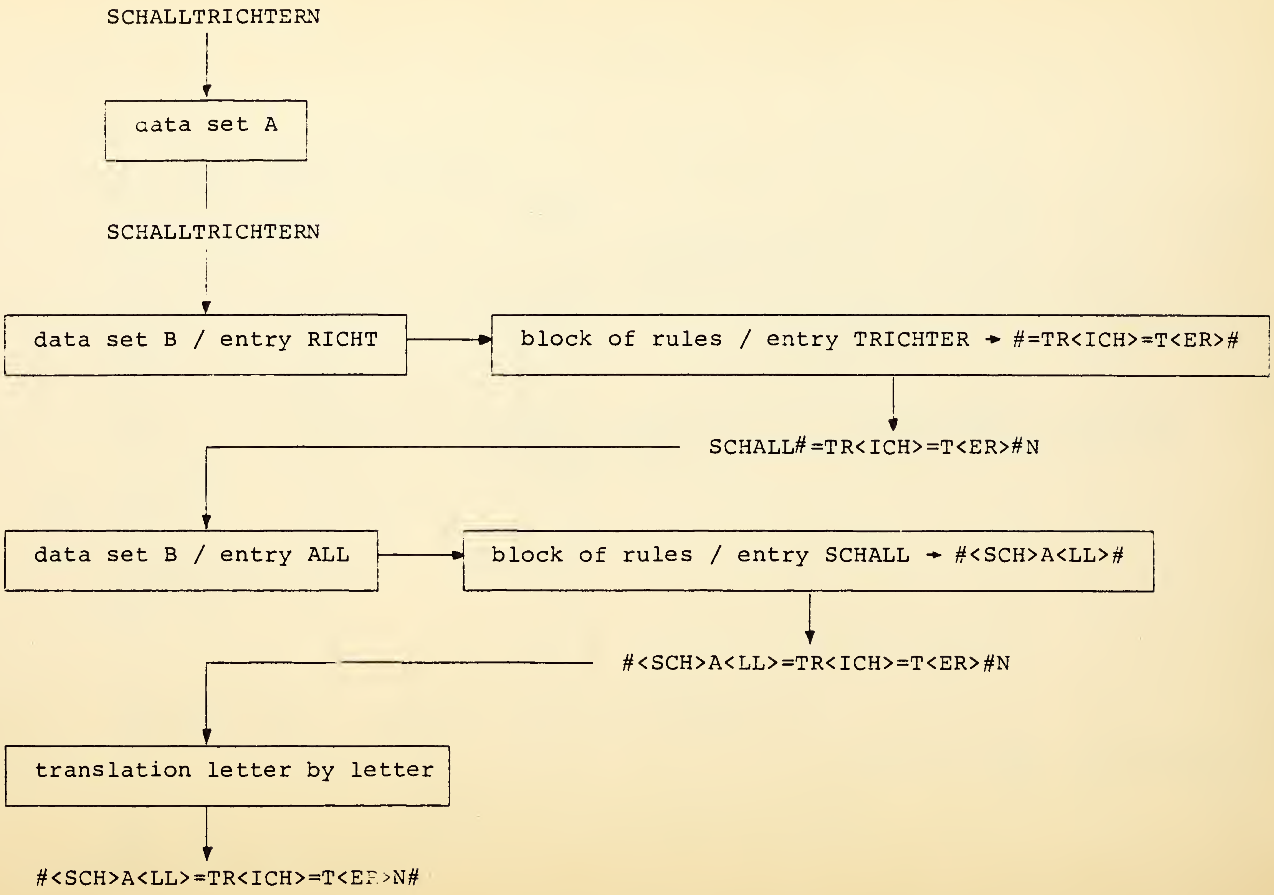
In an ideal case, all words of German language, or more precisely, all word forms of the German language would have to be taken into account, in which a sequence of letters exists that corresponds to a Braille contraction. According to the example mentioned above this means that all words or word forms in which the sequence RICHT is contained, would at first have to be compiled. This is of course impossible, as the vocabulary of a language cannot be counted up. Nevertheless all words have been considered which are listed in the German dictionary of LUTZ MACKENSEN which was available in a machine sensible form. For this purpose the word lists were printed in a quite special form with respect to the letters preceding and following the character strings which are possible contractions. Moreover certain word forms - for RICHT for instance BRICHT, belonging to the word BRECHEN - had to be taken into account. Above all the numerous number of compound words that are not listed in the dictionaries caused difficulties. For RICHT this was for instance the case for all compound words with the first part ending on -RICH and with an initial T at the second part. It was tried to find those words and take them into account among other by help of the reverse-word index of current German language by ERICH MATER.

In this way blocks of rules were compiled with altogether nearly 6000 translation statements. With these the program translated about 100 000 different German word forms that are based on a fund of 1.5 million textwords, into contracted Braille. After proofreading an error rate of about one per cent was found, the errors being chiefly in foreign words and names which are a special problem indeed. By updating and improving the translation statements these errors could be overcome. At present the error rate is less than a half per cent when translating unknown newspaper texts, as it is done every two weeks with the automatical production of a weekly paper in Braille.

It must be emphasized that a wrong translation can be corrected quickly and without problems. This is done by adding the correct translation statement to the data set A. After a sufficient number of updates has been assembled there, they can be substituted by inserting the correct translation statements into the corresponding blocks of rules. So far this control mechanism has a recursive component.

Finally it has to be mentioned gratefully that a group of students of our university has helped to compile the blocks of rules with the translation statements. I say 'thank you' also to the computing center of the university of Münster, especially to BERND EICKENSCHEIDT and Dr. HERMANN KAMP for their continuous advice and help; Mr. EICKENSCHEIDT has done the implementation of the computer program and will report himself about this aspect.


data set A
əA → #'<əA>#
EBENDEMSELBEN → #EB<EN>=<,DEM>=S<EL>=B<EN>#
data set B
GESELLSCHAFT
RICHT
ALL
SCH
EX
PATENTRICHTER → #=PA=T<EN>T=<,RICHT><ER>#
RICHT → #<,RICHT>#
block of rules
translation letter by letter
' ... Z → #'# ... #Z#


SCHALLTRICHTERN
data set A
SCHALLTRICHTERN
data set B / entry RICHT
block of rules / entry TRICHTER → #=TR<ICH>=T<ER>#
SCHALL#=TR<ICH>=T<ER>#N
data set B / entry ALL
block of rules / entry SCHALL → #<SCH>A<LL>#
#<SCH>A<LL>=TR<ICH>=T<ER>#N
translation letter by letter
#<SCH>A<LL>=TR<ICH>=T<ER>N#

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

AUTOMATIC BRAILLE TRANSLATION IMPLEMENTED AT THE COMPUTING CENTRE OF MÜNSTER UNIVERSITY

by

B. Eickenscheidt

Einsteinstr. 60
D-4400 Münster

Automatic Braille Translation implemented at the Computing Centre of Münster University

by B. Eickenscheidt

Before I give you some details of how the Markov algorithm for automatic Braille translation (described by Prof. JOCHEN SPLETT in this report) was implemented for actual production, I would like to stress once more what makes German Braille translation more complicated than many other languages Braille translation.

In English for example, the word CHEMO-THERAPY is usually written without the hyphen I used here, and as we heard there has been a book where GUIDE-DOG was spelled (mis-spelled?) without any hyphen or space between the components of the compound word. Both of those cases gave a computer program the chance of detecting an unwanted contraction, MOTHER, or ED-sign, respectively.

But I dare say that such cases can be regarded as rare exceptions, since in English language the general rule is to spell a compound word as <u>two</u> words with a space in between (or a hyphen if otherwise the first part might me misunderstood as an adjective or so), especially if the compound word is not a very common one or even is an ad-hoc construction.

This is different in German language. A word like LOHNSTEUER-JAHRESAUSGLEICH (yearly balancing of income tax) is usually spelled as shown here: no spaces and no hyphens in it. And this is only an example of a relatively "common" compound word: an ANTRAGSFORMULAR (request form) for LOHNSTEUERJAHRESAUSGLEICH is a LOHNSTEUERJAHRESAUSGLEICHSANTRAGSFORMULAR, where it is absolutely correct (though not mandatory) to use <u>no</u> hyphens at all. This way of chaining words up also applies to the most far-fetched ad-hoc constructed compound words and this will appear many times on every page of any given text which the computer is assumed to nevertheless translate correctly.

This should give you an impression of the huge amount of research work done by Prof. SPLETT and his student groups. Their rule system has made it possible to translate German texts into contracted Braille with a correctness well comparable to any good translation program for - excuse me - easier languages like English, French, or Spanish.

But let us come back to how it is practically done.

For three reasons, we decided to divide the process of text translation into several steps, where handling the "words" is just one of these steps.

First, as pointed out, to a higher degree than in other languages this is the "non-trivial" kernel of the whole job and is also of linguistic interest independent of the (certainly "normal") purpose of producing reading-material for the blind.

Second, the division into well-separated steps provides a very easy access to intermediate results which (usually need not but) can be inspected and/or manipulated at every stage of the process.

Third, in this way it is easy to exchange the word-translating module with one based on a quite different formal-language approach (see the paper of Dr. SLABY in this report) without duplicating the common actions in both programs.

The steps other than the word translation are not specifically German, so I will just mention them briefly.

The first phase in the process is to adapt the given input text to a standard coding: we for example process punched cards or material keytyped at a terminal, and also convert compositors' tapes, in particular several dialects of LINOTYPE punched paper tape.

Next step, over-simplifying, is to translate "everything" except the "words", i.e. handle numbers, punctuation, isolated letters, special signs like capitalization etc., or graphic effects to be expressed with Braille cells, and things like that. This step also does a lot of checking, for example whether misplaced or unusual punctuation would lead to misreadable output in Braille, and produces appropriate diagnostics. Further, if the formatting-codes interspersed into the text contain macros or alternate spellings, these are resolved or standardized in this step. Finally, what this step regards as "words" is specially marked, so that the next step can easily detect them.

The step of translating these "words" comes next. Later I shall give some details on programming techniques used there. But more important for now: If you look at the examples in the SPLETT paper, you find that the *output* of the translation, e.g. <SCH>A<LL>=TR<ICH>=T<ER>N, does not look like Braille at all! And this is *not* meant to transcribe the actual braille output for typographical reasons in this paper only, but it is exactly what the output looks like! The idea of this is that the non-trivial part of the translation, i.e. the decision which contractions to use and which not to use, *has been done*, but at the same time the output is still easily readable for sighted people with minimal or no knowledge of Braille and can be manipulated using *standard* editors on *standard* terminals. Our goal is certainly to make this feature absolutely unnecessary and, in fact, it usually is. But we like to have it. We call it Quasi-Braille.

Up to this phase, by "text" I always meant "endless" text where e.g. card or record boundaries have no importance, and where e.g. desired paragraphs are marked by a simple formatting code. But at the end of the process we usually have a device like a Braille printer or an embossing machine where the text appears as lines and pages. For formatting the "endless" text into the desired structure, also some more sophisticated actions than just "I'd like a paragraph here" can be expressed by appropriate formatting codes. These codes are executed by one of the next steps.

But then obviously the length of the actual Braille output must be known, and that is not easily derived from the Quasi-Braille appearance. Thus we have as next step a program which transforms Quasi-Braille into a one-one coding of actual Braille characters (this is a simple table lookup). Again we decided an at least theoretical possibility to edit this coding at an ordinary terminal could be valuable, so this one-one coding is not some machineinternal binary secret but uses two ordinary digits to octally encode one Braille character.

Next then is the formatting program. It has been told the line and page size desired, but that is usually all what is device-specific of its output. So if the braille output once is wanted on a printer and once on an embossing machine, all steps up to this nevertheless need not be duplicated.

The last step is, of cource, to adapt this output to the hardware requirements of the selected braille device. We have worked with the APH embossing machine, the fast Marburg/Heidelberg embossing machine, the SAGEM REM8BR, and with the IBM 1403 line printer with special Braille equipment. Another variation of this last step is a program that converts the formatted Braille back into some kind of inkprint, so that also the formatting can be checked by sighted people. Of course we also have a program which demonstrates the actual Braille dots in black on ordinary paper ...

The first step and the last one will be the most or only important ones for those interested in actual production of actual Braille from actual inkprint texts. And even we (as a research, not a production group) do have some experience what a lot of actual work is in these "trivial" steps; for example we have not yet found two printing houses using the same codes to control their typesetting-machines ... And in spite of our lack of manpower, we managed all these problems for some kinds of input source and some kinds of output devices so that our research did not remain mere theory. In fact, in cooperation with the "traditional" Braille printing houses in

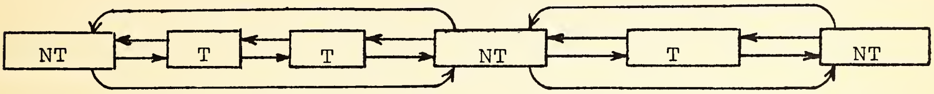
Germany, we produce a magazine of 52 pages every fort night which is distributed in more than 5000 copies; to my knowledge this is the second highest number of copies in Europe. But we will never be able to manage serious mass-production quite on ourselves, let alone on a commercial basis. We will always concentrate more on the research aspect of computerized Braille translation, and as we do give an example of its practical usefulness, we hope that you forgive us our somewhat theoretical standpoint.

From this theoretical view, the only non-trivial problem is to automatically find out which contractions are legal and which are not. One way we solved this is described in the SPLETT paper, and I am going to add some details on how his system was implemented in our program PUMA (Punktschrift-Uebertragung mit Markov-Algorithmus = Braille translation by Markov algorithm).

What is called "data set B" by SPLETT is too large to be held our current computer's memory and thus was implemented as a direct-access file on a disk. In order to achieve a reasonable block size with the disk I/O, the rules have been grouped into packages of ten each, but these packages have no relation to SPLETT'S very varying length "blocks" of those rules dealing with the same contraction. Some forty packages are always kept in memory, and a home-brew "paging" mechanism reads a new package from disk whenever required and not already "in"; the decision which package to "forget" then is done by standard "least-recently-used" technique. All the rest is in memory, along with an indexing system which finds the beginning of a block (in the SPLETT sense), so I shall mention the disk I/O no more.

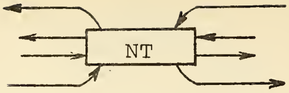
The "data set A" is simply an array, sorted by length first and then alphabetically. When a "word" has been isolated from the input text, its length is known, so that an indexing System can very rapidly reduce this array to the range of those words with the correct length, this range is then examined via binary search. In the case we find the word to be translated, we are done obviously.

If not then we must be prepared to do a lot of replacement steps in the word, and the positions where inside the word some part will be translated will jump around in unpredictable order. Therefore I decided to organize the intermediate results into a multiply-linked list whose elements are either not-yet-translated or translated fractions of the word:

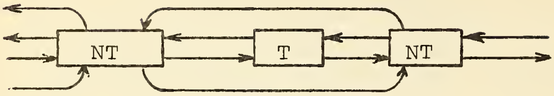


This design may appear rather complicated (and it is a lot of programming work in languages not directly created for list processing), but since most of the CPU time is spent for searching the not-yet-translated parts again and again, it is worth the effort.

The "list" is initialized, of course, to contain as its one and only element the whole not-yet-translated input word. Now the highest priority "entry" into "data set B" must be found; by the nature of the algorithm this search is sequential in the indexing system for "data set B", but each time the whole word must be searched for the "entry". As the "entries" are decreasing in length, we start, of course, with the first entry which is not longer than ML: = maximum length of the not-yet-translated fractions. (Initially we have ML = length of input word, of course.) Every "entry" which is not found saves us from the need of searching the word for the left sides of up to 250 actual translation rules dealing with that "entry". When a matching "entry" is found somewhere in the word, we know from the indexing system, where in "data set B" we have to start fetching left sides of rules which the word must be searched for, now. Although all of these left sides do contain the "entry" as a substring, we may forget where in the word we had found the "entry", because "entry" might occur more than once in the word, and the rules must be used in priority sequence. So again every time the whole word must be searched. But somewhere in the "block" belonging to "entry" we are certain to find a rule that is applicable. That means, we can change some



into



where the two new NT fractions are definitely shorter than the old one was. In fact it often happens that one or both of them are empty; in such cases we immediately simplify the linked list. After thus executing an applicable rule, a full Markov algorithm would require us to examine all rules again. But fortunately the SPLETT rules are arranged so that this is never necessary: of a translated fraction nothing will ever be changed again, and the not-yet-translated fractions won't ever be changed except by replacing part of them, but "finally" then. So we can safely resume the examination of "entries" where we left off; those which did not occur earlier cannot occur now.

But quite often we can even do better than that: if the old NT fraction had been of exactly length ML, it is well possible it was the only (i.e. the last) one of this length. In this case it is worth the effort to recompute the new "ML", as we might find it considerably smaller now so that we can skip over all "entries" that are too long now.

In the current system of SPLETT rules, all "blocks" belonging to "entries" of length 1 consist of the last rule only (which checks for "entry" with no context) and give a complete translation. (In fact, inkprint A is "translated" into Quasi-Braille A,B into B etc., i.e. nothing needs to be done at all.) Obeying priority here would not make a difference any longer, that means, the algorithm terminates not only when ML becomes zero, but already when it has been reduced to 1.
All what remains to be done is assemble the fractions of the "list" into the completely translated word.

To conclude, please allow me to add a word on Prof. JOCHEN SPLETT. When he started with this project, he was not yet used to electronic data processing very much, and as a consequence of that he did virtually no computer-aided testing of unfinished work (as I would have done it in his place). He used pencil and paper and his brain instead. But when he declared his system of rules complete, it worked! From the very first day on, it has seen almost no debugging at all, and today we are still running "Version 1" of it! I am not easily impressed, but that did impress me.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

APPLICATIONS OF PIAF PROGRAM

by

Blaise Mathieu

Labo Imag
BP 53X 38041 Grenoble Cedex
France

APPLICATIONS OF PIAF PROGRAM

Blaise MATHIEU
Labo IMAG
BP 53X 38041 GRENOBLE CEDEX
FRANCE

The Equipe Intelligence Artificielle from the IMA Laboratory in Grenoble has designed and realized a set of automatic treatment for natural languages. This program, PIAF system (or "Programmes Interactifs d'Analyse du Francais) offers to his users a set of informatic definition tools and manipulation of linguistic datum.

- definition tools allow to build or modify a dictionnary and a grammar of a selected language
- processing tools allow to choose the analysis and the treatment of linguistic datum in function of the grammar and the dictionnary already defined.

Following request of the "Brigitte Frybourg" Laboratory belonging to CNAM in Paris, we have been interested to the translation into contracted Braille of the text in French. The rules which allow the translation in contracted Braille belong to a standard type in the languages theory, it is very easy to use a finite state automaton to realize this translation. Therefore, our main job has been to formalized the rules of the contracted Braille in French to the PIAF language.

More, PIAF has already been useful to treat the data acquisition control (detection of an orthographical mistake), we mean to realize a complete string of treatment including a data acquisition controller and a contracted Braille translation.

Our paper will be developed through two points:

- 1) The PIAF program and the existing applications, more particularly the data of acquisition control and the contracted Braille translation

- 2) The new studies on hand: translation of editing tapes, translation of contracted Braille to its integral form.

PREMIERE PARTIE

I - GENERAL PRINCIPLES OF THE PIAF SYSTEM

The PIAF system includes two modules: the general transducer of states allowing to treat the morphological analysis and a set of algorithmes allowing to use an elimination selection process (tree structures).

The TGEF allows to realize the transduction of sequential datum from a dictionary and a grammar. This transduction can be the production of categories and grammatical variables (morphological analyser).

All these applications are possible from the same basic programs because of the linguistic parameters independance (dictionaries, grammar) in relation to these programs.

2 - THE DATA ACQUISITION CONTROL

The purpose of this program is to detect and correct operator encoding errors (editing, syntactic error) with respect to a language model during the input process. In case of mistake, the user is advised and can correct the text, index a new word in the dictionary or can even complete the model of the language.

Morphological analyser:

The dictionary contains invariable forms, roots, suffix and the finite states automaton controls the application for the concatenation (180 rules for the French) of various datum setting in a form: root plus suffix plus ending, and gives the transduction: production of grammatical categories and variables (verb 1st person plural...).

When no rule induce to a final state, a reverse-way track helps to look for other answer, otherwise an other mistake is pointed out.

The "Brigitte Frybourg" Laboratory finishes off the model of the French language that we had; then the morphological analyser is able to accept a rich typography (pointed letters).

Syntactical Analyser

The categories and the variables provided by the morphological analyser and the relation lists of dependency help the syntactical analyser to build one or several structures for a phrase. Once built, with a context-free rules, the analyser checks the concordance of gender and number (masc., singular, fem. plural...) between the different elements of the structure, this is the way to detect which are the words between these where is a concordance mistake.

III - TGEF APPLICATIONS TO THE CONTRACTED BRAILLE TRANSLATION

The analysis of contracted Braille rules shows off that the grammar is equivalent to a finite states grammar; then, the TGEF on his own is necessary to realize the contracted Braille translation.

The TGEF is a set of programs allowing the spliting of a phrase into words, or group of words, and the restitution of informations about this spliting. The spliting is realized with the help of the identification of various datum (base, suffix, termination) making up a word, these are found in a dictionary. The grammar checks the concatenation of these different elements.

In case of Braille, the different spliting is guided by the string of types to contract and the linguistic information is represented by the translation of different strings.

The fact that identification, through the dictionary, is realized making up the longer coincidence between elements of the dictionary and the entrance-string. The blank-type is not a separator only, it allows the direct translation of locutions such as: that is to say, here and now ... without any specific separator.

If the grammar refuses this spliting a back track technique allows to split the strings again.

If no solution is found, the user is aware of an error and he can change datum either on text, or grammar or dictionary.

Specific functions to translation have been inserted such as:

- to forbid the translation of some strings or force to begin a new paragraph or pass to a new page.

Following request of the user, the outcome can be done with dots.

SECOND PART

I - TGEF EXTENSION TO THE EDITING TAPES CONTROL

The very problem concerning the book printing in contracted Braille is not so much the automatic or assisted translation of a text as the data acquisition of this text. Then, we have looked for the best way to do this data acquisition in order to be the less expensive as possible. The solution seems to be the editing tapes of texts photocomposition machines. At the moment, some literary informations are stocked on these tapes, and if we could find the manner how to translate them in Braille, this will be the answer to get a book in Braille with a very low price. These editing tapes are used to instruct machines which produce photographic templets necessary to print offset plates. These contain two kind of informations: typographic information used for the page setting and the definition of the sorts of used types, and the literary informations. The typographic information is more or less rich according to the photocomposition machines.

The translation of these tapes sets gives a lot of problems ... the most important is because these tapes have got errors coming from the acquisition. Therefore it is important to be able to find out these errors in order to remove them. For the translation in Braille, it is just enough to leak out the necessary information for this translation (few typographic information in Braille) that is to say all typographic types respecting the text and the essential typography.

In function of these difficulties, it is not too much to strech our researchs to the entire analysis of these tapes and translate them. These tapes which are used to instruct a machine are obligatory analysable by a finite states automaton. The typography and the text are defined through completely different grammars. On the other hand, the typographic types and the types of the text are narrowly mixed up. The analysis of a tape can only be done by two automations working at unisum. Then, we have realized an extension of TGEF in order to be able to implant several grammars and dictionaries working at unisum. A surgrammar allows to control different automatons. This surgrammar keeping all PIAF caracteristics complete interactivity as well during the creation or the modification of the surgrammar as during the execution.

Such an extension of TGEF will allow to realize an entire string of treatment.

2 - TRANSLATION OF CONTRACTED BRAILLE TO ITS INTEGRAL FORM

We are interested in the contracted Braille translation to its integral form because this would allow to build a redaction tool for the blind people. Indeed, with the help of an adjusted terminal, for instance a terminal using the ephemeral Braille, blind could write a text in contracted Braille, correct it with an editor and afterwards print it on a normal and small printer. This would be a great help for communication between blind and seeing people. It is obvious that at the moment such a technique can only be experimented. The equipment to use is very expensive still, but with the evolution of mini and micro computers, such a solution can be looked in a near future.

For this application, and for others utilizations, we plan to implant PIAF on mini or micro computers.

CONCLUSION

The application to the contracted Braille can be done as well for an other language as French, without changing PIAF modules.

For a language of which contracted Braille would not be clearly defined, PIAF has got frequency counters of utilization which could help to decide help to decide some types strings interesting to be contracted.

At the moment the system is implemented on IMB 360/67 CP-CMS system or on IBM 370 CMS/VM.

The translation in contracted Braille is only a first step to the research of informatic solutions to blind's problems. It is important to develop the systems of recuperation of coded informations such as editing tapes and tools making easier the redaction and the research of informations.

World Council for the Welfare of the Blind

Committee on Cultural Affairs

International Conference on

"Computerised Braille Production - Today and Tomorrow"

London, 30 May - 1 June 1979

COMPUTER-AIDED BRAILLE TRANSCRIBING

by

Barry Hampshire

Synskadades Riksforbund

S-122 88 Enskede

BARRY HAMPSHIRE
SYNSKADADES RIKSFÖRBUND

COMPUTER-AIDED BRAILLE TRANSCRIBING

During the past three years, a research and development programme has been in progress to expand and improve the braille production facilities at the Swedish Federation of the Visually Handicapped (usually abbreviated to 'SRF'). As part of this programme, the first stage of a new production system for braille was introduced into SRFs printing house during March/April 1978.

An important design principle guiding this development has been to integrate as closely as possible the new equipment with the existing production system. The rationale for this was to try and minimise the negative effects of introducting a new technology - these being the loss of job satisfaction and job security by the existing staff, and disruption of the continuing production of braille during the introductory period.

Another aspect influencing the concentration on existing staff was that employment of staff is very expensive in Sweden. Thus, the introduction of equipment demanding new staff to key in text would not have been a very effective development from a cost/benefit point of view, especially in view of the fact that the present increasing cost of braille production is largely responsible for the decreasing quantity of braille being produced at the present time.

Another important factor influencing the development of the system was the pattern of use and needs for braille in Sweden. Analyses of these showed that the strongest areas of need for braille were for more specicalised material; study material, material for work, reference material, etc.. The medium of braille is particularly advantageous for this type of material with regard to ease of use. Furthermore, this area of use seems to be the most promising area of application for the braille display/cassette equipment now becoming available. As a result, to increase the efficiency of production of specialised material and to have this potentially available for distribution on braille encoded cassette, were significant design factors.

Introduction of computer aided braille production systems in other printing houses has usually only affected the production of ordinary prose books or magazines. Production of specialised or otherwise complex material has often continued, in these cases, to be produced in the conventional way using entirely mechanical equipment. In Sweden, production and distribution of talking books, magazines and newspapers is already at quite a high level so needs in this particular area do not claim quite such high priority as they might in other countries.

To summarise briefly then, the aim of this first stage of the development programme was to take full advantage of modern technology to increase braille production, but at the same time, to also develop and improve the work of the existing braille transcribers employed at the printing house. The new equipment should also be compatible with the development of more sophisticated techniques in the future. The new equipment introduced in the Spring 1978 is in accordance with this aim.

At the present time (March 1979) it consists of six main units - one braille encoding unit, two correcting/editing units which can also be used for braille encoding, a matrix printer (inkprint), a braille printer and a stereotype machine which can be controlled from a cassette reader (see Figure 1).

The encoding unit consists of a braille keyboard (corresponding in layout to a Perkins keyboard), a visual display unit and computer cassette machine for encoding the braille characters written on the keyboard. The visual display unit has an internal memory of ca 6,000 characters and text in this memory is available for checking and correction at any time. When this memory is full, the text is transmitted to a cassette and encoding continued. Although all three units used for encoding have correcting facilities, the use of these during transcribing is kept to a minimum. Correction of the text is carried out as a separate production stage after proofreading.

From the braille encoded cassette, proofreading can produced in either inkprint or braille. For inkprint copy there is a matrix printer which produces a single inkprint character corresponding to each braille character. Thus, word and letter-string contractions, which are single or double characters in braille are also printed as single or double characters in the inkprint copy. These same inkprint characters are used in the display to the encoding/correction units. In this way an inkprint format is produced which corresponds exactly to the eventual braille format. A braille printer is also available for production of proofreading copy in braille.

After proofreading, the text on the original cassette is corrected and edited, if necessary, on one of the correcting units . These correcting units are similar to the encoding units, the only differencebeing that there is an additional cassette machine which allows encoded text to be read directly into the visual display unit. Among the text processing facilities available with these units are automatic search of letter-strings and automatic adjustment of the text, i.e. re-formatting the line length and the number of lines per page to any set value.

All these units are based on standard equipment used in ordinary printing houses and newspaper composing rooms and are manufactured by Tele-ekonomi AB. This company also carried out all special modifications and interfacing which were necessary for braille applications. Two units exist for the actual production of braille which is embossed directly from the cassette obtained after correction has taken place. The first is a standard stereotype machine which has been modified for automated embossing.

For material required in single, or a small number of copies, a direct braille printer - a SAGEM REM 8BR - is used. Although this printer is not fast, it produces good quality braille and can emboss on both sides of the paper.

As stated earlier, the purpose of this new production system has been to increase the flexibility and efficiency of braille production, and, at the same time, continue to utilise the skills of the existing braille transcribers. There are a number of ways in which this system achieves this aim, of which a few can be outlined here.

One of the most important gains of the new system is the ability to correct, up-date, or otherwise edit, text, thereby avoiding a large amount of re-writing that is often necessary with the old system. It is this aspect which should create the greatest savings in production time over the old mechanical methods. This capacity also allows new types of production to be accepted, for example, handbook or reference works requiring frequent up-dating. Once the book is encoded, cassettes can be edited relatively quickly and easily. Such up-dating is not realistically possible with traditional braille production techniques or with recorded material. Furthermore, with respect to more complicated text requiring complex tables, footnotes and other types of special layout, transcribing is enhanced by the availability of a visual display and text processing facility.

A further consquence of the introduction of the new equipment is that all production will be on ordinary braille paper. That is to say that the production of small numbers of copies on plastic by vacuum-forming from a paper master will no longer be necessary. Most braille readers do not like reading braille when it is embossed in plastic and paper is considerably cheaper than plastic. Furthermore, it is possible to produce braille on both sides of the paper using the SAGEM printer. This will result in considerable savings in the bulk of text books for students which, with the old system, were always produced using just one side of the paper.

A number of developments are planned for the immediate future. For example, an ongoing project is the interfacing of an electro-mechanical braille display to the keyboards units. This, together with certain visual signals being substituted by available ones, will allow blind braille transcribers to use the new equipment.

The main development to be carried out during the coming months is, however, a facility for translating text directly from digital media. To begin with this will be applied to fairly straightforward prose books as general book production is rather low in Sweden. The philosophy to be followed regarding this development will be to design a system which will utilise the close involvement of the existing braillists at SRF. That is, only those contractions whose occurrence can be defined completely will be translated by the program. Therefore the aim will be to avoid using contractions incorrectly even if a number of letter-strings remain fully spelt when they could have been contracted as a result. (This aim can never be realised completely as there will always be the possibility that foreign words, for example, are contracted. These errors can, however, be eliminated or at least minimised by a pre-editing stage.) Any potentially "contractable" letter-strings which cannot be defined completely are recognised during the translation process and if found can be displayed on the visual display screen with some surrounding text. This will allow a skilled braillist to decide whether or not they can be contracted, and if they can, to insert the contraction manually via the keyboard.

At the end of this translation errors should be at an acceptably low level to allow proofreading and correcting to take place as normal. The main complications are foreign words and special text forms such as capital letters, italics, etc. which are only sometimes marked by special braille prefixes. It is possible that some of these can be recognised from codes in the compositor's tape. However, there will always be a need for a thorough proofreading after translation.

The same basic equipment will be used for this computer-aided translation as for manual encoding. Therefore existing braille transcribers who are already experienced in the use of text-processing techniques will also carry out the computer-aided translation.

The development of this basic facility for converting digitally encoded texts to braille opens a range of possibilities beyond just book production. The cassette equipment used is very amenable to interfacing to a range of other equipment, such as telex, text processing equipment, typewriters, modems, braille display/cassette machines etc.. The necessary interfaces are, in many cases, already available from Tele-ekonomi as much of their work involves interfacing their cassette equipment to a variety of text processing equipment.

This flexibility has particular relevance to braille users working in jobs requiring frequent monitoring of information or use of handbooks or reference works. Some examples of how this flexibility can be utilised is shown in Figure 2.

This shows input from a typewriter which simultaneously encodes the text being written on a cassette. This can be utilised in the context of local or regional production facilities (see below). Text processing equipment is becoming more and more common in offices and many of these machines have standard telecommunications interfaces. These allow a modem to be connected to the text processor and thereby a telephone link established between this equipment and the braille printing house. This would allow internal memos, reports, etc. to be translated to braille so that these can be available to blind employees.

With the development of braille display/cassette machines, a great many possibilities present themselves, especially with the more sophisticated machines which have search facilities. The most interesting area of application is for reference material which is long but only small parts of it are read at any one time, thus implying the need for efficient search facilities. There are many aspects to information provision for the visually handicapped person who requires information for his work, especially with regard to the editing of text material. These are discussed in more detail in Appendix A.

The flexibility and relatively low cost of the microprocessor-based cassette equipment allows a number of different system designs for de-centralised braille production facilities. One example is shown in Figure 3. This consists of a large-print typewriter with a cassette machine interfaced to it. Thus, the text is simultaneously encoded as it is being recorded. The cassette can then be sent to a regional centre with braille production equipment. Re-formatting of the original text, assuming it has a simple layout, and inclusion of whole-word contractions, can be achieved by the microprocessor in the cassette machine. For more complex texts or texts required in multiple copies, these cassettes can be sent to the central printing house, where the text can be edited and printed via stereotype machines.

Finally, synthetic speech equipment has been developed at the Department of Speech Communication at the Technical University in Stockholm. A project is currently being carried out to allow speech to be generated directly from cassettes used for braille production. This means that texts obtained from printing houses (and elsewhere) which are in ordinary print, can be used to produce synthetic speech on ordinary cassettes before being translated into braille. The present quality of the speech is such that only work-related information material can be considered for this type of production. However, this represents a considerable expansion in the distribution of material possible from the new text processing equipment.

The developments currently being pursued at SRF are, then, aimed towards achieving greater flexibility with regard to the type of material that can be produced. It is hoped, furthermore, that this flexibility will be achieved by making the braille production system compatible with as many other types of text processing equipment as possible. It is also hoped that the increasing availability of up-to-date material in braille will also motivate more severely visually handicapped people to learn to use braille.

FIGURE 1: Braille production system at SRF showing also planned development of computer-aided translation of compositors' tapes.

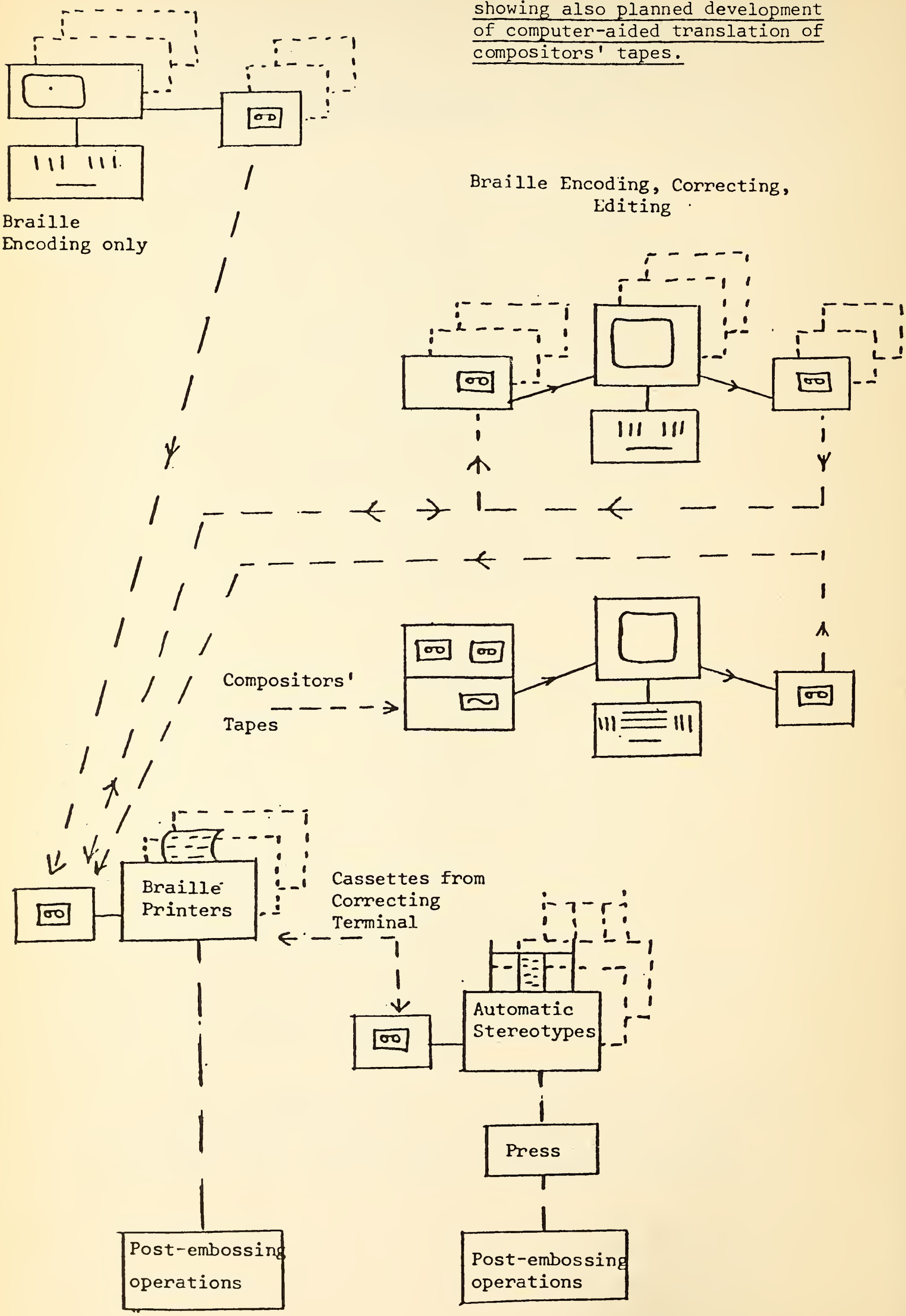


Figure 2: Some possible external inputs to a regional/central braille printing facility

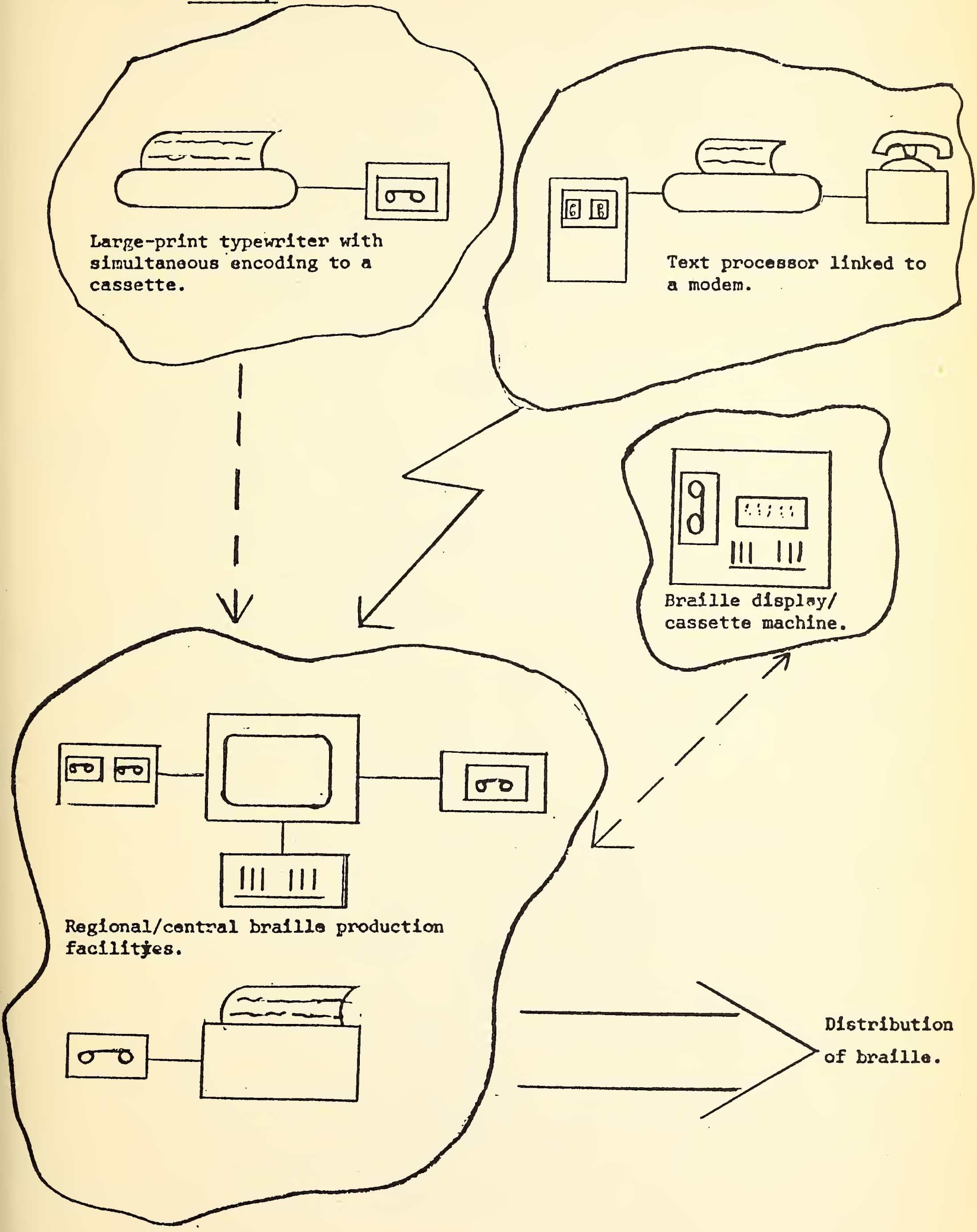
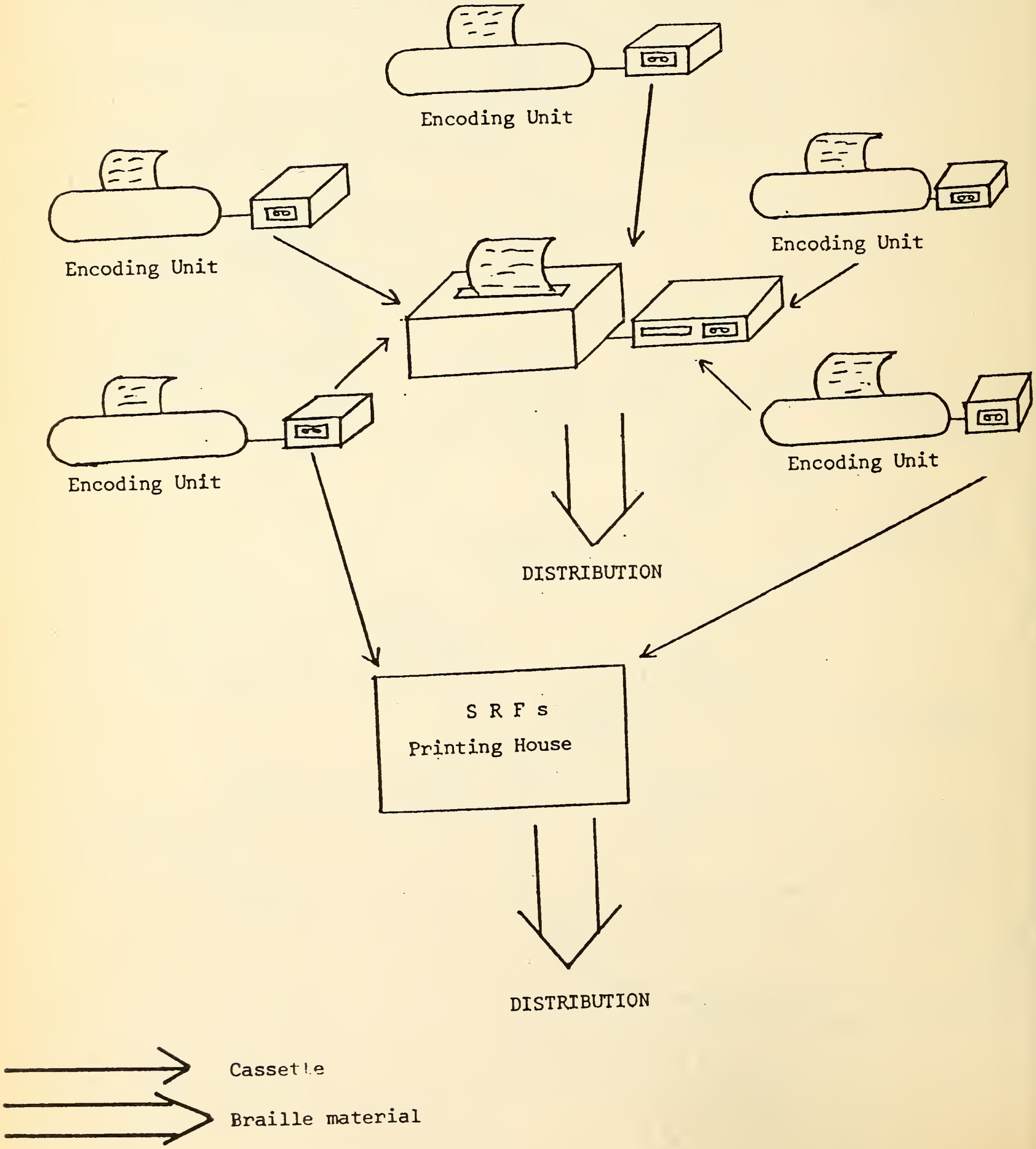


Figure 3: One possible organisation for a de-centralised braille production facility



APPENDIX A

Information Provision for Severely Visually Handicapped People at Work

Processing of information is becoming an increasingly important and common component of working life. Therefore, the possibility of the severely visually handicapped to gain access to material, and to search through and select those parts which are of particular interest, is becoming a correspondingly important factor in the ability of many visually handicapped people to obtain work.

Present provision from centralised facilities is very limited, and, with traditional production techniques, there is no real possibility to make any significant improvements. In most cases the severely visually handicapped are dependent on sighted colleagues or special helpers to provide the information and material they need to carry out their work.

Recent Technological Development

Developments in technology are now beginning to offer significant improvements in the provision and processing of information for severely visually handicapped persons. These development areas include the following:

- Conversion of Digitally-Encoded Texts: An increasing amount of print material is being produced via computer-aided phototypesetting, including newspapers, magazines, journals and books. Also, a great deal of reference material is now being stored in digital form. Utilisation of this digitally stored material for direct translation to braille or synthetic speech is now possible.
- Searching Within Stored Texts: Storage of text, using disc or floppy disc systems, permits fast searching through the stored material. The "text-base" can be

searched using a particular "search profile" which can define a special area of interest. This "search profile" can, for example, be built up from a number of keywords linked by a simple logic.

- Individual Aids for Storing and Reading Information: A number of pieces of equipment are now becoming available which consist of a braille keyboard, an electro-mechanical braille display showing from 20 to 40 characters, and a cassette recorder/player allowing recording of speech or encoding of braille from the keyboard. In addition, words or letter-strings can be written from the keyboard and these can then be searched for on a braille encoded cassette at high speed. Alternatively, spoken material can be indexed with braille thus allowing more efficient searching within this recorded material.

- Synthetic Speech: Equipment has now been developed which can generate speech synthetically from digitally encoded text, e.g. compositors' tapes. This means that text converted from compositors' tapes can be used for either braille or speech output.

Limitations of Braille and Speech

Neither braille nor speech can compete directly with print as an information medium. Therefore, straightforward conversion of print texts to one of these forms is not sufficient in itself for efficient information provision.

In the case of speech, this can be produced relatively quickly. However, if there is a lot of material and it is of the type that only needs to be referred to, rather than read from beginning to end, it can be very difficult, if not impossible, to use effectively. In the case of braille, this is expensive and takes a long time to produce if it must be keyed in specially for braille production. Braille, however, can be more convenient to use for certain types of material than speech.

Thus, the optimum utilisation of each of these media requires that the material can be appropriately edited and structured so as to aid the reader as much as possible. This, in turn, implies that, if the above listed technological developments are to be applied successfully, then their application must be carried out in conjunction with careful preparation of the text.

Design of Provision

At least two types of information need can be described in the present context:

- general material needed to keep up with everyday news, events and debates and background material to the particular area the person is working in;

- specialised handbooks or reference books containing much detailed information, although only a small part of it is ever needed at any one time.

General Material

There are at least two possible approaches to the provision of general material (see Figure A.1). Both involve the reading in, converting and storing of information from digitally-encoded material.

The first approach involves the selecting of headings and abstracts (if present) of all the available articles and distributing these either in braille or sythetic speech. The readers may then contact the printing house and order the complete article(s) that are of interest. The second approach is for each subscriber to the printing house to have a specific "search profile". A fairly well-defined interest area can be specified by linking several key concepts and words with a simple logic. Then, all available texts can be searched using each of these search profiles at regular intervals, once a week, for example. Those articles selected out can then be automatically distributed to the subscriber.

The latter approach obviously allows much quicker distribution of material than the former but, on the other hand, there is a risk of receiving irrelevant articles.

Reference Material

The braille display/cassette equipment now becoming available, allows an effective way of distributing reference material (see Figure A.2). This material can be distributed from the printing house in cassettes which the reader can use in his own braille display/cassette equipment. The search facilities available with the more sophisticated of these machines allow quick access to any part of the encoded material.

Updating of this material can be carried out at the printing house using the text processing facilities and a new cassette sent out. The original cassette can then be returned and

There are many other ways of organising the details of an information provision system. The purpose of this short appendix has just been to illustrate the possibilities that are becoming potentially available with regard to information distribution and to reinforce the particular usefulness of braille in this context.

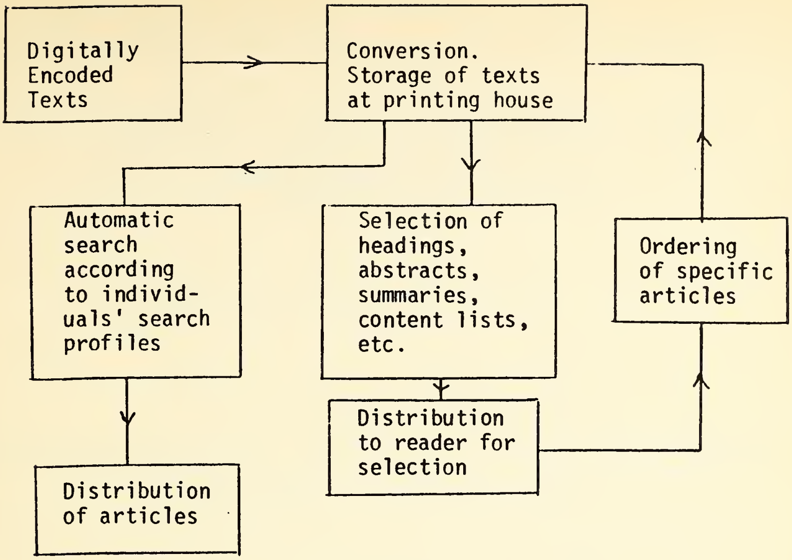


Figure A.1: Provision system for general material

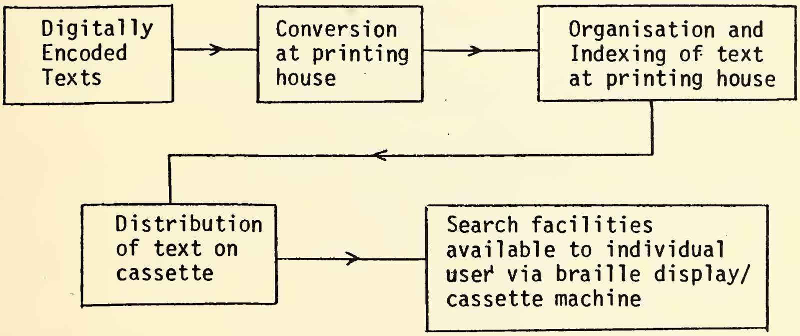


Figure A.2: Provision system for reference material

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

APPLICATIONS IN THE NETHERLANDS

by

R. van Vliet

Nederlandse Blindenbibliotheek
Zichtenburglaan 2660
2544 La Hage

and

M. Vliegenthart

Philips Research Laboratories
Dravikstraat 24
Geldorp

Paper to be presented at the Conference
"Computerised Braille Production - Today and Tomorrow",
30 may - 1 june 1979

Applications in The Netherlands
Part 1

Applications in The Netherlands

by Dr. M.J. Vliegenthart.

This paper is presented on behalf of the VNBW, the Association for and of the Blind in the Netherlands. This association is a society of: organisations of the blind, schools for the blind, libraries for the blind and braille and talkingbook producing organisations, workshops, fundraising and welfare organisations.

In the early senventies the Dutch Library for the blind in The Hague started a computerised production of braille. At that time the equipment, a D.E.C. PDP-8, only produced punched paper tapes for offline braille embossing on up to ten coupled braille embossers.

Last year a new system has been installed: a D.E.C. PDP-11 running under control of the UNIX time-sharing system. A variety of peripherals is connected, as there are hardcopy terminals, CRT-terminals, lineprinter, papertape equipment, a magtape equipment, floppy disks and two Triformation LED-120 braille lineprinters.

The braille printing house Sonneheerdt is using a set of word processors to prepare magnetic tape cassettes to control zincplate embossers. Another wordprocessor has been installed at the central office of the VNBW, it is used as an ordinary office tool. Magnetic tape cassettes are used to transfer information stored in the word processor to a LED-15 brailler, also available in the office. (c.f. the paper of mr. Spanjersberg to be presented at this conference).

Our solution to the translitteration problem proved satisfactory. Efforts are now directed to the improvement of the input/output systems.

Our objective is to free the keying personel from layout troubles and the application of braille rules. Therefore we have chosen for unformatted input, i.e. a merge of text and lay-out indicators. (c.f. the paper of mr. van Vliet below) The advantages are clear: input format has no impact on the output format; various output modes are possible, for instance contracted/uncontracted braille on paper of any size, normal and large print and output on cassettes for reading aides.

Of course a number of input terminals could be connected to the translitteration computer. Of the many disadvantages of this approach we emphasize the lack of portability of these terminals. The introduction of micro-electronics has enabled the development of a simple and cheap off-line text input unit (t.i.u). It is composed of a microprocessor to which a keyboard, a visual display and a cassette recorder are connected. A prototype of such a t.i.u. has been build and tested at the technical university of Delft. It seems that the edit facilities of most home computers will meet the requirements of a t.i.u.. We prepare a pilot experiment in this field. Developments like this will increase the need for fast, reliable braille printers.

As mentioned earlier our translitteration system is ready to produce tapes for reading machines. However, for which one? Currently there are at least six different systems, differing in baud rate, differing in signal definition, differing in code, differing in line width, differing in record lenght, differing in number of tracks used, differing in tape speed and so on. With great emphasis we underline the resolution of the Madrid conference (april 1978) on standerdisation of equipment. The potential users and customers have the right to claim progress in this respect and highly disapprove unwillingness and/ or reluctance of manufacturers to agree upon standards.

Paper to be presented at the Conference
"Computerised Braille Production - Today and Tomorrow",
30 may - 1 june 1979

Applications in The Netherlands
Part 2

On the processing of formatted text

by drs. R. van Vliet

Summary

In this paper we report on research concerning processing of formatted text. This research was performed in autumn 1978 at the Nederlandse Blindenbibliotheek.

Advanced text processing systems are based upon the use of unformatted input. This makes them flexible with respect to output format and output code. Computerised braille systems will develop in the same direction.

The appearance of optical character recognition (ORC) equipment opens the possibility of directly feeding inkprint text into a computerised braille system. However, these texts are formatted. This raises the problem of automatically recognising the logical structure of the text. A preliminary solution to this problem is presented.

1. Formatted/unformatted input

This distinction is important for almost every component of a computerised braille system. Therefore we first pay attention to the differences between these two types of input.

A. Formatted input

This is the input a typist is used to. The arrangement of text over the page (page layout) is directly controlled by the typist. Programs do not change it afterwards.

B. Unformatted input

In this case the typist needs not bother about the layout of the input text he is preparing. A subsequent program (layout program) will rearrange the text over the pages. Although the typist need not care for a neat layout, he must insert directives for the layout program into the text. Such directives are termed layout indicators. They are constructed such , that a comupter program (and of course human beings) may easely recognize them.

2. Advantages/disadvantages for computerized braille

Formatted input seems to be easier for the beginner. No magic in the machine may destroy his splendid text layout. However, the input text is to be transduced by a translation program into a braille-coded output text. If text layout is to remain unchanged after this transduction, the word size (the number of positions of the word on a line) must not be affected by the translation. When translating texts into braille this is generally not the case: sometimes signs are added (digit mark, capital mark) sometimes a group of letters is contracted to one braille sign. Problems are more evident when the transcription is to be made into contracted braille.

As a consequence when a braille transcription program uses a formatted input text a number of braille rules can not be brought in by the program. Namely, those rules that change the number of signs of a word. (Note: by a "word" we mean any part of the text inbetween two blanks.) In this case the typist must know all about these rules, and key in special signs (digit mark, capital mark etc.) as required. Nevertheless he will probably prefer typing this "ink print braille" over brailling by hand using a regular Perkins brailler. Moreover, once the text has been keyed in and transcribed, one can make as many braille copies as desired.

There is another disadvantage in the use of formatted input. We come to that when talking about text correction.

Altogether the use of formatted input for the production of braille via a computer is generally restricted to the production of uncontracted braille and not very wide-spread.

The use of unformatted input requires that the typist (or someone else) inserts layout indicators into the text. Today programs with a feasible set of layout indicators are available. A few hours experience is enough to learn the preparation of simple texts. Stuff like tables and indexes is more complicated. In fact this kind of text can only be handled by sophisticated layout programs, to be technical: by programs having a macro facility for the layout indicators. The difficulty in applying layout indicators is, that their affect on the structure of the text can not be seen until a proof-print has been made. As a rule we do not make such proof-prints. It turned out that a one-year's experience had taught out typists how to use layout indicators without making serious mistakes.

The use of unformatted input and a layout program offers the following additional possibilities: a text may be output in different formats (e.g., braille on pages, braille on display), the text may be output in contracted as well as uncontracted braille or in largeprint, reprints of a book may easily be edited form earlier versions of the input text.

3. Inserting layout indicators in a formatted text

Although we prefer to use unformatted input texts, we have to handle formatted input from time to time. As a rule it concerns old braille-coded tapes from which reprints in a different format are required. The introduction of OCR equipment will make an enormous amount of formatted text available. The idea has come about, that most troubles in treating formatted input can be avoided if we had a program that inserts appropriate layout indicators in the input text. In this way the input format itself can further be discarded.

The problems such a program has to deal with are:
- recognizing page headers and page footers;
- recognizing section headings etc.;
- recognizing paragraphs;
- recognizing the indent structure.

The algorithm we have found is not full-proof. But the highly interactive approach discussed in the next section may well be prefered above retyping or just correcting the formatted text.

The number of header and footer lines per page is specified to the program. These lines are skipped. Any information stored therein is lost.

Headers are treated as separate paragraphs. If required they may be recognized by their shortness (typically one or two lines). A left indentation is taken care of by the standard paragraph algorithm. Right indentations may cause troubles, especially in the case of centered headings.

The standard paragraph algorithm operates on all of the input text, except page headers and footers. Treatment of the left indentation must be included in it. The algorithm must be aware of the pagewidth of the input text, which must be specified by the user. The basic concept is, dividing the text into structural fragments, termed blocks. A block may be a paragraph, a heading, a part of an enumeration, or the like (see examples at the end of this paper).

Human beings may easily recognize the blocks in a text. Blocks always start at a line bound. Words cannot be moved over a block bound without affecting the readability or even the information contents of the text. On the other hand, words may be moved over a line bound inside a block at leisure. Readers would only complain about unaesthetic scarcely filled lines. It is worth emphasizing that a change in the left indentation need not imply the end of a block. For instance, a paragraph may well start with a single indented line.

How is the computer to recognize the blocks in a text? Some heuristic criteria:
- blocks may be separated by blank lines;
- all lines except the last of a block are filled, i.e. they end near the right edge of the paper;

- the last lines of a block tends to be short (unfilled) and as a rule is terminated by a punctuation mark.

The indentation within a block may be considered as indepedent of that of surrounding blocks. All lines except the first of a block have the same left indent. The indent of the first line may differ (see examples). The computer applies a hierarchy of criteria to locate block bounds. Empty lines imperatively indicate block bounds, as do changes in the left indent other than those occurring after the first line of a block. If these two criteria do not apply, the linelength is taken into account. Short lines terminated by a punctuation mark indicate the presence of a block bound, long lines not terminated by a punctuation mark indicate the absence of a block bound. The remaining two cases, short lines not terminated by punctuation and long lines terminated by punctuation, are questionable.
An interactive program should interrogate the user about what to do in such situations. In this regard some punctuation marks are stronger than others. E.g., question mark and colon are strong, whereas a comma is weak.

Once the computer has recognized a block bound, layout indicators are inserted to set the indent equal to the block indent, and to force a local indent for the first line of the block.

Note, the impact of significant right indent on this algorithm. Each line is recognized as short and hence indicates the end of a block. So each line is regarded as a block of its own, and will be given the left indent it had in the input text. Any difference in the input and output pagewidth will show up at the right edge of the paper. If the output pagewidth is reduced so much, that these lines no longer fit on the output lines, a dramatic and probably unwanted change in the layout of the output text will be enforced.

Implementing the algorithm discussed above is easy enough. It is sufficient to have a window showing two subsequent lines of the input text and a flag indicating wether or not a new block starts at the first line shown in the window. The window is shifted over the text line by line. Each time before shifting, it is decided wether or not a block bound must be located inbetween the two lines window. If so, the flag is set, otherwise the flag is cleared.

A block bound is allocated if:
- The first or second line in the window is empty.
- The left indent of the first and second line are unequal, whilst the flag is not set.
- The first line is recognized as short.

4. Text correction

Most computer systems are equipped with programs for text correction (text editors). The correction facilities offered by these programs are usually restricted to:

A. Easy means to indicate locations in the text where corrections are to be made. E.g., indicating page and line number, or specifying the text environment (surrounding text).

B. Facilities to change, delete, or insert text at spots indicated.

C. Facilities to repeat specific corrections. E.g., a correction program could change any occurence of "text 1" into "text 2" throughout the complete text file.

Such facilities are satisfactory for the correction of unformatted texts, but they are certainly inadequate for efficient correction of formatted texts (see below).

Correction of formatted text

To understand the problems related to the correction of formatted text, we distinguish between "local corrections" and "corrections having a global effect".
Examples of local corrections are corrections of typing errors and spelling mistakes. Corrections having a global effect are insertion or deletion of a word, a complete line, or even a complete paragraph. Such a correction may affect the contents of subsequent lines and pages.

Commonplace correction techniques only apply to local corrections. Note, that all corrections in unformatted texts are local. Hence conventional programs can be used to correct unformatted texts and local errors in formatted texts, but they are clumsy when used to bring in corrections having a global effect. Let us explain.

Assume, a word must be added to a line of formatted text. The line may then become too long to meet the format requirements. Some of the text on its right must be shifted to the left of the next line. This line may again become too long, so that text has to be shifted to a subsequent line, and so forth. This continues until the shifted portion of the text happens to fit on the next line. So, the insertion of a single word (even of a single character) may well imply the correction of a batch of lines. Things are even worse if the correction extends beyond page bounds. If, for instance, one single line must be shifted to the next page, this page will contain too many lines, which may imply shifting to a subsequent page, etc. This continue until a page is reached whereupon the shifted lines can be made to fit. Taking into account page headers and page footers further complicates the problem.

Deletion of course is simpler than insertion, although it too may require a fair amount of correction. Experience with formatted input texts has shown us, that corrections having a global effect occur quite often and may very well extend over several pages. Leaving on each page one extra empty line to be used for these corrections, is a limited improvement that enlarges braille books by about 4%.

To cope with the correction of formatted texts we plan to develop a special correction program. It is to operate on a highly conversational base. Its operation is outlined below.

First the location where the text is to be inserted is indicated. Next the insertion is keyed in. The computer now starts listing the corrected lines. Text shifting is performed automatically. Printing continues until the new text and the original text are in phase again. The operator may interrupt the listing as soon as he does not agree with the text shifts performed. He then indicates a better shift, whereupon listing is resumed.

Problems of importance for automatic text shifting are:

A. saving the left margin (how to distinguish between single indents and hanging indentation).

B. removing old hyphenation, bringing in new (the well-known hyphenation problem).

C. crossing the page bound (shifting over headers and footers).

This program will be based upon the concepts outlined in this document together with a set of standard modules of a layout program.

Examples of block bounds in a formatted text to be found by a program.

Example 1:

This is the first block of a formatted text that can easily be identified by a block identifying program.
This is the second block of a formatted text that can easily be indentified by a block indentifying program.

Example 2:

The problems such a program has to deal with are:

- recognizing page headers and page footers
- recognizing section headings etc.;
- recognizing paragraphs;
- recognizing the indent structure.

Each line of this example should be indentified as a separate block.

Example 3:

This is the first block of a formatted text in which it is difficult to identify separate blocks. This isn't often found in inkprint but we encounter this system in old types of braille books.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

COMPUTERISED BRAILLE PRODUCTION IN AUSTRALIA

by

John W. Berryman

Royal NSW Institute for Deaf and
Blind Children
Box 4120 GPO
Sydney, NSW2001

PAPER TO W.C.W.B. INTERNATIONAL CONFERENCE ON "COMPUTERISED BRAILLE PRODUCTION - TODAY AND TOMORROW"

"COMPUTERISED BRAILLE PRODUCTION IN AUSTRALIA"

SUMMARY

The Royal N.S.W. Institute for Deaf and Blind Children operates Australia's only computerised Braille production system. The installation is located at St. Leonards, a suburb of Sydney. Hardware includes a Data General Eclipse mini-computer and two Triformation Systems LED-120 embossers. The software was purchased from Duxbury Systems Inc. The system has been operational since May 1978.

Australia has some unique characteristics, some of which present problems in the area of blind welfare generally, and provision of Braille, more specifically. The total population, and the population of blind persons, is relatively small, and is widely dispersed. Each state has one or more agencies providing welfare services to the blind, and one or more Braille libraries. Co-ordination of services, especially Braille production and libraries, on a National basis has not been achieved. To date, no National Union Catalogue of Braille material has been compiled.

Initially the relative lack of financial support from Federal or State governments to Braille producers was a problem, but, the public purse-strings have now been loosened.

In addition to computerised Braille production, there are several centres throughout the country using manual production techniques. Most of the personnel involved are volunteers. This results in inefficient, but inexpensive Braille production, and potential customers often baulk at the prices asked for the computerised transcription service.

It has taken several months to gain patronage of the transcription service. However, sales are now being made to Welfare agencies, Braille libraries, and some private persons and to the N.S.W. Education Department.

THE INSTITUTE

The Royal New South Wales Institute for Deaf and Blind Children is the oldest children's welfare organisation in Australia, founded in Colonial Sydney in 1860. The Institute provides schooling and accommodation for five different categories of children - deaf, blind, educable deaf/blind, trainable deaf/blind and multihandicapped blind. A range of services is provided including medical, dental, and psychological care, vocational training and career education, parent counselling, library services and Braille transcription services.

Children from throughout N.S.W. and the Australian Capital Territory attend the Institute's Deaf and Blind Children's Centre, which includes modern schooling and residential facilities on 16 hectares of land. The N.S.W. Department of Education is responsible for the provision of teachers at the school for deaf children, and the school for blind children, which includes the educable deaf/blind. The staff at the deaf/blind trainable school and the multi-handicapped blind children's special school/nursing home are employed by the Institute.

The Institute is governed by a Board of Directors presided over by Sir Garfield Barwick, who is the Chief Justice of Australia. Annual recurrent costs amount to almost A$2,200,000 (about £ 1,250,000), of which about half is met by gouvernment subsidy. Subsidies for major items of capital expenditure can be as high as 80% of total cost.

THE INSTALLATION

The Institute operates Australia's only computerised Braille production system. The installation was established in May 1978. The initial configuration of the Data General Eclipse C330 minicomputer included 96K bytes of main memory, 10 megabytes of disk, a tape drive, a 60 character per second printer terminal, two visual display units, and two Line Embossing Devices, (LED-120's), the last supplied by Triformation Systems Inc. of Stuart, Florida. At the time of writing we had ordered and were awaiting delivery of a further 32K bytes of main memory and a 50 megabyte disk unit to replace the current 10 megabyte unit. (This upgrade was completed with minimal disruption to operations towards the end of March this year.)

The application software was purchased from Duxbury Systems Inc. of Stow, Massachusetts. Others to use the Duxbury software are the Candian National Institute for the Blind, Toronto, and the Clovernook Printing House for the Blind, Cincinnati, Ohio.

DESCRIPTION OF THE PRODUCTION SYSTEM

The transcription process consists of three parts, viz:

1. The entry of text;
2. The translation of text into Grade I or Grade II Braille;
3. The embossing of the translated text.

Text entry is performed by operators typing the text at the two visual display units. Text entry is controlled by the Duxbury Text Editor programme[1], which has numerous advantages over the various Data General text editor programmes available. The Editor is a re-entrant subroutine, permitting editing to take place simultaneously at the two terminals. It is "line-oriented"[2] and operates directly upon a directaccess disk file. The operator is able to add, insert, and delete lines by reference to line numbers. It is also possible to locate text by searching, without knowledge of the corresponding line number. Alterations can be made on specific lines, groups of lines, or throughout the text.

For example, the command '15 TO 20 REPLACE TOM BY HARRY' would alter lines 15 to 20 in such a way that wherever the three-character sequence "TOM" used to appear, "HARRY" will appear instead.

The Editor also provides a facility for creating a sequential file from the direct-access file. This sequential file is used by programmes other than the Editor itself. A direct-access file can also be reconstituted from the sequential file. Once a file has become stable, the direct access file is deleted and set up again only temporarily for editing as necessary.

The second stage, translation, is performed by the Duxbury Braille Translator. This programme, written in Fortran IV, is an outgrowth of two earlier programmes DOTSYS II and DOTSYS III developed in the early 1970's at the M.I.T. Research Establishment. The Translator reads the sequential disk file produced by the Editor (or produced by another programme, such as a Compositor's tape convertor). The input text is in the form of variable-length lines (records) of ink-print text images. The output, another disk file, is the sequence of Braille signs equivalent to the input text.

The Translator's author, Joseph Sullivan, describes it as being almost completely table-driven, i.e., details of the translation algorithm are determined by tables read in at execution time rather than by the programme itself.[3] Tables exist for several languages, and there are different tables for 'American' English Braille and 'standard' English Braille. On the Data General Eclipse as configured at the Institute, the rate of translation is 1500 words per minute. Translation errors do occasionally occur. These can be corrected either by the insertion of special "editor's symbols" in the original text, or by updating the table, an action which will prevent recurrence of the error. The Translator expects to receive text just as it appeared in the original ink-print, with a few exceptions. For example, single quotes should have been typed as double quotes, even though a quotation may be included in another. A dash in ink-print should have been typed as two hyphens.

Additionally, the Translator recognises certain insertions in the text (editor's symbols) which are used to perform particular functions, either the special representation of certain punctuation and print conventions in Braille, or the formatting of the Braille pages. Examples of the former include capitalisation, italics, accent marks, and the correct sequencing of combinations of parentheses, brackets and italics. Examples of formatting include new paragraphs, new lines, the skipping of one or more lines, skipping to a new page from other than the last line, tabulation (e.g. for columnar information), hanging indentation and centred headings. The Translator automatically prevents the splitting of words at the end of lines, and looks after paging and page numbering.

Other functions which can be controlled through editor's symbols include switching between Grade I and Grade II Braille, the inclusion of ink-print page number in the Braille for cross-referencing and the automatic inclusion of a title on every page. As mentioned previously, situations may arise that cause problems with the automatic translation of Braille, usually deriving from the sound or meaning of the text. To overcome these, symbols may be included which will either enforce or prevent the contraction of nominated letter-groups as required.

For the most part, the insertion of editor's symbols is performed by the data entry operators as part of the data entry process. In some cases, where layout of the Braille is less straightforward (e.g., Braille calendars), the text is subjected to a pre-input step which involves hand-writing the required editor's symbols onto the source text before data entry.

The third stage of processing is embossing. The translated text can be transferred directly from disk to a Line Embossing Device (LED - 120) or can be printed out as ink-print Braille. At the Institute we use the former option and the embossed Braille is then proof-read.

Proof-reading is performed by a blind reader reading aloud from the Braille and a sighted reader checking what is read against the original source of the text. The Braille reader also uncovers spelling, contraction, and layout errors. Required corrections are indicated on the computerprinted text listing, which is returned to the text entry operator.

Corrections are made via the V.D.U.'s and the text is retranslated. The final corrected version is then stored on magnetic tape, from which Braille copies can be produced as required.

Each of the two LED - 120's can produce 2,000 pages of Braille in an 8-hour working day. The paper used is continuous sprocket-holed fanfold stationery weighing 148 grams per square metre.

The computer services staff trims, separates, punches, and binds the brailled sheets into volumes of about 60 pages. The embossing programme is flexible with regard to page length and spacing between lines of Braille.

Transcriptions

In the ten months, July 1978 to April 1979, 61 titles, made up of 276 volumes, were transcribed and are held in the magnetic tape library from which Braille copies are produced for sale. These publications include, novels, plays and anthologies of poems or short stories. Almost half of the works were written by Australian writers. About half were written specifically for children or adolescents. Almost all appear on the N.S.W. Education Department's high school syllabus in the subject of English, History or Economics.

Additional transcriptions, totalling 32 original volumes, have been made on request for various agencies and associations of the blind. These include magazines, newsletters, agendas and minutes of meetings, information pamphlets, calendars, football fixture lists, tape recorder operating instructions and so on.

AUSTRALIA

Australia has some unique characteristics, some of which present problems in the field of blind welfare generally, and in the provision of Braille, more specifically. The total population, and the population of blind persons, is relatively small, highly urbanized, but widely dispersed because Australia's cities are so far apart. The nation has an estimated 750 Braille readers.[4]

Each of Australia's six states has several agencies providing welfare services to the blind. Only a few organisations operate in more than one state. Each state has one or more Braille libraries. Co-ordination of services, especially Braille production and libraries, on a National basis has not been achieved. Only recently has the National Library of Australia given an undertaking to compile a National Union catalogue of Braille and other special media materials.

Braille is acquired either through purchase from the Royal National Institute for the Blind or the American Printing House for the Blind or by production in Australia. About twelve centres of Braille production are in use, all but one of these using manual production techniques. Most braillists are unpaid volunteers.

The situation in Australia with regard to library services for the visually handicapped has been described as fragmentary and inadequate.[5] Special problems have been recognised; the almost complete lack of serial and foreign language Braille and the shortage of current, topical, and varied reading matter found in periodicals. However, in Australia there is a growing awareness of the needs of the handicapped generally[6] and the visually handicapped specifically. There is a widespread desire to improve and expand services. The installation of the computer to produce Braille is one initiative which illustrates that progress is being made.

THE BIGGEST PROBLEM - MONEY

Computer-produced Braille is expensive. In its report of November 1978, the Libraries Sub-Committee of the Australian National Council of and for the Blind stated "...(the Computer) system is now fully operational producing Braille of a high standard. However, because of the high cost ..... its effectiveness is limited, (and) many organisations still rely on manual transcription using voluntary labour."[7] The National Federation of Blind Citizens, in a submission to the Working Party on Library Services to the Handicapped, stated as a principle that "... the Braille reader should be able to buy books of his choice ... of a price no greater than the printed copy."[8] In the same submission they say "Its difficult to see how Braille production ought to be funded, to say who should bear the ultimate responsibility for cost. However, funds should surely be available through a joint response from the Federal and State Governments with appropriate contributions from agencies for the blind."[9] Unfortunately, the required joint contribution from Governments has not been forthcoming.

However, there is a bright note. The N.S.W. Education Department has for some years operated a braille transcription service employing paid transcribers and proofreaders. This service aims at meeting the Braille requirements of children at the Blind School at North Rocks in Sydney, and of blind children who attend integrated high schools under the 'mainstreaming' programme. The Education Department has recently decided to supplement their service by commissioning transcriptions from the Royal Institute's computer service. It is probable that the Department will allocate A$50,000 (about £ 26,000) annually for this purpose.

The role of the National Library of Australia in providing special media reading matter has been widely discussed. Those in Australia concerned with talking book or Braille production or with library services for the blind, look with admiration and some envy at the work done by the Library of Congress in Washington. At a National Library of Australia Consultative Seminar on Library Services for the Handicapped, convened in Canberra in August, 1978, a documented recommendation stated that

"The need (exists) to direct substantial resources into the production of materials. In this regard the National Library should commission works."[10] The Director General of the National Library, Dr. George Chandler agreed. In the National Library's report on services for handicapped people[11] he stated "The right to read is shared by all, and ways must be found to satisfy the reading needs of handicapped Australians." Unfortunately, faced with the need to provide numerous new services at a time when staff numbers have been reduced and budgets pared, the National Library has not been able to involve itself in this way.

The cost of Australia's Computer Braille installation including the recent upgrade was A$150,000 (about £ 80,000). This was met entirely by the Royal N.S.W. Institute for Deaf and Blind Children. Annual operating costs amount to about A$70,000 (about £ 40,000). This cost is made up mostly of the salaries of the six members of staff, paper costs, and equipment maintenance charges. The only source of government grant is through the Federal Government's book bounty, paid through the Department of Business and Consumer Affairs. The book bounty was introduced by Act of Parliament in 1969 to encourage the Australian publishing and printing industries by means of a subsidy. The bounty pays one-third of the total cost of production of eligible publications. Most Braille books are eligible under the Act. After bounty, annual costs (in £ Sterling) are about £ 27,000. At out anticipated level of production of 24,000 original Braille pages per annum the unsubsidized cost per page is equal to about £ 1.13. How should this cost be met? The imposition of a transcription charge by the Institute to cover this cost has resulted in rather limited patronage of the transcription service.

The cost of copies presents a less gloomy picture. The cost of materials for a sixty-leaf volume, including mailing container, is about £ 1.23. The Institute charges about £ 4.00 a volume for copies, and a number of Braille libraries and some private persons have availed themselves of this source of Braille.

Because of the Institute's heavy financial commitments in areas of child care, the budget for Braille production is limited.

Where, however, we have lowered the charge for "special" jobs, patronage has been good. We produce a quarterly magazine for the National Federation of Blind Citizens in 100 copies at £ 2.00 per copy. Another magazine, on the subject of cricket, sold 50 copies at £ 1.70. I imagine that it will take some time for potential customers, who are accustomed to using volunteer transcription services, to accept that the costs of efficiently, promptly and accurately produced Braille is high. I am keen to hear the experiences of the managers of other computerised Braille systems.

COPYRIGHT

I intend not to dwell on the question of obtaining copyright permissions, primarily because the procedure is likely to change with the passing of new Federal legislation. Briefly, until now permissions have been negotiated directly with copyright holders or their agents, not through any clearing house. This is a time-consuming process but is usually successful in acquiring permission without fee. The new legislation may result in the need to pay for permissions but through a much streamlined procedure.

MISCELLANEOUS PROBLEMS

Hardware and software maintenance have created some problems in our three-supplier configuration. The mine-computer, manufactured and maintained by Data General, has performed well. We lose only about three hour's production time per months.

Maintenance of the LED-120 embossers initially gave us problems, which we thought had been overcome. Unfortunately, with the heavy increase in usage following the recent large orders from the N.S.W. Education Department, new maintenance problems have arisen. In spite of daily attention, embossing quality has dropped. A new problem has arisen, that of the embossing pins piercing the paper.

This has apparently occurred because the embossing die has been roughened from the heavy usage. At the time of my departure from Sydney three weeks ago, advice was being sought from Triformation. Unfortunately time-zone differences add to communication problems; in Florida business hours are 11:00 pm to 7:00 am Sydney time. Triformation Systems are not represented in Australia and the maintenance is carried out, under contract by Honeywell Pty. Ltd.

The LED's require remedial maintenance, some preventive maintenance, and adjustment of the embossing pins. Down-time is as high as 50 hours per month.

I am particularly keen on learning from other installation managers about maintenance problems - and more importantly the solutions that have been arrived at.

Some software maintenace has been required, as can be expected with any software. My policy has been to have Duxbury Systems carry out that maintenace. Joe Sullivan knows his programmes much better than I know them. Unfortunately the separation of 20,000 kilometres between Joe and me has caused some difficulty with communication. Maintenance by mail is slow and prone to some level of error, and we have suffered some loss of productivity as we seek ways to overcome translation errors. I wish, however, to emphasize that the Duxbury translation package is an excellent product in the main, and the "bugs" we have uncovered occur only in obscure, and infrequently arising circumstances.

As previously mentioned, the translation programme relies largely on a table-lookup process. Over time and with the experience gained by Duxbury users, a rather complete and accurate table has been built up following American Braille rules. In Australia, however, British Braille is the standard, and we have met the need to fairly extensively modify the tables. More than 5% of table entries have been added or changed. We now have a fairly complete table. It appears though, that British Braille rules are less consistent than the American rules and we will always be required to force or prevent some contractions by using editor's symbols in the input text.

PLANS

The Royal Institute hopes to make use of the research developments of Braille producers to provide both new services and better production techniques. Should a demand for such services be indicated in Australia, we would seek to provide:

- Braille mathematical texts
- Braille music
- Braille Bank Statements

To improve production we are interested in developments which bypass the input keying process, such as the use of compositor's tapes and optical scanners such as the Kurzweil Data Entry machine.

CONCLUSION

Some of the problems faced at the Royal Institute with the production of Braille are technical. The installed system and our operational procedures satisfactorily produce Braille of good physical quality with few errors. We have some hardware maintenance difficulties.

Our problems in Australia have related to the difficulty of finding the Braille readers and in having them state their requirements.

Funding has been a major problem, with most potential users of the transcription service being deterred by the need to meet transcription costs. Copies of our Braille books do go into several libraries, with some persons buying books for private ownership. The specific assistance of Government Departments and Government-funded institutions was at first conspiciously lacking, but we now have a public patron, the N.S.W. Department of Education.

J.W. BERRYMAN

SYDNEY

February, 1979.

REFERENCES

1. SULLIVAN, J.E. (1978): "The Duxbury Editor User's Manual", Duxbury Systems, Inc., Stow, Mass., U.S.A.

2. ibid

3. SULLIVAN, H.E. (1976): "The Duxbury Braille Translator User's Manual", Duxbury Systems, Inc., Stow, Mass., U.S.A.

4. WILSON, J.W. (1977): Letter from Honorary Secretary, Australian National Council of and for the Blind to Mr. D.W. Croisdale, Chairman, W.C.W.B. Sub-Committee on Computerised Braille Production.

5. National Library of Australia (1978): "Report on the National Survey of Library Services for Handicapped Persons". National Library of Australia, Canberra.

6. ibid

7. Australian National Council of and for the Blind (1978): "Report of the Libraries Sub-Committee".

8. National Federation of Blind Citizens (1978): "Library Services for the Visually Impaired Print Handicapped Citizen".

9. ibid

10. WHITE, DIANNE B. (1978): Report by Executive Secretary, Australian National Council of and for the Blind, on the National Library of Australia Consultative Seminar on Library Services for the Handicapped, Cangerra, August, 1978.

11. National Library of Australia (1978): Op. cit.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

BRAILLE AT THE TOBIA CENTRE

by

Monique Truquet

Centre Tobia
Université P. Sabatier
31077 Toulouse, Cedex

Braille at the TOBIA* center

"French Braille (grade 2) and synthetic voice"

M. TRUQUET** - G. GOUARDERES** - B. CAUSSE**

This paper presents 3 parts :

I - Braille at the TOBIA center (M. TRUQUET)
II - The lexicon (G. GOUARDERES)
III - The applications (G. GOUARDERES)

---

* TOBIA : Transcription par Ordinateur en Braille Intégral et Abrégé
(M. TRUQUET - M. GALINIER)

D. LEVY and J. FRONTIN are two researchers who participate to the edition and the Braille research.

(TOBIA depends on the LSI : Langages et Systemes Informatiques).

** CERFIA : Cybernétique des Entreprises Reconnaissance des Formes Intelligence Artificielle (G. PERENNOU)

A. FATHOLAZADEH and G. TEP have participated in this research.

I - Braille at the TOBIA center

This center has two aims :

- to produce Braille

- to continue the research in this domain

I-a : The Braille production :

TOBIA was created on june 27th 1977 at Paul Sabatier University in TOULOUSE.

Since then it has produced about 35 thousand Braille pages which represent 450 volumes (1). These documents have been obtained on a SAGEM device which was built on our own initiative with the help of SESORI (IRIA)***

The areas of the translated documents are wast : education, teaching books, literature, novels, legislation, contents for exams, bank statements...

The process consists in three successive phases :

1 - Data input (punch cards, magnetic tapes, minicartridges).

2 - Data transcription : two programs are used by TOBIA. An integral program compatible with all large or small computers and an abridged program which conforms the grade 2 of the manual published in 1964 by the A.V.H****. This program is operated on an IRIS 80 CII computer. Both programs are written in FORTRAN IV.
Direct demonstrations were given during the SICOB 76 the SICOB 77 and ICEVH Congress in 1977 due to the Cyclades network with the computer of (IRIA).

3 - Braille print out : it is obtained on a paper weighing 160 grams. The Braille text is produced at the rate of 15 characters per seconds. The Braille text will be available later on cassettes for the DIGICASSETTE.

This Braille document is proof-read by a blind who signals the errors. To correct them we have two possibilities :

. to directly correct the mistake, on the Braille text by time sharing, when only one a word is erroneous.

. to obtain once again the Braille text (from the erroneous page to the end) when more than one word is wrong or forgotten.

TOBIA is an experimental center. It proves two things :

% - with only : one typist
one proof-reader
two "dark"-readers
and-only
one SAGEM device

---

(1) Please find in the following pages the list of some braille documents translated by TOBIA.

*** Institut de Recherche d'Informatique et d'Automatique
**** Association Valentin Hauy

we can produce an important number of Braille pages ; in 1978 we have produced 21 thousand Braille pages, so we can say that this experiment is conclusive.

% - with the letters we receive we realize the needs of the blind :

. in the field of literature

. in the field of information

Only a computerized production with real means can satisfy these needs.

I-b : The research

At the present time two researches are being undertaken :

- translating of Music into Braille

- Simultaneous obtention of synthetic voiced speech and French grade 2 Braille.

1 - Translating Music into Braille

This work is carried out by J.FRONTIN. He continues the previous works made by some researchers from the American University in Waxhington. He adopted the data codes previously used.

To obtain the Braille musical document,3 jobs are necessary :

- The first one is to checks if every beat is complete, and for the piano, if it is the same beat for each hand.

. It gives the results required to obtain music, clef.

. It adapts : the symbolic groups are transformed into records of 80 characters. Each character has a defined meaning : the location on the record ; the name is given by the character.

- The second job : its fonction is to generate the Braille characters.

- The third job : it permits the edition of the Braille piece of Music.

2 - Simultaneous obtention of synthetic voiced speech and French grade 2 Braille

This research will be presented in parts II and III which follow.

Braille Integral (grade 1)

- Utopia by L. NOLDUS - Thomas More (Home reading) (79 pages)
- English through England's history. G. Verhoven (107 pages)
- The murder of Roger Ackroyed. Agatha Christie (200 pages)
- A new reader (178 pages)
- Deutsche sprachlehre für ausländer.Schultz und Griesbach (770 pages)
- Démocratie Française V. Giscard d'Estaing (220 pages)
- Por el mundo hispanico, extraits (61 pages)
- Pueblo, extraits (66 pages)
- Novelas jemplares : La Fuerza de la sangré. Cervantes (70 pages)
- Bâtir une grammaire, Delagrave (373 pages)
- La mystérieuse aventure de la princeyse Marie-Ange J.M.Fabre (17 pages)
- Qui est là.Flammarion (14 pages)
- L'Oeil et la vision (33 pages)
- Qu'en savez-vous. Chantecler extraits (54 pages)
- Moi Pierre Rivière ayant égorgé ma mère, mon frère, ma soeur (187 pages)
- Adresses codifications pour A. N. P. E. (223 pages)

Braille Abrégé (grade 2)

- Histoire et géographie. A. Colin (250 pages)
- Montaigne sa vie, son oeuvre. Lagarde et Michard (182 pages)
- La jeune Parque F. Valéry (30 pages)
- Cours de psychopathologie (93 pages)
- Cours de psychologie (193 pages)
- Anatomie du système nerveux central (78 pages)
- Biologie du comportement P. Médioni (190 pages)
- Psychanalyse F. Dolto Winnicott (200 pages)
- Oeuvres de J. Rostand J.P. Sartre (85 pages)
- La justice, Le Particulier (467 pages)
- A. N. P. E. catalogues (170 pages)
- Démocratie Française V. Giscard d'Estaing (201 pages)

II - The LEXICONS

The lexicons (like dictionaries) can play a double part in every problem of studies and automatic processing of natural language : (PER 78, GOU 77 B)

- The first part consists of assigning a grammatical class, one or several meanings and different uses to every word (or lexical unity) ;

- The second part allows to simplify the grammar (and the rules) by an exhaustive list of exceptions and ambiguities.

1 - Automatic lexicons (GOU 77 A - GOU 79)

In this application, as opposed to dictionary, the term "lexicon" means a coherent set for the whole levels (acoustic, phonological syntactic...) Every element is coded in the same standard way, in terms of its nature, its content and its structure. The basic lexicon (or referent lexicon) includes all the intrinsic or contextual attributes of the word.

A derived lexicon is a rebuilt structure with integral or parts of attributes of the referent lexicon's whole word set.

A sublexicon is a limitation of the entries of the referent lexicon admiting one or several common properties.

In a more general context of applications in social sciences, we present a system of automatic lexicons.

This system is composed of :

- A referent lexicon for spoken French : its dimension and content are always developed (6000 entries at the beginning ; more than 37 000 identified morphosyntactic forms).

For this study, we derived a Braille-speech lexicon from the referent lexicon and the Braille dictionary.

Presently, the referent lexicon, the derived acoustic and Braille-speech lexicons, and the sublexicon of ordinal and cardinal numbers are implemented and used for several applications.

2 - Normalized description of the lexicon's structure

We present a normalized description of the lexicon's structure with the notion of a normal form (first form 1NF, second normal form 2NF, third normal form 3NF) and with the notion of functional dependence introduced by CODD (COD 72 & 78 - DAT 77) and also the notion of a fourth normal form introduced by FAGIN (FAG 76).

We describe the structure as if we had the description of a basic structure of relational datas (GOU 79 - CHR 78).

A - The referent lexicon (GOU 79)

A-1 Fields of relations :

1	RG#	CHARACTER(20)	Word or graphic radical
2	GR#	CHARACTER(2)	Pointer for rule of graphic termination
3	RPH#	CHARACTER(15)	Word or phonetic radical
4	PH#	CHARACTER(2)	Pointer for rule of phonetic termination
5	CP#	CHARACTER(2)	Phonological adaptor
6	G#	CHARACTER(1)	Gender
7	N#	CHARACTER(1)	Number
8	P#	NUMERIC(1)	Person T# = T1# U T2# U T3#
9	T#	CHARACTER(2)	Tense and mood T1#, T2#, T3# are disjoined
10	IDN	CHARACTER(11)	Class identifier
11	CS#	CHARACTER(2)	Syntactical category
12	CSEM#	CHARACTER(2)	Semantic category
13	PR#	CHARACTER(14)	Translation identifier
14..	Dij	CHARACTER(6)	Endings
24..	LiBk	CHARACTER(25)	Wordings

A-2 Relations

(1) Lexicon (RG#, RPH#, PH#, CD#, G#, N#, P#, T#, IDN, CS#, CSEM#, PR#) key : RG.

(2) Rules of termination :

RDGN	(GR#, N#, D10)	key : GR#, N#	Rules of termination for nouns
RDPHN	(PH#, N#, D20)	key : PH#, N#	
RDGA	(GR#, G#, N#, D11)	key : GR#, G#, N#	Rules of termination for adjectives
RDPHA	(PH#, G#, N#, D21)	key : PH#, G#, N#	
RDGVI	(GR#, T1#, N#, P#, D13)	key : GR#, T1#, N#, P#	Rules of graphic for verb-endings
RDGVIF	(GR#, D23)	key : GR#,	
RDGVP	(GR#, T2#, G#, D33)	key : GR#, T2#, G#	
RDPHVI	(PH#, T1#, N#, P#, D43)	key : PH#, T1#, N#, P#	Rules of phonetic for verb-endings
RDPHVIF	(PH#, D53)	key : GR#	
RDPHVP	(PH#, T2#, G#, D63)	key : GR#, T2#, G#	

(3) Syntatic categories

CATSYN (CSEM#, LIB2) key : CSEM#

(4) Other relations

NUMBER (N#, LIB-N) key : N#
PERSON (P#, LIB-P) key : P#
GENDER (G#, LIB-G) key : G#
MDTE (T#, MOOD, TENSE) key: T#

A-3 Operator of derivation

(1) The fields of the relations : FIELDS FROM 1 TO 24 OF A-1

and	25	MG#	CHARACTER(30)	Word : graphic form
	26	MP#	CHARACTER(30)	Word : Phonetic form
	27	REL#	CHARACTER(3)	Index of a relation
	28	REG(Dij)	CHARACTER(10)	Wording of the relation

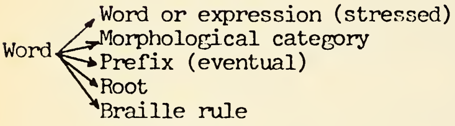
(2) Relation

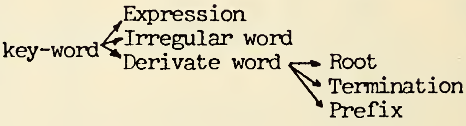
OPDER (MG#, MP#, REL#, RG#, RP#, GR#, PH#, CP#, G#, N#, P#, T#, CS#, CSEM#, PR#) key : MG# ou MP#

(3) Other relations

RELAT1 (REL#, REG (D13, D53)) key : REL#
RELAT2 (REL#, N#, REG (D10, D20)) key : REL#, N#
RELAT3 (REL#, G#, N#, REG (D11,D21)) key : G#, N#
RELAT4 (REL#, T#, G#, N#, P#, REG(D13, D33, D43, D63) key : REL#, G#, N#, P#

B - The Braille dictionary





B-1 Fields of the relations

25	MG#	CHARACTER(30)	Word writen in alphabetical characters
29	MTLE#	CHARACTER(30)	Word representative of a class
30	MIR#	CHARACTER(30)	Irregular word for braille rules
31	SOU#	CHARACTER(20)	Root or the word
32	EQB#	CHARACTER(20)	Braille equivalent to word form
33	DESij	CHARACTER(12)	Endings
34	CATP	CHARACTER(1)	Category of productions
35	LIB1	CHARACTER(30)	Wordings in characters
36	REGp	CHARACTER(3)	Rule for braille transcription

B-2 Relations

(1) Dictionary Braille (MG#, MIR#, MTCLE#, SOU#, EQB#) key : MG#

REGMOT	(MG#, MTCLE#, MIR#, CAT$_1$ LIB1)	key : MG#
REGMTR	(MG#, EQB#, CAT1, REG1)	key : MG#
REGMCL1	(MTCLE#, EQB#, DES10)	key : MTCLE#
REGMCL2	(MTCLE#, EQB#, SOU#, DES11)	key : MTCLE#
REGSOU	(SOU#, PRE11)	key : SOU#
REGPRO	(EQB#, LIB2)	key : EQB #

C - Derivate Lexicon Braille-SPEECH

(1) Fields of the relations : Fields

5 to 9 11 to 13	of A-1
25 to 26	of A-3
25 and 32	of B-1

(2) Lexicon relation : Braille to speech

LEXBP : (MG#, MP#, CP#, G#, N#, P#, T#, CS#, CSEM#, PR#, EQB#) key MG#

(3) Other relation

PRODUCT (EQB#, GRAPHISME, BRAILLE, PHONETIQUE, CATEG.SYNTAX) key : EQB#

This structure (according to Codd) presents several advantages.

- To be manipulated through a languages founded from the relational algebra, allowing to derive other relations (therefore new structures) from initial relations.

- To be manipulated more easily (DAT 77) than the traditional hierarchical relations of this sort of application (GRA 75).

- To allow it to be associated with constraint or integrity rules (assertions) which characterize the coherence of the fields' values.

- To assure a speedy and very accurate setting up (techniques used : hashcoding and virtual space : (Cf SOCRATE (CHR 78)).

Allowing :
- to minimize the memory space used by the lexicon
- A speedy access
- Some addings and extentions

III - Applications

In this article, we only present the subsystems about the braille and phonetic transcription of written texts. The other operational units of the project are given on the general map (fig 1).

The derivate lexicon, classed according to the acoustic form of the words, is given from the basic lexicon, of the A. R. I. A.* project (Fig 2). It contains, particularly, the irregular words of the braille dictionary. Its braille and phonetic equivalents and some morphosyntactic informations are associated with every graphic form. A graphic form can contain spaces or other specific characters (compound words, expressions, etc...). The written text is cut into words, the braille analyser identifies the expressions, the accomplished statements are described in fig. 3.

If the lexicon's consultation gives only one result, i-e : if the graphic form corresponds to only one braille representaiton and only one phonetic, we can generate the corresponding 4-uplet.

This 4-uplet is processed for the first part by the editor of braille and for the second one by the pronunciation modulus (Cf Fig 4).

This model gives particularly liaisons, elisions, nasalizations, stresses, pauses (TEP78) before the operation of the "Vocoder"**.If the lexicon's retrieval does not give a result, the word is unknown or mis-spelt. It is translated into braille by rules and into phonetics by a orthoepic translator (DIV 77). An other consultation of the lexicon from the phonetic argument given by the orthoepic translator, allows to disclose the spelling mistakes and to propose a correction.

As concerns the possible ambiguities for several braille forms of the same word, or several phonetic representations, we call the morphosyntactic filtering modulus. This program keeps a check on word context in accordance with gender, person, number, and phrase syntax.

For a single production we return to the general case (generation of a single 4-uplet) otherwise we can either generate the most frequently used form or call the "end-user" dialog.

Examples : "Les poules du couvent couvent"
"Nous relations nos relations".

1 - Here, ambiguities arise from both graphic and the phonetic forms. There are resolved by syntactic filtering :

(Couvent + N) → k u v ã
(Couvent + V) → k u v ə
(Relation + N) → r lasiõ
(Relation + V) → r latiõ

---

* ARIA : Analysis and Recognition of Acoustic Informations

** "Vocoder" : Synthetic voice production machine.

2 - "Il convient et ils convient"

This sentence illustrate the existence of a double ambiguity (graphic-phonetic and graphic-braille) for both forms of "convient". It needs a morpho-syntactic filtering :

(Convient + Sing) → K õ v i ɛ̃
(Convient + Plur) → K õ v i

3 - "Les fils de la couturière..."

In this example, we admit that the semantic nature of the ambiguity can only be resolved by the "enduser" dialog.
Our braille-to-phonetic transcription has no need for an understanding modulus.
In the state of our research, we think that an automatic text- understanding modulus involves the transcription system.

IV - Conclusion

Simultaneous braille and synthetic voice will permit a blind person to hear a document and obtain it in braille at the same time ; essentially they will have the possibility to choose which part of the document they want.
We hope that in the future the blind will have at their command a sort of braille self-service.

(TRU 73) M. TRUQUET
"Automatic Translation into French grade II Braille"
Première table ronde internationale pour la production automatique de textes Braille, Münster RFA, septembre 1973.

(TRU 74 A) M. TRUQUET
"The Automatic Transcription of French Ink print into Braille"
Research Bulletin, mars 1974.

(TRU 74 B) M. TRUQUET
"Some Remarks on French Braille Translation"
Conférence au Centre de Recherche de Münster, Septembre 1974.

(TRU 74 C) M. TRUQUET
"Automatic Translation into Braille in Basic FORTRAN IV"
Deuxième table ronde internationale pour la production automatique de textes Braille, Copenhague (Danemark). Septembre 1974.

(TRU 76) M. TRUQUET
"Braille grade II Translator Program"
National Computer Conference, New York, Juin 1976
- Rapport publié en juin 1976 dans la revue SIGCAPH de l'ACM n° 20
- Rapport publié dans la revue AFIPS 1976.

(TRU 78) M. TRUQUET
"Transcription par ordinateur en Braille intégral et abrégé"
Congrès AFCET Informatique, novembre 1978.

(COD72) CODD E.F.
Further Normalization in the Data Base Relational Model.
Data Base Systems, courant Computer Science Symposia Series, Vol 6.
Englewoods cliffs, N.J. : Printice-Hall (1972)

(FAG76) FAGIN R.
Multivalued Dependencies and a New Normal Form for Relational Data Bases
IBM Research Report RJ 1812, July 1976.

(DAT77) DATE C.J.
An introduction to Data base systems.
Second Edition, Addison-Wesley Publishing Company, 1977.

(COD78) CODD E.F., ARNOLD R.S., CADIOU J.M., CHANG C.L.
Rendez-vous Version I : An experimental English-language query Formulati
System For casual users of Relational Data Bases.
RJ2144 (29407) 1/26/78 Computer Science
IBM Research Laboratory San José, California 95135.

(PER78) J.P. HATON et G. PERENNOU
Reconnaissance automatique de la parole.
8ème école d'été d'informatique de l'AFCET, NAMUR, Juillet 1978.

(TEP78) G. TEP
Contribution a l'étude phonologique d'un système d'analyse de la parole
continue.
Thèse 3ème cycle UPS Toulouse 1978.

(DIV77) M. DIVAY - M. GUYOMARD
Conception et realisation sur ordinateur d'un programme de transcription
graphemo-phonétique du français.
Thèse 3ème cycle Rennes 1977.

(GOU77A) GOUARDERES G.
Organisation d'un lexique en vue de l'analyse de la parole en continu.
Diplome ingénieur CNAM. Janvier 1977.

(GOU77B) GOUARDERES G.
Saisie et représentation sur ordinateur d'un lexique pour le français
parlé.
Rapport interne CERFIA, Décembre 1977.

(GOU79) GOUARDERES G.
Les lexiques automatiques en reconnaissance de la parole. La version 3
du lexique du Projet ARIA. (A paraître)
10è Journées d'étude sur la parole. Grenoble Mai 1979.

(GRA75) E. GRANDJEAN
Conception et realisation d'un dictionnaire pour un analyseur interactif
de langues naturelles.
Diplome ingénieur CNAM. Grenoble 1975.

(CHR78) CHRISMENT C.Y.
Présentation du S.G.B.D. SOCRATE Gestion de la mémoire virtuelle.
Cours M.I.A.G., UPS, Toulouse III, Mars 1978.

Fonctionnement du système de transcription par ordinateur en braille intégral et abreg

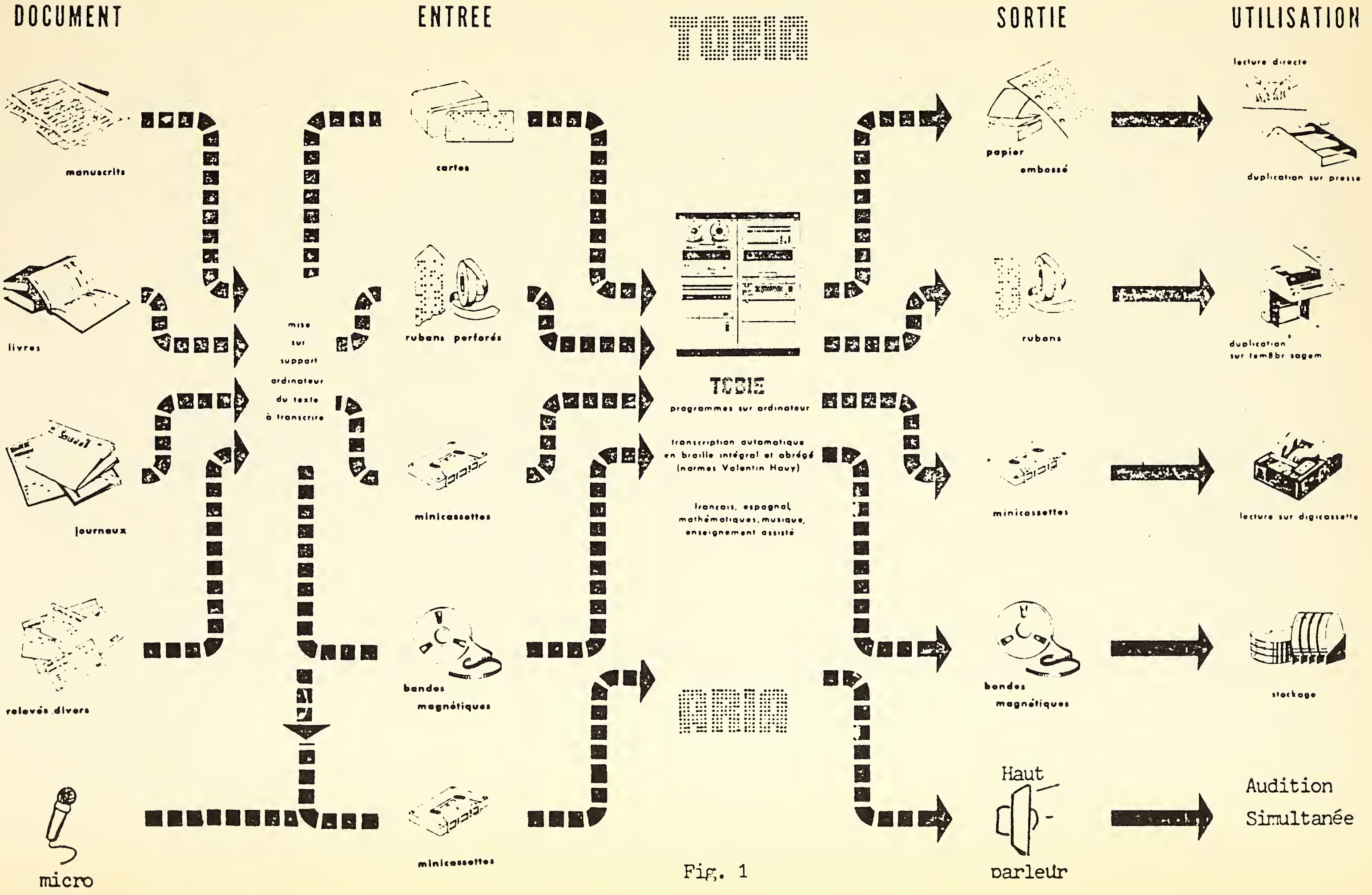


Fig. 1

FORME GRAPHIQUE	FORME BRAILLE	FORME PHONETIQUE	INDICATEURS MORPHOSYNTAXIQUES			
NID	NID	NI	MS	N		
NIE	NIE	NI	3PS	V	PERI	NIER
NIE	NIE	NI	1PS	V	PERI	NIER
NOEUD	N9UD	NQ	MS	N		
NOM	NW	N2	MS	N		
NON	N↑	N2		D		
NOS	NS	NO		Q	4NP1	NOTRE
NOUS	O	NW		R	P1	NOUS
NU	NU	NU	MS	J		
NUE	NUE	NU	FS	J		NU
OUI	8I	<I		D		
OUIE	(OU/E	<I	FS	N		
RAIE	R*E	R7	FS	N		
RANG	R,G	R4	MS	N		
RAT	RAT	RA	MS	N		
REIN	RE)	R5	MS	N		
RI	RI	RI		V	PAPA	RIRE
RIZ	RIZ	RI	MS	N		
ROND	R↑D	R2	MS	K		
ROUE	R8E	RW	FS	N		
ROUX	R8X	RW		K	M	
RUE	RUE	RU	FS	N		
YEUX	Y-X	YQ	MP	N		OEIL
BAIL	B*L	BAY	MS	N		
BAL	BAL	BAL	MS	N		
BAR	B"	BAR	MS	N		
BEL	BEL	B&L+	MP	J		
BOL	BOL	BOL	MS	N		
BORD	B+D	BOR	MS	N		
DEUIL	D-IL	DQY		D		
HIER	H<	YER		D		
LARD	L"D	LAR	MS	N		
LEUR	L#	LQR	S	T		
LEURS	L#S	LQR		Q	6NP1	LEUR
LOURD	LD	LWR	MS	J		
MAL	MAL	MAL	MS	N		
MAL	MAL	MAL		D		
MER	M<	M&R	FS	N		
MOEURS	M9URS	MQR	FP	N		
MORT	M+T	MOR	MS	K		
NERF	N<F	N&R	MS	N		
NORD	N+D	NOR	MS	N		
NUL	NUL	NUL	MS	J		
RAIL	R*L	RAY	MS	N		
RIRE	RIRE	RIRE		V	INFI	
DERNIER	DN	D&RNY&	MS	J		
GARDIEN	G"D.	GARDY5	MS	N		
MALGRE	M7	MALGR6		P		
MARBRE	M";E	MARBRE	MS	N		
MORDRE	M+4E	MORDRE		V	INFI	
MORDRA	M+4A	MORDRA	3PS	V	FUTI	MORDRE
MORDRAI	M+4*	MORDR7	1PS	V	FUTI	MORDRE

Fig. 2

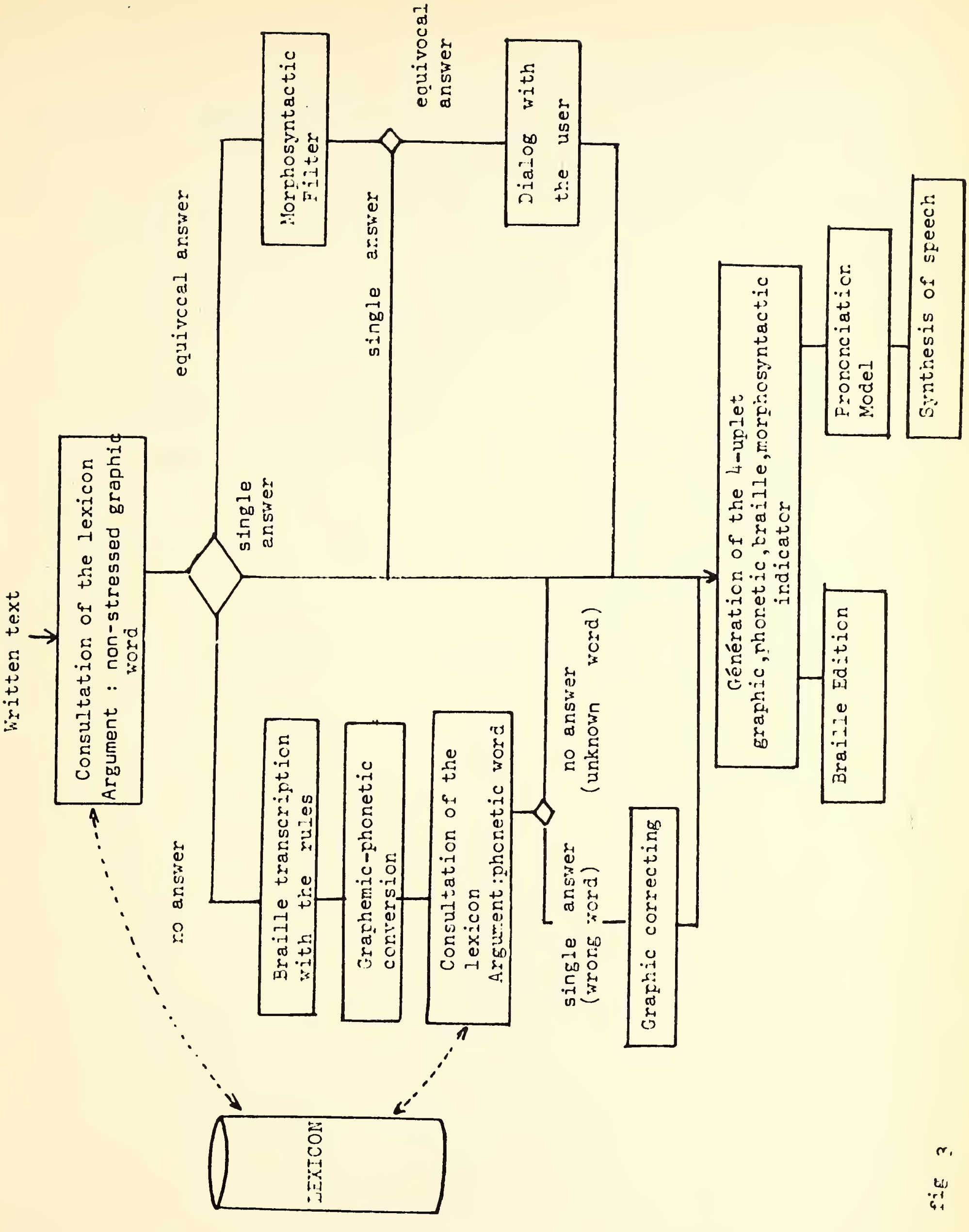

Written text
Consultation of the lexicon
Argument : non-stressed graphic word
LEXICON
no answer
single answer
equivocal answer
Braille transcription with the rules
Grapemic-phonetic conversion
Consultation of the lexicon
Argument:phonetic word
single answer (wrong word)
no answer (unknown word)
Graphic correcting
Morphosyntactic Filter
single answer
equivocal answer
Dialog with the user
Génération of the 4-uplet graphic,phonetic,braille,morphosyntactic indicator
Braille Edition
Pronunciation Model
Synthesis of speech

Fig 3

```
** EXECUTION **

ENTREZ VOTRE PHRASE ECRITE ?
?AU MOYEN AGE,LES HOMMES PREFERENT LES OISEAUX. @

## SUITE DES MOTS PHONETIQUES-SORTIE DU MODELE GRAPHEMES.PHONEMES :

O MWA*59 XAJE L6Z XQM PR6F7RET L6Z XWAZO

## PHRASE PHONETIQUE-SORTIE DU MODULE DE LIAISON.DENASALISATION :

O MWA*7NAJE L6ZQM PR6F7RE L6ZWAZO

## PHRASE PHONETIQUE-SORTIE DU MODULE DES PAUSES :

OMWA*7NAJE L6ZQMPR6F7REL6ZWAZO

## PHRASE PHONETIQUE FINALE-SORTIE DU MODULE DES ELISIONS :
(omwajɛnaʒə lezɔmprefɛrlezwazo) code A.P.I.
OMWA*7NAJE L6ZQMPR6F7RL6ZWAZO  ------------->  to the VOCODER

## SORTIE DU MODULE BRAILLE :

K #MOY?-#ICE & HMS !/F/R2 & ]SEKX  ---------->  to the SAGEM DEVICE

** FIN EXECUTION **
```

Fig. 4

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

NEW PRODUCTION SYSTEMS IN BRAILLE PRINTING

by

M. Harres

Forschungszentrum für Rehabilitation
und Prävention
Bonhoefferstr. 1
D-6900 Heidelberg

Presentation to the Braille Conference from May 30th to June 1st 1979 in London

New Production Systems in Braille Printing

At the moment the production of braille books in West Germany is undergoing a process of rationalisation in the large braille printing houses. This rationalisation involves the setting up of new production systems. Various German companies with their equipment are playing a part in the creation of this presentation is to describe the individual units produced by the manufacturers and their interplay within the production chains.

Before the planning of the new production organisation, the printing of braille books was carefully analysed and examined in its principal stages. Allow me to describe this analysis as background information.

1.0 Analysis of Braille Printing

Braille printing can be divided into two categories according to the costliness of production:

a) Literature in which special notation systems are used (e. g. mathematical, chemical, musical notation) as they occur in textbooks.

b) Literature consisting only of continuous texts (e. g. literary works, periodicals).

A further factor in braille printing is the widely differing size of editions, which can range from several thousand to editions of fewer than ten copies for textbooks. Independent of the type of literature and the size of the edition, the production of braille material falls into four principal stages:

a) Preparation of Text
b) Composition of Text
c) Correction of Text
d) Reproduction of Text

1.1 Preparation of Text

The blind person must be presented with the same information content as is available to the sighted person in a normal printed book. For this reason the black printed book must be revised before it is transcribed into braille. This revision is extremely costly, expecially in the case of textbooks (e. g. describing of pictures, graphs, etc.). It should be carried out by a specialist. With newspaper articles and literary works the preparation of the text can be performed by an experienced braille transcriber.

1.2 Composition of Text

This is the stage where the prepared texts are written. Under the old production system of braille printing the texts were typed in grade 2 braille on a braille keyboard by a braille transcriber and impressed into metal on an embossing machine. With single editions the composition is done by means of the page brailler on paper. Today the texts can be set up either in black print or in braille on a recording medium (e. g. compact cassettes or paper tape). Texts composed in black print must be transciribe into grade 2 braille by the computer before the next stage.

1.3 Correction of Text

Errors arising from the transcription of the texts into braille shorthand by the transcriber or from the computer program are in this stage corrected. Previously these errors have had to be painstakingly removed on the metal matrix. If however the braille is stored on a recording medium it is possible to correct the errors that arise directly and without difficulty on the recording medium.

1.4 Reproduction of Text

In the text reproduction stage the composed and corrected texts are transformed in such a way as to enable the blind person to read them. The corrected recording medium can be used to run various pieces of equipment, and the following possibilities for production are offered:

1.4.1 Metal matrices are produced on an embossing machine and used to run off a large number of copies.

1.4.2 For small runs the page brailler can be used to produce copies from the recording medium.

1.4.3 The recording medium can be read direct by means of reading instruments with tactile lines (e. g. Braillex, Braillocord).

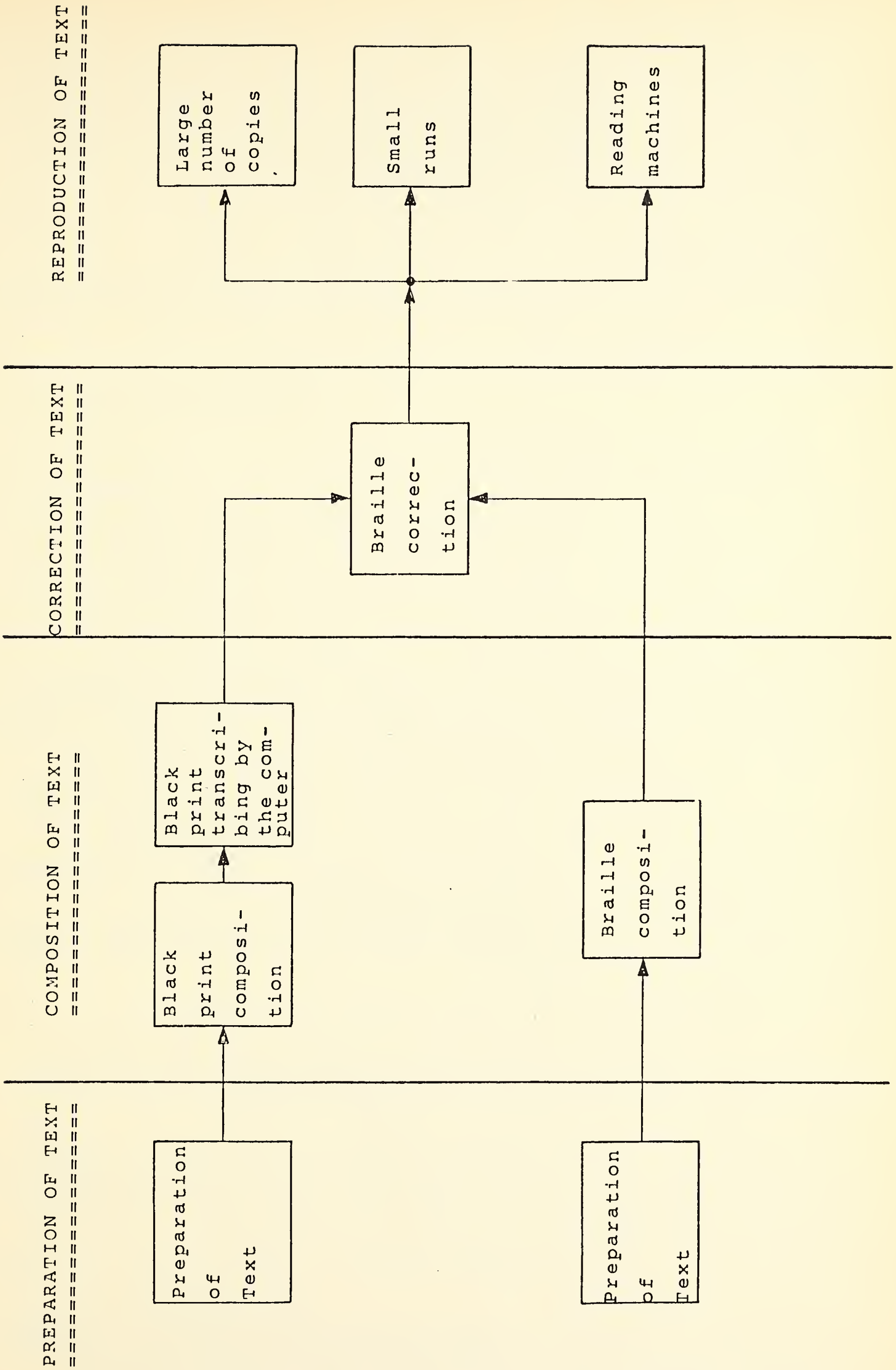
In the last case appropriate converters are needed to transform the recorded data since the reading instruments use different physical recordings than that used in braille printing.

2.0 New Production Systems

The stages of braille printing listed above can with the new production systems be executed rationally for the widest variety of braille literature. The two extremes of braille printing, namely the textbook with special notation systems on the one hand and printing without special notation systems on the other, led logically to the creation of two production systems for braille printing. As however there is neither a textbook without long passages of continuous prose nor a literary work without special characters, these two production systems are fully compatible. Compatiblity of the two production systems is ensured by a standardised recording medium. Figure 1 shows the production organisation for braille printing based on the University of Münster's programme.

The two production systems are identical apart from the composition stage. Whereas all continuous texts without special notation can be composed in black print and transcribed automatically into braille, all texts with special notation shall be composed in braille. The organisation illustrated above allows however for a book to be composed partly in black print and partly in braille. In this case text correction is extended to include merging and editing the two recording media.

3.0 Equipment for the Principal Production Stages


PREPARATION OF TEXT
COMPOSITION OF TEXT
CORRECTION OF TEXT
REPRODUCTION OF TEXT
Preparation of Text
Black print composition
Black print transcribing by the computer
Preparation of Text
Braille composition
Braille correction
Large number of copies
Small runs
Reading machines

3.1 Preparation of Text

This stage, as mentioned above, requires the experience of the specialist teacher or braille transcriber. This means that this stage will never be automated. It will continue to be an important cost factor in braille book production.

3.2 Composition of Text

As already mentioned, composition can be done in one of two ways:

1. black print composition
2. braille composition

3.1.1 Black Print Composition

Black print can composed efficiently by means of conventional text editing machines. It is important here that the machines deliver recording media that can be read by the computer. The translation of the black print into braille shorthand will then be performed in the computer.

The most efficient method of black print composition applied in West Germany up to now has involved the use of publishers' recording tapes. Today this method has already become standard in braille book production in many countries of the world. As well as the advantage that the black print is already composed on a data carrier, the use of publishers' tapes for braille printing in West Germany has revealed the following:

- There are great organisational difficulties involved in getting hold of compositors' tapes.
- The stored texts are as a rule not free of errors.
- The type of recording medium (paper tape, half-inch magnetic tape, compact cassettes, quarter-inch magnetic tapes) varies with the different composition systems.
- The recording codes are not uniform in the composion systems used.
- With some publishers special codes are also often arranged for each book to be printed.
- For the editing of compositors' tapes a versatile text editing system is indispensable.
- Extra material created at the time of prepartation of the text has to be incorporated into the main text on the terminal.

In West Germany however the trend is away from the use of publishers' tapes and towards an even more rationalised method. AEG-Telefunken are at present developing a character recognition machine which uses a television camera to detect black print which it recognises electronically. With this instrument the abovementioned problems arising from the use of publishers' tapes can be avoided. The character recognition machine can read the most important types of print and transfer them to a recording medium. Within the production system the recording medium is read by the computer, which translates the black print into braille shorthand.

3.1.2 Braille Composition

Braille Composition is this system is carried out with the aid of a data handling unit for braille developed by the STIFTUNG REHABILITATION, Heidelberg. This unit, which comes in versions for sighted as well as for blind operators, delivers a recording medium (ECMA 34 cassettes) on which the braille is stored. In addition to a light-action ergonomically designed keyboard, the unit offers correction facility within a line of braille (maximum line length 40 characters). As an alternative to the ECMA 34 cassettes, 8-position paper tape can also be used as the recording medium.

3.3 Correction of Text

The Federal Ministry of Research and Technology made funds available for a high-performance correction unit for braille, which has been developed at the STIFTUNG REHABILITATION, Heidelberg.

Here are some important technical data:

a) Microprocessor-controlled central unit
b) Storage of one page of braille with maximum 40 lines of 40 characters per line.
c) Within the page of braille every conceivable correction facility is offered.
d) The unit is equipped with a visual and a tactile display.

3.4 Reproduction of Text

With the standardised recording medium used throughout the system, machines for high-volume braille page production as well as machines built for small runs can be operated.

Large Editions

For large runs of braille books it is first necessary to produce a metal matrix. For this purpose a high-speed embossing machine of the German Blind Studies Institute, Marburg, is used, which embosses ten braille characters in one second (Puma 4). If large runs (over one thousand copies) are required, these are produced with the embossed matal matrix on the Marburg Blind Studies Institute's rotary press. This applies particularly in the case of periodicals. Alternatively the metal plate can be fitted to a flat bed press. This process is used particularly when high-quality printing is required, as is the case for textbooks.

Small Runs

Specialist literature is often only produced in small editions. In this case it is better to avoid the long and costly way via the embossing machine. In the production methods described here books are embossed in small runs on the (Darmstadt) Thiel-Company braille printer. This machine can also be controlled by ECMA 34 compact cassettes. It has a possible line length of 40 characters and embosses 3 lines per second.

Text reproduction is also possible outside the realm of printing. Braille recorders are available to the blind for this purpose. It is necessary thereby to convert the tape used in braille printing. In West Germany there are at the present time four instruments available:

a) The Braillocord, made by AID Electronic, Berlin
b) The Braillex, made by Papenmeier, Schwerte

With all two instruments it is possible to read braille books line by line, the braille characters being brought up a line at a time on an electromechanically controlled tactile display. None of the two instruments was designed purely for the reading of braille books. All 2 instruments have a multitude of additional functions, the discussion of which would be outside the scope of this presentation. We draw your attention to the detailed description of the individual instruments contained in the information files, and also to the exhibition, in which all models are on display. We would however like to say very briefly:

- The Braillocord is a versatile reading and writing instrument for Braille.
- The Braillex makes possible the independent creation and reproduction of ordered files of data, e. g. dictionaries; it could be described as an electronic private secretary.

Summary

On the last diagram you can once again follow the production chain and at the same time recognize the individual units in their positions within the chain. The indivudual specialists are available at all times in the exhibition hall to answer any queries you may have.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

BRAILLE BY TIME SHARING SYSTEM

by

Kenneth R. Ingham

ASI Teleprocessing Inc.
123 Water Street
Watertown, Massachusetts
USA

High Technology and the Blind

Perhaps the most significant development since World War II has been the creation of a new world culture based on the computer and its use in information processing. In western countries, the manufacturing of conventional products has taken a back seat to the development of high technology and service industries. This is especially true where the electronic and computer industries are concerned. This explosion of technology has generated a new population of "technocrats" or "computernicks". In addition, people working in such traditional vocations as law, medicine, and business, as well as those in such support categories as secretarial, must now learn about this new technology and must apply it in order to remain competitive.

Individuals working in these areas have, through the use of computer equipment, been able to manipulate much larger amounts of information than previously possible. In addition, new storage and transmission techniques have allowed for the implementation of systems where information is centrally located and at the same time, accessed and modified remotely. Thus companies may have extremely up to date knowledge concerning the flow of materials and goods, and about the people they employ, even though these companies may have branches widely distributed around the globe.

The difficulties of such a boom in technology arise out of the fact that almost a bewildering variety of devices and techniques are available and must be integrated. While some standards exist, most computer based systems differ both with regard to their characteristics, the operating systems they run, and in regard to the ways in which terminals are connected to them. Moreover, while many are working to establish universal standards, the proliferation of companies and software is working to increase specialization. Even the production of microprocessor systems has failed to reverse this trend since their low cost and variety has led to the implementation of many small and specialized systems.

The blind, seeking employment, must now be trained on these new devices and techniques. Means must be found which would allow blind persons to read the output from computers which appears on a wide variety of teleprinters, video keyboards, and other devices appearing in many job situations. While training is certainly possible, the cost of equipment to interface to existing systems is high. Not only must the visual displays be converted to a tactile or auditory one, but such non-visual displays must be software integrated to the employee's system.

Two approaches can be taken to solve these problems. The first is to design special purpose systems which solve specific problems. The second is to create a vehicle by which many problems can be solved while at the same time, preserving the ability to solve the specialized ones.

The ASI Quick Braille system is a special system which can be produced in quantities and which can help in solving the specific problem of producing quick turn around braille. On the other hand, a time sharing system can create a base of technical resources and personnel such that a systematic approach can be carried out. In addition, by time sharing the use of these resources among a large number of blind persons, the cost can be made acceptably low. This is true even when the cost of individually owned and operated microprocessors is considered.

In what follows, ASI and its ARTS Time Sharing System will be described. The Time Sharing System will be shown to be a vehicle for the solution of these problems. The Braille Services of the ARTS system will be covered, along with a brief summary of Quick Braille.

ASI Teleprocessing Inc.

ASI Teleprocessing Inc. is a manufacturer of teleprocessing systems, audio response devices, and is the creator of the Audio Response Time Sharing System (ARTS) for the blind and physically handicapped. The company began the development of the ARTS system in 1971 and this activity has continued, culminating in recent installations at Volunteer Services for the Blind in Philadelphia, Pennsylvania and Lighthouse for the Blind in New Orleans, Louisiana.

Since 1973, ASI has manufactured and installed a series of teleprocessing systems in major banks and retail organizations throughout the United States. These systems utilize audio response and packet switching techniques in the banking industry and provide the telecommunication support necessary for many common banking and retail terminals. Thus, ASI is a company providing computer based products and services for applications where data must be gathered from and information supplied to, geographically separated locations.

Combining its experience in teleprocessing and telecommunications with that gained in the development of time sharing systems, ASI has defined two standard configurations aimed at serving the needs of the blind. These are:

- Audio Response Time Sharing System (ARTS) which combines ASI's general purpose time sharing system supporting word processing, print formatting and braille translation in full conformity with Library of Congress standards, with ASI's powerful voice response system, Solid Voice.

- Quick Braille, which provides the word processing functions of the time sharing system and braille translation in an economical single user system.

ASI TSS

The ASI Time Sharing System which, without the audio response portion is called ASI TSS, is a computer system programmed to allow many persons to simultaneously use its functions. Built around a Digital Equipment Corporation PDP 11 computer, ASI TSS makes available a wide range of services which the blind may use.

These services known as ARTS services, are specially designed to be operated by beginning users who need know nothing about computers and their programming. On the other hand, a mode of operation is available where blind programmers may utilize tools which permit them to program either new services or programs required by the companies they work for.

The standard configuration offered by ASI is:

> Digital Equipment Corporation PDP 11/34 minicomputer with 96K Bytes of memory, which runs the ASI Time Sharing System programs.
>
> System console with printer which provides operator and spooling control.
>
> An 8 line multiplexor which provides the computer ports for user consoles and spooled devices such as the systems printer and braille embosser.
>
> A 10 megabyte disk system for storage of programs and text file.
>
> A dual floppy disk system which is a transfer device for files and diagnostics.
>
> One or more Digital Equipment Corporation VT100 video keyboards - user consoles.
>
> A medium speed printer for the production of office quality copy.

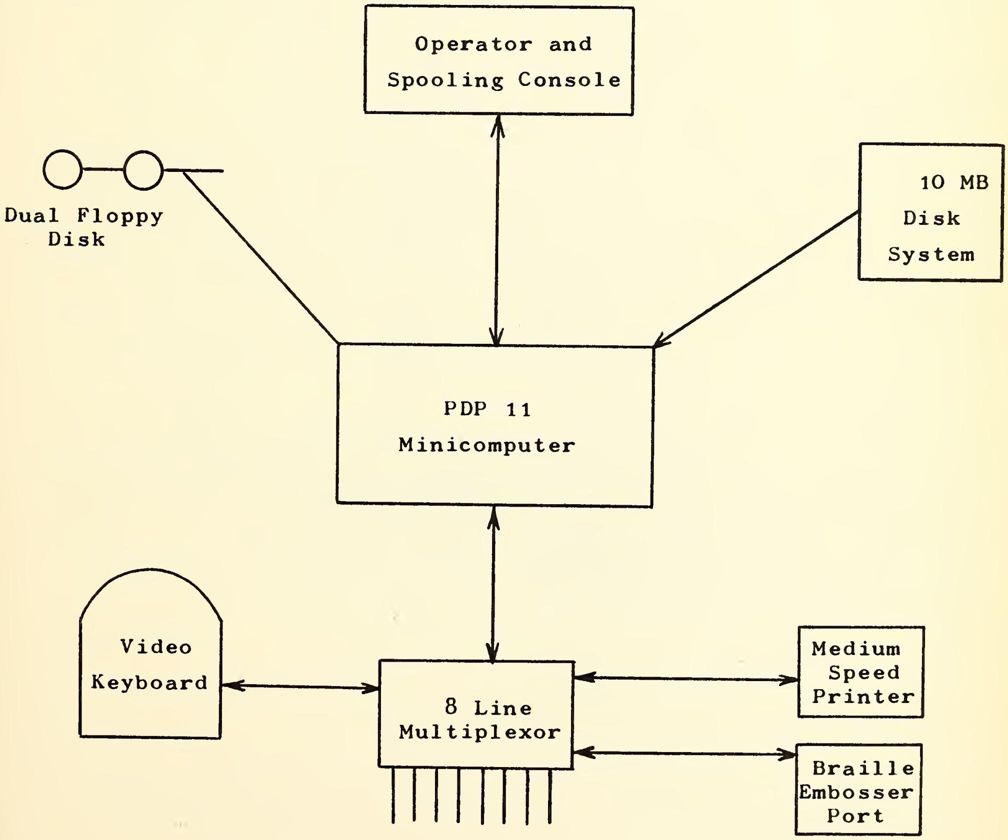
This system with its minimum 8 line mux will support, if two lines are reserved for printer and embosser, (six) 6 VT100 consoles local or remote. Additional consoles (users) can be supported if an additional 8 line multiplexor is added, along with additional and/or high speed swapping drum.

A large type capability is provided through the use of a Dataroyal printing system. This unit has a built-in type magnifier such that when text is transmitted to it from the computer, large type is produced. This unit can be particularly valuable to partially sighted individuals.

The Time Sharing System with audio response is typically used by sighted persons. The installations at VSB and Lighthouse for the Blind are configured with video keyboards as the standard user console. Sighted persons use these CRT's to input text which is intended for Braille production. In this configuration, ASI TSS is a valuable resource tool for an agency or school system. When the audio response function is added, the system features can be made available to blind individuals.

FIGURE I

STANDARD ASI TIME SHARING SYSTEM



Audio Response

The full ARTS configuration consists of the time sharing system just described, coupled to an audio or voice response system. ASI manufactures a system called SolidVoice which permits many persons to connect to the time sharing computer at one time. The blind person is supplied with a standard keyboard equipped with an amplifier speaker which he or she attaches to the ordinary telephone with an acoustic coupler. One dials the computer and after it answers, simply types commands to the computer. ARTS answers with spoken responses, informing the blind person about what he or she has typed. The audio response from SolidVoice is in the form of full words or sentences, based upon the storage of human originated speech sounds. This high quality form of audio response is available with vocabularies of 100 to 100,000 words. Figure II shows a SolidVoice unit as part of an ARTS system.

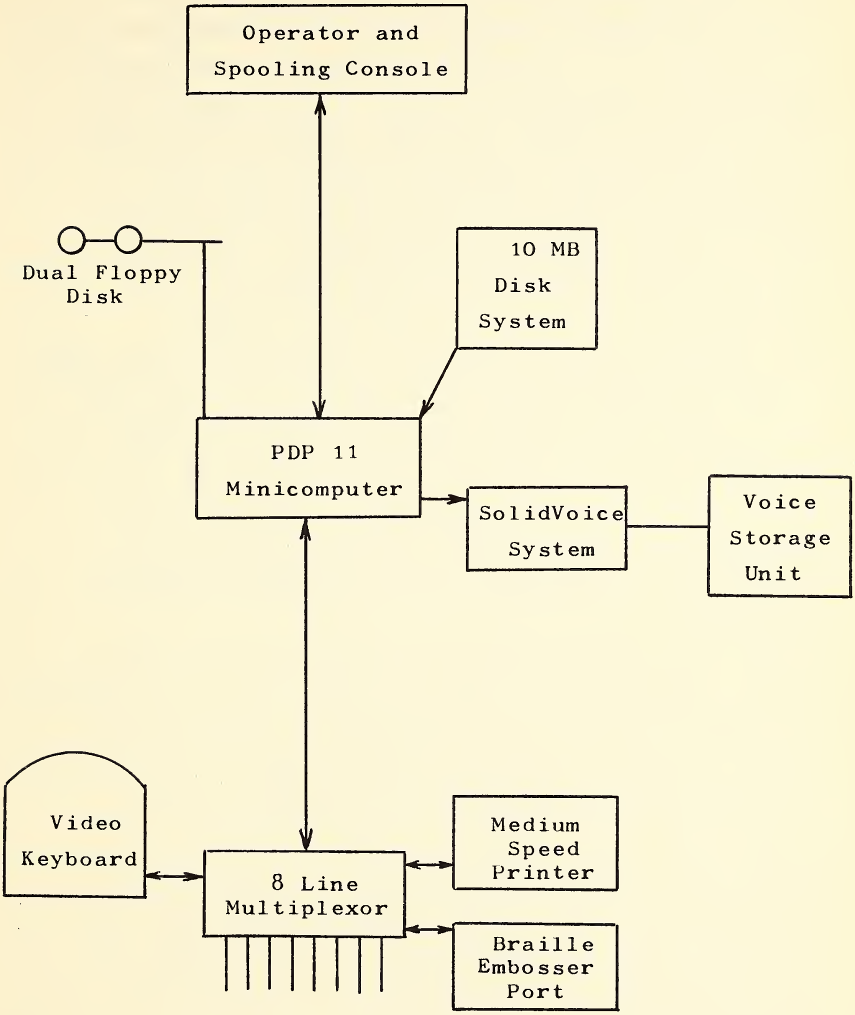
ARTS Services

The ARTS services available on the time sharing system consist of a growing number of highly useful programs. These are centered around a basic reading and writing service called Edit. With Edit, a person may type in correspondence, articles, programs, or indeed any other type of information. He or she may use Edit to proofread and/or to modify any part of the text. When the text to be typed is complete, it may be filed on the system for later recall.

A file on the system may be thought of in analogy to files in a file cabinet. Each contains information on a series of pages which are enclosed in a file folder on which is a label, or file name. With such a mechanism, a blind user may request the Edit service and at the same time, request by name a particular file which he or she wishes to edit.

FIGURE II

STANDARD ASI TSS SYSTEM



Other services which use files and either move or modify them include:

BRL - The Braille service is ASI's unique braille translation program which accepts typed material such as literary text, computer information, and limited mathematics, and converts these to the appropriate braille code. Foreign languages, special dictionary formats, and tabular material can automatically be translated and output by BRL. The program produces Grade I, Grade II, and limited Nemeth Code braille, along with the required textbook formats is full conformity with Library of Congress standards.

JUST - The Just service is the basic print formatting program. Commands inserted as part of the text are executed by Just so that runners, automatic page numbering, line centering, and indenting can be accomplished. In addition to these and many other functions available in Just, overstriking and the reservation of space for various types of diagrams are allowed.

COLUMN - The Column service accepts text or tabular material typed in the form of a long list and columnates it so that the material is organized into multiple columns on a page. Facility is provided to specify the width, height, and number of columns on a page.

PAGER - The Pager service is used to extract from a larger file, specific pages which are then printed or embossed. Pager may also be used to select specific pages of any non-standards height. This service is especially useful in conjunction with zinc plate embossing equipment.

COPY - The Copy service is used to copy one file into another or from one storage medium to another.

COMBINE - The Combine service is used to combine two files. The user specifies the two files in the order that they are to be combined and directs the results to a new output file.

EMBOSS - The Emboss service is used to send files to one or more braille embossers. The files are assumed to be already translated into the braille code. Embossers supported include Triformations LED 120, PED 30, and Sagem. Paperless braille devices will soon be available.

TECHTRAN - The Techtran service is used to transmit files between ASI TSS and Techtran cassette units. Two modes of operation are provided. The first reads or writes cassettes in standard ASCII code, while the second reads or writes cassettes on which braille code intendet for LED 120 or PED 30 operation is stored.

PRINT - The Print service is used to send files to the systems printer. The files are assumed to be already formatted with the aid of Just, since Print simply reproduces the contents of the file on the specified device.

The ARTS system provides these services via a voice response feature. With the ARTS keyboards, blind persons in a variety of occupational areas can interact with ASI TSS from their home, school or office. Material typed from any of these locations may be output in braille or in printed form. In addition, printing terminals can be located at the same site as that of the blind worker. Thereby the paper copy can be produced locally. However, the more that such equipment is used by a single individual, the higher the per capita cost will be. Naturally, a remote printer coupled with voice response is still significantly less expensive than this same equipment augmented with a local mini or microcomputer. With such terminals, a blind person can be trained to cope with a wide variety of employment situations. The staff of the ARTS center can develop carreer ladders which not only permit entry level success, but provide a promotion route for the handicapped person. It is vital in this connection that both individuated training and the development of work categories be evolved.

The ARTS system can also accommodate the growing number of telecommunications protocols. The ARTS computer can be connected to most existing computer systems and terminals. Thus, a blind person whose job involves the use of an existing computer system can tie into it via the ARTS system and can transfer files of information back and forth. Again, these techniques implemented for an individual may be quite expensiv. However, if they are carried out in the context of a resource center which develops a growing inventory of such techniques, their shared cost can be kept low.

Braille Embossing Systems

The Time Sharing System can be used to drive multiple braille embossers. Currently, the Sagem and Triformations machines are supported. At Volunteer Services for the Blind in Philadelphia for example, two LED 120 machines and one PED 30 are accomodated under the systems spooler. In this connection, we have found the Pager service to be essential. Operator errors and embosser failures can be overcome by using Pager to re-output specific pages.

In the near future, ASI will provide support for one or more of the available paperless braille devices. The Elinfa and TSI Versabraille devices are of chief interest to users in the United States. We hope that organizations owning ASI TSS systems will develop cassette libraries of recorded braille material. Here, the services may be used to maintain and produce library catalogues and subscriber lists, as well as to produce the braille material itself.

Perhaps the most important task before the computer operators and terminal designers is that of achieving standards. This is particularly true where braille embossers are concerned. Currently, none of the available suppliers uses a commonly accepted computer code to braille conversion. Thus, ASI TSS, in supporting the Triformations devices, the Sagem, and the various paperless machines, has been configured with software drivers that provide code conversion. The group working under Marjorie Hooper of the Florida

State University Braille Codes Committee, is working on defining a standard code set. Here, a ROM, or Read Only Memory device, is being defined which makes an 8 bit code into a braille representation. The 8 bit code set can be either ASCII, EBCDIC, or binary. A two cell output is provided that is compatible with the standard alphanumeric braille representations. We hope that such a device will be manufactured in quantity and then used by all of the braille embosser manufacturers so that machines can be plugged into any standard system.

Braille Production

The ARTS system without its audio response portion, that is to say ASI TSS, can be operated from CRT's or other visually operated terminals. VSB in Philadelphia and Lighthouse for the Blind in New Orleans, have impelemented systems where a sighted individual types in information using these CRT video keyboards. They operate the very same services used by blind persons and like them, they utilize the braille service to translate the text into braille. Thus, groups of volunteers, instead of painful transcription using such braillers as the Perkins, can utilize the standard typewriter keyboard to produce Library of Congress Standard Braille. Here, the emphasis is on the typing skill rather than on the braille translation. The typist need not know braille as the computer does the translation.

The source material, which is translated into braille, can be stored on the computer or on floppy disk media. Thereby, it is available for future braille production. Futhermore, the files so stored can be updated or re-edited as needed.

In the coming months, we expect these agencies to begin service to remote sites such as schools or other agencies. Such a capability will permit blind school children to receive quick turnaround braille of homework assignments, lessons or exams.

Another enhancement under discussion is the addition to the ASI TSS system of magnetic tape support wherein publishers tapes are read. Books and other material can thereby be made available in braille form almost as quickly as it is in printed form. These and other enhancements may be made owing to the programmability of the Time Sharing system.

Quick Braille

The second standard product produced by ASI for schools and agencies serving the blind is Quick Braille. Here, a large demand for a special purpose braille translation system was recognized. A low-cost microprocessor based system was the result.

Quick Braille is configured around a DEC LSI 11 microprocessor which is mounted within a DEC VT100 Video Keyboard. A dual floppy disk system is included in the configuration which is used to hold the systems programs, including the braille translator, and the users files. The system comes with an attractive desk-style cabinet with the video keyboard console on top. A braille embosser or paperless braille device is attached as the output terminal.

Many of the standard Time Sharing services are available on Quick Braille. However, they are used by a single individual in what is called a single-user stand alone mode. A sighted person keys in information on the video keyboard and when ready, commands output to the embosser or cassettee device. The available services on Quick Braille include:

Edit
Braille
Copy
Pager
Embosser
Techtran

It is hoped that many schools and small agencies will find this system a solution for Quick Braille production.

Conclusions

We have seen that, in most job categories and industries, computers and related systems are being used with increasing frequency. The blind, in order to compete for jobs worthy of their abilities, must learn to use these systems. Companies and agencies supporting the handicapped must develop and provide aids whereby the handicapped can interface to the data processing world.

ASI Teleprocessing, in view of the cost and complexities of computer systems, has implemented the ARTS system whereby the per user cost of data processing is significantly less than that resulting from a dedicated system. Furthermore, the ARTS system center should have associated with it both rehabilitation and technical personnel who can best marry the social service training required with the technology.

Special purpose systems where appropriate, should, of course, be implemented. Quick Braille, a single user braille translation system is one such device. However, people, like employment catagories, are ever changing and the support system for the handicapped should be so designed that they too can adapt to changing requirements.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

COMPUTER TRANSLATION OF GRADE 2 BRAILLE IN THE AMERICAN PRINTING HOUSE FOR THE BLIND

by

R. L. Haynes and J. R. Siems

American Printing House for the Blind
1839 Frankfort Avenue
P. O. 6085
Louisville, Kentucky
USA

American Printing House for the Blind
Incorporated
CARSON Y. NOLAN, Ph. D.
Vice President and General Manager

1839 FRANKFORT AVENUE
MAILING ADDRESS: P. O. 6085
LOUISVILLE, KY. 40206

Tel. 502-895-2405
TWX 810-535-3449

Computer Translation of Grade 2 Braille

American Printing House for the Blind
R. L. Haynes and J. R. Siems
March 1979

Summary

For the past fifteen years, braille has been translated by computer on a regular basis at the American Printing House for the Blind. The automated translation procedure begins with data entry operators, who do not know braille, entering the contents of novels, magazines, and educational texts on key disk machines. After keyverifying, the input is translated by computer. The speed of converting inkprint data to braille data is approximately 10,000 words per minute. Output from the computer is read by card-to-plate equipment automatically producing braille plates, from which an unlimited number of copies can be embossed. Experience has shown that the application of computer technology to braille translation is effective toward meeting the need in publications for the blind.

Introduction

As the largest publisher of literature for the blind, the American Printing House for the Blind (APH) has been embossing reading material for over 100 years. In the 1950's, APH, in cooperation with International Business Machines (IBM), began exploring the potential of automation for braille production. Consideration was given to a number of approaches and various types of equipment. Computer translation with keypunch input and card output for driving an automatic embossing machine seemed the most promising. A computer program to translate inkprint into Grade 2 braille was written by IBM. An automatic embossing machine built by APH was used for testing the results of translation.

In May, 1964, IBM made available to APH without rental cost a 709 computer to be used for braille translation and braille research. This made possible automated production on a regular basis. An added resource for development and research has been a U.S. Office of Education grant, awarded to APH for studies including exploration of computer translation of music, mathematics, and scientific literature into braille. Also, APH has been supported by the National Library Service for the Blind and Physically Handicapped (NLS) for development of computer techniques to assist in the translation of music notation into braille.

Since 1964, 2,155 titles have been translated by computer. These titles consisted of 848,921 metal plates which were used to print more than 60,000,000 pages of braille.

The majority of books translated are for general reading such as novels, biographies, and popular non-fiction works. Roots, Advise and Consent, For Whom the Bell Tolls, The Life of Tom Dooley, Portnoy's Complaint and a soup cookbook are examples of books of this type.

Magazine production involves a more demanding time schedule and more difficult format. The Reader's Digest, Fortune and Boys' Life are among the magazines translated into braille by computer on a monthly basis.

APH, as the government recognized agency for producing materials for the education of blind students in grades 1-12, has a responsibility for printing many types of textbooks. Automated translation has been of value in fulfilling this responsibility. Translation of textbooks requires capacity for handling page layouts of considerable complexity.

A number of titles have represented sizable works, particularly reference books. The largest title translated was Documents of American History having 25 braille volumes. Input for this book consisted of 100,000 records, and the result of translation was 85,000 braille output cards.

A sizable project was the translation of the King James Version of The Bible with the automatic insertion by the computer of inclusive page reference lines. Also, many tax forms for the Internal Revenue Service have been translated by computer into braille.

A by-product of the translation of 2,155 titles is the existence of a large body of text in machine readable form.

Text Preparation

After a title has been selected for publication and permission has been granted by the inkprint publisher, editorial decisions are made by an editor familiar with the braille code. The edited text is keyed character for character by an operator who does not know braille. The operator is responsible for the insertion of special symbols indicating changes in format. The input is entered on 3742 key disk machines (which have replaced 026 keypunches) as a continuous stream of data utilizing 60 characters of an 80 character record. One operator can produce the equivalent of 50-60 braille pages per day. The final step in input preparation for braille translation, is keyverifying of all material entered.

One major advantage of a computer system for braille production is that the operators preparing input do not have to know braille. A high school graduate with two years of typing can be trained to produce braille input in approximately one week. This compares with many months of training that is necessary for a braillist to become competent.

Consideration has been given to alternate input methods. One method that has been explored is the conversion of compositor's tape to translation input. The East Indiaman published by Little, Brown and Company was produced in braille by using this approach. Another method investigated was Optical Character Recognition (OCR). Input for The Unsubstantial Castle was the result of scanning the inkprint copy. Neither conversion of publisher's tapes nor OCR were proven to be a more effective method of braille input preparation than keyentering.

Computer Processing

When input is ready for processing, control cards are inserted. These cards specify the type of material to be translated, the size of the braille page and the page heading. At this time, the operator may insert control symbols for complex format such as footnotes or space to be reserved for drawings.

The first processing run is a pre-translation reading of the text by the computer. This run deletes from the text, words which have been translated in previous books. Potential translation problems are also indentified. The result is a condensed text containing unusual words and problem situations, of which a test translation is made.

The next step is a translation of the full text for format checking. The result of this run is a listing that simulates the appearance of a braille page. Braille cells are usually represented by two-digit octal numbers, but may be represented if desired, by printed periods corresponding to braille dots.

At this point, the test translation is checked for compliance with the braille rules. Also, the simulated braille page listing is checked for correctness of format. Revisions are made on tape based on problems which have been discovered in the results of the preceding runs.

The final processing run is a translation of the revised input data. Output from this run is a column binary card representing braille text and format.

The system used for processing is an IBM 360/65 with 512K storage, seven tape drives, an 1403 printer and a 2540 reader punch. Speed of translation is in excess of 9,500 words per minute. The 360/65 system replaced a 7040 system which had replaced the 709. The 709 was used originally to put braille translation on a production basis.

In day to day computer processing of braille, there are decisions that require an understanding of the braille code. The 360/65 operators are braillists who have also been trained on the computer.

As experience was gained from the translation of early titles, revisions were made to the original translation program. After the production of more than 100 titles, it was decided to develop a new program at APH. The new program provided for format control which made translation of textbooks and magazines possible. Included in the new format features were indexes, column material, glossaries, and different page numbering options. Also, a different approach was made in identifying and processing the successive units in translation of the text. The software support for braille translation now consists of more than 60 programs.

Plate Embossing

The punched card that controls the plate embossing process is a binary card with columns corresponding to the braille cells. Eight rows of the card are used to represent the eight-bit braille communication code - six rows for braille representation, one row for format control and one for parity. A biquinary number for insuring the correct sequence of the deck is contained in the cards.

A modified verifier is used as a braille card reader. The card reader is synchronized with the embossing mechanism of the plate making machines. Data is checked for parity by the reader and the card sequence validated.

The automatic plate embossing machine built at APH is controlled by data from the card reader. The embosser produces raised dots on folded metal plates at the rate of five cells per second or a page in three minutes and forty-five seconds. These plates are used on braille presses to produce interpoint braille. One operator can run three plate embossing machines. APH, at the present time, has 12 automated plate embossing machines in operation. Plates are proofread and any errors corrected in the metal. A major advantage of the automated braille system is that most errors are detected and corrected before the plate embossing process.

Observations

Years of braille production by computer has demonstrated the success of the automated approach. Keyentering offers an effective means for copying inkprint into computer readable data. Use of the computer in translating braille provides many advantages such as ease of making trial runs, job switching, reuse of input for repeated sections and automatic pagination. In the production of over 800,000 plates, the card-to-plate equipment has demonstrated a high degree of dependability and accuracy. In this automated environment, adjustments to change in volume of production can be made with a minimun of difficulty. Careful attention has been given to production costs. Records indicate the present system as installed at the APH, is efficient from a cost standpoint.

The system is applicable to the many types of publications which use the braille literary code. In Grade 2 braille translation, there are decisions which at the present time, cannot be made by computers. Attaining the desired result requires some human participation. One goal in development has been to increase the effectiveness and minimize the amount of human intervention.

The transfer of braille production to a computer is a significant step. It enables publishing for the blind to benefit from technological advances in computer hardware. It also creates a potentiality for capitalizing upon future breakthroughs in data retrieval and communication.

Appendix - A

A paper published in The International Journal for The Education of the Blind, December, 1965

A Frequency Count of the Symbology of English Braille Grade 2, American Usage*

By C. J. Kederis, J. R. Siems, R. L. Haynes

Braille is the medium for reading and writing used by the blind. It is a complex code that is read tactually and consists of 63 characters which are formed of from one to six dots in a two-column, three-row matrix called the Braille cell. The characters are used singly and in combination with one another to stand for letters of the alphabet, letter sequence, words, numbers, punctuation, and composition signs. The standard Braille literary code contains approximately 260 such assignments. Throughout the remainder of this paper these 260 items will be referred to as the elements, or constituents, of the code.

A factor important to the Braille code is the frequency with which the individual elements recur in the reading material of the blind. This factor is important in terms of (1) textual reduction through the use of contractions and abbreviations, (2) the amount of information that is conveyed by the elements of the code, and (3) the ease with which the code is read. It is with this factor of the recurrence of the elements of the code that the present paper is concerned.

There have been three such previous

---

*This study is part of a program of research supported by the United States Public Health Service grant, NB 05129-01, from the Institute of Neurological Diseases and Blindness.

Mr. Kederis is a Research Associate in the Educational Research Department, Mr. Siems is the Assistant Data Processing Manager, and Mr. Haynes is Data Processing Manager at the American Printing House for the Blind.

counts to which the authors could find reference. The earliest, Irwin and Wilcox (1929), was concerned with the difference in amount of space saved by Braille Grade 1½ and Braille Grade 2. (Braille Grade 1½ was a less-contracted form of the code officially used in the United States from 1918 to 1932; the present-day standard is Braille Grade 2). Their count consisted of 91,000 words taken from four literary works. Each of the four selections appeared to have been at the adult vocabulary and interest level. Concerning the validity of their results, Staack (1962, p. 89) stated, ". . . (the count) leaves much to be desired in the range of material sampled." This is true, but it should also be pointed out that the works were chosen because they were published in both Grade 1½ and Grade 2 Braille, enabling a comparison, the original purpose of the count. Another thing that affects the validity of the Irwin and Wilcox count concerns the fact it is 35 years old. While this is not a long time in terms of changes in written language, the present code differs on some 10 elements from the code used then; also, the rules for usage are not entirely the same today as they were then. In addition, their count did not include numbers or punctuation and composition signs.

The most recent of the counts was one by Staack (1962). As a good indicator of the frequency af the Braille elements, though, it is of little value because the sample size and sample material are entirely inadequate — 7000 words from patent applications. However, Staack's principal purpose in this undertaking was not to obtain an especially valid count, but to show that a computer is a useful tool for such an exercise.

The third count was one by Lochhead and Lorimer (1954). At the time of this writing, though, it was still not available to these authors. This count may have been the most valid of the three, but even so, it would also be subject to some objections on the grounds that it was done in the United Kingdom, and the use of the code there differs somewhat from the American code.

Thus, it should be obvious that another count is needed, a count which is more up-to-date, which samples a broader type of literature over a number of grade levels, and which includes all of the constitutents of the Braille code as they are used in this country.

Method

Sample

From a list of 45 books, two fiction and two non-fiction works were selected randomly, when possible, at each of three grade levels (4-7, 7-9, adult).* From each of the 12 books thus chosen, one Braille volume was used in the count. (An average-sized print book, 200-300 pages, requires four to five, 100-page volumes in Braille). Information about the sample is contained in Table 1.

Procedure

The frequencies were obtained on an IBM 709 computer. This computer is used at the American Printing House for the Blind as part of an automated process of producing Braille from ink print. The frequency count was accomplished by modifying part of the same program which is used in the translation step of the process. This is the Schack-Mertz IBM Braille Translation Program (1961).

A section of this program consists of a table of print symbols, each entry of which has three parts. Part 1 can be a letter, a letter sequence, a word or a punctuation mark. Part 2 is a code number indicating the circumstances under which the Part 1

---

*Due to time and financial considerations, only books on cards acceptable to a computer were used. Thus, only 45 books were available, and of these only five were rated as appropriate to the 4-7 grade level. All grade level ratings came from *The Book List* or CHILDREN'S CATALOGUE or STANDARD CATALOGUE FOR HIGH SCHOOL LIBRARIES.

items are used according to the rules of the Braille code. Part 3 is one or more three-digit numbers representing the 63 Braille characters into which the print symbols (Part 1) are translated.

TABLE 1

Sample Data

Publish. Date	*Title*	*Author*	*Type*	*Vocab. Level*
1957	Little Laughter	Love, K.	Verses & Rhymes	III-VII
1948	Yellow Fairy Book	Lang, A.	Fairy & Folklore	IV-VI
1964	Private Eyes	Kingman, L.	Fict.	IV-VIII
1963	Stormy: Misty's Foal	Henry, M.	Fict.	IV-VIII
1949	Hearts Courageous	Herman, W.	Biog.	VII-IX
1920	Children of Odin	Colum, P.	Myth.	VI-IX
1881	Prince and Pauper	Clemens, S.	Fict.	VI-IX
1963	Girl on Witches Hill	Lawrence, M.	Fict.	VII-IX
1963	Many Faces of Civil War	Werstein, I.	Hist.	IX-Adult
1964	Eighth Moon	Sansan	Soc. Sci.	Adult
1959	Advise and Consent	Drury, A.	Fict.	Adult
1962	Inheritors	Golding, W.	Fict.	Adult

For the frequency count, only Part 3 of the entries in the table was changed. Instead of the number codes representing the 63 Braille characters, a new identifying number was assigned to each of the 256 elements of Braille. For example, with Part 3 of the entries unchanged, the number 066 represents the Braille character with dots 2, 3, 5, 6. This character, depending on context, can be *gg*, *were*, or a *parenthesis sign* to the reader. However, for the frequency count, a distinct identifying number was given to *gg*, *were*, and the *parenthesis sign* separately. As the print input was read, instead of writing the numbers which represented the Braille characters, a total was developed for each of the 256 elements. The count was checked by comparing the total number of characters in a volume obtained by this method with a total previously obtained from the same volume by another method.

Results

The results consisted of the number of times each of the 256 elements of the Braille code appeared in the sample of Brailled reading material. The print counterpart of the sample contained approximately 291,000 words: grades 4-7 (84,000), grade 7-9 (94,000) and adult (113,000). It is necessary to specify Braille reading matter, because the frequencies of some of the elements, due to the rules of Braille usage, are not the same as in print, albeit the correspondence is close. For example, if the *ea* and *ed* letter sequences were being counted in print, their appearance in the words "uneasy" and "reduce" would add to their totals. However, in Braille, they would not be counted because their use would violate one of the rules of the code (English Braille, 1962, p. 33).

The number of times each of the Braille elements appeared in the sample is shown in Table 2. Similar information also exists for each of the grade level divisions and each of the individual volumes of the sample. To those who are interested, this material is available through the American Documentation Institute.*

Also included in Table 2 is the number of ink-print letters saved by each of the contractions and abbreviations in the Braille code. The total letter spaces saved, over the total number of ink-print alphabet letters used in the sample, showed a letter-space saving of 31%. Thus, the contractions and abbreviations of Braille reduced by 31% the use of the ink-print alphabet letters.

Other items contained in Table 2 are the punctuation marks, composition signs, and numbers. The punctuation and special signs (decimal, dollar, etc.) constituted 6½% of the total material. The special Braille composition signs represented 4½% of the total. Numbers accounted for less than 0.1%.

TABLE 2

Occurences of the Braille Elements and the Ink-print Letters Saved by the Contracted Forms†

Elements	*Occur.*	*Letters Saved*
E	61,862	
A	53,515	
S	43,211	
L	41,491	
O	41,032	
Cap. sgn.	40,849	
T	37,887	
I	36,498	
R	32,861	
N	27,978	
the	24,810	49,620
D	22,006	
M	21,786	
P	19,649	
Period	18,437	
Y	18,400	
C	18,346	
Comma	17,645	
U	16,456	
H	16,301	
ed	15,100	15,100
and	13,060	26,120
er	12,425	12,425
B	12,155	
W	11,807	
G	11,417	
in	11,321	11,321
K	11,047	
F	10,962	
en	9,098	9,098
ing	8,850	17,700
ar	8,490	8,490
st	7,973	7,973
V	7,708	
to	7,360	7,360
of	7,117	7,117
Quot (LD)	5,963	
Quot (RD)	5,932	
th	5,455	5,455
ea	4,829	4,829
Apost.	4,804	
sh	4,789	4,789
ow	4,233	4,233
ou	4,159	4,159
was	4,082	8,164
wh	3,582	3,582
ch	3,564	3,564
for	3,174	6,348
it	3,148	3,148
his	3,143	6,286
that	3,143	9,429
with	2,513	7,539
gh	2,337	2,337
you	2,266	4,532
one	2,178	2,178
had	2,139	2,139
as	1,834	1,834
but	1,812	3,624
be	1,758	1,758
said	1,636	3,272
him	1,613	1,613
Qu. Mark	1,565	
Hyphen	1,514	

---

*Frequencies by book and grade level have been deposited with the American Documentation Institute. Order Document No 8562 from ADI Auxiliary Publication Project, Photoduplication Service, Library of Congress 20540. Remit in advance $2.25 for microfilm or $5.00 for photocopies and make checks payable to: Chief, Photoduplication Service, Library of Congress, Washington, D. C.

†In order to avoid confusion on the point of textual reduction resulting from the encoding of Braille, it should be noted here and elsewhere that the savings or reductions referred to do not include the letter spacings saved between words as a result of rules governing the use of some of the whole-word contractions (*to, into, by, and, for, of, the, with*).

Elements	*Occur.*	*Letters Saved*
not	1,480	2,960
J	1,440	
were	1,338	4,014
from	1,274	3,822
Excla. Mk.	1,254	
X	1,243	
so	1,174	1,174
there	1,138	3,414
out	1,130	2,260
ound	1,105	2,210
com	1,074	2,148
ble	1,052	2,104
this	1,024	3,072
ever	986	1,972
would	976	2,928
ong	971	971
Z	968	
Dash	968	
have	964	2,892
could	947	2,841
by	908	908
Semicolon	811	
con	803	1,606
some	753	1,506
into	751	1,502
their	745	2,235
day	726	726
ment	680	1,360
ful	653	653
time	641	1,282
ation	634	1,902
do	629	629
Q	610	
more	610	1,830
Italic sgn.	595	
about	595	1,785
will	585	1,755
dd	584	584
like	582	1,746
ought	572	1,716
tion	567	1,134
little	521	2,084
ally	506	1,012
know	491	982
can	490	980
right	485	1,455
where	484	1,452
which	451	1,804
your	438	876
go	419	419
just	413	1,239
gg	411	411
ity	411	411
Colon	410	
ness	402	804
again	400	1,200
ff	397	397
ance	396	792
people	396	1,980
before	392	1,568
under	391	1,173
work	390	780
dis	389	778

Elements	*Occur.*	*Letters Saved*
great	389	778
after	385	1,155
good	361	722
here	359	718
ount	350	700
ence	343	686
through	323	1,615
very	322	966
still	300	1,200
much	294	588
himself	292	1,168
Number sgn.	287	
word	268	536
first	266	532
sion	263	526
upon	257	514
less	252	504
father	247	988
cc	245	245
part	245	490
because	241	1,205
bb	240	240
its	240	240
many	238	476
these	236	708
must	232	464
every	228	912
us	226	226
young	211	633
against	209	627
friend	205	820
shall	197	788
name	196	392
should	190	760
1	189	
world	187	561
always	182	546
such	181	362
0 (zero)	170	
across	165	495
mother	159	636
quick	157	471
Ellipsis	150	
Accent sgn.	148	
between	136	680
together	136	680
behind	134	536
those	128	384
enough	118	590
beside	116	464
also	111	222
quite	109	436
children	107	642
lord	99	198
almost	98	294
8	95	
above	94	188
herself	92	368
already	85	340
6	83	
although	80	320
question	78	468
rather	76	380
themselves	70	420

Elements	*Occur.*	*Letters Saved*
child	69	276
either	67	268
2	66	
Quot. sgn. (LS)	66	
Quot. sgn. (RS)	63	
receive	60	240
letter	59	236
cannot	58	236
whose	57	171
afternoon	55	330
neither	53	212
today	52	156
tomorrow	51	306
5	48	
myself	48	144
3	47	
immediate	46	276
perhaps	46	184
itself	45	180
beyond	44	176
4	42	
9	42	
spirit	42	168
Paren. sgn. (L)	39	
Paren. sgn. (R)	39	
yourself	39	195
below	37	111
tonight	34	170
declare	26	104
according	25	175
7	24	
character	24	168
paid	24	48
beneath	22	110
o'clock	22	88
blind	21	63
Brckt. sgn. (L)	16	
Brckt. sgn. (R)	16	
knowledge	16	128
Dollar sgn.	15	
Termin. sgn.	13	
rejoice	13	52
Letter sgn.	12	
necessary	12	72
afterward	11	66
ourselves	10	50
Asterisk	9	
perceive	9	36
thyself	9	36
deceive	6	24
Decimal pt.	5	
rejoicing	5	25
conceive	4	20
perceiving	3	15
receiving	2	10
yourselves	2	12
altogether	1	7
oneself	1	4
Ampersand	0	
Fraction line	0	
Percent sgn.	0	
braille	0	0
conceiving	0	0
declaring	0	0
deceiving	0	0
TOTAL	924,760	392,521

To obtain a total for all the print letters used in the sample, it was necessary to reduce the frequencies of the Braille elements to those of the 26 letters of the alphabet. Thus, a frequency count of English print letters was obtained. This information is shown in Table 3.

TABLE 3

Number of Occurences of the Alphabet Letters in English Print

Ltr.	*Occur.*	*Ltr.*	*Occur.*	*Ltr.*	*Occur.*
E	158,080	D	62,364	P	21,113
T	114,903	L	52,565	B	19,997
A	101,990	U	35,011	K	12,705
O	94,832	W	31,877	V	10,468
N	85,359	M	30,201	J	1,871
H	83,907	C	28,462	X	1,243
I	80,010	G	28,186	Z	968
S	76,412	F	26,061	Q	954
R	70,741	Y	25,127		
				Total	1,255,407

Comparing the order of the print letters in Table 3 with those for Braille in Table 2 showed one of the effects of the contractions in Braille. The position of *t* changed from second place in Table 3 to sixth place in Table 2. *L* changed from 11th to 14th place, and *h* from 6th to 16th, etc.

In order to determine the frequency of recurrence of the 63 Braille characters, the frequencies for the constituents were combined wherever one of the characters was used. The results of this reduction are contained in Table 4.

TABLE 4

Number of Occurences of the 63 Single-cell Braille Characters

Char. Dot Nos.	*Occur.*	*Char. Dot Nos.*	*Occur.*	*Char. Dot Nos.*	*Occur.*
1-5	63,995	1-2	14,743	1-4-6	5,176
1	56.200	1-2-4	14,406	1-3-4-6	4,676
2-3-4	48,089	2-4-5-6	14,406	1-5-6	4,574
1-2-3	44,567	1-2-4-5	14,336	3-6	4,524
2-3-4-5	43.500	1-2-3-4-6	13,060	1-6	4,239
1-3-5	43,233	1-2-4-5-6	12,575	2-4-6	4,233
6	42.084	3-5	12,090	5-6	4,039
2-4	36,706	1-3	11,711	2-3	3,931
1-2-3-5	35,030	5	11,177	4-5-6	3,409
1-3-4-5	31.401	3-5-6	10,985	1-2-3-4-5-6	3,174
1-4-5	28,726	2-3-6	10,737	4-6	2,966
2-3-4-6	26.999	2-3-5	9,762	1-3-5-6	2,802
1-3-4	25,583	2-6	9,216	2-3-4-5-6	2,513
2	22,474	3-4	8,980	1-2-6	2,337
1-3-4-5-6	22,374	3-4-6	8,850	2-4-5	2,041
1-2-5	22,035	3-4-5	8,490	2-3-5-6	1,859
1-4	20,481	1-2-3-6	8,290	2-5	1,462
1-2-3-4	20,372	1-2-3-5-6	7,117	3-4-5-6	1,339
2-5-6	19,425	1-4-5-6	7,019	1-2-3-4-5	954
1-3-6	17,330	1-2-5-6	5,871	4-5	946
1-2-4-6	15,100	3	5,368	4	148
				Total	966,235

those of ink-print, including numbers, punctuation, and special signs, it was found that Braille used 26.5% fewer characters. Here characters refer to all the symbols in written English for which there are Braille counterparts. Thus, Braille diminished the use of the ink-print letters and symbols by approximately 26.5%.

One further reduction of the data was made to determine the frequencies of the individual dots within the Braille cell. This information is contained in Table 5.

This reduction showed that the dots in the various positions within the Braille cell did not recur with equal frequency, but that the dots on the left occurred 7% more often than those on the right, and the top

The grand total for all the characters is also contained in Table 4. Comparing the total number of Braille characters with

TABLE 5

Number of Occurences of Dots by Positions Within the Braille Cell

Occur.	*Dot Positions*		*Ocur.*
583,487	1	4	453,028
491,320	2	5	452,659
499,690	3	6	303,298

dots were 8% more prevalent than the bottom dots. Moreover, dot 1 occurred almost twice as often as dot 6. The total number of dots counted over the total for the characters give 2.88 as the average number of dots for the characters.

Discussion

Even though the present count was suitable for the purposes of the authors, as a count reflecting the frequencies of Braille symbology in the reading matter of the blind, it was somewhat limited. According to Josephson (1961) only a small percentage of blind persons continue to read books by means of Braille after leaving school. By inference, then, a large, perhaps the largest, amount of the material read in Braille is of the textbook or periodical type, and none of this material was included in the sample for the count. Hence, although the present count was more extensive and covered a wider range of material than either of the two mentioned earlier, it was still not as adequate as it should have been.

One of the reasons for the count was to determine whether relationships existed between the frequency of recurrence of the characters, dots, etc., and these factors as they influenced perception of the Braille characters. Nolan and Kederis, in Nolan (1964) established orders of legibility of most of the single-cell Braille characters. Significant correlations, coefficients up to .46, were obtained between these orders and the order of frequency of recurrence of the characters. They also found that the Braille characters with fewer dots were more easily recognized than those with many, and that the dots on the right and at the bottom were missed more often than those on the left and at the top. These findings, also, corresponded well to the results of the frequency count. Approximately 70% of the characters that appeared in the sample for the frequency count were of three and fewer dots, and the occurrence of dots in each of the six cell positions was in a direct inverse relationship with the frequency of missed dots by cell position. Thus, some of the factors that are important to the perception of the characters appeared to be partially related to their frequency of appearance in reading material.

Another interesting aspect of the results of the count was the change the contracted forms of Braille produced in the order of frequency of recurrence of the letters of the alphabet. Thus, in studies such as the one by Mayzner, Tresselt, and Adler (1964), wherein they found high correlations between subject-generated letter frequencies and the frequency of recurrence of the letters in English, different relationships might have obtained had the subjects been Braille readers. Conceivably, such differences could even reflect on the way in which Braille is perceived.

BIBLIOGRAPHY

ENGLISH BRAILLE AMERICAN EDITION 1959 (REVISED 1962) Louisville: American Printing House for the Blind, 1962.

Irwin, R. B. & Wilcox, Ruth E. A COMPARATIVE STUDY OF BRAILLE GRADE ONE AND A HALF AND BRAILLE GRADE TWO New York: American Foundation for the Blind, 1929.

Joesephson, E. "Some Insights Concerning Reading Problems as Refined in AFB Leisure Study." IN REPORT OF PROCEEDINGS OF CONFERENCE ON RESEARCH NEEDS IN BRAILLE. New York: American Foundation for the Blind, 1961.

Lochhead, H. M. & Lorimer, J. REPORT OF THE SURVEY OF THE FREQUENCY OF ALL THE CONTRACTIONS OF STANDARD ENGLISH BRAILLE GRADE 2. Craigmillar Park, Edinburgh, Scotland: Scottish Braille Press, 1954.

Mayzner, M. S.; Tresselt, M. E.; Adler, S.; & Schoenberg, K. M. "Correlations Between Subject Generated Letter Frequencies and Observed Frequencies in English." *Psychonomic Science*, 1964, 1 (10), 295-296.

Nolan, C. Y. "Cues in the Tactual Perception of Patterns." Progress report, NIH grant no. NB 03129-01. Louisville: American Printing House for the Blind, 1964 (mimeo).

Schack, Ann S. & Mertz, R. T. BRAILLE TRANSLATION SYSTEM FOR THE IBM 704. New York: IBM Data Systems Divisions, 1961.

Staack, G. F. A STUDY OF BRAILLE CODE REVISIONS. Master's thesis, Massachusetts Institute of Technology, 1962.

* * *

Appendix - B

An early report of computer translation indicating the various areas involved in developing a foundation for an automated braille production system

American Printing House for the Blind
INCORPORATED
FINIS E. DAVIS, VICE PRESIDENT AND GENERAL MANAGER
1839 FRANKFORT AVENUE
MAILING ADDRESS: P. O. BOX 6085
LOUISVILLE, KY. 40206

COMPUTER TRANSLATION OF GRADE 2 BRAILLE
AMERICAN PRINTING HOUSE FOR THE BLIND
MAY, 1964 - JUNE, 1968

In May, 1964 International Business Machines made available to the American Printing House for the Blind without rental cost a 709 computer to be used for Braille translation and Braille research. A computer program for translating inkprint into Grade 2 Braille had been written and tested in a production environment by IBM in cooperation with APH. Developments during the four years in which the system has been in operation are presented in this report. An added resource for development and research has been a U.S. Office of Education grant awarded to APH in June, 1966 for studies including computer translation of music, mathematics and scientific literature into Braille.

Production

In the four-year period, 288 titles or 809 Braille volumes were translated by the computer. These titles consisted of 114,910 metal plates which were used to print 6,300,000 pages of Braille. The largest title was a reference book having 25 Braille volumes. Input for this book consisted of 100,000 punched cards, and the result of translation was 85,000 Braille output cards.

As the number of titles translated has increased there has been a parallel extension of the range of materials covered. The application has been broadened to include textbooks and magazines. Indexes, glossaries, tables of contents, title pages, and columned material have been formatted by the computer. Also, books in poetry form throughout were translated.

The production procedure begins with an operator, who does not know Braille, producing input by keypunching. One operator can produce the equivalent of 50 to 60 Braille pages per day. After keyverifying, the input is translated by the computer. The speed for converting inkprint data to Braille data is about 2,000 words per minute. Some additional time is required for peripheral operations. Output from the computer is handled by card-to-plate equipment which was developed by APH and IBM. This equipment reads punched cards and automatically produces Braille plates, from which an unlimited number of copies can be embossed. Four card-to-plate output units are in operation. One operator, using these machines, can make 300 plates per day.

Accuracy of the system was good from the beginning and has progressed to a level where the vast majority of problems in Grade 2 Braille translation are encountered and solved before plate-making, thus reducing to a minimum the difficult task of correcting metal plates.

System Development

As further testing of the original program was made along with production a number of areas were identified as requiring program revisions. These included refinements for controlling format at the end of a page, and further techniques in translating prepositions. A Braille sequence routine was added to the program so that in some cases translation could be based upon a longer succession of characters. Since the original program was written for an IBM 704, input and output instructions were rewritten in 709 language. An intermediate table was inserted to serve as an index to the Grade 2 table, permitting the program to operate at an increased speed.

The computer translation procedure was designed originally to produce Braille for general reading. In 1965 work was done toward programming the translation of Braille textbooks. An attempt was made to accommodate the 1964 adoption of rule changes covering textbook format. Among features of this format is a requirement for indicating inkprint page numbers. This adoption was for the benefit of blind students attending classes with sighted students, and using the same texts. The textbook program was added to the original translation program and provided for ink-print and Braille pagination, optional interlining, and other text arrangements required in educational materials. This development in format control made possible the translation of Braille magazines and also provided features, such as right margin adjustment, which were useful in literary books.

After 18 months experience in Braille translation, which included the translation of a textbook series, it appeared that a new approach to translation might be desirable. A procedure for computer translation employing new principles was developed. The procedure uses a somewhat different basic unit of translation and separates the processing of general rules from the processing of exceptions. This provides for greater ease in updating the program as new exceptions are encountered.

Within the last year at APH a completely new Grade 2 Braille translation program has been written for the 709. This program incorporates the principles of the new approach and includes the textbook format features. Other features have been added to overcome various translation difficulties and to provide greater flexibility and ease of program operation.

Progress was made during the first two years in reducing the amount of human intervention required in the translation procedure. This is particularly true in regard to that phase of the procedure in which a prooflisting of the results of translation is read by a Braillist. Some reduction in the length of the listing was achieved by eliminating the shorter words (five Braille characters or less) which rarely contained errors. A more effective approach was the development of a one-time translation routine by which a word was translated only the first time it appeared in a volume. Later a program to eliminate common words was added to the system. By making a test translation of a volume omitting occurrences of the 10,000 most frequently used words and translating only the first occurrence of the remaining words, prooflist checking was reduced by about 90%. The present procedure is to make a test translation of a condensation of the input. The resulting prooflist shows words which have not been translated in previous books and problematic character combinations, with their approximate location in the ink-print book. This requires a reading of perhaps only 300 words per book by a Braillist. Format is checked by scanning a full prooflisting in which Braille cells are represented by two-digit octal numbers.

A keypunch manual was written before the second year of production began. This manual serves as a reference for keypunch symbols used in preparing input for Braille translation. A supplement to this manual includes symbols for indicating diacritical marks.

A number of auxiliary programs have been developed to support the Braille translation system. Among these are a program to search input data, a program to list the exception table, and a Braille card to printer program. The latter simulates the automatic embossing machine.

A Braille output program has been written as a corollary to the new Braille translation program. This program punches Braille cards from tape. It also provides for the printing of the results of translation. The Braille cell may be represented either by dots or by an octal number. If desired, the output program can also print beneath each Braille line the inkprint equivalent of the Braille signs.

Related Developments

During the last two years progress has been made under the Office of Education grant toward computer translation of mathematical texts following the rules of the Nemeth Code. Considerable attention has been given to the creation of a procedure for keypunching input. Computer techniques have been developed for producing Braille mathematical notation. Work on this project has been discontinued while the Nemeth Code is in the process of being revised.

Because the present Braille translation system does not hyphenate words at the end of a line, a study was undertaken to determine the space-saving value of hyphenation. The results of the study indicate that the amount of space saved by hyphenation is insufficient to justify a major programming effort at the present time.

An important by-product of the Braille translation system is that for the first time a large quantity of Braille literature is in machine readable form. Also, a large amount of ink-print copy is similarly available for processing. Counts have been made of Braille dot frequency, Braille character frequency, occurrence of Braille signs, and ink-print letter sequences. Programs were developed at APH for the above studies. Anticipating further research in Braille symbology all entries in an unabridged dictionary along with their syllabication were punched into cards.

During the four years, systems support programs were written to facilitate the development of applications and research programs. Also, with a view toward exploring future hardware needs, programs were written to enable the 709 configuration to simulate System 360 program assembly and execution. A test program for solving a typical Braille translation problem was written in 360 language. The program was assembled and debugged on the 709 and later run on a 360 system.

Observations

Four years of Braille production by computer has demonstrated the success of the automated approach. Keypunching offers an effective means for copying inkprint into computer readable data. Use of the computer in translating Braille provides many advantages such as ease of making trial runs, job switching, reuse of input for repeated sections, and automatic pagination. In the production of over 100,000 plates, the card-to-plate equipment has demonstrated a high degree of dependability and accuracy. Careful attention has been given to production costs. Records indicate the present system as installed at APH is efficient from a cost standpoint. The system is applicable to the many types of publications which use the Braille literary code.

The transfer of Braille production to a computer is a significant step in that it enables publishing for the blind to benefit from technological advances in computer hardware. It also creates a potentiality for capitalizing upon future breakthroughs in data retrieval and communications.

In Grade 2 Braille translation there are decisions which, at the present time, cannot be made by computers. Attaining the desired result requires some human participation. One goal in development has been to increase the effectiveness and minimize the amount of human intervention.

The IBM 709 computer has provided the necessary capability for a Braille translation system. Results of a benchmark indicate that the data format and instruction set of IBM System 360 is even more applicable to this complex task. It appears that a 360/40 would have the capacity and speed for efficient Braille translation.

Requests for information which have come from many groups indicate that there is a wide interest in Braille translation by computer. Many hours have been spent in furnishing programs and technical assistance to representatives of organizations from foreign countries and throughout the United States.

Appendix - C

An APH study in OCR distributed in the research publication of the American Foundation for the Blind

CURRENT RESEARCH NOTES

PROCESSING OPTICAL CHARACTER RECOGNITION (OCR) OUTPUT INTO INPUT FOR BRAILLE TRANSLATION

Robert L. Haynes
Data Processing Manager
American Printing House for the Blind
Louisville, Kentucky

In June, 1970, the American Printing House for the Blind (APH) agreed to investigate the feasibility of Optical Character Recognition (OCR) as a means of producing input for a braille translation system. The first book to be published as a result of this project is *The Unsubstantial Castle* by Robert Rushmore.

Conferences were held with two OCR companies in which text reading was discussed in detail. Dissly Systems in Easton, Pa. was willing to scan the inkprint copy of *The Unsubstantial Castle*. The other company declined to participate. Dissly furnished a tape that was not error free but was a reasonable representation of the inkprint copy. The cost of the Dissly tape was somewhat higher than the APH cost of keypunching and verifying.

The APH is appreciative of the interest and cooperation shown by Dissly Systems in making this project possible.

The following report describes the detailed procedure for converting scanner output to translation input.

PROCEDURE

The 9-track output tape was copied to 7 track. Since the size of the character set exceeded the capacity of 6-bit representation, the 8-bit characters of the 9-track tape were converted to 12-bit characters on the 7-track tape.

The text was scanned to identify all the characters used and to determine the significance of each by locating the first occurrence of each. Besides numerals and upper and lower case alphabetic characters, 19 special characters were used in the text.

A run was made to change the tape record format. Numbers indicating the line number of the inkprint page were removed.

1. Page number records were deleted and page number was inserted into a fixed field in each record. Numbers were adjusted to correspond as closely as possible to the inkprint page numbers. Because of the blank pages in the inkprint text, numbers on the scanner output tape varied from 2 to 7 pages from the inkprint copy.

2. On the scanner output tape, special format features such as indentation or centering were indicated by an asterisk at the beginning of the record. Since the special feature in the majority of cases was a new paragraph, the paragraph symbol was inserted at these points.

3. The text was put into records having 72 characters.

4. Hyphens marking word division at the end of the inkprint line were removed in this run. The scanner output used a different character to represent the hyphen of a compound word at the end of a line. These hyphens were retained.

Text data was converted as far as possible to the symbols used in braille translation input. For example,

colon	to	#.
question mark	to	$Q
A	to	#A
b	to	B

Processing up to this point was primarily done by the computer with a minimum of human intervention.

Further work was done primarily of an editorial nature but using the computer a considerable amount to locate and print out various text features and to rewrite the text incorporating the editorial changes.

In the scanner output, fully capitalized words or word series were enclosed by asterisks. These words were located. In this particular book, most of these words needed to be changed to lower case. (14 changes)

Italicized words or word series were enclosed by percent signs in the scanner output. These words were listed in context so that appropriate single of double italics signs could be inserted. Some of these words also required Grade 1 symbols. (141 changes)

Ellipses, spaced periods in the scanner output, were examined to determine use of ellipsis only, ellipsis preceded by period, or ellipsis followed by period. (59 changes)

Dashes were checked for spacing between the dash and the word following. (26 changes)

Chapter and section headings were located and symbols were inserted to control chapter and heading titles. (30 changes)

Hyphenated words were noted and the $H (non-line-ending hyphen) symbol was inserted where required as in stammered words. (50 changes)

Format symbols for features such as poetry lines and blank lines, obvious on the printed page but not indicated in the scanner output, were inserted. (9 changes)

Need for a few changes such as letter sign was discovered inadvertently. (3 changes)

In the process of making editorial changes a number of scanning errors or omissions were noticed.

Corrections were made for these errors. (78 changes)

p 4	ln 22	Sheridan, a	space between words omitted
p 10	ln 29	cry	for cry.
p 13	ln 1	Two	special format of chapter title not indicated
p 20	ln 28	"Auratum	right quotes for left quotes
p 20	ln 28	platyphytllum	for platyphyllum

p 27	ln 18	”Oh?“	for “Oh?”
p 29	ln 23	Francis--"“	for Francis--"
p 34	ln 27	the pain	space between words omitted
p 35	ln 12	Nothin	for Nothing
p 37	ln 1	Three	special format of chapter title not indicated
p 38	ln 12	possible---while	for possible--while
p 39	ln 21	"I-I	wrong hyphen at end of line
p 39	ln 22	didn't--	for didn't--"
p 43	ln 13	No	not placed in italics
p 49	ln 1	réclame	accent not indicated
p 51	ln 25	'A	apostrophe instead of left single quote
p 51	ln 25	skin,'	apostrophe instead of right single quote
p 51	ln 26	'It'll	apostrophe instead of left single quote
p 51	ln 26	somewhere.'"	apostrophe instead of right single quote
p 51	ln 26	somewhere.'" But	extra space between words
p 52	ln 27	. . . those	extra space between ellipsis and word
p 53	ln 2	Francis	for Francis.
p 53	ln 3	"Yes.	special format of beginning of paragraph not indicated
p 54	ln 17	I	not palced in italics
p 55	ln 30	fills	for filles
p 62	ln 17	cied	for cried
p 69	ln 14	bizrre	for bizarre
p 74	ln 2	I	not placed in italics
p 75	ln 15	men."	left quotes for right quotes
p 78	ln 23	hard,.	for hard,
p 81	ln 1	Four	special format of chapter title not indicated
p 94	ln 6	sir?"	left quotes for right quotes
p 96	ln 15	hungry-	for hungry."
p 98	ln 5	horse?"	left quotes for right quotes
p 99	ln 2	pâté	accents not indicated.
p 99	ln 3	pétillant	accent not indicated
p 99	ln 3	Neuchâtel	accent not indicated
p 103	ln 30	just	for Just
p 104	ln 24	. . .	for . . ."
p 110	ln 29	sarcastically, "What	space between words omitted
p 113	ln 28	affectionate, if	space between words omitted
p 114	ln 9	knocks. He	space between words omitted
p 116	ln 6	Rhy-Jones's	for Rhys-Jones's
p 119	ln 21	I	not placed in italics
p 119	ln 31	I	not placed in italics
p 120	ln 24	"He	right quotes for left quotes
p 122	ln 2	wish	not placed in italics
p 123	ln 1	Six	special format of chapter title not indicated
p 124	ln 1	Rhy-Jones	for Rhys-Jones
p 132	ln 25	giv-- ing	for giving
p 135	ln 15	Musical?'	for Musical?"
p 135	ln 21	uncle."	left quotes for right quotes
p 144	ln 20	Requiem	not placed in italics
p 145	ln 27	"Killed---?"	for "Killed--?"
p 148	ln 19	too--"	left quotes for right quotes
p 149	ln 6	couple, attractive	space between words omitted
p 152	ln 26	pride	not placed in italics
p 156	ln 1	Seven	special format of chapter title not indicated
p 161	ln 3	I	not placed in italics
p 163	ln 22	scene, guilty	space between words omitted
p 165	ln 26	Francis walked	special format of beginning of paragraph not indicated
p 166	ln 29	cro-- quet	for croquet
p 168	ln 3	two	not placed in italics

p 175	ln 16	die."	left quotes for right quotes
p 178	ln 22	hands	not placed in italics
p 182	ln 15	they've---they've	for they've--they've
p 187	ln 7	"Hey!""	left quote annexed to right quote
p 187	ln 15	in,"	left quote for right quote
p 193	ln 1	Eight	special format of chapter title not indicated
p 196	ln 18	you---"	for you--"
p 200	ln 4	bigger."	left quote for right quote
p 200	ln 12	voice,asking	space between words omitted
p 201	ln 31	how---?"	for how--?"
p 203	ln 7	Sherriden/ speeking	for Sherriden/speeking
p 203	ln 8	go /away	for go/away
p 204	ln 21	you	not placed in italics
p 207	ln 10	"Long	right quotes for left quotes
p 207	ln 17	William?"	left quotes for right quotes

Computer processing time was 5.5 hours. This does not include time for compilation and testing of programs or for punching cards from tape.

Cards punched from the tape were keyverified. An additional 150 text errors were discovered in this process.

Editor's Note: As the reader may infer from the above details, the prospect of using optical character recognition techniques as a method of generating input to an automatic braille transcription process is not encouraging. Optical character recognition (OCR) systems that can "read" a large number of differing fonts, and can cope with variable pitch printing, tend to be very expensive. And the output from the best of such systems yields text which is only 92 to 99 percent correct. For many purposes, a one percent error rate due to machine faults may well be acceptable; but as an input to an automatic braille transcription scheme, this error rate is so high that it would be better to prepare input in other ways. Note that the errors encountered, and tabulated, in the above report were left after a thorough proofreading by the OCR device's manufacturer.

It would seem clear, from this one instance at least, that the future of OCR systems in making input for braille transcription systems would depend on the further technical refinement of OCR techniques in the areas in which they are now being used with success, that is, in postal applications, in publishing, in machine translation, and the like. But we would be interested in the reaction of readers, both from those who have used OCR systems, or are working with them, and from those who are now developing next-generation systems which we can consider as potential candidates to generate input to an automatic braille transcription system.

Appendix - D

APH developments in braille music

SAMBA and RUMBA: Systems for Computer Assisted Translation of Braille Music

by William Watkins, John Siems,
and Robert Haynes

INTRODUCTION

In May, 1971, the American Printing House for the Blind (APH) undertook a project, in cooperation with the Library of Congress, Division for the Blind and Physically Handicapped, to develop a system for computer assisted translation of Braille Music. To date, progress has been made toward the implementation of two systems for this purpose--SAMBA (Systems Activated for Music Braille Automation), and RUMBA (Representations Utilizing Music Braille Alphmerics).

SAMBA

Input

In collaboration with Music Print Corporation of Boulder, Colorado, APH has developed a music typewriter-to-card punch machine, a modification of an IBM 826 typewriter-to-card punch, capable of encoding standard music notation of virtually any complexity. In addition to alphamerics, the type font contains 38 music graphics which can be used singly, or can be superimposed, to produce computer-identifiable symbols. The output code is identical to the 47 punch combinations of the 826 keypunch with codes reserved for horizontal and vertical spacing. Forward and backspacing generate codes, as does movement up or down the music staff. In this way, a two-dimensional array is represented by a succession of punched codes and a typed copy of the music text is produced for subsequent proofreading.

Programming

The programs written for the SAMBA System are, at present, limited to tanslation of relatively simple single-staff music with a significant amount of human intervention. While the programs are sizable, requiring some 350,000 words of storage, some of this area is reserved for routines to be added for translation of music of higher degrees of complexity. This notwithstanding, it is felt that these are pioneer efforts and that eventual full production programs will have to be rewritten. As a result, the problem of running time has not been an overriding consideration in the design of initial programs.

Programs, written in assembler language, were developed for each of five phases of processing between the music typewriter device and the Braille music code.

Input Device to Music Characters

The input device produces a typed copy of the music text and punched cards. Codes in the punched cards represent music typewriter graphics, rotation of the platen, and forward and backward spacing of the carriage. The first computer program essentially reconstructs the two-dimensional music text image from the data in the punched cards. It identifies each music character along with its vertical and horizontal coordinates. Provision has also been made to handle error corrections included by the operator in the input.

Music Characters to Music Units

The program for the second phase attemps to interpret the meaning of the inkprint music graphics. Alphabetic items are separated from music signs. Basic forms and values are identified and modifiers are associated with items modified.

The output of the program consists of a representation in words of the music text. It is designed to be easily understood and provides a point at which changes may be made, if desired, before further processing.

Music Units to Edited Music Units

Information implied in the original music notation is stated more specifically in this phase. The pitch value of a note, for example, implied by its position on the staff and the nearest preceding clef sign, is inserted in the text. Other editorial additions specify details about features, such as octave numbers which are particularly related to Braille representation.

Edited Music Units to Braille Sign Names

The fourth program converts the edited music text to a series of abbreviations which stand for the Braille music signs. The terms correspond to the meaning of the Braille characters. Currently, about 1250 sign names are used.

Another opportunity for intervention is available at this point, since changes may be made in the Braille sign names before going on to the last phase.

Braille Sign Names to Braille

The fifth program is designed to run simultaneouly with the APH Grade 2 Braille translation program. Together the two programs process alphabetic and music data and set up the format of the Braille page. One of the functions of the music program is to effect the modifications required as the test is arranged into lines.

Auxiliary Programs

Several programs have been written for peripheral operations. One of these is a program to display music input on the printer. Music typewriter graphics not available on the printer are represented by two-letter symbols. This two-dimensional display may be compared with the inkprint copy to determine whether or not a section of music text has been correctly entered into the computer.

Output

The automatic stereograph machines developed at APH for literary translation are usable for music Braille, as the same cell configuration and page format are common to both codes. From computer-prepared instructions, these machines emboss metal plates which are subsequently used in conjunction with presses to produce multiple paper copies.

Discussion

The high degree of complexity of music notation in general, and the Braille music code in particular, which ramify in computer factors such as lengthy run times and high volume internal storage requirements, render inconclusive the possibilities for the practical application of a highly automated system for the production of Braille music--further research is indicated.

As a point of interest, one of the largest problems of computer translation of Braille music from the standpoint of storage, seems to be the ramifications of a Braille page change. To minimize the necessity for cross-page referencing on the part of the Braille music reader, beginning a new Braille page embodies the potential of having a dramatic effect on the configuration of the music notation.

For example, a doubling presumed to be in effect: (1) can be continued by restatement on the new page; (2) cannot be begun on the old page but must be started on the new; (3) can be in effect on the old page but must be terminated on the new; or (4) can disappear entirely, becoming illegitimate on both pages.

Braille repeats and page changes are potentially an even more complex problem. A series of measures of music represented as a repeat by half a dozen cells or so can, with the occurrence of a page change, subsequently require many lines of Braille. These and similar problems require tremendous computer storage capacity; the phenomenon of the potential page change requires that many decisions can be finalized only at the last possible moment in translation, prohibiting (or at least greatly qualifying) the use of common programming techniques which conveniently allow the components of items, one calculated, to be dropped from storage.

RUMBA

The first step in the operation of RUMBA consists of the conversion of music into the RUMBA language by an input specialist of whom a thorough knowledge of the Braille music code is required as well as certain fundamentals of Grade 1 literary Braille. Because most titles, credits, and literary footnotes are translated into Grade 2 Braille, it is usually necessary to have them tranlated prior to the encoding step in order for the input specialist to maintain an accurate cell count on the Braille line.

The input specialist indicates the RUMBA language on coding sheets consisting of rank and file numbered blocks corresponding to the cells on a Braille page; this allows for an accurate line and page count to be maintained.

After completion by the input specialist, the coding sheets are keypunched into cards or magnetic disks.

Processing

If the music to be processed is vocal music (and therefore contains lyrics), it requires extra initial processing to edit the lyrics and thereby prepare the data for the basic music translation programs, RUMBA 1 and RUMBA 2. If the data contains no lyrics, the file goes directly to RUMBA 1.

Prooflisting

In addition to the availability of standard prooflisting procedures from the APH literary translation system, special Braille music prooflisting programs have been developed. Along with the Braille dots, the RUMBA input is also indicated on the prooflisting.

Input Revision and Editing

Changes in the resulting Braille music output can be effected by the input revision procedures. These programs index the input file; that is, each item of input data is given a specific, unique, and serialized label enabling an editor to address any item(s) and substitute virtually any other item(s).

Programming

The programs for RUMBA, written in assembler language, have as their function the implementation of the language used by the input specialists. In its present form the RUMBA system requires two runs to transform the coding sheet terminology into Braille symbols. The first run analyzes the data for indications of the various modes and recognizes the material as musical or literary, page text or footnote, parenthetic or non-parenthetic, etc. It then edits for translation. The second run converts the music symbols to the appropriate signs of the Braille music code and translates literary material to Grade 1 or Grade 2 Braille.

For the second run the APH Braille translation program is required along with an additional RUMBA program. The latter effects the recently adopted music format feature of placing the inkprint page number in the upper left corner of the Braille page and also handles certain aspects of spacing which are characteristic of music.

A table of signs used in RUMBA processing is contained in a file of cards. The file is arranged in logical sections with headings. At the beginning of the first run the table file is read. The headings are ignored; and the entries are sorted for look-up. The same file may be read by a printing program which lists the symbolic names and shows the dot configurations of the corresponding Braille cells. By means of the print program, the file becomes the basis for a reference manual.

Attention has been given to programs to assist in making corrections and revisions in input data. A number of edit programs form part of the APH Braille translation operation. Some of these are applicable to music. However, additional editing approaches especially related to text structure typical of music have been completed. Input files consist of 80-character records having actual text in colums 1-72 and the inkprint page number or other identifying information in colums 73-80. The revision program for RUMBA, along with updating the text, also maintains this marginal page information as the file is rewritten. The is expected to be more significant as longer works having more inkprint pages are attempted.

In order to accommodate vocal lyrics, a program was written which makes modifications before the data is read by the other RUMBA programs. Music text containing lyric lines is read by the edit program and changed as necessary so than the repeat signs and words containing "in" will be processed correctly by later programs. The RUMBA lyrics program recognizes the input symbols which distinguish both from the mnemonic code. Provision is also made for preserving the applicable status as the data alternates between page and footnote material.

In developing the lyric program, a support program was written for text manipulation. Thes support program might be used for other processing of RUMBA input or of intermediate data files of RUMBA or SAMBA. It provides for maintaining, as a file is read and rewritten, a continuous text work area and for performing within this area deletions and insertions and other text operations. Programming for additional modifying functions which may be needed will by facilitated by this set of routines.

The RUMBA programs, including the lyrics and literary Braille translation programs, require approximately 200,000 words of storage.

Output

Automatic stereograph machines are utilized in both the RUMBA and SAMBA Systems.

Discussion

The RUMBA System is complete, is in production, and is not limited in the types and comlexities of Braille music it can produce. Training of input specialists requires approximately eight weeks, based upon a 40 hour week, which contrasts with the two-to-three years apprenticeship usually necessary for a manual transcriber. The input language shares the problems in use and proofreading of all music input languages which attempt to represent two-dimensional music notation with one-dimensional alphamerics. While much less convenient to use than the SAMBA input system, it is a viable alternative to the manual process of producing metal plates at a stereograph machine, and it does not require an in-depth knowledge of Grade 2 literary Braille.

Appendix - E

PROSPECTS FOR UTILIZATION OF COMPOSITOR'S TAPE IN THE PRODUCTION OF BRAILLE - Grete Grunwald

A Revolution in Braille Publishing Procedures?

A goal for many workers for the blind has been to make braille as abundantly available to the blind as inkprint is to the seeing.(1) The obvious main obstacle in achieving this goal is the relatively small number of copies required in braille, and the corresponding high cost per copy.

Could not commercial printing give a hand? For instance, "...A text that is automatically revisible and transformable along planned lines would allow regional and minority publishing."(2)

Compositor's tapes (that is, tapes prepared for automated typesetting) contain text in automatically revisible and transformable form. Such tapes, therefore, seem to be useful for minority publishing (as would be any text which is machine-readable). The problem is to find sources for such material.

Is then a Braille Revolution coming, based on compositor's tape? That would depend on whether there is enough tape available (especially of straight matter); that means enough to assure a braille publisher fast and convenient access to a wide choice of titles and to keep the overhead from spiraling as a consequence of information-hunting.

Furthermore, not every compositor's tape is equally suitable for transformation; Rice, in the quote above, rightly stresses transformation along planned lines; he lists, for instance, careful indexing and correct coding as well as "flagging" of points at which decisions may be necessary.

In order to weigh the chances for finding a firm base for braille in materials used primarily for commercial printing and publishing, it is necessary to understand what is involved up to the stages of printing and publishing literature.

In 1964 approximately 28,000 titles were published in inkprint in the U.S.A.(3). Editing and printing is the combined business of the publishing and printing industry.

This huge industry is in great pains through a challenge which started some 30 years ago, when rapid advances in printing techniques occurred. Mechanization, automation and specialization, together with entirely new techniques, and finally, within the last decade or so, the use of computers in composing rooms caused revolutionary changes. Individual firms adjusted, converted; new mutations evolved. Traditions were shaken. (These statements are derived from the body of quotations in the appendix.)

One of the most unsettling events in this whole uproar was undoubtedly the introduction of electronic data processing. Here are a few corresponding news items:

> "...There are approximately 100 to 125 commercial printers, publishers and trade typesetters utilizing the computer for composition functions."(4) October 1966.
>
> "...Fantastic growth of computerized typesetting installations: from 77 in 30 countries in 1964 to 1,093 worldwide in 1969."(5)
>
> "...installations jumped 54% in 1967 and 33% in 1969; in U.S.A. about 900 composition typesetting operations."(6)
>
> "...The first 20 RCA Video Comp., which went on line between 1967 and 1969 are capable in themselves of setting all of the books published in the U.S.A. during the full year."(7)
>
> "...There are now well over 1,000 computers being used for typesetting."(8) January 1970.

On the other side of the coin there has been much publicity and many unsupported claims. Many references to this kind of "information" are to be found in the appendix. Lamparter says for example: "...Superficial review articles are appearing with increasing frequency in both the business publications and in the popular press. Unfortunately this outpouring of information has too often been incomplete and oversimplified."(9)

Regarding the peculiar aspects of braille publishing, consider that traditionally manuscripts are submitted to the publishing industry to print books, and eventually some of these books get transcribed into braille -- relatively inefficient and at high cost. Lately, a few titles were keyboarded for computer processing of the text into Braille Grade II. The advantage of this method is that any of the many trained keyboarders can handle English without having to know the rules of Braille Grade II; these rules can be supplied by a computer later. However, the cost for this work was found to be excessive for the production of a relatively small number of braille copies per titles.

So, it was natural that people concerned with publishing braille got excited about the possibility of using existing tapes -- produced by large commerical printers for their own composition -- to generate braille.

The idea was to obtain compositor's tapes inexpensively after they have been used for commercial printing, where the commercial printer would absorb the high cost of keyboarding -- certainly an intriguing thought.

This thought has been pursued for many years and by different people. However, as far as we know, only in one place has compositor's tape been used to produce braille on a continuing basis, and even this operation has again dropped most of the automatic data processing in its operation.(10)

Why, one asks, has the idea of "generating braille from compositor's tape" not been more successful?

In looking for answers and in order to understand the attitudes of professionals in the field of printing and publishing in the U.S.A. as well as abroad, we searched numerous publications from government, industry and information centers (see Appendix); we asked for response to a questionnaire (sample in Appendix, p. III-I-68); we visited and interviewed directly a number of experts.

As a result we found much that is disputed, but general trends seem to emerge.

> We give here a selection of brief direct quotations taken from publications. (For more detailed and extensive listings the reader is referred to the appendix.)
>
> ..."We know that straight keyboarding of copy, regardless of the high speed system following, is handcuffed to manual operator speed."
>
> The Honorable J. L. Harrison, p. III-I-30
>
> ..."working to bring about a solution to the conflicting needs for speed in composition and desirable printing practice."
>
> J. F. Haley, p. III-I-31
>
> ..."Opinion regarding markets and production technology...based on few or no data."
>
> H. C. Lamparter, p. III-I-33
>
> ..."The application of computer controlled ultra high-speed typesetting to book production -- principally textbooks -- is an area of some controversy."
>
> ..."the high powered gear saves nothing on keyboarding or proofreading which, with page make-up, is at least 80% per page composition cost in book work."

..."to a machine cost of $6 per page one adds at least $8 - $9 for keyboarding and proofing."

..."page rate on a trade book goes from $6.50 - $8 via conventional methods to at least $12 - $15 on an ultra high-speed typesetter."

..."In-house data creation at the publishers-- its immediate advantages lie in control of data creation and production schedule."

Sedgwick, p. III-I-43,III-I-44

..."Complete automation of typesetting will probably remain a technologist's dream."

..."Any discussion of automatic typesetting seems to flounder at the input stage."

..."typesetting demands a preponderance of data preparation and relatively little computer processing."

..."a consensus of published experience suggests that on a single <u>column news (newspaper)</u> the improvement in productivity will be about 15 to 25 percent. A book printer employing wider measures should not expect much more than 7 to 10 percent."

..."40 to 50% of caseroom time (goes for)...corrections, page make up (takes) a good proportion of the remainder."

Wallis, p. III-I-48,III-I-50

..."Computerized composition has not yet made much of a dent in traditional typesetting of 'straight matter' such as book work."

..."What is required of the publishing industry is standarization of format."

Rice, p. III-I-57

..."Computer typesetting in U.K.: ...already a multiplicity of equipment, which pose problems to the printer who feels he ought to be in."

> ..."In most instances the basic reason has been (the hope) to reduce costs."
>
> ..."Most of the schemes have involved the retraining of linotype operators, many of whom have not yet reached their potential speed on the new keyboard."
>
> Pira, p. III-I-63,III-I-64

Results from the inquiry with our special questionnaire were as follows:

Question 1: Tape used: material(s), format(s), code(s).

Question 2: Systems used.

It is particularly striking -- and it has, of course, far-reaching consequences -- that each of the respondents uses different formats and systems. We therefore think it important enough to quote the corresponding answers verbatim.

1 6 channel punched paper, TTB code.
7 track 556 B.P.I. with BCD blocked in 1000.
as character records.
2 H 1200, 65K, 6 tapes + 2 discs.

1 perfected from 8 level paper tape by Frieden 2501 keyboards
9 track 800 BPI
variable records.
2 RCA Spectra 70-35 with Videocomp 822 using page 1 software.

1 6 channel TTS paper
special keyboard layout
unjustified
2 Hot metal, Electron mixer, automated via Line-a-Sec. for line measurement.

1 800 BPI mag. tape, EBCDIC code.
9 track
2 IBM 360/30 or 360/40 with ABC composition system.

1 6 level TTS paper tape and MT/ST (extended keyboard-k Pak).
We produce tape for composition and computer manipulation.
2 Photon 713-20, IBM MT/SC, AM725, Electrons, Harris fototronic CRT, IBM CRT (Alphanumeric).

1 1/2" Magnetic tape, EBCDIC code
7 track 556 B.P.I.
2 Originally designed/implemented SIGMA 2 equipment.

1 6 channel punched paper, converted to 1/2" magnetic tape
7 track film
2 Videocomp RCAbase system, Spector 70 computer; Poole Brothers' RCA system, Spector 70-35 computer.

Question 3: How clean are tapes? What is the error rate? Are all corrections carried back to data base?
3a: Are tapes complete?

Especially disturbing is the variety of responses to the items "error rate" and "clean-up" carried out on tape. We give therefore again a complete listing.

3 1 error per 1000 characters
3a We correct data after it has passed "raw" data stage. (clean?)

3 100% clean
3a Yes, 100% carried back; tapes are complete except for front and some back matter.

3 95% clean
3a Corrections are not carried back to data base; tapes incomplete.

3 over 90% correct; (clean)
3a Yes, corrections are--at present--brought back to data base; tapes complete but may be on multi-reels.

3 -----(no answer)
3a Yes, (back to data base); tapes complete.

<u>3</u> -----(no answer)

<u>3a</u> Yes; all corrected; complete in galley form.

<u>3</u> All paper, as well as magnetic tapes are corrected; "clean".

<u>3a</u> -----(not answered).

Question 4a: Reformat to specified code?

Some houses can and are willing to do so. Others could do this on magnetic tape only -- or, "Have no present program for this, but it could be done."

4b: Price per title? (reformatting to specified code)

"Would depend on amount of change. MT/ST cost per hour is $7.85." "Estimated cost on first basis $1,000"; others say: "program and machine time $200."

Question 5: How long are tapes kept after completion?

"Hold as directed by publisher, or (else) not kept at all."
"Some titles kept not at all, others 4 to 5 years, depending on subject."
"One year, thereafter asking for storage charge."
"If not otherwise specified, will be erased and reused as soon as safely past photographic plate making."

Question 6: Release of tapes (or copies) for braille production:

Only one respondent replied: "free of charge." Other responses vary from $15 to $200; or "Free--if costs for eliminating machine commands, etc. are recovered."

Question 7: Scope--titles per year--produced with compositor's tape.

	(Range of estimates by respondents).
Straight English	From 20 to 500 (average 325)
Scient. & College	From 40 to 150 (average 83)
Element. & High Sch.	From 6 to 50 (average 18)

Question 8: Projected usage of compositor's tape by 1980.

Straight English	From 50 to 100% (average 78%)
Scient. & College	From 25 to 80% (average 64%)
Element. & High Sch.	From 75 to 80% (average 77%)(11)

For first-hand information we have visited or contacted the following publishing houses, printing plants and computer composition companies:

Bantam Books, Inc., CompuScan. Inc., Computer Typesetting, "Black Dot", Dell Publishing Co., R. Donnelly & Sons, Follett Publishing Co., Hadden, W.F. Hall Printing Co., MacGraw Hill Book Co., The McMillan Company, Poole Bros., Prentice Hall, Rand MacNally & Company, Random House, Roccappi Computerized Composition, "Tape Type", Vermont Photo-Tape serv., Westcott & Thompson, and Western Publishing Co.

The Goss Co. in Chicago kindly allowed me to use their excellent library.

The interviews largely reinforced information obtained through the literature research and the questionnaires. The following observations are worth mentioning separately:

a. The research director of a large publisher said: "We use data systems' Keymatic Console (somewhat like MTST, but directly computer compatible output). One should do much clean up at this stage (before composition); keyboarding accounts for approximately 45% of total composition cost."(12)

Advantages of automatic composition are: "The resulting print page is superior to manual composition; furthermore, there is fast turnaround. --Automated composition is mostly attractive for material which is subject to change; but better product and reduction of printing cost (because of optimal page utilization) make straight matter also competitive." -- "newspapers are switching to offset; this causes the hot metal technology to wither, because of loss of their biggest market."

b. A compositor: ..."expects to see hot metal technique to disappear in a few years, to be replaced by optical (film) composition. But tape availability is nevertheless not assured: Only ∿ 5% of the text are complete, fully corrected tapes and are retained (for future printing, etc.). Most tapes are kept only a short time." And we may add: often as a jumble of short sections of paper tape which are thrown in cartons without organization. Many corrections are made by hand on the final (film) copy and therefore never get back to the tape!

"Tapes are not necessarily complete; they may for one book be made by several compositors; sections of one book may be done directly in hot metal, whereas others are done with tape, etc. Many different systems are in use and publishers use different codes, thus making some available tapes incompatible with others."

c. A very large publisher told us that: "straight matter (novels) is usually printed only once";

that: "They use MTST and convert its output into computer readable form".

"...perhaps as much as half of our compositions for books of general interest are presently being composed from tape".

(however) ..."At the present time, we have very little material available on which a 'perfect' tape exists."

..."Unfortunately, most if not all of the conventional composing systems and the earlier photo-composing systems do not result in a correct 'perfect' tape. In the case of all of these systems, the keyboard tape is used to create galley proofs of hot-metal type, or photographically-created type images. All identified errors are then corrected directly in metal type or by stripping corrected copy on the film output. It is only in the case of the more advanced computer composition systems that a 'perfect' tape, embodying all of the corrections, is created."

"permission to use a compositor's tape would not only have to be obtained from the publisher but also from the author."

d. The president of a large compositor: he confirms that ..."at present composition of 'straight English text' is rarely produced by computer because so far it is more expensive." (Do you expect that to change?) "Definitely!" (How will it become more competitive?) "Prices will come down with volume." (You mean if a publisher contracts with you for 100 book titles instead of 10, you would lower your price drastically?) "Absolutely!" (When do you expect that automated composition will take 100% of this (straight matter) market?) "Never; about 50% in 10 years." "As far as the next six months are concerned it is difficult to predict what we may produce."

Conclusions

At the beginning of this paper we asked the question whether braille production can be revolutionized by utilization of compositor's tape.

Techniques for producing braille in this way do work. But, what does this do to enhance--in Professor Mann's words--the availability of braille? Unfortunately, the answer to this question is "at present next to nothing."

Relatively few titles get on tape;
relatively few of them get on tape completely;
not many of these are "clean" of errors;
and not all of these are preserved for any length of time after having served in the printing operation;
and the final residue which is useful for the braille producer is formatted in numerous, different and noncompatible ways.

So, the braille producer has to pick out a few grains of wheat from what is tons of chaff to him. Furthermore, offerings of tapes by the industry are concentrated in a few special fields, such as textbooks, dictionaries, etc., as a consequence of economic factors. At least this is so at present, and it seems likely that these factors will continue to operate in the foreseeable future.

The resulting effort in finding usable material is bound to discourage the braille producer.

Furthermore, because of the low volume of braille demand the braille producer is hardly ever specialized with respect to type of literature; thus, even if compositor's tape were relatively more abundant for some limited part of his production, he would not seem to be much better off: To take advantage of this opportunity would make life for him more complex and difficult and he may well wonder whether the results are worth all that much.

What can be said about the future?

Generally our search has not identified causes for drastic change from the present situation; some agreement seems to exist that more tapes and better tapes will gradually become more available--but hardly anything like complete conversion is predicted by the experts. As a matter of fact there can be seen some factors at work which may counteract such a development, for instance, the proliferation of competing and not compatible equipment. It would seem, then, that the future does not hold any revolutionary promise for compositor's tape utilization by braille producers.

If and when much--or all?--information (including novels, poetry, etc.) will be stored in machines, one would of course not have to worry about sources for machine readable input any more; braille would simply become one of many possible outputs of "the machine," and complete availability would result.

If and when......

---

(1) Robert W. Mann, "Enhancing the Availability of Braille." American Foundation for the Blind, 1963, p. 413.

(2) Stanley Rice, "What's the Future of 'Straight Matter'?" Appendix p. III-I-57.

(3) The World Almanac and Book of Facts. New York World Telegram and Sun, 1966, p. 512.

(4) William Lamparter, Economist at "Battelle" (worldwide 'contract' research), "Impact of Electronic Composition on Commercial Printing and Appendix, p. III-I-34.

(5) Spencer A. Tucker, Consultant Director, Composition Information Service, Los Angeles. Printing Magazine, March 1970, p. 37.

(6) Ibid.

(7) Ibid.

(8) McLean Hunter, Computer Information Center, Inc., Los Angeles, "Survey of Computerized Typesetting," British Printer, January 1970, p.16.

(9) William Lamparter. Appendix p. III-I-34.

(10) The joint venture of the University of Muenster, with a German magazine publisher and a social agency for the blind, does issue regularly braille versions of several magazines.

(11) Several of the respondents stressed the speculative nature of their own estimates. Furthermore, the most optimistic estimates came in several cases from respondents who had only recently started automated typesetting. In the light of the literature quoted, it is our opinion that the estimates given here, especially for straight matter, are far too optimistic.

(12) Editors' intervention concerns itself mostly with formalism (not much with meaning)--thus unimportant for braille use even if missing in available tape.

A paper on using compositor's tape is braille production by Grete Grunwald in _The Argonne Braille Project_, Argonne National Laboratory, 1977

The following articles (not included here) are appended by Mrs. Grunwald in support of her conclusions:

1. Stevens, M. E., and Little, J. L., "Automatic Type Composition", National Bureau of Standards
2. Proceedings of a Symposiom, "Electronic Composition in Printing", National Bureau of Standards
3. Mollman, Peter, "Programming CRT'S First General Book", _Printing Magazine/National Lithographer_
4. Sedgwick, Henry, "CRT for Straight Matter", _Printing Magazine/National Lithographer_
5. "McCall and The Computer", _Modern Lithography_
6. Wallis, L., "Computer Setting Methods, Cost and Performance", _British Printer_
7. King, Howard, "What We Learned in 16 Years of Photocomposing", _Printing Trades Journal_
8. Rice, Stanley, "Computerized Composition", _Modern Lithography_
9. Griffin, R. F., "Don't Let Us Kid Ourselves About Computers", _Printing Equipment and Materials_
10. Cooper, D. L., and Nield, C. D., "An Analysis of Computer Typesetting Systems in the UK", _Pira Report_
11. "Eastern European Printing Equipment", _Printing Equipment and Materials_
12. "Study Predicts Phototypesetting to Rise 300 Percent in Next Five Years", _Computer World_

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

A DOUBLE-SIDED BRAILLE EMBOSSER

by

Knut Grimnes

Foundation of Scientific Industrial Research
University of Trondheim
Aud 48
N-7034 Trondheim

A DOUBLE-SIDED BRAILLE EMBOSSER

K. Grimnes

SINTEF,*) Division of Automatic Control

Summary

The embosser is developed at SINTEF and the prototype is used in braille production at Tambartun School. The prototype embosser is producing one sheet of paper, or two pages of text, in 20 secs. The production model will at least work at double speed, i.e. one sheet in 10 secs. The embosser must be used in conjunction with a text-processing system.

1. A BRIEF HISTORY

The first thoughts on the embosser emerged during a seminar on braille printing in 1969. At that time, SINTEF had been engaged in the development of computerbased typesetting systems. Such systems seemed to be very suitable for braille production also and a system, including the present embosser, was outlined. One of the problems, at that time, was that the Norwegian grade II braille was very complicated to translate by a computer. The rules were, however, to be revised and further work was postponed while waiting for this revision. The plans were then more or less forgotten.

In the autumn of 1976, SINTEF was contacted by Tambartun School, a centre of education for the visually impaired. This school has the responsibility for producing the necessary textbooks for blind children at elementary shool level. According to a new Norwegian law, handicapped children should - if at all possible - be allowed to attend ordinary schools. As these "integrated" pupils needed to have the same textbooks available to them as the other children had, this new law resulted in a sudden increase in the demand for new textbooks in braille, a demand that far exceeded the capacity of Tambartun.

---

*) The foundation of scientific and industrial research at the University of Trondheim.

They planned, therefore, to install a computer-based, text-processing system, but they were in need of an output embosser able to emboss on both sides of the paper. So, we went back to work. The embosser was designed and built, originally as a test model. It appeared to work surprisingly well and, after some necessary modifications, ordinary book production at Tambartun started at the beginning of 1978. The embosser worked flawlessly for some months but then a problem in one part of the mechanism started occuring with increasing frequency. This part of the mechanism was, therefore, modified and since then the embosser has worked excellently.

From the time it began work until the writing of this paper, the embosser has produced approximately 14000 sheets of paper or 28000 pages of text. It is no longer the embossing that is causing the bottleneck at Tambartun.

2. THE OUTPUT EMBOSSER

The basic principle of the embosser is shown schematically in fig. 1. The paper is fed stepwise between two oscillating beams. The two beams have a row of embossing pins arranged as shown in fig. 2 as full drawn circles. The pin spacing is equal to the point spacing in one horizontal row of points, e.g. all points 1 and 4 of all the characters in one line. The embossing pins in the upper beam match recesses in the lower beam and vice versa. The recesses are shown as dotted circles in fig. 2. The embossing pins in the two beams are positioned relative to each other so that the correct interpoint positions are secured.

The printing cycle goes as follows: when the paper stops in the correct position to have one row of points on each side embossed, the two beams will close on the paper and hold it firmly. The appropriate embossing pins are then actuated to emboss the points. When the embossing pins are withdrawn from the paper, the two beams open to allow the paper to be transported forward to the next row to be embossed. Three cycles are thus necessary to emboss one complete line on each side of the paper, or 80 characters, assuming a line length of 40 characters.

The embossing pins are actuated through levers mounted on bearings in the beams, as shown in fig. 3. The beams are divided into two main parts: the lever part, where the levers are mounted, and the pressure part, which is the part that presses upon the paper during the embossing phase of the cycle.

The beams are driven up and down by means of an eccenter mechanism. The pressure part is springloaded to a position about 2 mm in front of the lever part and, as the beams are closing on the paper, the pressure part of the two beams will meet and hold the paper firmly while the lever parts will move a further 4 mm closer to each other. It is in this part of the cycle that the embossing takes place.

The actuating levers are mounted such that one end is directly above the embossing pins, while the other end protrudes from the beam, as shown in fig. 3. If the protruding end of the lever is free to move up and down, nothing happens and no point will be embossed. If, however, the downward movement is restricted, the lever will push the embossing pin into the paper and emboss a point. The downward movement of the lever can be restricted by a locking pin, also shown in fig. 3. This locking pin is actuated by a magnet, and as the force needed to move the locking pin is very moderate, it is possible to use relatively small actuating magnets. Nevertheless, in order to get enough room for the magnets, every second lever is brought out on each side of the beam and every second magnet is placed above and below the locking pins. In this way, we have 12 mm available for each magnet.

The force on the locking pin is carried by a prestressed spring, so that when the embossing force exceeds approximately 20 N, (2 kg), the spring will start to bend and limit the maximum force.

The paper used is manufactured by a Swedish paper mill especially for braille embossing. Other paper qualities may certainly be used but they have not been tested. In the prototype embosser, the paper has the same dimensions as one of the standards used for paper in computer peripheral printers, as shown in fig. 4. The paper is fed by a sprocket belt driven by a stepping motor. Each step represents approximately 0.5 mm, (more exactly: 0.496 mm). The distance between points within a character is 2.5 mm, i.e. 5 steps. The distance between lines may be chosen at will to the nearest 0.496 mm. 21 steps will, for instance, give a line spacing of 10.4 mm = 410".

The paper-feed motor and the printing magnets are controlled from a dedicated microprocessor. This processor receives signals from the paper-feed mechanism in order to synchronise to the top of each page. It also receives a signal from the embossing beams' eccenter drive, and this is used to syncronise the printing magnet drive to the printing cycle.

The processor also takes care of the communication to the text-processing computer. Text is transferred two pages at a time from the text-processing system to a text-buffer in the embosser microprocessor memory. The microprocessor converts the ASCII characters to braille, and disassembles this to the rows of points to be embossed. As the two pages are embossed, the next two pages are simultaneously read into the microprocessor.

The prototype embosser at Tambartun operates at a speed of 10 cycles/sec, but it has worked up to 20 cycles/sec during a test run. In the prototype printer at Tambartun the beams are not divided in the two parts described above. Because of this it is only possible to emboss one side at a time. The paper is therefore kept stationary for two cycles in order to emboss both sides. The embossing time for one page (29 lines) is thus approximately 20 sec. on this prototype embosser. Conservatively, we estimate the production model also to operate at 10 cycles/sec. This means one sheet, or two pages of text in approximately 10 secs, allowing for some extra time for pageshifting. The maximum line length will be 42 characters but, assuming a line length of 40 characters, (80 characters on both sides), the minimum embossing speed will be 266 characters/sec. With some experience, we hope to increase this speed to at least 500 ch/sec.

The embosser will be manufactured by the Norwegian firm "Egil Railo Verktøy-industri". At the time of writing, the firm is working to finance this production and hope to be ready to start production in April 1979.

3. THE TEXT PRODUCTION SYSTEM

As the embosser prints on both sides simultaneously, it is necessary to have ready one full page and at least the first line on the next page, before the printout starts. It is, therefore, necessary to use the embosser in conjunction with a text-processing system. There is a multitude of such systems available on the market and the embosser may be interfaced to any of these.

In the Tambartun system, a rather advanced text-processing system, from the Norwegian firm "Norsk Data", is used. The main part of this is an unmodified text-processing system as used in newspaper and printing offices. Added to this standard system are some special features:

- automatic translation according to the unambiguous parts of the braille rules,
- presentation of "visual braille" on the screen, or at the output printer,
- communication with the embosser.

The production of a book is divided into three main steps:

- text input
- text processing
- printout and binding

Input text may come from one of the computer terminal keyboards, or through paper tape. Paper tape may again be punched on a freestanding punch with ordinary or braille keyboard. More interesting, however, is the possibility to get a computer-readable copy (e.g. paper tape copy) of the text used in production of the inkprint book from a printing office. In this case, the text is loaded into the text processing system without any need to rewrite it. Other input methods may, of course, be used, as, for instance, OCR or magnetic tape but this is not yet used in the Tambartun system.

In the text processing phase, the text is converted to a form suitable for printout in braille. Some of the text must be rewritten, e.g. text referring to pictures, and some must be modified to braille rules as, for instance, the addition of symbols for capital letters. The unambiguous parts of this conversion are done automatically in the computer and the manual parts are done interactively on one of the computer terminals. At this stage, a proof copy is made either in inkprint or in braille on the embosser. Errors found by the proof reader are again corrected interactively at the terminals and the corrected text is stored in the computer memory, ready to be embossed.

The book may now be embossed in the required numbers. One complete book at a time is printed out, on Z-fold paper. The book comes from the embosser as a stack of papers with feed holes in both edges. One edge is glued together with the cover, while the other three sides of the stack are cut clean. The feed holes are thus kept in the back of the book where they act as a kind of hinge to facilitate reading of the book.

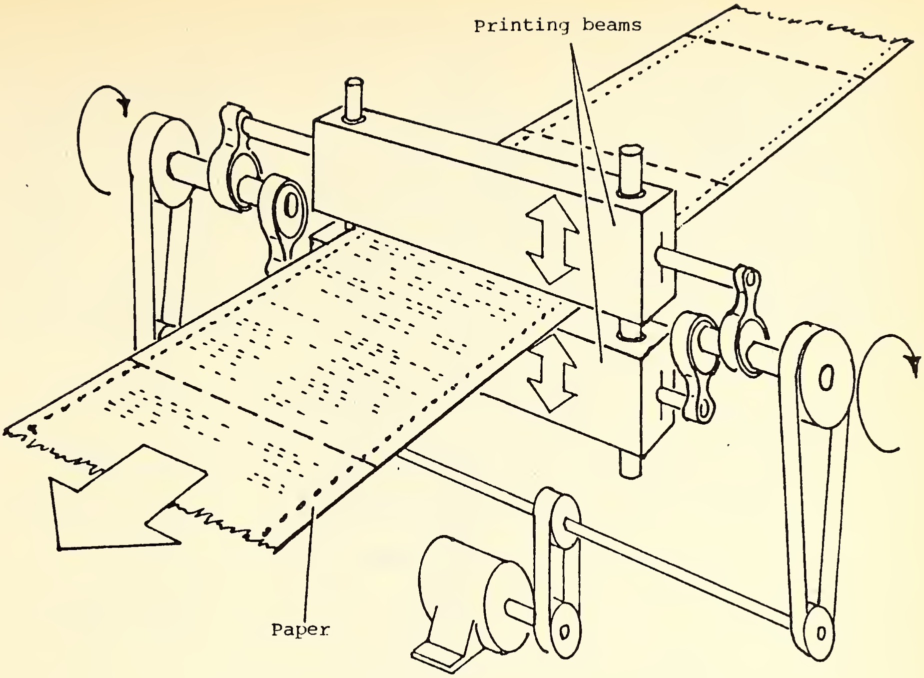


Fig. 1. Principle of the embosser.

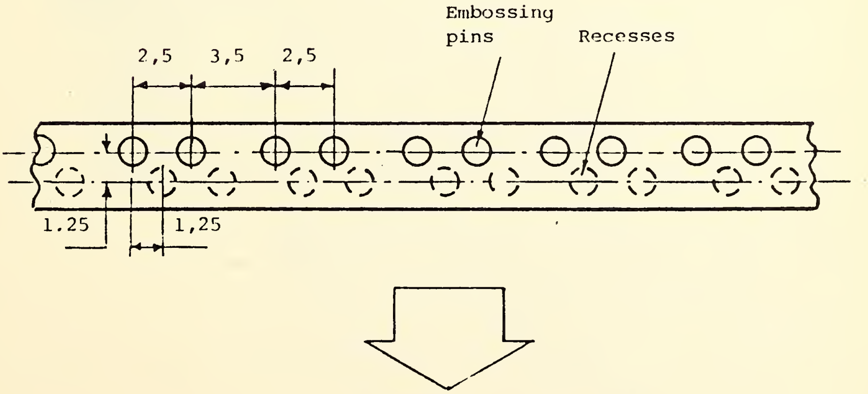


Fig. 2. Embossing pins.

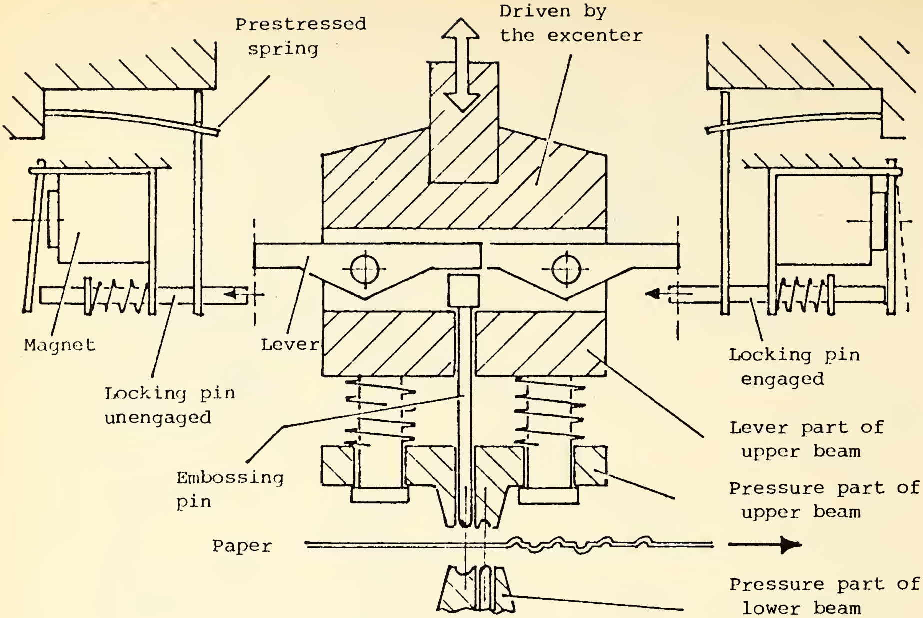


Fig. 3. The embossing mechanism.

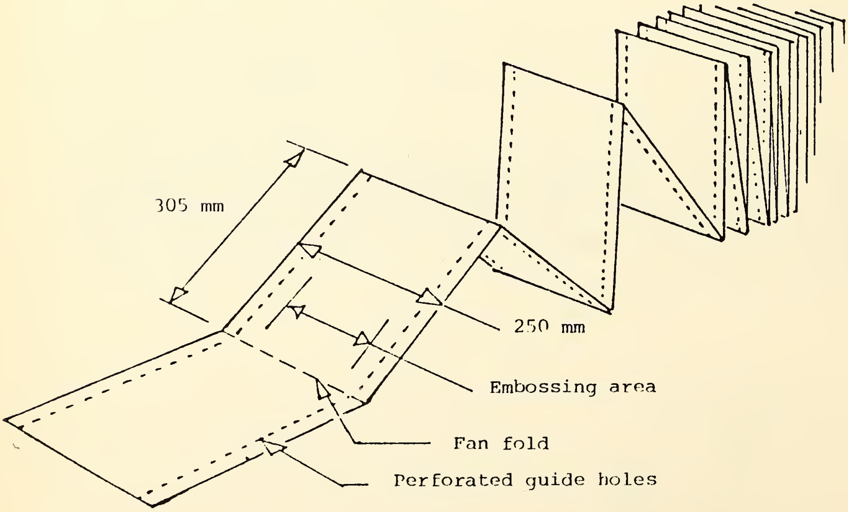


Fig. 4. Paper dimensions.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

PORTUGUESE BRAILLE BY COMPUTER

by
Ricardo J. Machado

IBM of Brazil
Rua do Ouvidor 107
Caixa Postal 2106
Rio de Janeiro-RJ-20.000

PORTUGUESE BRAILLE BY COMPUTER

Machado, Ricardo J. - IBM
Facco, Edno - FLCB
Nunes, Marcus V.V. - IBM

ABSTRACT

The current process for braille book production presents difficulties such as the high cost for training stereotypists, slowness of translation and frequent errors, necessity of detailed revision and the complexity of metal plates correction made manually dot by dot.

To solve these problems the "Fundação para o Livro do Cego no Brasil" jointly with IBM is developing a computer system to produce automatically grade two braille books in portuguese language.

The system is formed by 4 modules:

1. Portuguese to Grade Two Braille Translator

   This module reads the text punched in cards, translates it word by word recording the output in a disk file, does abbreviations and contractions, detects and reports several types of errors, prints an image of the source text for visual verification.

2. The Pre-composer module composes the text and prints a previous image of the braille book allowing to verify the adequacy of the format. Several types of format inconsistencies are detected and reported.

3. Correction Module

   Punching errors as well as format inadequacy may be detected analysing the reports produced by modules 1 and 2.

   Correction commands for insertion, substitution or deletion on the text are read, interpreted and recorded in an error disk file.

4. Page Composer Module

This module composes the text in its final form obeying the desired dimensions for the page as well as the format control commands introduced in the source text.

It uses an inputs the Braille and Error Disk Files, executes the correction commands, makes hyphenation for portuguese words, margination if required, automatic page numbering, and special processing for titles and indexes.

The following benefits are expected with the implementation of the system.

- To decrease the time and effort spent in the production of a new braille book.
- To diminish the number of errors and to make correction easy.
- To increase the efficiency of stereotype machines.
- To increase significantly the number of titles of braille books in portuguese language available to the community.

1. INTRODUCTION

The grade 2 braille code for the portuguese language has been established by an agreement between Brasil and Portugal in 1963.

In Brazil, the production of braille books is made by "Fundação para o Livro do Cego no Brasil" (FLCB), a public utility entity that offers its service to the brazilian community without profit objectives. It delivers 30000 volumes yearly and has produced since 1949 563 titles of braille books in portuguese language.

However the current process for braille books production presents difficulties such as the high cost for training stereotypists, slowness of translation and frequent errors, necessity of detailed revision and the complexity of metal plates correction made manually dot by dot.

As a consequence we have long periods of time to produce new titles or to update them, and an insufficient number of titles available to the community affecting negatively the education of brazilian blinds.

To solve these problems the "Fundação para o Livro do Cego no Brasil" jointly with IBM of Brazil have developed a computerised system to produce automatically grade two braille books in portuguese language.

The system is formed by four modules:

1. Portuguese to grade two braille translator.
2. Pre-composer module
3. Correction module.
4. Page composer module.

2. SYSTEM DESCRIPTION

The process of production of a braille book by this system can be divided in 8 phases showed in figures 1A and 1B and presented next:

2.1 PREPARATION

Firstly the book to be produced must be prepared by the editor that will analyse it and introduce some command symbols to the computer programs. (Figure 1A phase 1).

The format characteristcs of the final braille book are largely determined by these commands. As examples of commands available to the editor we have: compulsory end of page, compulsory end of line, paragraph, beginning of title, end of title, beginning of index, end of index, information of source book page number, margin size, non-contraction command, blank page, compulsory space, end of text, low case signal, upper case signal.

The editor may optionally divide the book in several segments to be translated separately and linked later during the composition phase.

2.2 PUNCHING PHASE

The prepared book is punched like a string in 80 collumns cards. In this phase all the format characteristics of the original book are lost. (Figure 1A phase 2).

Considering the accentuation of characters in the portuguese language as well as the command symbols of the system we may have 104 different characters that exceeds very much the capacity of Keyboards in conventional data entry equipment.

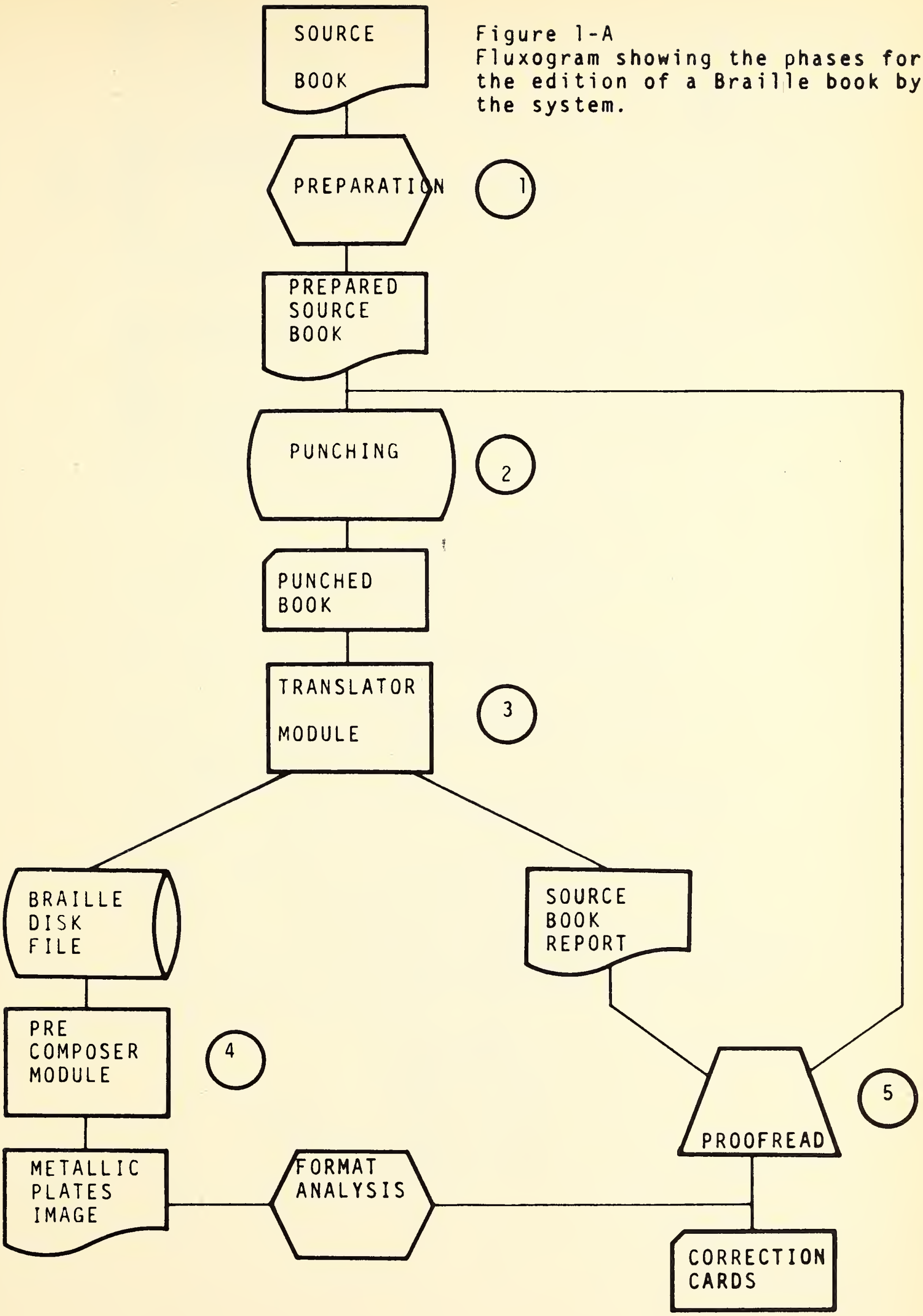


Figure 1-A
Fluxogram showing the phases for the edition of a Braille book by the system.

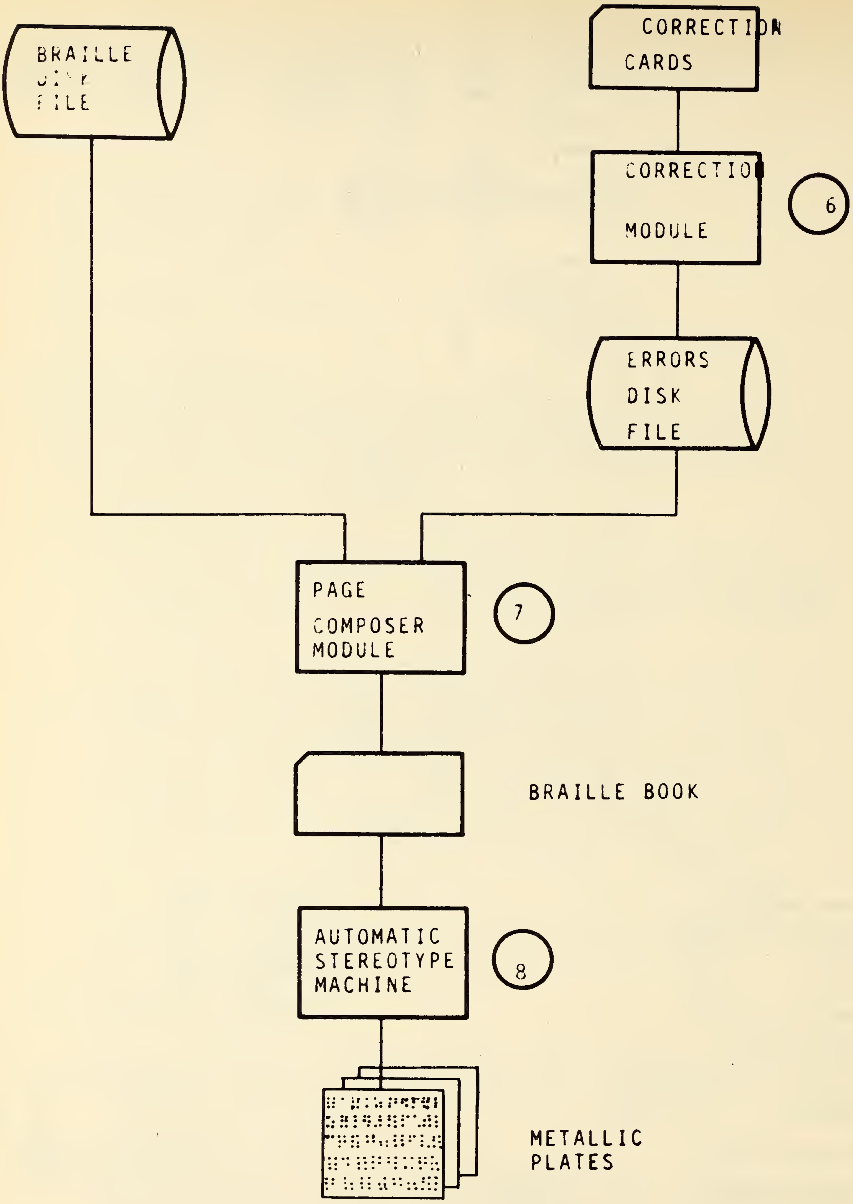

BRAILLE DISK FILE
CORRECTION CARDS
CORRECTION MODULE
6
ERRORS DISK FILE
PAGE COMPOSER MODULE
7
BRAILLE BOOK
AUTOMATIC STEREOTYPE MACHINE
8
METALLIC PLATES

Figure 1-B

To solve this problem we use compound symbols with the format:

¬XY

where XY is a mnemonic representing the symbol in consideration.

These compound symbols will be detected by the translator module and converted to the EBCDIC code.

2.3 TRANSLATION

This phase is executed by the Portuguese to Grade Two Braille Translator (Figure 1A phase 3).

This module reads the text punched in cards, translates it word by word recording the output in the Braille Disk File, does abbreviations and contractions, detects and reports several types of errors, prints an image of the source text for visual verification as shown in figure 2.

A detailed description of the algorithm employed in the translator module is presented in the Appendix A.

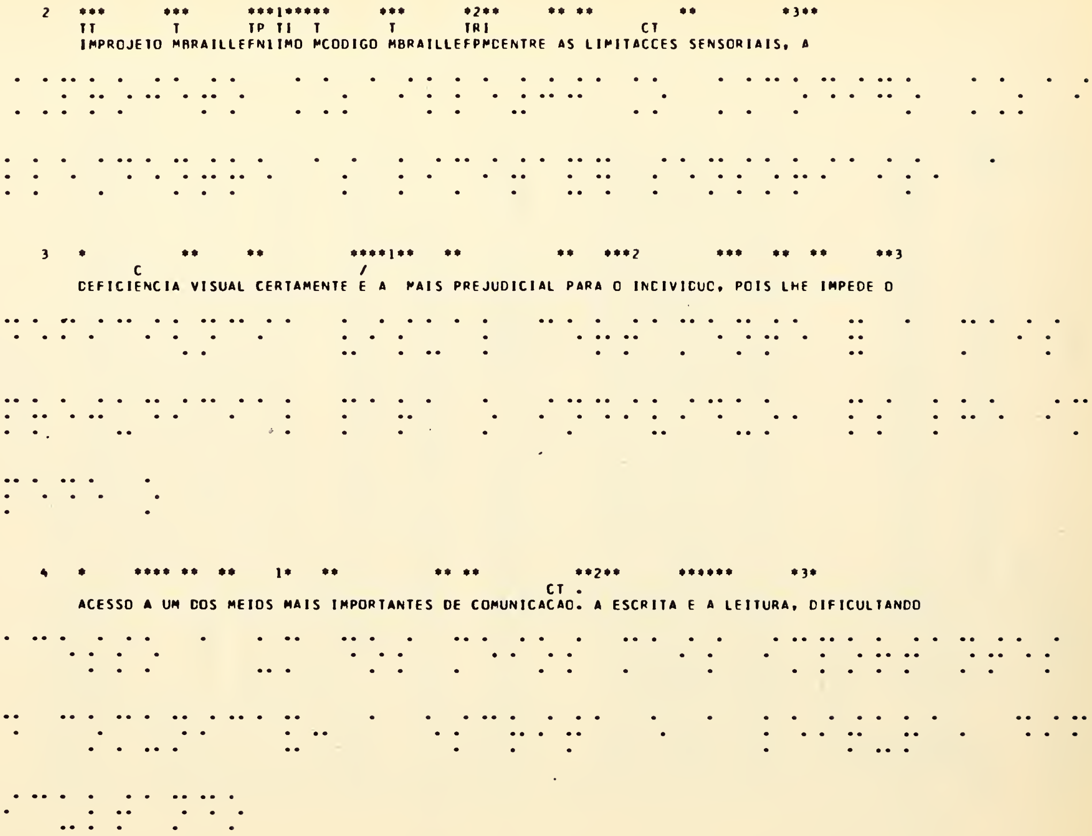
2.4 PRE-COMPOSITION PHASE (figure 1A phase 4)

The pre-composer module composes the text and prints a previous image of the braille book shown in figure 3.

It allows to the editor to verify visually if the book fits the desired format and to determine which adjustments must be made.

Several types of format inconsistencies are automatically detected and reported by the program.

To make easy the analysis, the text to be printed is reconverted and printed in the form of alphabetic characters avoiding the need of braille knowledge by the editor.



```
2   ***      ***      ***1*****     ***      *2**     ** **        **         *3**
    TT        T       TP TI  T       T       TRI               CT
    IMPROJETO MBRAILLEFN1IMO MCODIGO MBRAILLEFPMCENTRE AS LIMITACCES SENSORIAIS, A

3   *          **     **         ****1**   **          **   ***2        ***  **  **     **3
          C                       /
    CEFICIENCIA VISUAL CERTAMENTE E A  MAIS PREJUDICIAL PARA O INCIVICUC, POIS LHE IMPEDE O

4   *     **** **  **   1*    **         ** **          **2**      ******      *3*
                                                     CT .
    ACESSO A UM DOS MEIOS MAIS IMPORTANTES DE COMUNICACAO. A ESCRITA E A LEITURA, DIFICULTANDO
```

Figure 2 - Source book report produced by the translator module. The source text and the braille representation corresponding to each record of the Braille Disk File are shown.

The pre-composer and page composer modules employ the same algorithm so that the report of figure 3 is an exact image of the pages of the final braille book.

2.5 PROOFREAD AND CORRECTION

Errors made during the punching phase may be detected comparing visually the source book with the translator report (Figure 1A phase 5). Format distortions are detected studying the pre-composer report.

Correction commands for insertion, substituion or deletion of words on the Braille Disk File are punched in cards and processed by the correction module. (Figure 1B phase 6)

The general format for the correction command cards is:

T, L = x, P=y, (text)

where T = type of command (I for insertion, S for substitution, D for deletion)

x = number of translator report line which contains the wrong word.

y = position of the wrong word within the line.

text = new text to be inserted or substituted in the Braille Disk File.

The position must be indicated by the ordering of the wrong word within the line, since there is no correspondence, character to character, between the source text and the grade 2 braille text.

2.6 COMPOSITION PHASE

The processing of the book is completed by the page composer module. (Figure 1B phase 7).

This module composes the text in its definitive format obeying the desired dimensions for the page as well as the format control commands introduced by the editor in the source text during the preparation phase.

It uses as inputs the Braille and Error Disk Files, executes the correction commands contained in the Error Disk File, makes hyphenation for portuguese words, margination if required, automatic page numbering, generation of headline for odd pages consisting of the number of the page in the source and braille texts as well as of a book identification.

Special processing for titles and indexes is executed. Linkage among different segments of a same book is provided if necessary.

The output of this module is a multipunched card deck containing the composed braille book. The cards contain braille cells represented in the heights 1 to 6 of the collumn, spaces represented by blank collumns, end of line command represented by a hole in height 7 and end of page command represent by a hole in height 8.

This deck of cards will be read later by a card reader coupled with a stereotype machine that will automatically emboss the metallic plates for the braille book press (Figure 1B phase 8).

3. SYSTEM CHARACTERISTICS

The system was developed for an IBM-1130 computer and was programmed in FORTRAN and ASSEMBLER.

This choice has been made considering that FCLB intends to establish agreements with brazilian universities to produce braille books, and the IBM-1130 is the most dissemined computer in these universities.

The sublevel 3 of grade 2 braille for portuguese language was implemented. We have measured the following processing times for a 100 pages book:

Translation:	2 hours 33 minutes
Pre-composition:	1 hour 40 minutes
Correction:	0 hours 58 minutes
Page-composition:	1 hour 30 minutes

4. BENEFITS

The following benefits are expected with the implementation of the system:

- Decrease the time and effort necessary to the production of a new braille book.
- Avoid the need of braille expertise to operate the system.
- Diminish the number of errors and to make correction easy.

- Eliminate the need of correction on the metallic plates.
- Increase the efficiency of the stereotype machines.
- Increase significantly the number of titles of braille books in portuguese language available to the community contributing by this way to a best education for the blinds.

5. BIBLIOGRAPHY

- Sistema Braille Grau 2 da Lĩngua Portuguesa.
  Ministẽrio da Educação e Cultura, 1970.
- Gramãtica Normativa da Lĩngua Portuguesa
  Lima, Rocha
  Livraria Josẽ Olympio, 18ª. edição, 1976
- Fundamental Algorithms
  Knuth, D.E.
  Addison Wesley, 1972
- SELBRAC - Operation and System Manual
  Machado, R.J.
  To be published by IBM of Brazil

APPENDIX A - PORTUGUESE TO GRADE TWO BRAILLE TRANSLATION ALGORITHM

A.1 INPUT AND OUTPUT DATA

The algorithm receives as input data a string of characters representing the source book after the phase of preparation and punching.

The algorithm produces as outputs the Braille Disk File and the Source Text Report. (Figures 1A and 2).

The Braille Disk File is a sequential disk file with records of fixed size containing the unformatted braille book as a string of characters (one per byte). Bytes with the 7th bit "off" contain braille cells whose dots are represented in the first 6 bits. Bytes with 7th bit "on" represent format commands. The 8th bit is used to indicate if the character in consideration occupies the last position of a syllable in the word.

Before every word there is a byte containing a symbol called "Beginning of Word Indicator". This information jointly with the syllable indication are useful for the composition modules. To facilitate the correction process, words are not broken between consecutive records.

The source text corresponding to the content of each record of the Braille Disk File is printed in the Source Book Report (Figure 2).

Three lines are used for each record allowing to represent all the 104 types of characters and to indicate the beginning of each word.

A.2 THE WORD CONCEPT

The translation is done word by word. The words are identified sequentially from the source book and aggregated to the Braille Disk File.

We understand by "word" a group of consecutive characters logically related. The concept of word is fundamental and we can classify the words in the following classes:

1. Alphabetic character
2. Numeric
3. Special
4. Blank
5. Upper case signal
6. Non contraction signal
7. Source book page number
8. End of text
9. Unidentified character
10. Garbage
11. Alphabetic word

A.3 TRANSLATION AUTOMATON

One word dependind on its class may influence the processing of the following word. To choose adequately the processing for a specific word a finite state sequential machine is employed. Its output and next state functions are shown in the tables of figure**3** .

The automaton depending on the word class and its state determines its next state and chooses one of 12 processing options presented in the translation algorithm description (A.5).

A.4 AUXILIARY ALGORITHMS

The main subroutines that can be called by the translation algorithm are:

		ALPHABETIC CHARACTER	NUMERIC	SPECIAL	BLANK	UPPER CASE SIGNAL	NON CONTRACTION SIGNAL	SOURCE BOOK PAGE NUMBER	END OF TEXT	UNINDENTIFIED CHARACTER	GARBAGE	ALPHABETIC WORD	
		1	2	3	4	5	6	7	8	9	10	11	STATES
OUTPUT TABLE	1	2	5	2	2	10	2	2	3	9	7	1	NORMAL
	2	2	5	2	2	10	8	8	3	9	7	2	NON CONTRACTION
	3	2	5	2	4	10	2	2	3	9	7	1	BLANK
	4	8	6	8	4	8	8	4	3	9	7	8	SOURCE BOOK PAGE NUMBER
	5	11	12	12	12	12	2	12	12	9	7	1	UPPER CASE
		1	2	3	4	5	6	7	8	9	10	11	
NEXT STATE TABLE	1	1	1	1	3	5	2	4	1	1	1	1	NORMAL
	2	1	1	2	1	5	2	1	1	1	1	1	NON CONTRACTION
	3	1	1	1	3	5	2	4	1	1	1	1	BLANK
	4	1	1	1	4	1	1	4	1	1	1	1	SOURCE BOOK PAGE NUMBER
	5	1	1	1	1	1	2	1	1	1	1	1	UPPER CASE

Figure 3 - Output and next state tables of the translator algorithm automaton

- WORDIDENTIFICATION (WORD, BRAILLE, CLASS, IDENTIFICATION STATE, CHARACTER UNDER ANALYSIS).

  The objective of this algorithm is to identify from the input text which is the next word, to supply its grade 1 braille translation obtained by transliteration and its class.

- HYPHENATION (WORD, SYLLABLES)

  The objective is to determine the position of the ends of each syllable in a portuguese language word.

  The algorithm is based on the use of acyclic directed graphs and allows to obtain a success index of 99.77% having efficiency proportional to the word size.

- ABBREVIATION (BRAILLE)

  The objective is to search if the word in consideration belongs to an abbreviation dictionary structured as a tree and in positive case to determine which is its abbreviation.

- CONTRACTION (BRAILLE, SYLLABLES)

  This algorithm verifies if there are segments of the word that can be contracted according to the grade two braille restrictions.

  The contractions are done only between the limits of one syllable.

- CONTRACTIONSEARCH (BRAILLE, BEGINNING, DIRECTION, LIMIT, CONTRACTION, VALIDITY PLACE)

  This algorithm analyses the word departing from a specified position (BEGINNING) following a specified direction (DIRECTION) searching this segment against a contraction dictionary structured as a tree. If there is a match the contraction and its validity place are delivered.

- OUTPUT (WORD, BRAILLE)

  The source word is printed in the Source Book Report (Figure 2) and the braille word preceded by a Beginning of Word Indicator are recorded in the Braille Disk File.

A.5 TRANSLATION ALGORITHM

```
A:  BEGINNING PROCEDURE

B:  WORDIDENTIFICATION (WORD, BRAILLE, CLASS, IDEN-
    TIFICATION STATE, CHARACTER UNDER ANALYSIS)

C:  DECISION←OUTPUT TABLE (STATE, CLASS)

D:  STATE←NEXT STATE TABLE (STATE, CLASS)

E:  GO TO DECISION

1:  COMMENT: COMPACT THE WORD FOR GRADE TWO BRAILLE
    ABBREVIATION (BRAILLE)
    IF BRAILLE WAS ABBREVIATED
    THEN MARK THE LAST CHARACTER
         GO TO F
    ELSE HYPHENATION (WORD, SYLLABLES)

F:  OUTPUT (WORD, BRAILLE)
         GO TO B

2:  COMMENT: EMPLOY THE INTEGRAL BRAILLE TRANSLATION
    OUTPUT (WORD, BRAILLE)
    GO TO B

3:  COMMENT: END OF TEXT
    OUTPUT (WORD, BRAILLE)
    CLOSE FILES
    STOP

4:  COMMENT: ELIMINATE THE CURRENT WORD
    OUTPUT (WORD, EMPTY)
    GO TO B

5:  COMMENT: NUMERIC WORD
    BRAILLE←NUMBER SIGNAL, BRAILLE
    OUTPUT(WORD, BRAILLE)
    GO TO B

6:  COMMENT: FORMAT THE SOURCE BOOK PAGE NUMBER
    TRANSFORM THE SOURCE BOOK PAGE NUMBER CONTAINED
    IN BRAILLE TO STANDARD FORMAT
    OUTPUT (WORD, BRAILLE)
    GO TO B
```

```
7:   PRINT THE MESSAGE: ERROR DURING
     WORD IDENTIFICATION
     OUTPUT (WORD, BRAILLE)
     GO TO B

8:   PRINT THE MESSAGE: LOGIC INCONSISTENCY
     OUTPUT (WORD, BRAILLE)
     GO TO B

9:   PRINT THE MESSAGE: UNINDENTIFIED CHARACTER
     OUTPUT (WORD, BRAILLE)
     GO TO B

10:  GO TO B

11:  OUTPUT (WORD, LOW CASE SIGNAL)
     GO TO B

12:  PRINT MESSAGE: ERROR DUE TO AN UPPER CASE SIGNAL
                    FOLLOWED BY A NOT ALPHABETIC WORD
     GO TO B
```

A.6 WORD IDENTIFICATION ALGORITHM

The characters are drawn sequentially from the source book one each time. The character class is determined and serves as input to a finite state sequential machine whose output and next state functions are presented in figure4 .

Depending on the character class and the automaton state, the character in consideration is aggregated to the word in formation or the processing is interrupted being the content of the vector WORD delivered as the word identified. The algorithm is the following:

```
     PROCEDURE WORDIDENTIFICATION (WORD, BRAILLE,
     CLASS, IDENTIFICATION STATE, CHARACTER UNDER
     ANALYSIS)
     COMMENT: IDENTIFICATION OF THE NEXT WORD IN THE
              INPUT TEXT

A:   WORD SIZE ← 0
     BRAILLE SIZE ← 0
     BRAILLE (1)←BRAILLE CODE OF CHARACTER UNDER ANALY-
     SIS
     WORD (1)←EBCDIC CODE OF CHARACTER UNDER ANALYSIS
     CLASS←CLASS OF CHARACTER UNDER ANALYSIS
```

```
B:    WORD SIZE ← WORD SIZE + 1
      BRAILLE SIZE←BRAILLE SIZE + SIZE OF THE
      CHARACTER UNDER ANALYSIS CONVERTED TO BRAILLE
      READ THE NEXT CHARACTER FROM THE SOURCE BOOK
      TO CHARACTER UNDER ANALYSIS
      SEARCH (CHARACTER UNDER ANALYSIS,
      CLASS OF THE CHARACTER UNDER ANALYSIS, BRAILLE-
      CODE)
      IF (WORD SIZE > 30) ∨ (BRAILLE SIZE > 30)
      THEN GO TO C
      ELSE DECISION←OUTPUT TABLE (IDENFICATION STATE,
           CLASS OF CHARACTER UNDER ANALYSIS)
           IDENTIFICATION STATE←NEXT STATE TABLE
           (IDENTIFICATION STATE, CLASS OF CHARACTER-
           UNDER ANALYSIS)
           GO TO DECISION

1:    CLASS  11
      GO TO B

2:    STOP

3:    GO TO B

4:    CLASS  10
      GO TO B

5:    CLASS  11
      STOP

C:    CLASS  10
      IDENTIFICATION STATE  3
      STOP
```

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	STATES
		E, O	A,Ā,E,Ō,Ā,Ā...Ũ	Ī, Ū	Y,I,Ï,Ī	U,Ū,Ũ	Ā, Ō	H, I	R, L	G, Q	B,C,....K,-	S	N	NUMERIC	SPECIAL	BLANK	PREFIX	NON CONTRACTION SIGNAL	SOURCE BOOK PAGE NUMBER	END OF TEXT	UNINDENTIFIED CHARACTER	
OUTPUT TABLE		1	1	1	1	1	1	1	1	1	1	1	1	4	2	2	2	4	2	2	2	1. ALPHABETIC WORD
		4	4	4	4	4	4	4	4	4	4	4	4	3	2	2	2	4	4	2	2	2. NUMERIC
		3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	3	3	2	2	3. GARBAGE
		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	4. OTHER
		1	1	1	1	1	1	1	1	1	1	1	1	4	5	5	5	5	5	5	5	5. ALPHABETIC CHARACTER
NEXT STATE TABLE		1	1	1	1	1	1	1	1	1	1	1	1	3	4	4	4	3	4	4	4	1. ALPHABETIC WORD
		3	3	3	3	3	3	3	3	3	3	3	3	2	4	4	4	3	3	4	4	2. NUMERIC
		3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4	3	3	4	4	3. GARBAGE
		5	5	5	5	5	5	5	5	5	5	5	5	2	4	4	4	4	4	4	4	4. OTHER
		1	1	1	1	1	1	1	1	1	1	1	1	3	4	4	4	4	4	4	4	5. ALPHABETIC CHARACTER

Figure 4 - Output and next state tables of the word identification algorithm automaton

A.7 CONTRACTION ALGORITHM

This algorithm identifies and compacts segments of the word that can be contracted. According to the rules of grade two braille, one contraction is valid only in specific places of the word (beginning, middle or end) and must be necessarily within the limits of a syllable.

Several types of scanning are done on the word to seek final, initial and middle contractions.

The compacted word is set up while the successives scannings are done. The algorithm is the following:

```
        PROCEDURE CONTRACTION (BRAILLE, SYLLABLES)
        COMMENT: WORD CONTRACTION

A:      LIMIT←BRAILLE  SIZE
        END← EMPTY
        OUTPUT←EMPTY

B:      COMMENT: SEEK OF FINAL CONTRACTIONS
        BEGINNING← BRAILLE SIZE
        CONTRACTIONSEARCH (BRAILLE, BEGINNING, FROM
        THE RIGHT TO THE LEFT, LIMIT, CONTRACTION,
        VALIDITY PLACE)
        IF (CONTRACTION IS POSSIBLE) ∧ (VALIDITY PLACE=
        END) ∧ (SEGMENT OF BRAILLE IN CONSIDERATION
        BELONGS TO ONE SYLLABLE) ∧ (SIZE OF THE SEGMENT
        BRAILLE SIZE)
        THEN END← END, CONTRACTION
            LIMIT← LIMIT SIZE OF THE CONTRACTED SEGMENT
            MARK THE LAST CHARACTER OF END

C:      COMMENT: SEEK OF INITIAL CONTRACTIONS
        BEGINNING←1
        CONTRACTIONSEARCH (BRAILLE, BEGINNING, FROM THE
        LEFT TO THE RIGHT, CONTRACTION, VALIDITY PLACE)
        IF (CONTRACTION IS POSSIBLE) ∧ (VALIDITY PLACE=
        BEGINNING) ∧ (SEGMENT OF BRAILLE IN CONSIDERATION
        BELONGS TO ONE SYLLABLE)
        THEN OUTPUT← OUTPUT, CONTRACTION
             BEGINNING← BEGINNING + SIZE OF CONTRACTED
                        SEGMENT - 1
             MARK THE LAST CHARACTER OF OUTPUT IF THE
             END OF THE SEGMENT IS AN END OF SYLLABLE
             GO TO D
        ELSE
           E: OUTPUT← OUTPUT, BRAILLE (BEGINNING)
             MARK THE LAST CHARACTER OF OUTPUT IF BRAILLE
             (BEGINNING) IS AN END OF SYLLABLE
             GO TO D
```

```
D:   COMMENT:  SEEK OF MIDDLE CONTRACTIONS
     BEGINNING←BEGINNING + 1
     IF BEGINNING>LIMIT THEN GO TO F
     ELSE CONTRACTIONSEARCH (BRAILLE, BEGINNING,
          FROM THE LEFT TO THE RIGHT, LIMIT
          CONTRACTION, VALIDITY PLACE)
          IF (CONTRACTION IS POSSIBLE) ∧ (VALIDITY
          PLACE=MIDDLE) ∧ (SEGMENT OF BRAILLE IN
          CONSIDERATION BELONGS TO ONE SYLLABLE)
          THEN IF (END=EMPTY) ∧ (BEGINNING +
               SEGMENT SIZE=LIMIT)
               THEN GO TO E
               ELSE OUTPUT← OUTPUT, CONTRACTION
                    BEGINNING←BEGINNING+SIZE OF
                    CONTRACTED SEGMENT - 1
                    MARK THE LAST CHARACTER OF
                    OUTPUT IF THE END OF THE SEGMENT
                    IS AN END OF SYLLABLE
                    GO TO D
          ELSE GO TO E

F:   BRAILLE←OUTPUT, END
     BRAILLE SIZE←BEGINNING + END SIZE - 1
     STOP
```

World Council for the Welfare of the Blind

Committee on Cultural Affairs

International Conference on

"Computerised Braille Production - Today and Tomorrow"

London, 30 May - 1 June 1979

THE RNIB'S COMPUTERISED BRAILLE SYSTEM

by

G.F.S. Bennett

Royal National Institute for the Blind
224 Great Portland Road
London W1N 6AA

" The RNIB′s Computerised Braille System "

Summary

This paper to be presented at the conference "Computerised Braille Production - Today and Tommorow", is a presentation of the computerised braille production system currently installed at the Goswell Road, EC1 premises of the Royal National Institute for the Blind. The ideas leading to the inception of the system are first discussed, followed by the principles leading to the choice and fundamental system design of the current system. A description of each system module is presented, with a short section detailing current operational experience with the system as it stands. A few speculations are made as to the likely developments in the near future. Appendices are supplied detailing the hardware and software configuration of the RNIB system.

1. Introduction

It is fashionable these days to invent quaint acronyms by which to label a computing system which has been built for a particular application. The custom is as old as the electronic computer itself; one of the very first was given the name ENIAC - Electronic Numerical Integrator And Calculator. ENIAC was huge - occupying several large rooms, weighing many tons (the floor had to be specially re-inforced to hold it), and consuming hundreds of kilowatts of power. I cannot remember how many thousands of thermionic valves ENIAC contained, but the Mean Time Between Incidents was measured in minutes, and the replacement of faulty components was a full-time job!

Nowadays, of course, the situation has much changed. Computers of greater powers than ENIAC are contained on a small piece of silicon - the ′chip′ beloved of the popular press - less than one quarter of an inch square. As the machine has become smaller, so the field of applications has grown tremendously. I started my computing life around 1970 in the field of artificial intelligence. In those days (less than ten years ago) the average ′computer′ was still thought of in ENIAC proportions - that is, occupying large air-conditioned rooms, and consuming amounts of power which generally required a motor alternator producing a three-phase stable supply. Minicomputers were just beginning to make serious inroads into the market for dedicated applications (usually in the laboratory), and microprocessors had not even been heard of outside research establishments. Machine intelligence workers talked a lot then about the ′combinatorial explosion′. In broad terms this refers to the fact that the amount of storage and processing power needed to resolve problems is not in linear proportion to their complexity; the relationship is a power law. Researchers into chess-playing programs, to take a very famous case, were continually hampered by the fact that making a computer ′look ahead′ more than a few moves required enormous amounts of processing time and storage.

Computing applications seem to be undergoing a similar combinatorial explosion at the present time. Which of us would dare to speculate what areas of life may or may not be invaded by the microprocessor or even later technologies in fifteen or twenty years time. We desperately need a 'heuristic' by which to govern our development work if we are not to enthusiastically run headlong to our fate into the explosion of ideas and possible applications which our rapidly-advancing technology creates for us. This problem faces most would-be project managers and systems designers - and there are no easy answers. The path that we at the RNIB have chosen is familiar enough, and in the remainder of this paper I shall describe the RNIB Computerised Braille Production System in some detail. To make reference again to the opening sentence: I make no apology for not composing a suitable acronym for the RNIB system; no catchword could possibly pay tribute to all the dedicated effort that has put into it by all the RNIB Committee members, staff and advisors over the past few years.

2. Requirements for the RNIB computer system

One of the more obvious traps to be avoided in seeking to implement a high-technology computer system is that of computerisation for the sake of computerisation. Two questions should be asked here:

1). Why is a computer system being considered in the first place ?
2). What overall benefits would be required from a computerised system ?

Truly honest answers to these questions are prerequisites of an effectively designed computer system, for they provide the directing yardstick against which all further analysis and design can be measured.

There are a number of reasons why the RNIB, two to three years ago, was considering the introduction of computing power on a large scale into the braille production process.

i) The need for expansion of braille production capacity
ii) The need to shorten the time between the appearance of a publication in inkprint, and the subsequent publication in braille (which can be years)
iii) The chronic shortage of trained braillists, and the difficulty of recruiting new personnel to train as braillists
iv) The very desperate need to provide much better editing facilities than the hammer and punch used to correct the final zinc plates
v) The need to overcome the massive storage problems associated with keeping the zinc plate braille masters for subsequent reprints.

Our feelings that a properly-configured computer system would go a long way toward solving a lot of these problems stem from the largely successful results of an experimental system based on a second-hand IBM 1130 computer and punched cards, and the knowledge that other sites (there are many in the USA, and in the U.K. there is the Warwick University Research Unit for the Blind) had successfully used machine translation techniques for the production of braille. To answer question (2) above, what benefits did we expect to accrue from the computerisation of our braille production techniques? It goes without saying that the five criteria mentioned above should be satisfied. There are certainly other advantages to be had from a computer-based system, too, and I shall briefly outline some of these.

i) Elimination of noisy work environments

Up to the present time, braille transcription and translation, either by person or machine, has been a fairly noisy business. The manual embossing machines that have been in use (and still are, to a limited extent) are extremely noisy, and each requires a sound-proofed room to itself to prevent unacceptable noise levels to those working around. The punched-card equipment that was in use for the experimental computer system is also inherently noisy, as well as being wasteful of consumables.

ii) Improved management information

It is certainly possible to develop software to gather statistics on almost every function that is mediated by the computer system; therefore more precise information should be available to management on the performance of the system as a whole.

iii) Greater system reliability

The previous systems for braille production at RNIB have involved a lot of mechanical devices (such as card punches; manual embossers) which contain a large number of moving parts. Although such devices can never be completely eliminated in the production of braille, there was certainly scope for a considerable reduction. Replacement of some machines would in any case have been necessary, since they were nearing the end of their serviceable life.

iv) Greater frugality of expendables

Punched cards are becoming an expensive commodity; and zinc plates are even more so. The computer system envisaged would be capable of completely eliminating the former, and considerably reducing wastage of the latter (by correcting almost all errors before a job reaches the final embossing stage).

3. The Choice of the RNIB Computer System

To return to the point that I made in section 1; with such a bewildering array of rapidly-advancing technology around one, where does one begin in the specification and selection of a 'computer system'? On the one hand there would be a large, centralised mainframe system, of enormous power and flexibility in operation, but great cost. At the other extreme, there would be a 'turnkey' system designed around desk-top microprocessor-based computers, at comparatively low cost, but with little spare capacity and less flexibilty in operation. In between, there is a whole spectrum of solutions based on minicomputers of differing shapes and sizes, with similarly varying capacities, flexibilities - and cost!

The RNIB has opted for this middle course, basing its computer system around two fairly powerful minicomputers (GEC Computers' 4070s) which support a large range of standard and non-standard peripheral devices. The two computers together offer the power and flexibility of a small mainframe, at much less than half the cost. The total cost of the entire computer system was in the region of £300,000 of which the GEC Computers' equipment accounts for approximately £180,000.

Why did we adopt this particular course? Let us look again at our requirements in more detail, this time from the point of view of system functions.

i) Data Capture

The most important criterion to be applied to the new systems of data capture was that of quietness of operation. The most effective way to achieve this with present-day technology is with a key-to-magnetic-medium of some kind. The input from the English text entry system (as opposed to direct brailling) also needs to be verified in order to ensure absolute correctness. About the most flexible quiet, 'soft-copy' device available for such use at the present time is the visual display unit (VDU) for which there are standard computer interfaces, and which are easy to program and use.

ii) English-to-Braille translation

Obviously the most important criterion for this is that it should be as accurate in translation as possible! The RNIB aims for an error rate arising from each part of the system (before proof-reading) of 0.1%, which can be roughly summarised as one wrongly-spelt or wrongly contracted word in one thousand. This criterion must also be applied to the braille translation program. The only available programs which can achieve this level of accuracy at the present time are large, table-driven context-sensitive processors which are greedy on just about every conceivable overhead from core storage and c.p.u. time to disc transfer activity. (The one we use is a modified version of DOTSYS III-F, as provided for us by the Warwick University Research Unit for the Blind.) Such programs are always written in high-level language (as they should be) - in this case Fortran-H.

iii) Proof-reading

Once the braille exists within the computer, it is obviously important to be able to get it out again - not only for the final, polished product, but also for proof-reading purposes. There is a need for us to produce, online, both embossed braille for blind proof-readers, and a printed simulation of braille for sighted proof-readers.

iv) Editing

On-line editing of braille text was considered to be one of the most important design features of the new system. Not only was the old method of attacking mistakes on zinc plates with hammer and punch to be eliminated, but the provision of on-line facilities was to enable us to provide sophisticated text-processing operations similar to those found in computer typesetting systems in commercial print-houses.

v) Output

Two channels of output were considered necessary. The first - direct embossing of braille onto paper online for the speedy turnaround of short document and low-volume work; the second - the ability to make 'perfect' zinc plates for use on the printing presses.

vi) Archiving

The most obvious way to alleviate the ever-worsening problem of the storage of zinc plates is to store the material on magnetic tape instead. About 60 braille volumes can be held on one conventional 800 bpi magnetic tape - imagine what the equivalent in zinc plates would be like!

vii) Statistics

Management information concerning the performance of the various system modules - and the personnel who operate them - would be a feature of the new system.

The sheer complexity of the features that the new system required seemed to rule out the possibility of using any kind of micro-based, 'desk-top' system (although some were considered - the IBM 5100, for instance). The cost of any kind of mainframe system would have been prohibitive, however - especially in view of the back-up facilities required of the new system.

Our eventual choice then was for a system based on large minicomputers. The requirement for system reliability (2.iii) persuaded us that we should look a little way down the road to 'distributed processing'. It was eventually decided to base the system on two processors which were linked, and as similar in configuration to each other as was practicable.

As well as transferring information from one system to the other by means of the interprocessor link (the normal method), the working discs from one system can be immediately mounted and used on the other, if the need should arise.

The contract for the supply of the main computing hardware was placed with GEC Computers, Ltd., of Borehamwood, U.K.; That for the supply of video terminal equipment with Lynwood Scientific Developments Ltd., of Alton, U.K.; that for the supply of the On-line braille embossing equipment with Triformation Systems, Inc. of Stuart, Florida. U.S.A., and that for the remainder of the special-purpose hardware with Sigma Electronic Systems, Ltd., of Horsham, U.K.

4. Description of the RNIB Computerised Braille Production System Modules

i. Data Capture

i.a. Text Entry

The (letterpress) text entry system was designed to incorporate the best features of the old card-based system, such as continuous typing ability (no need to press 'feed' or 'return') and ease of verification. We considered the ease of use by the operator of a video terminal to be of prime importance, and evaluated almost every terminal on the market at the time. We came to the conclusions that:

- A separate keyboard and monitor is absolutely essential where a video terminal is in use for any length of time
- 'flicker' or 'waver' of the display is not to be tolerated
- a green (slower) phosphor is more restful to the eye, as well as helping to reduce flicker
- at least twenty-four lines of text should be available for viewing
- the keyboard should be as free as possible from excess clutter of so-called 'special function' keys.

The software we designed for text entry had to meet conflicting requirements. On the one hand, every character to the system should be checked for integrity (especially during verification); on the other hand, very fast response time is necessary to cope with keyboard input at peak speeds. Thanks to the flexibility of the GEC operating system, we were able to instal the text entry subsystem at a higher system priority than most of the operating system itself (only the real-time clock, the disc driver and the communications multiplexor driver, all of which the text entry system uses anyway are at higher priority), and we can guarantee a 5 millisecond terminal response time, which we find works very well in practice.

The software incorporates a simple text input system, a verifier, and a full-scale screen editor.

i.b. Braille input

The manual braille zinc plate embossers are progressively being replaced by a key-to-cassette system, which is almost noiseless. The key-to-cassette units have a small six-line monitor on which simple editing operations can be carried out, and a separate, Perkins-like keyboard. The cassettes are blocked at four braille lines (=168 bytes) per block; a convenient, if somewhat odd, size, since it allows a largish braille volume to be just comfortably contained on one side of a cassette. Cassettes thus produced can be read onto the computer's backing store, from where more sophisticated editing procedures can be invoked (see below). The key-to-cassette units (Braille-Cassette-Terminals - "BCT"s) are completely 'hard-wired', and are not programmable in any way.

ii. English-to-Braille Translation

At RNIB, we use the DOTSYS III-F translation program, originally developed at The Warwick University Research Unit for the Blind in Fortran, from a Cobol program first developed at M.I.T. DOTSYS is a fairly large program (in minicomputer terms), being about 60k in size. It is also extremely demanding in other overheads, especially c.p.u. time. The program is at basis table-driven, the main table of course containing the necessary contractions, others containing context tables. The translation speed of DOTSYS on the GEC 4070 is approximately 3,500 words per minute (elapsed time). We have done a fair amount of modification to the original DOTSYS program, in order to improve its facilities, and also its accuracy. We have had quite considerable problems in modifying the Fortran-H source to be suitable for the machines we have used. In particular, we used to use the ICL 1900 machines at Queen Mary College, University of London (to whom we are most grateful for allowing us to use their machine time free of charge). The ICL Fortran compiler offers very little in the way of Fortran-H type non-ANSI facilities, and this hampered our development for quite some time. However, after about 1.5 man-years work on the program and its tables, we now have an extremely reliable version of DOTSYS, which easily meets the requirements for accuracy in translation mentioned in Section 3.ii. We also modified DOTSYS and its tables to emulate the translation program which we used on the IBM 1130, which has allowed us to transfer all the work which was being processed on the 1130 to the new system, and thereby cease production use of the 1130, which was becoming extremely unreliable.

iii. Proof-reading

iii.a. Braille embossing

A custom-written despooler has been inserted into the spooling system, allowing the despooling of braille files to the Triformation LED-120 embossers. These (asynchronous) devices have in fact proved quite a headache to program, since the models we have have a small internal buffer, which is easily overflowed by incoming data during, for example, a page-advance sequence, which takes up to four seconds on the LED-120. Quite sophisticated timing has therefore had to be built into this despooler, to allow for the disparities between line speed and ′LED-speed′. Copies of braille files thus printed can be proof-read by blind proof-readers.

iii.b. Sighted braille simulation

Another custom-written despooler prints (on the line printer) matrices of full-stops which serve as ′visual′ braille cells for sighted proof-reading. Setting the line printers to print at eight lines to the inch provides quite a close approximation to the normal braille cell proportions. Although the line printers are capable of printing at 300 lines per minute, braille simulation can take up to three times as long as this because, of course, it is necessary to print three lines (of full stops) for each braille line in the file!

iv. Editing

As stated above, the provision of on-line braille editing facilities is one of the most significant advances of the RNIB system. As well as what may be termed 'normal' text editing facilities, The RNIB braille editing system offers sophisticated word-processing ability, approaching commercial typesetting systems in its manipulative power. We felt it worth the extra money and machine overheads to instal a full refreshed-screen display system for this work, using Sigma Electronics' QVEC graphics hardware. We also employed a consultant who specialises in this work, and who has written a number of computer typesetting systems for industry, to design and write our system for us. The result has been a system that can edit and reformat braille text quickly and easily to almost any layout that we require. In fairness, it should be pointed out that the overheads involved in using this type of text editor are very considerable; apart from the 'display file' (the area of main storage from which the 'picture' on the editing screen is maintained) which must obviously always be resident in main store, the rest of the editing subsystem also has to be store-resident, in order to ensure the fast response time needed. The total resident size of the editing subsystem is about 25 kbyte, which makes a very noticeable 'hole' in a 192kbyte machine. If this were not a problem enough in itself, some editing operations involve a lot of text manipulation, and need a fair amount of central processor time to execute. Since the subsystem runs at fairly high operating system priority (in order to ensure fast response time), this has the unfortunate effect of 'locking out' the rest of the system! Careful programming is necessary to avoid this problem, and we are still in the process of ironing out the last wrinkles from it.

v. Output

As well as direct braille embossing (mentioned above) for the production of short document and limited-run work, ECMA 34-standard cassettes are made from the braille files, which can be used on the RNIB's automatic plate embossers. Provided reasonable care is taken (cassettes are not a particularly robust medium) no correction of plates need ever take place now, since all the errors have been trapped further back in the system, and the final cassettes are (or should be!) completely 'clean'.

vi. Archiving

As mentioned above, the main stream of archiving of RNIB-produced braille is now to magnetic tape. This means that we could if we wished archive everything that we produce, and not be likely to run into storage problems for a very long time indeed, since at least 50 volumes can be stored on one magnetic tape. The system has two tape drives, to allow tape-to-tape editing, should this ever prove necessary.

vii. Statistics

At present, we gather statistics on the following aspects of the system:

- Text entry operator performance.
  As well as being of obvious interest to the data prepration management, we have found that the operators themselves like to know how they are getting on, and so all the figures (concerning speed and error rates) are made available to them.
- Translation program performance.
  As stated above, current performance of the DOTSYS program is about 3,500 words per minute. We have no particular wish to improve on this at present, but it is an interesting statistic, insofaras when modifications are made to the program, it is very easy to see whether they have significantly altered the overheads that the program imposes on the system.
- Total work throughput.
  This statistic is simply calculated by measuring the number of finished volumes that find their way onto an archive tape per unit time! (At present, we archive all our work).
- Processor utilisation.
  Although the best computing circles view the measurement of central processor utilisation with some scorn these days, we have found it useful, not only in determining the relative load of the various system modules, but also in system tuning, and in preventing ʹdeadly embraceʹ situations. For example when processor utilisation appears high but work is being processed but slowly and there are great numbers of disc transfers taking place, this indicates that the system is overloaded and is ʹthrashingʹ. (This is an unpleasant condition which can afflict virtual-memory computers; The paging system becomes overstretched and in an effort to keep up with demands which are being made of it, ʹthrashesʹ segments rapidly between main store and backing store, without leaving them in main store long enough for any useful processing to take place.)

5. Current Operational Experience

Subsequent to the report by Lawes(1), the first ideas for the new computer system were born in late 1975. I joined the staff of the RNIB in February 1977 to oversee the systems side of the project. The order for the main computer system was placed in September 1977, and the bulk of the hardware was installed and commissioned by June 1978. Systems design and program development had started with my appointment to the RNIB in February 1977, but reached full development from February 1978, when University College Computer Centre were kind enough to allow us the use of their newly-installed GEC 4000 computers. Pilot runs of most system modules were commenced during August 1978, and the system finally went ʹliveʹ on 2nd. October 1978 - only three weeks behind schedule.

We have now been in production for about six months, and in that time have processed over 140 braille volumes, that is over 20,000 braille pages. This compares very favourably with the previous year's production, and certainly it seems that there will be no difficulty in meeting the target set for the first year of the new computer system - to equal the final year's production on the manual and IBM 1130 systems. Equally, the stage is set fair to reach next year's target of double production, since at present the text entry staff number only four (when the full complement is ten), and most of the transcribing staff are still new to the B.C.T. machines.

(1) Lawes, W.F.
"The Feasibility Study into the Usage of Computers by and for the Blind."
The Royal National Institute for the Blind. London, England. January 1975.

6. Future Plans

And what of the future? The committee members and officers of the RNIB are actively looking into a number of possibilities. One to which we are firmly committed is the processing of compositors' tapes. Clearly the re-keyboarding of text that is currently done [in the text entry stage] could be drastically reduced, if some way of processing the magnetic tapes produced in composition for letterpress printing could be utilised. The technical problems which are associated with dealing with the format-codes on these tapes (and converting them where necessary into DOTSYS format-codes) are by no means inconsiderable, but some preliminary work carried out over the last six months has encouraged us sufficiently to make formal approaches to some U.K. printing houses with a view to processing tapes on a production basis.

Another is to attempt to do some braille music translation by computer, following up the excellent work done at the Warwick University Research Unit for the Blind by Mr. J. Humphreys. We are still evaluating the likely difficulties that we would encounter in trying to mount his program suite on our system, but suffice to say that we are sufficiently encouraged to acively pursue the matter.

There are many other areas of interest to the RNIB in the application of computerised techniques to the problems of braille in general, but these are still very much at the 'ideas' stage, and hence beyond the scope of this paper. I think it can be seen, however, that the RNIB intends to stay at the forefront of research and development, in order to ensure that the braille-reading population of the United Kingdom benefits in every possible way from the application of digital computers, both large and small.

G. F. S. Bennett   R.N.I.B.   London.   April 1979

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

THE SEGMENT TRANSLATION SYSTEM: A NEW FORMALISM FOR COMPUTERISED CONTRACTED BRAILLE TRANSLATION AND PHONEMISATION

by

Wolfgang A. Slaby

Computing Centre University of Münster
Einsteinstr. 60
D-4400 Münster

The segment translation system: a new formalism for computerised contracted braille translation and phonemisation

by Wolfgang A. Slaby, Computing Centre,
University of Münster, Germany

Summary

In connection with the design and implementation of an algorithm V that generates a formal translation system by analysing a set of only a few examples of the underlying translation process the new model of a *segment translation system* has been developed, which is very suitable for formalising translation processes of comparatively low complexity (e.g. braille, phonology, hyphenation). Thus the segment translation system has been used as a formal framework for the research project "German braille translation", where a new algorithm for the translation into German contracted braille has been implemented as a special segment translation system S_{gb} consisting of 3977 rules for the moment [SLABY(1979)]. In translating German texts this system yields a correctness rate of more than 99,2 % on different tokens. Furthermore a segment translation system S_{gp} is being developed by F.-P. SPELLMANN and the author as a formalisation of the translation process "German phonemisation".
[SLABY/SPELLMANN (1979)]

The segment translation system: a new formalism for computerised contracted braille translation and phonemisation[1]

by Wolfgang A. Slaby
Computing Centre, University of Münster, Germany

Introduction

The main goal of our research work on computerised braille translation at the University of Münster was to develop algorithms producing high quality contracted braille. But in order to design *good* algorithms of this kind, in my opinion, it is absolutely necessary to take an adequate linguistic model of a translation system as the formal framework for the translation algorithm. Thus the first step I did consisted in analysing existing models of formal language theory with respect to their applicability to the problem of translation into contracted braille. The main and in some cases mutually conflicting criteria for this evaluation were the following:

1. *suitability*
   - the problem of contracted braille translation can be completely formalised by this model
2. *adequacy*
   - the complexity of the model is only as high as required by the translation problem
3. *universality*
   - the model is independent of special languages and their respective definitions of braille contraction rules
   - the model is applicable to problems of a comparable level of complexity (e.g. hyphenation,phonemisation)

---

[1] The segment translation system has been developed in connexion with the author's doctoral thesis [SLABY (1977)], University of Heidelberg, Department of Applied Linguistics (Prof. Dr. K. Brockhaus)

4. *ease of implementation*
   - a concrete translation system of this type formalising a concrete translation problem (e.g. German inkprint ⟶ German contracted braille) can be built up without greater difficulties
   - the translation system can be easily implemented as a computer program
5. *efficiency* of the resulting translation algorithm

After a thorough analysis of the existing models of a MARKOV system of translation rules, a syntax-directed translation scheme, and a finite state transducer I suggested in 1973 to try the *MARKOV system of translation rules* [SLABY (1974, 1978a)] as a formal model that especially meets the criteria of *suitability*, *universality*, and to some extent *ease of implementation*, and is able to canonically incorporate the notion of priority. The tedious work of building up a special MARKOV system for the translation into German contracted braille was performed by J. SPLETT, the implementation being made by B. EICKENSCHEIDT [for further details of this algorithm see the papers of B. EICKENSCHEIDT and J. SPLETT in these proceedings].

But in the light of the more than ten years experience with the DOST/WERNER program for German braille translation, one of the first programs for computerised braille production in the world [DOST/WERNER (1969)], I suspected that translating by the priority of rules and not sequentially from left to right would lead to a low efficiency. Therefore, I tried to develop a new model retaining the method of translating sequentially from left to right but at the same time providing a high flexibility in formalising braille contraction rules. For this purpose I performed a thorough analysis of the problem of translation into contracted braille and similar translation processes of low complexity like phonemisation and hyphenation. The results of this analysis suggested the hypothesis that the translation z of a string w can be looked upon as being built up by a concatenation of strings that are translations of fixed segments of w. Therefore, a formal translation system that is to perform such a translation process automatically, should consist of rules with a basic component like u→v, i.e. "segment u must be translated into segment v". But in order to determine automatically the correct translation of segment u if there are alternatives $v_1, v_2, \ldots, v_n$ (consider, for example, the translation of the segment "ea" with respect to English phonology where there are at least the phonemes [iː] (least), [ɛ] (pleasure), [ɛː] (heard), [e] (head) as possible translation), to each basic component there must be attached an additional condition that controls the correct application of the rule. Under the hypothesis that this decision can be uniquely made (with the exception of homographs) by taking into account the context of symbols surrounding the segment u in the string under consideration, one defines a *segment translation rule* by $u \rightarrow v\ [x,y]$, meaning: in a string w the segment u is to be translated into v, if it occurs in w with segment x left and segment y right of it.

The segment translation system

In order to give a precise definition of a segment translation system it is presumed that the reader is familiar with the basic terms and definitions of formal language theory like *alphabet* Σ, *finite string* over Σ, *nullstring* ε , *concatenation* vw of two strings v and w, *length* $\ell(v)$ of a string v, *segment* or substring u of a string w, *prefix* $h_n(w)$ and *suffix* $t_n(w)$ of length n of a string w, *formal language* over Σ , concatenation of formal languages, etc. that can be found, for instance, in AHO/ULLMAN (1972). In addition to this for two formal languages L_1 over Σ and L_2 over Δ, $w \in \Sigma^*$, and $(w_1, w_2, \ldots, w_n)$ n-tuple of strings over Σ one defines:

1. $(w_1, w_2, \ldots, w_n)$ *decomposition* of w with respect to L_1 :

   $\Leftrightarrow$ $w = w_1 w_2 \ldots w_n$ and $\forall i (1 \leq i \leq n)$: $w_i \in L_1$

2. T *translation process* from L_1 into L_2 :

   $\Leftrightarrow$ $T \subseteq L_1 \times L_2$ and

   for each $w \in L_1$ there exists $z \in L_2$: $(w,z) \in T$

3. z *translation* of w with respect to T :

   $\Leftrightarrow$ $(w,z) \in T$

Definition 1: [*segment translation system*]

A quadruple $S = (\Sigma, \Delta, \blacklozenge, R)$ is called a *segment translation system,* if

1. Σ , Δ are alphabets
2. $\blacklozenge \notin \Sigma \cup \Delta$ [$\blacklozenge$ the *frontier symbol*]
3. $R \subseteq \{ u \rightarrow v\ [x,y] \mid u \in \Sigma^*,\ u \neq \varepsilon\ ;\ v \in \Delta^*$
   $x \in \{\blacklozenge, \varepsilon\} \cdot \Sigma^*;\ y \in \Sigma^* \cdot \{\blacklozenge, \varepsilon\} \}$

   R finite, non-empty.

   Each $u \rightarrow v\ [x,y] \in R$ is called a *segment translation rule*
4. For any $u \rightarrow v_1\ [x_1, y_1]$, $u \rightarrow v_2\ [x_2, y_2] \in R$:
   $(x_1 = x_2$ and $y_1 = y_2 \Rightarrow v_1 = v_2)$
5. Let $\ell s(R) := \{ u \mid u \in \Sigma^*$ and there exists $u \rightarrow v\ [x,y] \in R \}$.
   Then for each $u \in \ell s(R)$ there exists $v \in \Delta^*$ such that
   $u \rightarrow v\ [\varepsilon, \varepsilon] \in R$ [the *base rule*] .

   If in addition S satisfies
6. $\ell s(R) \subseteq \Sigma$

   then S is called a *character translation system.*

To generate a translation of an input string w the application of a segment translation system S to w consists of two steps.

First, w must be decomposed into a sequence of segments with respect to $\ell s(R)$, i.e. into segments which occur as left sides of segment translation rules of S. If such a decomposition of w does exist - a fact which can be guarantied by the assumption that Σ is a subset of $\ell s(R)$ - in general this decomposition is not unique. On the premises that the translation of w should be unique, one decomposition of w must be marked *canonical.* One possible candidate for a canonical decomposition is that resulting from iteratively splitting off the longest possible segment from the remainder of the string w. The second step of the application of a segment translation system S to w consists of determining for each segment u_i of the canonical decomposition $(u_1, u_2, \ldots, u_n)$ of w with respect to $\ell s(R)$, for which segment translation rule $u_i \rightarrow v_i\ [x_i, y_i]$ w satisfies the context-condition $[x_i, y_i]$, i.e. for which $u_i \rightarrow v_i\ [x_i, y_i]$ x_i is a suffix of $\blacklozenge u_1 \ldots u_{i-1}$ and y_i is a prefix of $u_{i+1} \ldots u_n \blacklozenge$.

Among the different rules $u_i \to v_{i_1}[x_{i_1},y_{i_1}],\dots, u_i \to v_{i_s}[x_{i_s},y_{i_s}]$ the context-conditions of which are satisfied by w, that rule with a context-condition of maximal length $\ell(x_{i_\nu} y_{i_\nu})$ will be provided with highest priority. With these standardisations the translation of w with respect to S now can be defined exactly by the following algorithm:

Application of a segment translation system

```
input:   segment translation system $S = (\Sigma,\Delta,✦,R)$,
         input word w

method:  variables:  word, result, left, right   of type string,
                     lsides, rules              of type list
         word := w , left := ✦ , result := $\varepsilon$ , lsides := $\ell s(R)$, rules := R
         while (word $\neq \varepsilon$) do
                 1. search for longest possible string u of lsides
                    with   $u = h_n(\text{word})$
                 2. right := $t_{\ell(\text{word})-n}(\text{word})$.✦
                 3. search for segment translation rule
                    $u \to v[x,y]$ of rules with
                    $x = t_{\ell(x)}(\text{left})$   and   $y = h_{\ell(y)}(\text{right})$
                    and  $\ell(xy)$  maximal under these conditions
                 4. left := left$\cdot$u ,  word := $t_{\ell(\text{word})-n}(\text{word})$,
                    result := result$\cdot$v
         end

output:  S(w) := result ,
         the translation of w with respect to S
```

To give an impression of concrete segment translation systems

1. the subset of R_{gp} ("German phonemisation") consisting of all rules with left side "sch" (compiled by F.-P. SPELLMANN), and
2. the subset of R_{gb} ("German contracted braille") consisting of all rules with left side "stell" (the root of the German verb

"stellen" [to put]) are presented here arranged according to increasing priority:

$R_{gp,sch}$ = { sch → <sch> [ε,ε],
→ s<sch> [ε,ef] (Stabschef)
→ sk [ε,ao] (Verkehrschaos)
→ s<ch> [ε,ato] (Eschatologie)
→ s<sch> [ε,anc] (Überlebenschance)
→ <sch> [ε,eff] (Scheffel)
→ sk [ε,rom] (Gaschromatograph)
→ sk [ε,ron] (Unheilschronik)
→ sk [ε,or*] (Werkschor)
→ sk [ε,öre] (Werkschöre)
→ s<ch> [ä,en] (Bläschen)
→ s<ch> [ö,en] (Höschen)
→ s<ch> [ε,urie] (Ischurie)
→ sk [ε,rist] (Unionschristen)
→ s<ch> [ε,ines] (Festlandschinese)
→ sk [ε,arak] (Verkehrscharakter)
→ sk [ε,lor*] (Reinigungschlor)
→ s<ch> [än,en] (Hänschen)
→ s<ch> [äu,en] (Häuschen)
→ <sch> [lö,en] (löschen)
→ <sch> [wä,en] (Kopfwäschen)
→ sk [ε,orges] (Werkschorgesang)
→ s<ch> [ε,irurg] (Transplantationschirurgie)
→ s<ch> [die,en] (Radieschen)
→ <sch> [frö,en] (Fröschen)
→ <sch> [täu,en] (täuschen)
→ <sch> [räu,en] (Geräuschen)
→ sk [ε,orleit] (Kreischorleiter)
→ s<ch> [kreb,en] (Krebschen) }

$R_{gb,stell}$ = { stell → <,stell> [ε,ε],
→ <st><el>l [ε,eu] (Apostelleuchter)
→ <st>e<ll> [ε,at] (Konstellation)
→ <st>e<ll> [pa,ε] (pastellfarben)

→ <st><el>l [ε,esu] (Epistellesung)
→ <st>e<ll> [ε,are] (quasistellare)
→ <st>e<ll> [ε,ar*] (interstellar)
→ s<te><ll> [a,er] (Glasteller)
→ s<te><ll> [lau,ε] (Nikolausteller)
→ <st>e<ll> [*ka,ε] (Kastell)
→ s<te><ll> [cht,er] (Weihnachtsteller) }

Compiling a concrete segment translation system

In order to compile a concrete segment translation system as an implementation of a special translation process at first you have to decide which segments to take as appropriate left sides of segment translation rules. With German contracted braille this decision was not so difficult, since all those segments of German words, for which there exist immediate contractions, are best candidates to become left sides of segment translation rules. One can suppose this fact holding true also for the braille definitions of other (European) languages. With respect to German phonemisation we decided to take in a first step only those segments as possible left sides which in some word have a representation as a single phoneme. [In a segment translation system for English phonemisation, for instance, under this criterion the segment "ea" will be taken as a suitable left side for a segment translation rule.] Since this method yields a set $\ell s(R)$ of only very short left sides of segment translation rules, one might consider widely used prefixes of suffixes of German words as a completion of $\ell s(R)$. This alternative will be tried in a later stage of our research project on German phonemisation[1].

A further problem in connexion with the selection of appropriate left sides for the segment translation rules originates from the technical fact that by the application of a segment translation system a translation of a word is generated sequentially from left to right.

---

[1] This project is financially supported by the German Federal Ministry for Research and Technology (BMFT Grant o1 VJ o97 - ZA/NT/MT 267)

Therefore the situation may occur that two left sides overlap in a word [e.g. "Apostelleben" (life of apostles) where the segments "stell" and "leb" overlap by the letter "l"] and not the first but the second segment is responsible for the correct translation. In this case two different strategies to overcome this problem are possible:
1. that string of minimal length containing both overlapping segments [in the example, the string "stelleb"] can be taken as a new left side of a segment translation rule;
2. the concept of segment translation system can be extended to allow a suffix of the left side of a rule remaining untranslated when this left side is considered during the translation process; by this way those characters of the first segment [in the example, the character "l"] belonging also to the second segment can be left untranslated, so that the second segment as a whole will be at your disposal in the next step of the translation. Both of these methods have been applied in compiling the concrete segment translation systems S_{gp} for German phonemisation an S_{gb} for German contracted braille.

When the set $\ell s(R)$ of segments for possible left sides of rules is determined, then for each segment $u \in \ell s(R)$ one has to build up a system of segment translation rules with common left side u that at best in all circumstances in which u occurs as a part of a German word will give the correct translation of u. Therefore one has to analyse as many German words as possible, u is a part of which, in order to create the appropriate context conditions. For this purpose several collections of more than 15o ooo different inflected German words in machine-readable form have been used to produce for each $u \in \ell s(R)$ a special list of all words containing u as a segment, sorted according to the context surrounding u in the word, and printed with u always starting in the same column. A thorough analysis of this list renders possible a decision which of the alternative translations of u should constitute the base-rule $u \rightarrow v[\varepsilon,\varepsilon]$, and which context conditions have to be generated in order to guarantee the correct translation of u in all words of the list. To pay attention to the German preference for building compound words that are not yet listed in a dictionary a further step of controling the completeness of the set of segment translation rules for u is performed. Systematicly the segment u is decomposed into two segments u_1 and u_2, and by means of a reversed word index and normally sorted dictionary all the words w_1 with suffix u_1 and w_2 with prefix u_2 are discovered that could be combined to a correct German compound word $w_1 w_2$. [Actually the procedure is more complicated because you have to pay attention to the fact that in many cases an additional boundary morphem is inserted between the two components of a compound word.] By this method the words "Glasteller", "Nikolausteller", "Weihnachtsteller" and "Apostelleuchter", "Epistellesung", for instance, have been discovered each of which gave rise to a new segment translation rule for the segment "stell".

The results

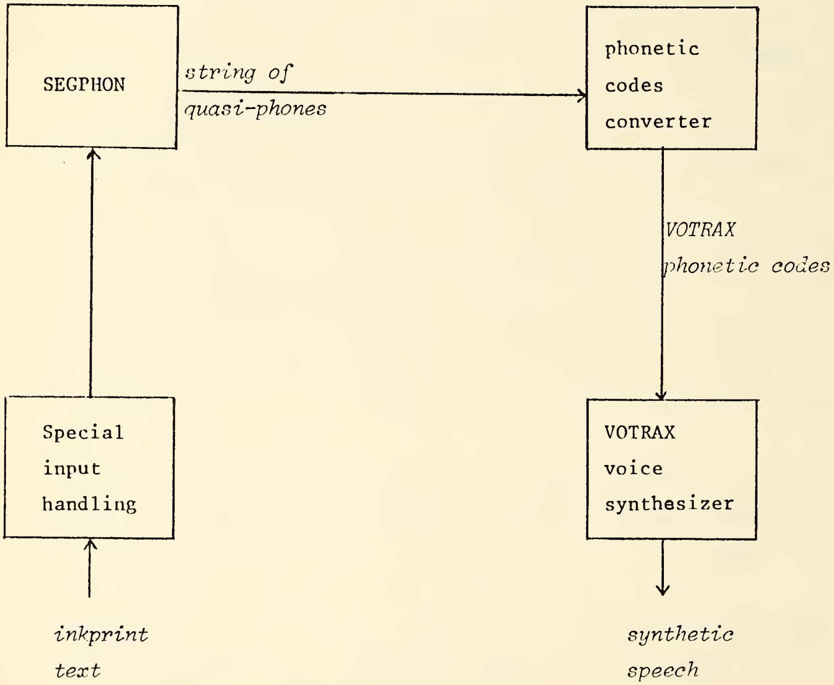
As one of the practical results of this research work on formal models for translation processes the algorithm SEGBRA for the translation into braille has been developed, the underlying segment translation system S_{gb} for German contracted braille consisting of 3977 rules for the moment. Alternating with the braille translation module PUMA (the details of which are given in the paper by EICKENSCHEIDT/SPLETT in these proceedings) this algorithm is in practical use for the regular production of our biweekly braille magazine with cuts of the German weeklies "Die Zeit" and "stern".

Although a thorough analysis of the characteristics of these two translation modules is still in preparation, some preliminary observations can be reported:

characteristics of	PUMA	SEGBRA
1. formal model:	MARKOV system of translation rules	segment translation system
2. size of the system:	~ 6 ooo rules	3977 rules
3. speed of translation:	~ 7 5oo words	~ 22 ooo words
	of average length 11.6 chars/word per minute CPU on an IBM 3o32	

4. correctness rate: < 1 error ~ 1 error
per 5o braille pages
(28 lines of 36 characters each per page)

Furthermore, based on a special segment translation system S_{gp} developed by F.-P. SPELLMANN and the author, an algorithm for the automatic production of synthetic speech from German inkprint texts is being implemented. The following diagram gives the details of the complete system:



special input handling

- spelling vs. writing in full of abbreviations
- transforming numbers into the corresponding textual numerals
- marking of word boundaries

translation module SEGPHON

- formal model: segment translation system
- size of the system: 585o rules (for the moment)
- stage of development: system not yet completed,
  being implemented and tested

phonetic codes converter

- converting of quasi-phones (corresponding to the international phonetic alphabet) into groups of VOTRAX phonetic codes that are the appropriate input symbols for the VOTRAX voice synthesizer in order to produce the correct speech output of the different quasi-phones.

Current developments

We have started to implement our different algorithms on a microcomputer system that will finally consist at least of

- a video display with keyboard
- a floppy disk unit
- two units for reading/writing ECMA-34 cassettes
- an 8o8o-based microprocessor with about 128 K bytes of RAM/EPROM storage capacity
- a SAGEM braille printer

The costs for the complete braille translation system (hardware and software) are estimated to be about 5o ooo DM.

Literature:

1. AHO/ULLMAN (1972):

   Aho, A.V./Ullman, J.D.:
   The Theory of Parsing, Translation, and Compiling
   Volume I: Parsing
   Englewood Cliffs, 1972

2. DOST/WERNER (1969):

   Dost, W./Werner, H.:
   Automatisierte Herstellung von Blindenschrift mit Hilfe einer Datenverarbeitungsanlage
   IBM-Nachrichten 194 (1969), pp. 594-599

3. SLABY (1974):

   Slaby, W.A.:
   Automatische Übersetzung in Blindenkurzschrift
   EDV in Medizin und Biologie 5 (1974), pp. 111-116

4. SLABY (1977):

   Slaby, W.A.:
   Automatische Erzeugung formaler Übersetzungssysteme aus endlichen Mengen von Beispielen
   Dissertation Heidelberg, 1977
   Schriftenreihe des Rechenzentrums der Universität Münster, Nr. 24, Juli 1977

5. SLABY (1978a):

   Slaby, W.A.:
   The Labeled MARKOV System:
   A concept to formalize translation processes
   in: Werner, H. (ed.): Computerized Braille Production, Proceedings of the 2nd International Workshop in Copenhagen (Denmark), September 1974
   Schriftenreihe des Rechenzentrums der Universität Münster, Nr. 3o, June 1978, pp. 12-18

6. SLABY (1978b):

Slaby, W.A.:
1. Zwischenbericht zum Forschungsvorhaben:
"Entwicklung eines Programmsystems zur automatischen
Umsetzung von Texten des Deutschen in gesprochene Sprache"
Kennzeichen: o1 VJ o97 - ZA/NT/MD 267 -
Münster, Februar 1978

7. SLABY (1979):

Slaby, W.A.:
SEGBRA -- a universal translation system for grade 2
braille translation
to appear in: Schriftenreihe des Rechenzentrums der
Universität Münster, 1979

8. SLABY/SPELLMANN (1979):

Slaby, W.A./Spellmann, F.-P.:
Automatische Übersetzung von Texten in gesprochene Sprache
to appear in: Schriftenreihe des Rechenzentrums der
Universität Münster, 1979

Address of the author:

Dr. Wolfgang A. Slaby
Computing Centre
University of Münster
Roxeler Straße 6o
D-44oo Münster
Fed. Rep. of Germany

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

FRENCH BRAILLE, LARGE PRINT, AND VOICE SYNTHESIS

by

D. Keeping, M. S. Doyle, and P. A. Fortier

University of Manitoba
Winnipeg, Manitoba
Canada

INTRODUCTION

Since the winter of 1965, when our Computer Programming course for the Blind started, considerable activities in the computer field on behalf of the blind has been carried out at the University of Manitoba.

A course in education of blind computer programmers has been in progress yearly since 1965. We have graduated upwards of 100 students and most of them are presently employed in the programming field. In the past year, the course has been completely revised employing new teaching methods and visual aids and brailling equipment which have been developed over the past ten years. Our course has been chosen by several large American business firms as a vehicle for training their blind computer personnel.

The activities we are reporting on here fall into three categories.

(A) Computer Braille

Since 1973, the University of Manitoba has been developing and producing both English and French computer Braille. The original Dotsys III program for English Braille has been greatly enhanced to produce nemeth code mathematics, uncontracted French Braille and pronunciation symbols, double spacing for smaller grades, Braille and print table of contents and many other formatting improvements. We have also developed a pre-processor which accepts as input, text which looks like the print page. For the past 5 years, all blind students in Manitoba have had all the Braille they need when they need it.

(B) Large Print Production

We have been using a VERSATEC electrostatic printer/plotter to produce texts in large print. This machine works on a principle similar to that of the xerox copier, and as a result can produce a very wide and flexible range of output.

(C) Voice Synthesis

We have been using the Votrax S11 voice synthesizer which was developed by the Federal Screw Works in United States and run on a PDP11 mini-computer. In the summer of 1975, we developed a 1400 word vocabulary as well as some basic grammatical work for the system. Presently, we are examining the possibility of converting the system to micro-computer.

I. BRAILLE

In 1973, we faced two cases in which traditional production methods did not fulfil a pressing need. On examination, we realised that there were three bottlenecks in the system: 1) the training required for manual keying, 2) the slowness of this operation, 3) the necessity to proofread after the Braille had been produced.

We were able to set up production of Braille avoiding these bottlenecks. Texts were input to a standard disc-editing system. Proofreading was done on listings of the input, and corrections made before the text was transcribed. This eliminated one bottleneck--proofreading. We produced uncontracted French, German and Spanish Braille using a simple reformatting program which we wrote, and Grade II English Braille using an existing program, DOTSYS III. Translation into Braille codes was done at machine speeds, and actual production via a 120 character/second Braille embossing terminal. This eliminated the bottlenecks caused by manual keying of the Braille characters.

With computerised Braille production humming along quite satisfactorily, we turned our attention to the programming side of the activity. We expanded the capacities of the existing English Grade II Braille translation system; the original DOTSYS program has been modified to improve output formats, to handle uncontracted French, German, Latin and Spanish Braille, as well as to produce the special codes needed for phonetic transcription, and for nemeth code mathematical notation for all levels, from primary to university.

The following is a list of some of the other modificiations made to the Dotsys program at the University of Manitoba.

(a) Augmented heading capability-- a page eject automatically takes place when a heading is required and there are less than 3 available lines on the current Braille page.
(b) Programming mode--capability of dropping into one-for-one computer terminal mode for transcribing computer programs. No capitals are permitted and number signs are dropped.
(c) Guide dots--in the table of contents (copied directly from the Atlanta School Board System)
(d) Braille and print table of contents--provides the page numbers in Braille and print in the same table.
(e) Running top titles--occurs at the top of each Braille page; may be changed or deleted as necessary throughout the volume.
(f) Running print page numbers--occur at the top of each Braille page indicating the current and the last print page on the current Braille page e.g. 57-8, 99-100.
(g) Double spacing--for junior grades; may alternate with single spacing throughout the Braille volume.
(h) Preliminary pagination--separates the preliminary material including the table of contents from the text by using a different sequence for the Braille page numbers e.g. p1.

Next we turned our attention to Grade II French Braille. Grade II Braille was devised at the end of the 19th century, using the linguistic principles accepted at the time, and with particular emphasis on mnemonics. Naturally enough, the rigorous logic and total unambiguity required for easy computer processing were not taken into account when these rules were elaborated. The French Grade II Braille rules are a case in point.

About forty syllables in French Grade I Braille are replaced by a single character in Grade II. For example, the diphthong 'ui' is replaced by a colon, when followed by a consonant, unless the combination 'br' follows immediately, in which case the 'br' is replaced by the colon but 'ui' is left unchanged. As if this type of rule were not complex enough, the contraction systems has been elaborated in terms of syllables, not strings. Thus in 'oublient' the last four letters are taken to be the letter 'i' followed by the verbal ending 'ent', and contracted 'i-ent', but in 'vient', the last four letters are taken to be the syllable 'ien' followed by the letter 't', and contracted 'ien-t'. This leads to the classic problem of 'convient' which is contracted 'i-ent' if it is the third person plural of 'convier', and contracted 'ien-t' if it is the third person singular of 'convenir'.

There are about 2,600 words in French Braille which have Grade II versions falling outside these rules, either because they are exceptions to the rules, or because they have discrete word abbreviations. These word abbreviations are also used as part-word contractions, when they form part of a longer word without changing meaning. The problem is that they must not change meaning. So the word 'logique' has a special word abbreviation, which is also used for the string 'logique' in 'illogique', and through a relaxation of the rules in 'biologique' as well. But although 'nombre' has a word abbreviation, the string 'nombre' is not abbreviated in 'pénombre'.

French Grade II Braille rules are more elaborate than the ones used for English, but not as complex as those governing German. But they are quite complex enough to cause serious problems for computer processing. The semantic and/or pronunciation element calls to mind the type of problem incurred in computer aided parsing of natural language texts.

The solutions proposed for Grade II Braille production are not dissimilar from those implemented in parsing. The widely used DOTSYS III system for English Braille is a table driven program, which takes each word in turn from the stream of text and, through a series of tests and branchings, arrives at the place in the table showing the word contraction, or part word contractions applicable to the word, translates it into Grade II codes, outputs it, and reads in the next word for processing.

French Braille has about the same number of syllable contractions as English, but about twenty-five times as many word contractions, plus the use of full word contractions as part word contractions in special circumstances. The rule testing or table look up approach used for English would be very difficult and very costly to implement for the much more extensive French Grade II Braille system. It is our opinion that the nature of text data suggests a more fruitful approach not only to French but to English Braille as well.

Computational linguistics studies have demonstrated that the curve shown in figure 1 is typical not only of the 20th century French novel texts from which it was derived (Engwall 1974), but of the distribution of words in natural language texts in general (Allen 1970, Juilland 1970, Kucera 1967). This information can be used in analysing present Grade II Braille production programs. The hundred most frequent words at the left of the graph comprise 50% of a natural language text. When their contraction is computed each time that the word appears in the stream of text, as is done in present programs, this computation has to be performed 10,000 times for a short text comprising 20,000 words. An approach which grouped all similar words together before computing the Grade II version would perform the same computation 100 times. Processing a specially ordered text would also promote more expeditious handling of problem words, and facilitate re-use of contractions already computed. These advantages have made themselves evident in computational linguistics applications, notably in the grammatical tagging of texts (Duro 1973, Fortier 1973, Zampolli 1973).

These considerations led to the development of the system structure shown in Figure 2. BRAILLE1 and BRAILLE5 are the input and output modules respectively of the system. We decided to implement them first, because elegant and easily readable output is a *sine qua non* for a successful system in this type of application, and, of course, an output module pre-supposes an input module which supplies data to it.

BRAILLE1, the first program in the system, accepts the text in the form of a continuous stream, with flags to control output format. This program translates the text into Grade I Braille codes, either in French or in English, depending on what has been signaled to it via control flags. The output from BRAILLE1 is divided into word records, each record containing the word, identification of its position and a control field. BRAILLE1 also checks the validity of format control commands before passing them on for further processing.

Output from BRAILLE1 can go directly to BRAILLE5 if Grade I or uncontracted Braille is desired. BRAILLE5 accepts a file of word records in text order, and reformats it for final output via the Braille embossing terminal. Each Braille page has its own page number, and page numbers from the ink print text are also shown if desired. Footnotes appear at the bottom of the Braille page in which the reference to them is found.

Experience developing BRAILLE5 dictated modifications to BRAILLE1 on many occasions. Develpment of these two programs proved to be the most meticulous and demanding part of the system development, and we were glad we accomplished it at the beginning of the project.

We had originally planned to add to BRAILLE1 a routine which would translate multi-word expressions and those hundred very frequent words which make up 50% of any natural language text. For simplicity, and in order to keep the memory requirements of BRAILLE1 within reasonable limits, we decided to put these functions in a separate program, BRAILLE2, which also divides the text up into three files. One contains format control words, words and expressions already translated, mathematical expressions, punctuation, and words signaled to be left in Grade I or uncontracted Braille codes. This file goes directly to the merge step preceding BRAILLE5, since it requires no further processing. Words flagged as French and not yet translated go into a separate file; they are sorted into alphabetical order, and translated into contracted Braille codes by BRAILLE3. Words flagged as English undergo similar processing, which is not explicitly shown in the flowchart.

After translation, the file of French and English words is sorted back into text order, merged into a single file along with the file of words which did not need translating, and passed on to BRAILLE5 for final formatting. BRAILLE4, which translates English words, is a small PL/1 program using equivalency tables taken from DOTSYS. It was virtually impossible to modify these tables, because they have approximately 5,000 entries, or conversion rules, expressed in a highly symbolic manner. We were thus facing a complex special-use computer language--on with sketchy documentation and no built-in debugging aids.

With our knowledge of DOTSYS firmly in mind, and realising that the historic trend is for machine costs to decrease and programming costs to increase, we decided to make BRAILLE3 as easy as possible to maintain or modify, while keeping operating costs within reason. The basic element of this program is a relatively simple routine which determines the abbreviated version of a word by comparing the file of known contractions to the text file. Since both files are in alphabetical order, this process is quite fast and cheap. When a text word is in the file of known contractions, a straight substitution is made as many times as the word occurs, and the entry from the known contractions file is passed on to a new file of known contractions. At this point garbage collection is also performed to remove from the file contractions which haven't been used in a specified number of runs of the program. When a word is not in the file of known contractions, the above process is carried out virtually unchanged, except that control branches to the largest section of the program, which computes the valid contraction for the word.

The first thing is to compare the word to a list of 795 possible endings. These endings are classed by length--from 2 to 19 characters--so that a check of the length of the word permits limitation of the number of possible endings which need to be examined. Once this is done, the uncontracted part of the word is processed for syllable contractions. We do this using the INDEX function in PL/1 to determine if a contractible string is in the word. If it is, control branches to a labelled position in the program, where tests are made on the environment of the string to see if it can indeed be contracted. After appropriate action has been taken, testing of the word continues until the whole array of 48 possible syllable contractions has been compared to the word.

To facilitate debugging and subsequent maintenance of the program, we have written into it a trace feature which can be turned on or off using a run-time parameter. When this feature is turned on, BRAILLE3 produces a file ready for output through the Braille terminal. Each different word in the text produces one record in the file. This record is a line of Braille containing a word in uncontracted Braille codes, three commas, and the same word in contracted Braille. Parallel printed output contains the same information plus identification of words found in the file of known contractions, identification of endings computed by the program, and a list of possible syllable contractions found. When the contraction was actually made, an asterisk signals this fact (see Figure 3).

The first material run through the program was a data base of 42,283 forms drawn from 1,191,627 words of literary and school texts recorded for other purposes, plus a second data base of 25,367 forms originally recorded for theme study. This latter file contains a large number of learned and specialised words. The results produced by the trace feature of BRAILLE3 were sent to the Reverend Rolland Campbell, Director of Les Editions Braille du Quebec, for checking and comment. This exercise proved both dismaying and gratifying.

We were dismayed to discover that a few dozen input errors, which had escaped all previous checking, were still cluttering the files. We watched with dismay as the basic file of known contractions--containing word abbreviations defined by the Braille system, as well as anomalies or exceptions beyond the capacities of the program--grew from an original 1,500 entries, past the estimated maximum of 2,600 to a total of 3,125. We were also disappointed that it was possible to reduce the endings list only from a maximum of 2,000 to a final size of 795 entries.

We were touched and gratified by the endless store of devotion and hard work expended by Father Campbell and his staff in checking Braille output. This checking process also allowed a practical solution of the "ient" ending problem. This ending is abbreviated "i-ent" or "ien-t" depending on its pronunciation.

To handle this anomaly, we originally planned a hand update stage after BRAILLE3. Our experience has now shown that virtually all words ending in "ient" have only one possible abbreviation. We have found only two exceptions--"il convient, ils convient; un expédient, ils expédient". Thus it is possible to treat these two exceptions as material for flagging on input, and to avoid the necessity for human intervention in the midst of the system's operation.

Most gratifying of all is our certain knowledge that our French Braille translation program works accurately on an extremely large sample of text.

Figure 4 shows the cost of running the system. It was generated from information gathered from production runs on Camus' l'Etranger and Gide's l'Immoraliste. The overhead was estimated by passing a twelve word text through the system. The production costs have been normalised to 10,000 words. Thus the figures for BRAILLE1 of 26.4 seconds of CPU time must be multiplied by 3.498, and the overhead of .42 seconds must be added to arrive at the total cost of running this program on the complete text. The verification of format control commands explains the relatively high cost of this program. Most gratifying of all is the relatively low cost of BRAILLE3, even when one adds in the cost of sorting required by its design. It will be noticed that the Gide text is slightly more expensive per ten thousand words than the Camus text. This is because Gide's text has a considerably richer vocabulary than Camus'. On the other hand, the file of known contractions output by the Camus run was used as input for the Gide text. By doing this we reduced the CPU time requirement of BRAILLE3 by 2.33 seconds per ten thousand words, vis-a-vis the cost if the basic file had been used. The use of the updated file thus produces a savings of roughly 25%. These figures justify our basic strategy for the contraction process.

II. LARGE PRINT

Many people who are legally blind do have a certain minimal amount of sight. They cannot see sufficiently to read an ordinary book, but if the print is large enough, these people can read it. Texts in large print are frequently used in teaching, but they are almost as difficult to obtain as Braille. The reason for this is easy to discover.

Books in large print are produced in very small numbers, using techniques which are almost as labour intensive as those used for Braille production. As a result, the problems resulting from the high costs of materials and of high wages, as well as the decline of volunteer labour, have produced a similar dearth of material in large print.

We have been using a VERSATEC electrostatic printer/plotter to produce texts in large print. This machine works on a principle similar to that of the xerox copier, and as a result can produce a very wide and flexible range of output.

The program which produces texts in large print accepts data in exactly the same format as the programs used for Braille. Thus our staff do not have to learn two sets of rules for preparing data, and, more important, once a text has been put into machine-readable form for Braille or for large print, it can be produced in the other form at virtually no added cost.

The flexibility of the machine which we use allows us to tailor the text in large print to the needs of the intended user. Both the size and the shape of letters can be modified so that they are entirely comfortable to read, but not overly large. Thus, the texts, which tend to be bulky, are no bulkier than they absolutely have to be. If three people have different levels of sight, we can produce texts in three different sizes to accomodate them with no trouble, and hardly any added expense, since the greatest cost arises from putting the text into machine-readable form and proofreading it. The actual printing of the finished product is the cheapest part of the process.

III. VOICE SYNTHESIS

People who work with computers tend to collect vast amounts of printed output. This is because the development of a working and efficient computer program is a trial and error process, usually requiring at least ten or twelve versions and revisions of the draft program. Someone who visits a programmer's office usually comes away with a strong impression that every flat surface is covered by a pile of computer output at least a foot high.

If one were to multiply this volume by five, in order to compensate for the greater bulk of Braille, one would have some idea of the amount of Braille output a blind computer programmer might need to carry on his profession. It should not be forgotten that the heavy paper on which Braille is produced costs about one dollar for 25 pages.

Fortunately, the blind tend not to accumulate written--even Braille--material to the same extent as people who can see. But it can be understood that computer output in Braille is not a practicable solution for the needs of a blind computer programmer in his day to day work.

In response to this need, the standard solution has been the employment of a sighted reader to help the blind programmer. This solution has two flaws. First is the cost, which tends to be open ended and subject to the constant upward pressure on wage rates. Second, and more important, such a solution creates an absolute dependency of the blind programmer on the reader. Not only is the blind person at the mercy of any errors made by the reader, but he is also constantly reminded of his handicap, hardly a happy work situation.

For these reasons, an inexpensive audio response unit was acquired in 1974. This device can produce 64 different phonemes approximating the basic speech sounds, and can utilize one of four different inflections with each phoneme. It is driven by a computer program which specifies a sequence of inflected phonemes which the device produces in order to simulate speech output. During the summer of 1975, two staff members created a phoneme vocabulary corresponding to approximately 1100 words which are frequently used in our computing environment. A mini-computer program which takes as input the character representation of a particular word, finds the matching entry in the vocabulary table, and transmits the corresponding sequence of inflected phonemes to the voice synthesizer was also developed. This program also pronounces standard endings such as "ing", "ed", "s", etc. when these are encountered for words in the vocabulary. Any word which is not found in the vocabulary is decomposed into single characters, each of which is then pronounced individually to spell out the word.

The mini-computer in which the program executes is connected between the blind user's typewriter terminal and the University of Manitoba's main computer. It acts as a filter by transmitting typewriter input unchanged to the host computer, but translates output from the main computer to audio output which is sent to a loud speaker next to the terminal. The blind user can also use a tape recorder to record audio output produced while he is absent, providing a facility roughly equivalent to the use of printed output by a sighted person.

The availability of this audio output capability has effectively freed the blind programmer from dependence on either sighted assistance or Braille output, and has dramatically increased his productivity and independence. Due to the years of experience with this form of computer output, the limited intelligibility of the speech output is quite acceptable to the blind user, but would not likely be acceptable to a person with no training or experience in the use of this facility. It should also be noted that no changes whatsoever were necessary to the software of the main computer system in order to implement this facility.

It is most unlikely that very many installations which were prepared to employ a blind computer programmer would also have such a mini-computer and audio response unit available, and the acquisition cost would be prohibitive in most cases. A very simple and flexible audio response unit has therefore been connected to a micro-computer to produce a prototype computer terminal with audio output capability which communicates with a computer system via standard interfaces. A basic software system similar to that utilized by the mini-computer facility has been constructed, although the limited storage available with the micro-computer presents a serious constraint. The intelligibility of the speech output produced by this prototype is comparable to that produced by the mini-computer, although the flexibility of the former may allow more complex software to produce significantly more comprehensible speech.

The University of Manitoba does not have the facilities to begin manufacturing portable computer terminals with audio output capabilities, and has therefore made information available to social agencies assisting the visually handicapped, and to organizations which manufacture equipment to assist these persons. It would now appear that such a device will be announced shortly by one of these manufacturers, and it is our hope that the cost will be sufficiently low that the social agencies could reasonably be expected to provide such a terminal to each blind computer programmer during the training period.

CONCLUSION

At the University of Manitoba, we have been able to draw on facilities and expertise in a wide range of field. We feel that this has permitted us to make a credible start towards providing the means for the visually handicapped to take their rightful place as active and productive members of society.

REFERENCES

Allen, S. (1970)
Nusvenskfrekvensordbok baserad pa tidningstext I.
Stockholm: Almqvist and Wiksell.

Duro, A. (1973) Elaborations électroniques de textes effectuées
par l'Accademia della Crusca.
Linguistica Matematica ∈ Calcolatori,
pp. 53-75 (ed. Zampolli, A.). Firenze: Olschki.

Engwall, G. (1974)
Fréquence et Distribution.
Stockholm: Skriptor.

Fortier, P. (1973) and McConnell, J. C.
Computer-aided Thematic Analysis of French Prose Fiction.
The Computer and Literary Studies,
pp. 167-81 (eds. Aitken, A.J., Bailey, R.W. and
Hamilton-Smith, N.). Edinburgh: Edinburgh University Press.

Gerhart, W. K. (1971), Millen, J. K., and Sullivan, J. E.
DOTSYS III: A Portable Program for Grade 2 Braille
Translation.
Bedford, Mass.: MITRE Technical Report MTR-2119.

Juilland, A. (1970), Brodin, D., and Davidovitch, C.
Frequency Dictionary of French Words.
Den Haag: Mouton.

Kucera, H. (1967) and Francis, W. N.
Computational Analysis of Present-day American English.
Providence, R. I.: Brown University Press.

Zampolli, A. (1973) La Section Linguistique du CNUCE.
Linguistica Matematica ∈ Calcolatori,
pp. 133-99 (ed. Zampolli, A.) Firenze: Olschki.

FIGURE 1:

DISTRIBUTION OF WORDS IN A FRENCH TEXT

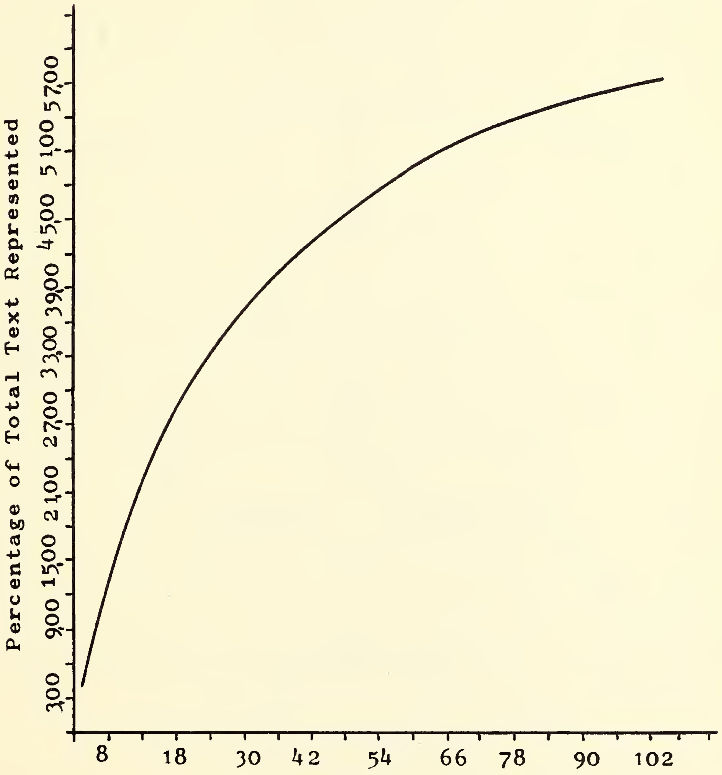


FIGURE 2

THE BRAILLE SYSTEM

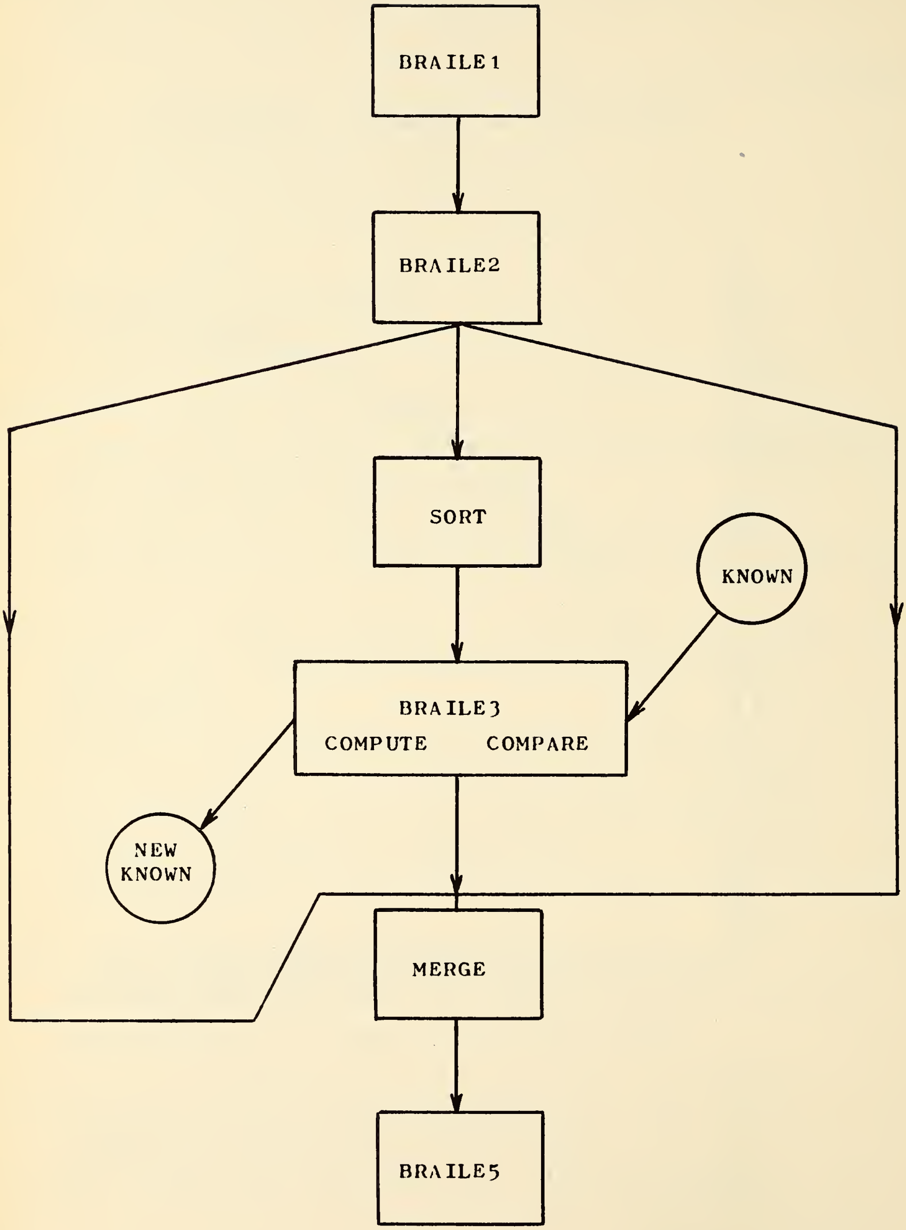


FIGURE 3

BRAILE3: PRINTED OUTPUT

```
OBJET ,,,OJ            **KNOWN CONTRACTION**

ORATEUR ,,,ORAT.       ENDING(EUR) (1)OR

ORCHESTRE ,,,;(ESOE    (*)(1)OR (1)RE (*)(1)TR (1)ES (*)(1)CH
```

FIGURE 4

COST FACTORS

	MEMORY	OVERHEAD		CAMUS		GIDE	
	K BYTES	CPU SECS	I/O	CPU SECS	I/O	CPU SECS	I/O
BRAILE1	120	,42	251	26,4	188	26,8	194
BRAILE2	88	,46	288	14,79	300	14,94	289
SORTS	256	1,73	565	1,75	646	1,87	654
BRAILE3	180	3,17	528	8,39	140	9,13	168
BRAILE5	116	,39	216	5,15	125	5,22	124
TOTALS		6,17	1848	56,48	1399	57,96	1429

World Council for the Welfare of the Blind

Committee on Cultural Affairs

International Conference on

"Computerised Braille Production - Today and Tomorrow"

London, 30 May - 1 June 1979

NATURAL LANGUAGE ASSISTED PROCESSING:
PROJECT PIAF

by

E. Grandjean

Ingénieur de recherche

Laboratoire Imag

BP 53X Grenoble Cedex

France

NATURAL LANGUAGE ASSISTED PROCESSING

PROJET PIAF

E. GRANDJEAN

Ingénieur de recherche
Laboratoire Imag
BP 53X GRENOBLE Cedex
FRANCE

This research topic has been developped since 1971 by l'Equipe Intelligence Artificielle du Laboratoire d'Informatique et de Mathématiques Appliquées de GRENOBLE (IMAG).

In a parallel fashion, this research group is also interested in Robotics problems. The group consist of 15 persons: professors, researchers and engineers from the University and the C.N.R.S.

Computer processing of text information is basically different from other types of processing like administrative or scientific computation for three main reasons: the nature of the information to be processed, the difficulty to formalize composition rules for signs or group of signs, and the volume of data to be stored.

We have observed that the needs of handicaped users, like the blind and others raise problems of a very similar nature.

A - DEVELOPMENT OF BASIC TOOLS

We have tried to define interactive computing tools for the processing of natural languages for non-computer users. The basic processing tools constitute the building blocks necessary for long-term applications. These tools control the data input, the segmentation of sentences in elementary units. The translation of strings, the development of a dictionary and the construction of a grammar.

Application fields are as follows:

- information retrieval (creation and interrogation of a documentation data base)
- text editing and photocomposition
- translation (in basic Braille or phonetics)
- mail processing
- linguistics, philology and lexicology
- assisted instruction
- artificial intelligence

We observed that we use the same basic processes for different types of applications.

These tools must operate in an interactive mode because the data contain human encoding errors which must be corrected during the automatic recognition phase. Furthermore, new words will affect the machine vocabulary build-up and the linguistic model development. This interactive approach requires a new design of internal data structure and a definition of corresponding algorithms and programs.

The programs we have developped, have been kept as general as possible by separating the algorithmic part (the mechanisms) from the linguistic part (the parameters). This approach renders possible processing of different natural languages by replacing only the linguistic parameters without modifications in the programs.

Finally, since in most applications users have no computer experience, one must use a type of hardware adapted to teleprocessing and a software (PIAF) with a simple control language constituting a guide to the users.

Design of the PIAF project

We have defined two classes of problems for which we have developped application programs:

- processing tools for character strings
- processing tools for tree structures.

1) The strings processing is subdivided in procedures:

- recognition or segmentation while refering to a character string dictionary
- control and validation of this segmentation with a finite-state automaton
- string generation and translation.

2) Tree processing is subdivided in procedures:

- construction of sentence dependancy structure with a set of relations on grammatical categories
- validation or filtering of these structures with respect to concordance rules for grammatical variables
- transformation of these structures (paraphrases)
- semantic evaluation (concept extractation).

3) A particular application will guide the choice of the necessary tools and the definition of the linguistic parameters. An artificial language has been developped to describe the various parameters and the following procedures:

- definition and consistency control
- verification by generation
- general application (files).

4) The finite state automaton includes a set of models as reference to the dictionary and a set of rules with validation and saturation leading to a final state and/or the search of other segmentations.

The rules allow the transduction of grammatical properties of strings from the string under analysis.

B - DEVELOPMENT OF APPLICATION PROJECT AND PROTOTYPE CONSTRUCTION

All operational applications, or under development, can be accessed with terminals directly connected to the University de Grenoble computers (CII - HB and IBM) or through a network (CYCLADES). Various types of terminals can be used, either printing of display.

We also use a terminal with Braille embossing (SAGEM) and a key-board terminal (with 6 Braille keys) with temporary Braille dot output which enable blind users to interact with the computer, since there is a Braille dot representation for each alphanumeric character.

1 - At the request of CNAM (Conservatoire National des Arts et Métiers), we have developped an application for translation in basic Braille, grade II, (PIAFRDV). For this application, we have used the string processing tools, a dictionary of word segmentation and a set of rules formalizing the condensed French writing in the PIAF language.

This program could be used with a dictionary and abreviation rules for another language.

The texts to be translated are input from a key-board terminal (SAGEM) with all accented vowels or from a magnetic file. During this input operation, errors can be introduced (hélas!) and we had to develop a control program (PIAFCTL) with a dictionary of accented French and corresponding grammar. This program controls the data input before translation.

2 - At the request of la Documentation française, a government organization, we have developped an operational version of a data preprocessor for the construction of a documentary data base (PIAFDOC). For this application, we had to take care of the following:

- control of documents to be input in computer storage
- maintain a systematic indexation with a key-word spelling

standardization (descriptors) for the documentary data base description.

Furthermore, it is advisable to use the same tools for question analysis (interrogative sentences) and for document retrieval with the same key-words.

3 - Grammars have been written for phonetic transcription in two research projects:

a/ automatic correction of phonetic spelling errors (invariant) (PIAFCOR)

b/ matching of surname files with three levels of serial transcriptions (PIAFPHO).

4 - For our own purpose, we have used the string-processing procedures (PIAFTRAD) to translate our programs from an IBM language to a CII - HB language.

C - DEVELOPMENT OF RESEARCH PROJECTS

It has become evident, after our several operational applications, that there exists a need for new developments and new basic tools based on real data instead of limited examples.

1) The translation experiment in basic Braille has illustrated the need to be able to control the input of extended character sets. This linguistic model will have useful applications in automatic documentation and text processing.

2) Programs and parameters for basic Braille could be adapted to microprocessors to facilitate its utilization by several local organizations in various cities.

3) To facilitate written information communication between blind and non-blind people, we plan to develop an inverse translation with the same grammar rules and an inverse algorithm. Multiple answers will be filtered by the extended character set model.

4) We are planning to develop an automatic index generator by concrete information extraction. This will extend our research in semantic evaluation, and will become an important component of text processing.

5) The CNET (Centre National d'Etudes des Télécommunications) is considering the use of the tree processing tool for the recognition of prosodic markers necessary for speach synthesis.

6) We would like to cooperate with teachers to develop a natural language teaching system for adults with use of the string and tree processing tools:

a/ for physically handicapped or deaf people
b/ for the blind (basic Braille teaching)
c/ for the teaching of a second language.

CONCLUSION

The development of software for natural language assisted processing will play an important role in easing access to test information stored in computer memories and in editing this kind of information.

Present heterogeneous computer networks will increase the use of terminal for consulting various documentary data bases.

Specialized terminals and microprocessors will help in the development of shortterm projects in the areas previously mentioned.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

A COMPUTERBASED SYSTEM FOR PRODUCTION OF BRAILLE MUSIC

by

John Humphreys

Warwick Research Unit for the Blind
University of Warwick
Coventry CV4 7AL
England

Summary

A computer-based system has been developed for production of braille music without the use of a skilled braillist. The system is designed to be used by an operator having no knowledge of braille or computers. Some ability to read printed music is helpful, but not essential, and no expert musicianship is required.

The input to the system is a coded representation of music notation typed on the keyboard of a graphical display unit by an operator reading printed music. Syntax checking is performed by the program during input to help reduce the number of errors made by the operator. The music is displayed graphically on the screen as it is typed, in a form similar to the printed copy, giving the operator immediate feedback and an opportunity to correct any mistakes.

The music, once stored, may be edited until correct using the keyboard of the terminal in conjunction with a joystick or light pen. Any part of the stored music may be rapidly accessed and displayed on the screen or a digital plotter for checking purposes. The stored music may be archived temporarily or permanently on magnetic tape for subsequent retrieval, in both print and braille forms.

The translation algorithm performs automatic translation to a close approximation to the international braille music code. Although the system is intended primarily to be used without the aid of a braillist, it does allow access to, and editing of, the braille before final output.

The output from the system is the braille music notation corresponding to the input, suitable for embossing directly on an braille terminal connected on-line to the computer. The quality of samples of music of various types produced in braille using this system has been evaluated with the assistance of a number of braille music readers.

The system is suitable for implementation at a braille printing house, and incorporates a general-purpose editor for music notation which may also be useful for other musicological applications.

A Computer-based System for Production of Braille Music

by John Humphreys

Introduction

This paper describes a computer-based system for production of braille music without the aid of a skilled braillist. It has been designed to meet the following primary objectives:

(1) the user (the person encoding the printed music) should need to have no knowledge of braille or computers and, as far as possible, little expertise in music;

(2) the braille produced should correspond as closely as possible to that which would be produced by a manual transcriber; this corresponds in general to the international braille music code defined in the official manual (Spanner 1956), as amended by the 1975 addendum to the manual (APH 1975);

(3) the system should be portable, since it must be installed at one or more braille printing houses if it is to be of practical value.

There are three major stages involved in automating braille music production (see figure 1). Firstly, the music notation as printed must be entered into the computer and represented in a digital form. Secondly, this representation must be converted into a corresponding braille representation. Finally the digital representation of the braille must be converted into tangible braille.

The work described below is concerned with the first two stages. The third stage is common to automation of all forms of braille and presents no additional problems specific to braille music. Several models of braille computer terminal are already available for this purpose. In the prototype system, two of these (Triformation LED120 and SAGEM TEM 8BR) have been used to emboss the braille music, but other devices, including stereograph machines for embossing zinc plates, might equally well be used in a production version of the system.

The major part of the system has been developed as a single computer program, CIMBAL, which performs all operations from initial input of printed music to the creation of a file representing the final braille output, suitable for copying to an on-line braille embosser. The program also includes facilities for temporary or permanent archiving of both print and braille representations on magnetic tape. The prototype is implemented on a Sigma 5 computer.

Background

A number of computer programs have been developed in various countries for translation of text to contracted literary braille. Considerably less work has been done on automation of the specialised braille codes, including music. This is partly because of the greater complexity of the task and partly because of the small demand for the specialised codes. The number of readers of braille music in the United Kingdom, for example, is estimated at 500 to 1000, compared with about 12,000 readers of all forms of braille.

The introduction of computers for translation of text to literary braille has resulted in an increase in production capacity, while braille music production remains dependent on skilled transcribers whom braille printing houses find are becoming increasingly difficult to recruit. The decline in the number of skilled music braillists led the American Printing House for the Blind (APH) in 1971 to investigate the possibility of using a computer for braille music production (APH 1972-5, Watkins & Siems 1976). The RUMBA (Representations Utilising Music Braille Alphamerics) system, one of two which emerged from this work, is the only computer-based system which is currently used for regular production of braille music. Music is encoded by input specialists who require four months training and a considerable knowledge of the braille music code.

Other computer systems for braille music production have been developed at the American University, New York (Patrick & Friedman 1975, Patrick & Patrick 1976), at the Association RITM-Braille, Paris, and at the Hatfield Polytechnic, England (Wilkinson 1975, 1976). Each of these projects is experimental, or has not been completed because of a lack of funds.

These systems have generally been described by their developers as performing "computer-assisted" rather than "automatic" braille translation, because in each case the user must have some knowledge of the braille music code in order to produce a good quality transcription. None of them meets all of the three objectives mentioned above. In view of the gradually diminishing number of skilled music braillists, there remains a need for a system which will satisfy these requirements.

Input

The input of text to a computer for automatic translation to literary braille is fairly straightforward, using standard keyboard devices such as the card punch, teletype or visual display unit. The only significant problem is the method of specifying how the braille is to be laid out. There is no analogous natural and widely-used method for input of music notation.

There has been interest in digital encoding of music since the early 1960's, largely in the United States, for various purposes including computer-aided composition, music analysis, automated printing of music, and cataloguing of musical themes.

The music has usually been encoded using intermediate input languages based on the character set of a standard keypunch. One such language, DARMS (Digital Alternate Representation of Music Symbols), has achieved a more widespread use than the others (Erickson 1977). Being originally designed for automated music printing, it represents western music notation comprehensively. A proposal has been made to use music encoded in the DARMS language as a source for braille translation (McLean 1976).

Most users of the alphanumeric encoding method have found that, although use of the keypunch is cheap and convenient, it is not entirely satisfactory. In particular, they have found that proof-reading is extremely difficult. Mistakes in encoding are easy to make and harder to identify subsequently because the encoded form is not merely a copy of the original print.

Several other methods for input of music have been developed experimentally by various researchers, although none of these has yet been used to create a large corpus of music in digital form. These include: music typewriter, on-line sampling of an organ-type keyboard, optical character recognition, and graphical display unit (GDU).

The music typewriter is similar to an ordinary typewriter, with music symbols instead of, or as well as, ordinary characters. The music typewriter produces a digital encoding on punched cards or paper tape as well as a printed copy of the music typed. It is slow to use, since notes and other symbols are built up from their individual printed components.

The use of a digitally sampled keyboard is attractive from the point of view of speed. The keyboard is played by the user as if it were a piano. The position of each key (depressed or not depressed) is sampled at a typical rate of 20 to 100 times a second. This method is, however, almost entirely limited to specifying the pitch and duration of notes. There is much additional information required in braille music which would have to added by another method.

Optical character recognition would, in principle, be an ideal method, at least for music already in print, because it eliminates human intervention, giving less scope for error. At least one experimental system has been developed for recognition of printed music using a flying-spot scanner (Prerau 1975), but it recognises only a restricted set of symbols, which excludes letters, and appears to need excessive computation time. The development of such a system is a major task beyond the scope of the work on braille music described here. However, if and when the state of the art of optical character recognition is sufficiently advanced, it will provide an effective supplement or replacement for the method of input which has been developed for the braille music system.

On-line graphical display units have been used to specify music notation in a more natural form by selecting symbols from a displayed menu, and moving symbols around the screen using a light-pen or similar device. This is slower than alphanumeric coding for an experienced user. Use of a refreshing display limits the amount of music which may be displayed at one time.

The method of input which has been developed for the braille music system is based on a hybrid of the alphanumeric encoding and graphical methods. This combines the potential speed and comprehensiveness of alphanumeric encoding with the easier checking and correction allowed by on-line graphics. It also enhances portability of the system by avoiding the non-standard hardware necessary for the other methods mentioned above.

Some form of graphical display is highly desirable in any case for adequate proof-reading of music, which is essentially a graphical notation. Either a GDU or a digital plotter could be used for this purpose, although the latter is comparatively slow. However, extending the use of the GDU to include initial input of music as well as subsequent proof-reading simplifies the detection and correction of errors at the time they are made. This eliminates the additional time necessary to locate the position of an error when it is corrected at a later stage.

The input language is similar to previous such languages insofar as it is based on a largely mnemonic coding. It differs from them in that it is interactively interpreted by the computer program, which is thus able to check the input as it is typed and inform the user immediately of detectable errors. The program detects all syntax errors and some semantic errors, a syntax error being a character which, in a specific context, cannot be interpreted within the formal definition of the input language, and a semantic error one which is meaningful but musically incorrect. The prime example of the latter is a wrong total duration of notes within a measure.

Although the user is assisted in this way by the program to avoid many potential mistakes, there are still many errors which cannot be automatically detected. It is essential in any computer system involving human input that adequate allowance should be made for correction of such errors. Comprehensive editing facilities are therefore an important aspect of the braille music system.

As with computer text editing systems, there are two stages in error correction. Firstly the position of the error must be specified, and secondly the correction must be made. The measure is taken as an intermediate unit in identifying the location of an error, by analogy with the line of text in a text-editing system. A measure, specified by its bar number and (for multi-stave music) stave number may be displayed on the screen for editing purposes. Individual items within the measure may then be identified by use of a joystick, light pen or equivalent device for addressing positions on the screen. Changes are specified using the ordinary input language in conjunction with a set of editing commands (for example, "insert", "delete").

The use of the measure as a basic unit for storage of the music leads to a number of problems, but there is no other more suitable unit, especially for braille translation, in which repetition is based on multiples of the measure.

Braille translation

The problems of automatic translation of music to braille differ in several ways from those involved in translation of literary braille which may, for the most part, be performed adequately by a single sequential scan of a linear text, the translation at any point depending only on a limited context on each side of that point and a small set of state variables dependent only on translation of earlier text.

Braille music translation cannot be treated in this sequential manner because of the nature of the abbreviation devices used in the braille music code. The two major devices used to achieve space saving in the braille are doubling of signs and the use of repeat signs. Both of these take advantage of the repetetive nature of music notation, rather than the contraction of frequently occurring elements as in literary braille.

The principle of doubling is used in literary braille for the italic sign. This has usually been dealt with in automatic translation by instructing the user to specify the doubling explicitly in the input text. This is justified by the infrequency of the italic sign and the additional complexity needed in the translation algorithm to determine when to double the italic sign. In braille music, where signs such as staccato marks often apply to several consecutive notes, doubling of signs occurs much more frequently. It is unreasonable to expect the input operator to learn the complex rules involved, so doubling must be evaluated automatically.

This could be achieved by buffering of input or output while the program determines whether a doubled sign or separate single signs are required. In practice, it has been found more convenient to identify the range of a potentially doubled sign at its first occurrence.

The use of repeat signs for whole measures or sequences of measures makes it essential to use a random-access form of input. A repeat sign may refer back to any part of the music already translated. To identify repeated passages automatically using a purely sequential method would require a prohibitively large number of passes through the input.

The use of repeat signs is additionally complicated by the fact that they do not necessarily correspond to exact repetition in the printed notation. For example, where a passage is repeated with different dynamic markings it may be permissible to use a repeat sign preceded by the new dynamic marking. Similarly, where a dynamic marking is not repeated in print because it is still in effect, a braille repeat sign may still be used.

In deciding whether or not to use a repeat sign, the transcriber has a considerable degree of discretion in achieving a balance between space saving and ease of reading, which may conflict with each other. In the automatic translator, two relevant factors have been used in trying to achieve this balance: distance between the original passage and its repeat, and the amount of space saved.

The two-dimensional nature of music notation introduces problems in defining a sequence for translation. The interrelation between the different staves makes it desirable to translate all staves in parallel. On the other hand some aspects of the braille music code, for example part-measure repeat signs, doubling, and grouping of notes, make it desirable to consider the music on one stave together. The compromise which has been adopted is to translate the staves in parallel, one measure at a time. The context for each stave is saved while the remaining staves are translated.

Major problems arise from the way in which braille music is laid out. The rules of braille music may be divided into two groups: those defining the braille signs and their sequence, and those defining the layout of the braille on the page. A two-pass translation algorithm has been developed to reflect this division. The intermediate "unformatted" representation of the braille produced by the first pass is available to the braille music expert for editing if this is considered necessary. This allows changes having a major effect on the layout of the braille which could not easily be made in the final formatted output.

The two-pass approach has the additional advantages that overall program size is reduced, extra internal buffering of braille is avoided, and the formatting module may be replaced to provide alternative layouts with only minor alterations to the main translation algorithm. It does, however, introduce some problems because the sequence of braille signs and their layout are interdependent to a certain extent. For example, a note may need an octave sign if it is the first note of a braille line but not otherwise. The dependence of the layout on the sequence of cells is handled by the inclusion of format controls in the unformatted braille. The dependence of the sequence of cells on the layout is mainly resolved by including alternative forms in the unformatted braille, the correct form being chosen at the time of formatting.

"In-accord" parts (in which two or more parts are shown on a single stave) cause problems because the parts appear in parallel in print but sequentially within each measure in braille. This may lead to the omission of accidentals shown in print or the inclusion in the braille of accidentals not shown in the print. The ordering of signs for in-accord parts has not been implemented automatically but instead passed on to the user in the form of rules for order of input.

Where accidentals marked in print are redundant, they are normally omitted from the braille to save space. This is achieved in the translation algorithm by comparing the true pitch of each note with the pitch it would have if no accidentals were shown, and omitting any marked accidentals from the braille if these two pitch values are the same.

Apart from the problems mentioned above, there are a number of minor difficulties which have been solved by more or less straightforward means. In general, the rules of the braille music code are diverse in nature; they do not readily lend themselves to the type of neat table-driven algorithms which exist for literary braille translation.

There are many situations where the formal rules are inadequate to define a correct transcription. In such cases the manual transcriber uses previous experience and general principles to decide on a suitable braille rendering. Because it would be extremely difficult to ensure that an automatic translator could produce correct braille music in all circumstances, the system includes comprehensive facilities for editing the braille after translation, so that a braille expert can if necessary modify details in the braille.

Evaluation

The input side of the system has been tested by a typist with no knowledge of music or braille and a group of nine musicians with no knowledge of braille. Individual members of this group had up to 20 hours overall experience in input and editing of music, in sessions averaging two to three hours.

Users were each given a few minutes training followed by supervision during an initial session. Users who had more than one session worked alone after the initial session, noting occasional points of difficulty for later clarification. All users were able to produce almost accurate stored music after proof-reading and correcting their own errors. The mistakes which they did not detect were due mainly to insufficient instructions concerning the order of certain items. These mistakes would probably disappear with further training and practice.

The rate of input was slow for all users in the preliminary session but gradually increased as they gained experience. The non-musician was averaging the equivalent of about 8 braille lines per person-hour (LPH) after five to ten hours experience. Comparison with her early use of a computer text editing system used for a pilot computer-based short braille document service suggests that this rate would increase significantly with further practice. The musicians' average rate was higher, increasing from an average of about 10 LPH during an initial session to about 18 LPH after some 20 hours experience. For comparison, braille music transcribing staff working in pairs at the Royal National Institute for the Blind (RNIB) transcribe at an estimated rate of about 12 LPH.

My own average rate, with reasonable though not complete familiarity with music notation and extensive experience of entering music into the computer, was up to 30 LPH. This gives some indication of the speed a regular user with good knowledge of music might be expected to achieve, although this average was taken over sessions of only about one hour each.

The quality of the braille transcription was evaluated by two methods. Firstly, samples of braille music produced using the system were distributed to braille music readers to determine the acceptability of the braille to typical readers. Secondly, test pieces were compared with equivalent braille produced manually by the RNIB or given as examples in the braille music manuals.

It emerged from the comments of readers that there are some differences of opinion between experts on the finer points of the braille music code, particularly where layout is concerned. It also became clear that in some cases the manual gives insufficient guidance on the correct transcription, or that modern usage differs from that defined in the manual.

However, most readers agreed on most points. In the later stages of development of the system they were generally satisfied with the standard of transcription for the simpler forms of instrumental music. No detailed feedback has yet been obtained from readers on the more complex keyboard music or vocal music, except from one or two readers.

The portability of the system has not yet been fully tested, but the program has been successfully compiled with minor modifications on a second computer (Burroughs B6700). No major problems in transfer are anticipated since no non-standard hardware is needed apart from a graphical display unit and braille terminal and the bulk of the software is written in standard FORTRAN. The program has been designed for use on a small or medium-sized system, the overall size of the prototype program being 115K bytes, overlaid to run in 34K bytes. The use of single character input is the only feature likely to cause difficulty on more sophisticated systems.

Conclusions

The small-scale evaluation suggests that the major objectives have largely been met. Users have successfully input music of a fairly complex nature without any knowledge of braille. A few braille rules concerning the order of signs have, however, been effectively passed on to the user in the form of instructions regarding the order of input. Although a person with minimal musical knowledge can use the system successfully, possibly with occasional annotation of the printed music by a musician, musicians are capable of typing music more efficiently. They can achieve an overall transcription rate up to twice that of manual transcribers.

The quality of instrumental braille music produced is good. Test pieces of single-line music for a solo instrument have been transcribed with no known errors. Simple keyboard music has been produced to a standard which authoritative readers of braille music have found acceptable. There are some minor flaws which remain to be corrected in the transcription of more complex instrumental music. The transcription of vocal music is at present less satisfactory and requires further work to improve the layout of the braille. The main problem here concerns the interaction of words and music rather than simply the contraction of literary text, for which many acceptable algorithms already exist.

The system now appears suitable with minimal modification for use at a braille publishing house, supplementing existing braille music transcription resources. A long-term evaluation and development of the system is recommended in parallel with regular production in order to identify and eliminate any remaining deficiencies.

Acknowledgments

The work described in this paper was financed by the Science Research Council of Great Britain and carried out between 1976 and 1978 at the University of Warwick. The paper is based on a thesis submitted to the University of Warwick for the degree of Doctor of Philosophy (Humphreys 1979).

References

American Printing House for the Blind. Computer-Assisted Translation of Braille Music; Progress Reports: May 10, 1971 - May 9 1972; June 28, 1972 - June 27, 1973; September 1, 1974 - November 30, 1974; December 1, 1974 - February 28, 1975; March 1, 1975 - May 31, 1975; June 1, 1975 - August 31, 1975. Louisville, Kentucky, 1972-5.

American Printing House for the Blind. Revised International Manual of Braille Music Notation 1956, 1975 American Addendum. Louisville, Kentucky 1975.

Erickson, Raymond F. DARMS: a Reference Manual. Queens College, New York, 1977.

Humphreys, John. A Computer-based System for Production of Braille Music. Ph.D. thesis submitted to the University of Warwick, May 1979.

McLean, B. Translating DARMS into Musical Braille. Braille Automation Newsletter, no. 2, pp. 65-67, August 1976.

Patrick, P. Howard and Patricia Friedman. Computer Printing of Braille Music Using the IML-MIR System. Computers and the Humanities, vol. 9, pp. 115-121, 1975.

Patrick, P. Howard and Rosalind E. Patrick. Computers and Music Braille. Braille Automation Newsletter, no. 2, pp. 52-61, August 1976.

Prerau, David S. Do-Re-Mi: A Program that Recognises Music Notation. Computers and the Humanities, vol. 9, pp. 25-29, 1975.

Spanner, H.V. Revised International Manual of Braille Music Notation 1956, Part 1, Western Music. World Council for the Welfare of the Blind, Paris, 1956.

Watkins, William and John Siems. SAMBA and RUMBA: Systems for Computer Assisted Translation of Braille Music. Braille Automation Newsletter, no. 2, pp. 47-51, August 1976.

Wilkinson, Sally. Investigation into the Feasibility of Computer-Assisted Transcription of Music into Braille Notation. M.Sc. Project Report, School of Information Sciences, The Hatfield Polytechnic, September 1975.

Wilkinson, Sally. Computer-Assisted Transcription of Braille Music. Braille Automation Newsletter, no. 2, pp. 68-71, August 1976.

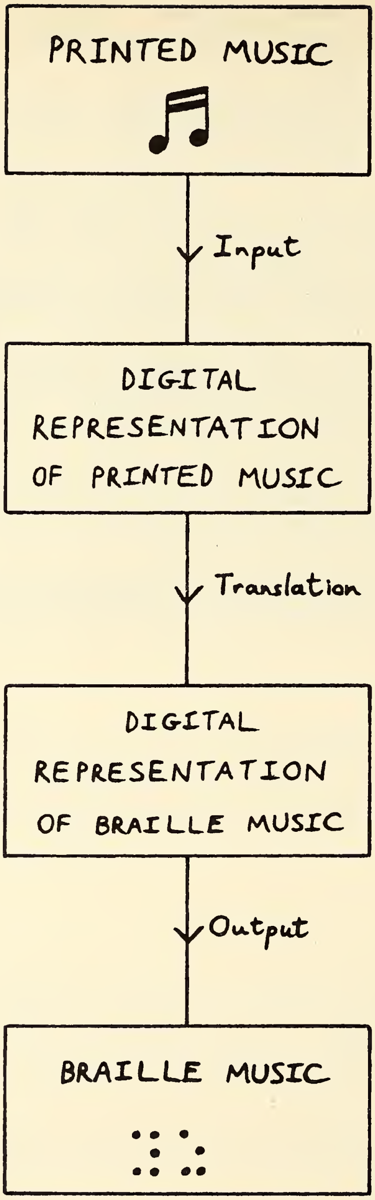
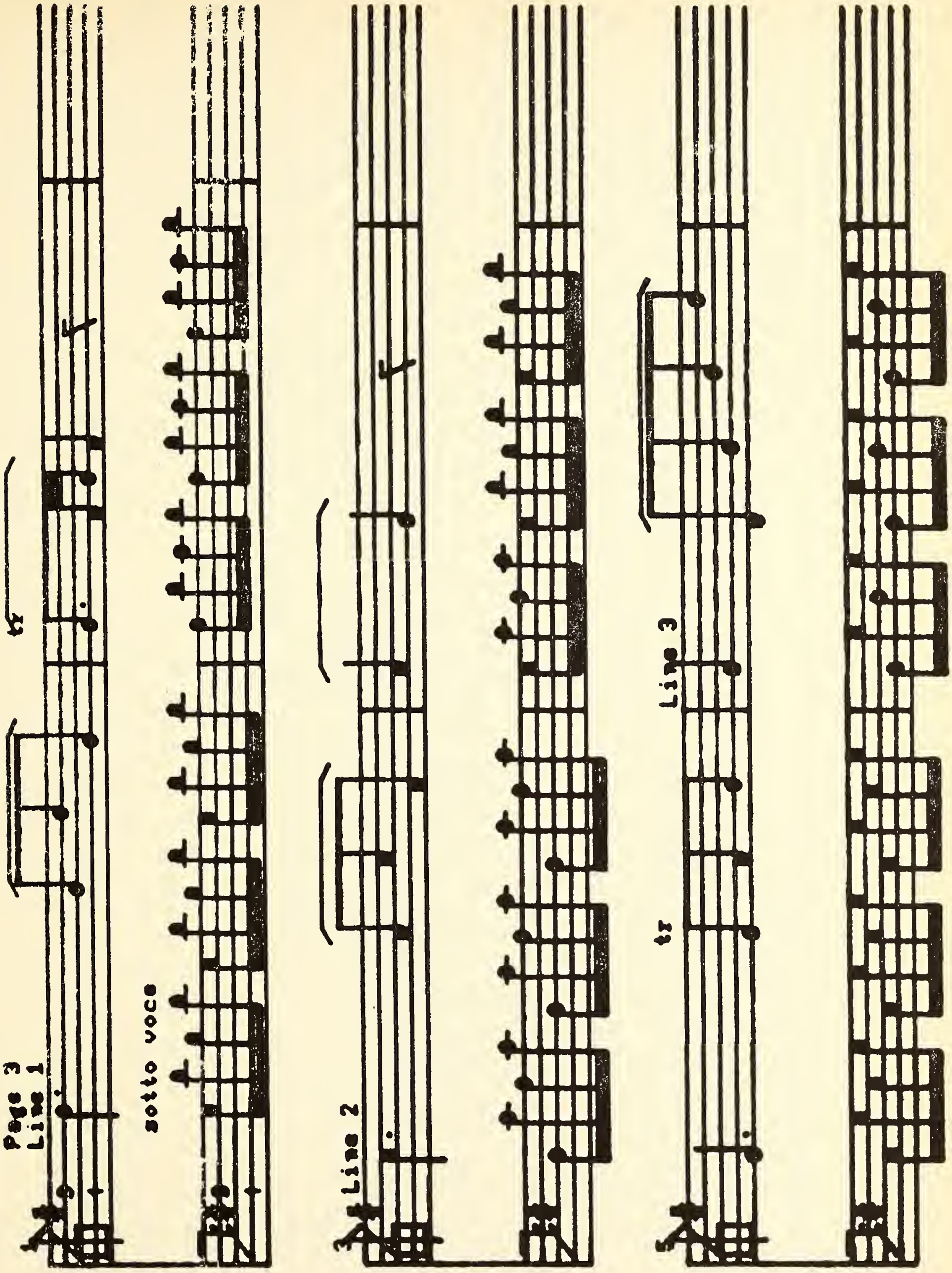

PRINTED MUSIC
Input
DIGITAL REPRESENTATION OF PRINTED MUSIC
Translation
DIGITAL REPRESENTATION OF BRAILLE MUSIC
Output
BRAILLE MUSIC

Figure 1



Example of printed music displayed on the screen of the graphical display unit

Example of braille music displayed on the screen of the graphical display unit

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

BRAILLE TRANSLATION BY MICROCOMPUTER

by

Ian Davidson

Dept. of Electrical and Computer Eng.
University of Cincinnati
Mail Loc. No. 30,
Cincinnati, Ohio 45221
USA

Braille Translation by Microcomputer

Ian Davidson *

Abstract:

The use of a microcomputer system is being studied for carrying out the conversion between English in the form of individual characters and grade 2 Braille.

The short-form words and the special rules have been studied, and it is expected that they can be implemented in about 4K (1096) bytes of ROM (Read Only Memory). A special feature is the method of compressing the data within the memory space. Also in order to achieve a reasonable speed of operation, a high speed search system is being considered. This is based on a hashing method of finding entries in the table which is considerably faster than a linear search. It is hoped to be able to achieve real time grade 2 Braille with normal typing speed input of the data.

Introduction:

Braille is a very important means of communication for blind people, representing a very powerful method for reading printed material. However the translation from script to grade 2 Braille by manual means requires a very specialized knowledge, and cannot be justified unless the document is very important or is required for a large circulation.

It has been recognized that it would be worth while to utilize computers to carry out the translation between normal written English and grade 2 Braille. A number of programs[1,2] have been written to do this. During the last few years there have been major changes in the availability of computers, especially with the advent of microcomputers. The cost of computing has been reduced to a point where it is economical to have a computing system dedicated to the translation process. The concept is that a sighted person could type in English script, without necessarily any experience of knowledge of grade 2 Braille. The computer would look for contractions and exceptions and take care of the special punctuation and annotations with the result being output

---

* Assistant Professor, Department of Electrical and Computer Engineering
Cincinnati, Ohio, U S A

on a Braille embosser. With such a system it becomes economical to translate individual documents, notices, lecture notes, memos etc., and bring the availability of Braille material to a much wider population.

Grade 2 Braille and the use of computers:

A Braille character, consisting of a pattern of up to six raised dots, can represent 64 different symbols (including the space, for which no dots are present). In the original concept of Braille, there was a one-to-one correspondence between ink-print symbols and their Braille counterpart. However this has evolved to a contracted form, known as grade 2 Braille, where there are 189 contractions and short-form words, which have been chosen to correspond to the words and syllables that are most frequently used. It should be noted that although there is a well-defined list of contractions and short-form words, there is a less well-defined list of exceptions. Grade 2 Braille is based on the phonetic use of the language, and in many cases the grouping together of the letters is dictated by the use of the language. This will vary from country to country, as well as varying with time to keep in step with language evolution. The present accepted standard for use in the U S A has been published[3] and the interpretation of this document by computerized means will be discussed. This is a standard form of Braille that can be learnt by blind people, but there are many instances where individual people have developed their own abbreviations that they have found to be more suitable.

Definition of the codes:

The set of 64 Braille characters can be represented in a computer by 6 binary bits. The format used corresponds very closely to that for asynchronous data communication, often referred to as the ASCII (American Standard Code for Information Interchange) standard. This is also used by Triformation Systems Inc.* in their printer unit. The complete set of codes is given in table 1 and in table 2. Table 1 lists the codes in terms of the 6 bit codes, expressed as 2 octal + digits, corresponding to the right hand six

---

*Triformation Systems Inc. - Model LED-1 Braille terminal

+The octal representation of a number uses digits between 0 and 7, each digit showing the binary value of three binary bits.

bits (2 digits) of the ASCII code. The Braille pattern is given for each code, together with the ASCII symbol. In table 2 the codes are listed in terms of the conventional grouping of Braille patterns. In this it can be seen that, except for the alphabetic characters, the ASCII characters do not necessarily correspond to the ink-print characters.

Computer system:

The computer system for translating English text into grade 2 Braille is shown in figure 1. The data would normally be entered into the system via the input terminal, although the system could be operated with information that has previously been entered into a storage unit, and is made available as input to the processor. The processor will perform the following operations:

1. Translate the input characters into the correct format for the corresponding Braille patterns.
2. Search for character strings that can be represented by the grade 2 Braille contractions.
3. Generate the correct translation and format for punctuations and annotations.
4. Look for any exceptions to the rules, changing the format to fit the normally accepted grade 2 Braille output.

The signal is then sent to the Braille printer, to produce the embossed output. The CRT screen of the terminal can be used to verify the input information, and to show the operator the translated Braille output. Any manual changes can be made at this time, to be sure that the output is in the correct form.

It is possible for a computer to take care of items 1, 2 and 3 above, and in most cases item 4. If the translation, the contractions and the conversion of punctuation etc. can be defined by non-phonetic rules, then the process can be computerized. Most of the 'exceptions' can also be placed in this category, but there are some which it is virtually impossible to process without operator intervention. The following are two examples of such operations:

1. When considering the abbreviations that are associated with numbers in the form of 'coinage, weights and other special symbols' (Rule VIII), the sentence 'the table has 4 feet on the ground' would be ! TA# HZ #F FEET ON ! GR.D

or in Braille patterns ⠮ ⠞⠁⠼ ⠓⠵ ⠼⠋ ⠋⠑⠑⠞ ⠕⠍ ⠮ ⠛⠗⠨⠙ Whereas 'the room is 12 feet long' would be ! ROOM IS FT#AB L;G or ⠮ ⠗⠕⠕⠍ ⠊⠎ ⠋⠞⠼⠁⠃ ⠇⠰⠛ The identification of the word 'feet' as being the thing on the end of the legs as distinct from the measurement of length can only be made by the meaning of the sentence, and cannot be defined by computer-identifiable rules.

2. The sentence 'you should mind your ps and qs' should be translated as Y %D M9D YR ;P'S & ;Q'S or in Braille patterns ⠽ ⠩⠙ ⠍⠔⠙ ⠽⠗ ⠰⠏⠄⠎ ⠯ ⠰⠟⠄⠎ The words ps and qs are abbreviations (for pints and quarts) and are therefore written in the form shown. There is no method, except by having a table of all possible abbreviations, by which the computer could pick these out as being abbreviations as distinct from the words 'ps' and 'qs'.

Software structure:

The software flow chart for the computer program is shown in figure 2. The input data is parsed to pick out words in the form of strings of alphabetic characters. It should be noted that ' (apostrophe) is included as an alphabetic character so that words such as 'o'clock' can be abbreviated.

The search for abbreviations is carried out as follows: For words having more than ten characters, the first ten characters are compared with a table of abbreviations, and if a match is found, then this group of letters is extracted and output in Braille form. If no match is found, the process is repeated with the last ten characters. If still no match is found the process is repeated with a string of nine characters at the front and back of the word, and so on until the string has been reduced to two characters only. If this is unsuccessful, the first and the last letters of the word are removed and translated into Braille, and the process repeated until all the characters have been translated. As an example, the translation of the word 'showering' is shown in appendix I. This word can be abbreviated as (sh)(ow)(er)(ing) or %[]+ or ⠩⠪⠻⠬

Words can therefore be broken down into abbreviations, making sure that the maximum length abbreviation is used, with preference being given to the beginning and end of the word. Exceptions can be taken care of by the use of additional words in the look-up table. For example, as well as 'ch' or * or ⠡ and 'cc' or 3 or ⠒ there would be an additional word 'cch' or C* or ⠉⠡ which would give preference to 'ch' over 'cc'. For example 'saccharine' would be SAC*>9E or ⠎⠁⠉⠡⠜⠔⠑ and not SA3H>9E or ⠎⠁⠒⠓⠜⠔⠑ It should be noted that the ending 'ing' can be either incorporated as part of the word or be given as a specific abbreviation. For example 'conceiving' would be found as a short-form word as 3CVG or ⠒⠉⠧⠛ , whereas 'blinding' would be BL+ or ⠒⠇⠬

Search process and the abbreviation table:

The critical part of this software program is the search routine for abbreviations. As can be seen above, many searches may be required for one word, for example the word 'showering' required a total of 24 searches before the word could be translated. It is assumed also that the system must only have a small memory, so that the direct look-up of all words would be impractical. Furthermore the search must be made in the shortest possible time. A process known as 'hashing' is used to find the likely position of the word in the table. It has been found that if the first two letters of all the abbreviations are extracted and grouped as shown in table 3, each group only contains a small number of words, so that the required word can be found quickly. The table shows the groupings for the first and second letters, and the number of abbreviations and short-form words in each category. This table was generated by heuretic means. The largest category, with 14 entries, has the first letter STUV and the second letter GH and has entries 'sh','th', 'the', 'that', 'this', 'shall', 'their', 'there', 'these', 'those', 'should', 'through', 'thyself' and 'themselves'. The number 62 comes from STUV being group number 6 of the first letters, and GH being group number 2 of the second letters, and it will point directly to the position in the overall list of words where the required grouping of words are to be found.

Structure of the abbreviations table:

Within each section of the table, the abbreviations are arranged in ascending order of length. A string of characters need not be searched beyond the point where the table has more letters than the input word. The following example can be used to illustrate the operation of the search process. Consider a search for a match to the input sequence 'thy'. The following steps would be taken:

1. 'thy' having 'th' as the first two letters would have a hash value of 62, and it would therefore go to this section of the table.
2. 'thy' would be matched with 'sh', but the comparison would be rejected as the number of letters in the two groups are different.
3. 'thy' would be matched with 'th', but this would be rejected for the same reason.
4. 'thy' would be matched with 'the', and the matching process can continue as both have three characters.
5. The 't' of 'thy' is compared with the 't' of 'the'. These are the same so the process can continue.
6. The letter 'h' of the two words are compared, and the process continues.
7. The 'y' of 'thy' is compared with the 'e' of 'the'. No match is found, so the process will be repeated with the next word.
8. 'thy' is compared to 'that'. However now that 'thy' has less letters than the word 'that', the search is abandoned with an indication that the word 'thy' is not in the abbreviations table.

Associated with the entry of each abbreviation is data that gives the required output Braille. There must also be codes that give the length of the words as well as an indication that all entries in a given section have been tested. Each entry in the table consists of one byte (8 binary bits, usually represented by 3 octal digits) and they can be categorized as follows:

a. 01 0xx xxx Alphabetic characters, for example the letter Z would be 01 011 010 or 132 (octal). Note that in this table the ' (apostrophe) is assigned the code 01 000 000.

b. 00 yyy yyy Braille character, with 6 variable bits corresponding to the 64 possible Braille patterns. For example & has the code 46, i.e. 00 100 110, with a Braille pattern ⠯

c. 01 1zz zzz This is an identification before each abbreviation, giving the number of letters in the abbreviation. The value given by the z bits is one less than the actual number of letters in the word. For example, the word 'should' would be prefixed with 145 (octal) or 01 100 101.

d. 10 000 000 This code indicates that the end of the abbreviations in a given group has been reached, and therefore that no match has been found.

e. 1w www www This indicates that the end of the section of memory allocated for a given group has been reached, and that the remainder of the group is available elsewhere. The number given by the w's gives the magnitude of the jump forward in memory that is required to continue searching this set of abbreviations.

A section of the abbreviations table is shown in Appendix II. This is the section with the initial letters CD-'ABCD, and which contains 5 words. The entry point is given by the hash number 10 (see table 3). The first byte, 141, indicates that the following word contains two letters, and the next two bytes show these as 'cc' by the codes 103 and 103. If this abbreviation is required, then the next byte, 063, gives the Braille code as 3 or •• The process is continued until the last byte is reached which has a value of 200.

Hardware implementation:

Figure 3 shows in more detail the structure of the hardware. The processor is connected to the terminal and the Braille embosser via the interface units. It is also connected to the RAM (Random Access Memory) and the ROM (Read Only Memory). The RAM is used as a scratch pad area, and includes buffer areas for the input and the output. The ROM is divided into two parts, the program containing the software instructions and the table areas.

The system has been implemented on a Digital Equipment Corp. PDP11 computer, and although the initial work was carried out on an 11/60 system, it is designed to be transferred to an LSI 11 or 11/03, this being the microcomputer version of the 11 series. The memory requirement when installed on this type of machine is as follows:

1. RAM (buffer area) approximately 216 bytes
2. ROM (for program) 480 bytes (excluding the I/O routines)
3. ROM (for tables) Hash number table - 54 bytes
   Abbreviations table - 4096 bytes

It should be pointed out that the amount of program ROM would have to be increased if a more structured approach is made to the 'exceptions'. Also the abbreviations table in its present form can be contained in 2048 bytes, but it has been expanded to the larger size to allow for the entry of more words so that as many exceptions as possible can be taken care of.

It is planned to implement the program in other microprocessors such as the Motorola M6800, the Intel 8080 or 8085, the Zilog Z80 and the Texas Instrument 9900. In all cases the RAM buffer area and the ROM tables will be unchanged. However most machines will require more ROM for the programs, but all are expected to be less than 1K (1024) bytes in size.

Current work on the system:

The hardware-software combination described in the paper is at present being evaluated, with changes being made to improve its performance in the generation of grade 2 Braille. However the use of the abbreviations table as a solution for finding all 'exceptions' is being questioned. An alternative method, where an identification number is generated alongside the abbreviation or short-form word, is being considered. This number would be used in a heuristic algorithm to check to see if the abbreviation is required, or in the case of two alternative abbreviations, which one should be chosen. When this comparison is complete, the better of the two methods will be incorporated in the final system.

One application of the translation program is in the generation of grade 2 Braille on a paperless Braille strip. For this purpose real time operation is essential as the input would be from a standard typewriter keyboard. The system is based on the use of a Texas Instrument silent 700 terminal, which contains a 9900 microprocessor. The conversion of the software to this unit will present no real difficulties as there will be no changes in the logic flow of the software.

Conclusion:

A description of a software program for the conversion of English text into grade 2 Braille has been given, that is suitable for implementation on a microcomputer. A hashing method has been proposed for a rapid search for abbreviations, based on the distribution of their first two letters. The resulting system is not only fast, but fits well with the memory availability of low-cost microcomputing systems.

Although most of the 'exceptions' for grade 2 Braille can be accomplished, further investigation is required into other methods in an attempt to improve its performance. It is hoped to be able to produce a high quality output document in Braille without the utilization of specialized operators.

This work was made possible from funds made available by the University of Cincinnati, and by the use of the LED-1 Braille terminal from the Bureau of the Services for the Blind in Cincinnati. I should also like to acknowledge the help and guidance of Josch O'Wan for useful discussions on the correct form for the Braille output.

References:

1. Schack, Ann S and Mertz, R T. Braille Translation Systems for the IBM 704, New York: I B M Data Systems Division 1961
2. Gerhart, W R, Millen, J K and Sullivan, J E. A portable Program for Grade 2 Braille Translation. The Mitre Corp. Bedford, Mass. 1971
3. English Braille - American Edition 1959 American Printing House for the Blind, Louisville, Ky. 1977

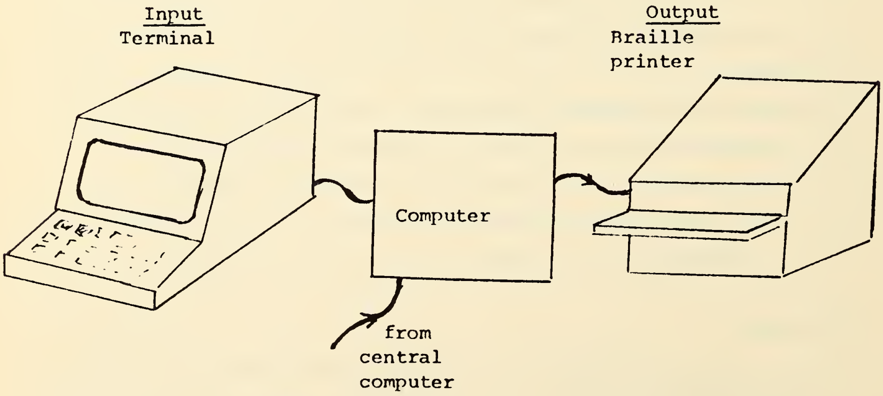


Figure 1: Braille Translation Computer System

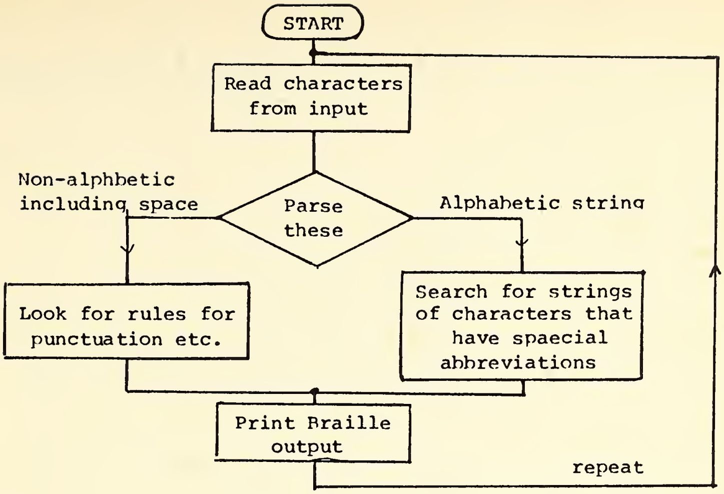


Figure 2: Software flowchart for the translation process

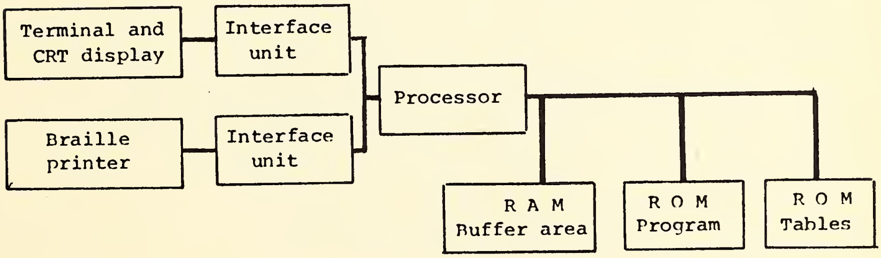


Figure 3: Hardware structure

Table 1: ASCII symbols and Braille patterns arranged by code numbers

ASCII	code	symbol	Braille	ASCII	code	symbol	Braille
100	00	@	⠈	040	40	space	
101	01	A	⠁	041	41	!	⠮
102	02	B	⠃	042	42	"	⠐
103	03	C	⠉	043	43	#	⠼
104	04	D	⠙	044	44	$	⠫
105	05	E	⠑	045	45	%	⠩
106	06	F	⠋	046	46	&	⠯
107	07	G	⠛	047	47	'	⠄
110	10	H	⠓	050	50	(	⠷
111	11	I	⠊	051	51	)	⠾
112	12	J	⠚	052	52	*	⠡
113	13	K	⠅	053	53	+	⠬
114	14	L	⠇	054	54	,	⠠
115	15	M	⠍	055	55	-	⠤
116	16	N	⠝	056	56	.	⠨
117	17	O	⠕	057	57	/	⠌
120	20	P	⠏	060	60	0	⠴
121	21	Q	⠟	061	61	1	⠂
122	22	R	⠗	062	62	2	⠆
123	23	S	⠎	063	63	3	⠒
124	24	T	⠞	064	64	4	⠲
125	25	U	⠥	065	65	5	⠢
126	26	V	⠧	066	66	6	⠖
127	27	W	⠺	067	67	7	⠶
130	30	X	⠭	070	70	8	⠦
131	31	Y	⠽	071	71	9	⠔
132	32	Z	⠵	072	72	:	⠱
133	33	[	⠪	073	73	;	⠰
134	34	\	⠳	074	74	<	⠣
135	35	]	⠻	075	75	=	⠿
136	36	^	⠘	076	76	>	⠜
137	37	_	⠸	077	77	?	⠹

Table 2. ASCII symbols and codes arranged by Braille patterns

Braille notation	Code	ASCII code	ASCII symbol	Ink-print interpretation
	40	040	space	Space
(patterns formed with dots 1,2,4 and 5)				
⠁	01	101	A	A or 1
⠃	02	102	B	B, 2 or BUT
⠉	03	103	C	C, 3 or CAN
⠙	04	104	D	D, 4 or DO
⠑	05	105	E	E, 5 or EVERY
⠋	06	106	F	F, 6 or FROM
⠛	07	107	G	G, 7 or GO
⠓	10	110	H	H, 8 or HAVE
⠊	11	111	I	I or 9
⠚	12	112	J	J, 0 or JUST
(patterns formed with dot 3 added to patterns with dots 1,2,4 and 5)				
⠅	13	113	K	K or KNOWLEDGE
⠇	14	114	L	L or LIKE
⠍	15	115	M	M or MORE
⠝	16	116	N	N or NOT
⠕	17	117	O	O
⠏	20	120	P	P or PEOPLE
⠟	21	121	Q	Q or QUITE
⠗	22	122	R	R or RATHER
⠎	23	123	S	S or SO
⠞	24	124	T	T or THAT
(patterns formed with dots 3 and 6 added to dots 1,2,4 and 5)				
⠥	25	125	U	U or US
⠧	26	126	V	V or VERY
⠭	30	130	X	X or IT
⠽	31	131	Y	Y or YOU
⠵	32	132	Z	Z or AS
⠯	46	046	&	AND
⠿	75	075	=	FOR
⠷	50	050	(	OF
⠮	41	041	!	THE
⠾	51	051	)	WITH

Braille notation	Code	ASCII code	ASCII symbol	Ink-print interpretation
(patterns formed with dot 6 added to patterns with dots 1,2,4 and 6)				
⠡	52	052	*	CH or CHILD
⠣	74	074	<	GH
⠩	45	045	%	SH or SHALL
⠹	77	077	?	TH or THIS
⠱	72	072	:	WH or WHICH
⠫	44	044	$	ED
⠻	35	135	]	ER
⠳	34	134	\	OU or OUT
⠪	33	133	[	OW
⠺	27	127	W	W or WILL
(patterns formed with dots 2,3,5 and 6)				
⠂	61	061	1	, (comma), EA or final letter combination
⠆	62	062	2	; (semi-colon), BE, BB or final letter combination
⠒	63	063	3	: (colon), CON or CC
⠲	64	064	4	. (period), DID, DD or final letter combination
⠢	65	065	5	EN or ENOUGH
⠖	66	066	6	! (exclamation), TO or FF
⠶	67	067	7	) or ((brackets), WERE or GG
⠦	70	070	8	" or ? (opening double quotation or question mark) or HIS
⠔	71	071	9	IN
⠴	72	072	0	" (closing double quotation), WAS or BY
(patterns formed with dots 3,4,5 and 6)				
⠌	57	057	/	/ (bar, oblique stroke, fraction line sign) WAS or BY
⠬	53	053	+	ING
⠼	43	043	#	# (number sign) or BLE
⠜	76	076	>	AR
⠄	47	047	'	' (apostrophe) (note: internal code of 00)
⠤	55	055	-	- (hyphen) or COM
(pattern formed with dots 4,5 and 6)				
⠈	00	100	@	accent sign
⠘	36	136	^	initial letter combination
⠸	37	137	_	initial letter combination
⠐	42	042	"	initial letter combination
⠨	56	056	.	. (decimal point or italic sign)
⠰	73	073	;	letter sign
⠠	54	054	,	capital sign

Table 3: Abbreviation grouped by first letters

(number of abbreviations in each group)

second letter first letter		0	1	2	3	4	5	6	7
		'ABCD	EF	GH	IJKL	MN	OP	QRST	UVWXYZ
0	'AB	4	13	2	8	2	0	4	2
1	CD	5	4	4	1	0	6	0	0
2	EFG	3	1	2	2	3	3	4	3
3	HIJ	2	2	0	4	4	0	4	0
4	KLMN	2	6	0	2	2	4	0	3
5	OPQR	4	9	0	1	3	0	0	10
6	STUV	1	1	14	3	1	9	3	1
7	WXYZ	1	1	4	2	0	9	0	0

Appendix I

The analysis of the word 'showering' The letters underlined indicate the string that is being compared with the abbreviations table. Unless otherwise indicated, no match is found. Note as there are nine letters in the word (i.e. less than ten) the search is started with a search for all nine letters.

showering
showering
showering
showering
showering
showering
showering
showering
showering
showering
showering
showering
showering

A match is found, the string 'ing' being denoted by + or ⠬

shower
shower
shower
shower
shower
shower
shower
shower

A match is found, the string 'sh' being denoted by % or ⠩

ower
ower
ower
ower

A match is found, the string 'ow' being denoted by [or ⠪

er

A match is found, the string 'er' being denoted by] or ⠻

The final reduction of 'showering' is (sh) (ow) (er) (ing)

i.e. % [ ] + or ⠩⠪⠻⠬

Appendix II

An example of part of the abbreviations table, showing the entries for group 10, with the initial letters CD - 'ABCD (5 entries)

	Octal	binary	Interperetation
Entry point	141	01 100 001	The following abbreviation has 2 characters
	103	01 000 011	Letter C (abbreviation CC)
	103	01 000 011	Letter C
	063	00 110 011	Braille code 63, i.e. 3 or ⠉
	141	01 110 001	2 characters to follow
	104	01 000 100	Letter D (abbreviation DD)
	104	01 000 100	Letter D
	064	00 110 100	Braille code 64, i.e. 4 or ⠙
	142	01 100 010	3 characters to follow
	103	01 000 011	Letter C (abbreviation CAN)
	101	01 000 001	Letter A
	116	01 001 110	Letter N
	003	00 000 011	Braille code 03, i.e. C or ⠉
	142	01 100 010	3 characters to follow
	104	01 000 100	Letter D (abbreviation DAY)
	101	01 000 001	Letter A
	131	01 011 001	Letter Y
	042	00 100 010	Braille codes 42 and 04, giving the initial
	004	00 000 100	letter contraction "D or ⠐⠙
	145	01 100 101	6 characters to follow
	103	01 000 011	Letter C (abbreviation CANNOT)
	101	01 000 001	Letter A
	116	01 001 110	Letter N
	116	01 001 110	Letter N
	117	01 001 111	Letter O
	124	01 010 100	Letter T
	037	00 011 111	Braille codes 37 and 03, giving the initial
	003	00 000 011	letter contraction _C or ⠸⠉
	200	10 000 000	End code, i.e. end of group

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

BRAILLE TRANSLATION BY MICRO-COMPUTER AND A PAPERLESS BRAILLE DICTIONARY

by
Osamu Sueda

Dept. of Biol. Engng.
Faculty of Engng. Sci.
Osaka University
Toyonaka, Osaka 560
Japan

1. Introduction

The biggest problems of Braille books compared to the originals are volume and weight. For example, the Braille dictionary, translated from a concise English-Japanese dictionary of about 300g, becomes about 50kg and 71 volumes. Because of these problems, a dictionary has been rarely used by most blind students. The fact that a blind student attends a foreign language class without pre- and post-study with a dictionary is very irregular and it shows that education of the blind has been left in a poor state.

The study of a new Braille system at Osaka University originated from the development of the tactile television system for the blind in 1971. The trials of Braille display on vibrotactile television Mark II and the reading tests of the blind subjects were carried out in 1974. The results were satisfactory: the blind user could easily understand with short-term training, although every blind person was not able to read.

The development of the paperless Braille system was begun as a regular topic of research in 1976. The research consisted of two themes, one was the production of paperless Braille books and the another was the development of reading aids for the paperless Braille books. The former theme has become larger as it includes the reproduction of conventional paper Braille books.

2. Coding of Braille Character

In Japan, sighted people use two kinds of phonetic symbols, 46 Hiragana characters and 46 Katakana characters, and a large number of ideographs, 2,000 - 3,000 Kanji characters, whose origins were the old Chinese characters.

On the contrary, blind people in Japan use only one kind of phonetic symbols, 47 Kana characters, as regular Braille characters.

Japanese contains many words of the same pronounciation with different meanings. Therefore, Japanese Braille sentences of phonetic symbols are sometimes misunderstood and the explanation of a word in English-Japanese dictionary sometimes cannot be well understood.

In order to improve the imbalance between sighted and blind persons, Kanji Braille characters of 8 dots were proposed and trial teaching of them has been carried out since 1970 by groups of schools for the blind. In order to ease the introduction of the new Kanji Braille system in the near future, we decided to design the system for both the conventional 6 dot cell and new 8 dot cell.

Two kinds of coding methods are in existence: one is direct mapping, in which each dot of a Braille character is mapped to a binary signal, and the other is numerical coding, in which each Braille cell is named as a numerical string.

2-1. Direct Mapping

Direct mapping code for a 6 dot cell where 6 dots have been mapped onto the lower 6 bits of an 8 bit code also has been reported in Japan.

The new eight bit code is defined as follows: the left four dots in a cell are mapped on the higher four bit nibble and the right four dots in the cell are mapped on the lower four bit nibble as shown in Fig. 1. The first to the sixth dots are same as a conventional 6 dot cell, the dot above the 1st dot is named the 0th dot and the dot above the 4th dot is named the 7th dot. The 0th dot is used as the beginning indicator of a Kanji character and 7th dot is used as the end indicator of the Kanji character, which is constructed by one, two or three 6 dot cells.

It will become difficult to treat data files in an 8 bit computer system if all of the 8 bit codes are defined as data codes. Fortunately, in the new Kanji Braille system, the three codes of just 0th dot, just 7th dot and just 0th and 7th dots have not been defined as characters. Four codes, these three and the null code of non-dot, are defined as special codes as follows:

(i) Space code of Braille is defined as (08) in hexadecimal, which corresponds to the code of just 7th dot.

(ii) New line code, or carriage return code and/or line feed code, is defined as (80) in hexadecimal, which corresponds to the code of just 0th dot.

(iii) Hexadecimal number (00), or null data, is not defined as a Braille code and this code (00) is used as leader, trailer or separator of a file in the computer system.

(iv) Hexadecimal number (88) which corresponds to the code of just 0th and 7th dots is defined as a special code and/or command header, and this special code is terminated by two contiguous spaces, (08)'s, such as the following:

```
(88)(X1)(X2).......(Xn)(08)(08)
 |   └ Command String    └ Terminator
 |
 └ Command Header
```

The string (X1)(X2).......(Xn) is used as a command by the user of the paperless Braille to communicate with the device, and the command is terminated by the following two space codes.

Two commands are specially defined.

[1] (88)(08)(08) : This set of code is used as item separator or item header, which is used in an item indexing system, such as a dictionary.

[2] (88)(80) : This set of code is used as form feed or page separator.

2-2. Numerical Coding

In the numerical coding method, a Braille character is expressed by a two digit number in octal or hexadecimal and a comma is used as a character separator. For example, "a" ⠁ is expressed as "40,", "z" ⠵ is as "53," and space ⠀ is as "00,". In the numerical coding system, new line and new page codes are defined the same code as in ASCII, "CR" and "FF", respectively.

This numerical coding system is convenient and reliable for transmitting a data file by telephone line or paper tape and easily error correctable in the 8 bit of 7 bit ASCII + Parity, although the data file is three times bigger than the direct mapping file.

3. Braille Key Board

3-1. Two Hand Key Board

In order to input a draft to the computer system, the use of a Kanji-Kana key board which is constructed with several thousands of keys is one manner, and the usage of a Braille key board of 6 or 8 keys is another manner. The latter is easily operable, although the learning of several thousands of Kanji Braille characters would be neccessary.

How to arrange the 8 keys is fundamental and very important. Several key boards for the 8 dot Braille cell were produced and it was discovered in many tests that the simultaneous keying of the 4th finger with the other three fingers was very difficult.

The new Braille key board is basically constructed with eight Braille dot keys and one delete key, of which two keys are placed in the thumb positions of both hands as shown in Fig. 2. All the keys are placed in slightly bent finger tip positions and the two keys for the thumbs are designed as multi-function keys; the right key being used as the 7th dot key in combination with (an)other dot key(s) and the space key when it is used alone, and the left key being used as the 8th dot key in combination with (an)other dot key(s) and the new line key when it is used alone. These key configurations and function definitions reduce the number of keys and increase convenience of operation. The use of long key tops and their rotations of about 17 degrees from right angles, make the positioning of fingers easy for many operators whose sizes are different, and permit considerably imprecise touch on key tops.

The sampling of key-in is started at the time when any one key is pressed, and continues until all keys return at reset position. The bit corresponding to each key pressed during the sampling period is set one.

3-2. One Hand Key Board with Tactile Feedback

In order to input data with one hand while reading the draft with the other hand, a one hand Braille key board has also been developed for a blind operator. Key tops are arranged as a Braille cell rotated 90 degrees, and the additional new two dots for an 8 dot cell are a little separated from conventional 6 dots, to prevent mistouch of the new keys. A character is keyed into the computer system by pressing the 1st entry key with the right thumb after setting, i.e. keying, any dots, usually by the 1st, 2nd and/or 3rd fingers. The second entry key is pressed for keying the special code, as CTRL + A in ASCII, with setting dots of "A". The data register for each key has a toggle function and the status of the register changes by the initial key pressing and returns to original state by a 2nd key pressing, to correct missetting before entry key pressing, and all registers are reset by entry key pressing.

A tactile feedback mechanism to notify the operator of the status of each data register is installed in this key board. It is evident that for the feedback of the key status, it is most natural and intuitive to situate the signal on the same finger tip used to press a particular key. A pin of about 3 mm in diameter is installed at the center of each data key top and it appears about 1 mm above the key surface, pressing into the finger tip at the reset state of the data register and it is pulled in by a magnet beneath the key surface when the data register is set.

4. Braille Display for the Blind

The biggest task in the development of the paperless Braille system is how to display a Braille character to the blind user. If we used a Braille display device with the form similar to a conventional Braille book, 4,320 pins and drive mechanisms to push them up and down individually would be neccessary to display one page of a Japanese book with 18 lines per page and 30 characters per line. Just this display unit would cost about $13,000, if we estimate the cost of each pin with its driving circuit at three dollars. How to design a covenient Braille display unit inexpensively can be easily understood as the first obstacle to making the paperless Braille system useful.

Braille display methods would be classified as the following two: the first is static display, in which an user reads the Braille by sliding his finger(s) on the display surface of one page, several lines or several characters, the same as conventional Braille book reading, the other is dynamic display, in which a user puts his finger(s) on the display surface and reads the sentences displayed as the sentences flow, without sliding his finger(s) on the device, similar to the reading of a light news board by a sighted person.

If we consider the reading of Braille, it is evident that the area to be recognized by the finger tip(s) at any one moment is a small part of a page.

In order to decrease the cost of the display unit and in accordance with our experiences with dynamic display of patterns from the study of tactile television, we chose the dynamic display method as the first area of study.

As a disply unit, we used 4 x 16 = 64 pins of the image display unit of tactile television RAB-III, which has 16 x 16 = 256 vibrating pins arranged with a 2 mm pitch in two dimentions similar to OPTACON of TSI.

Experiments with the following three kinds of dynamic display methods were made, with 10 kinds of texts of random characters and sentences, by 5 blind subjects during about a total of 60 hours:

(i) Text automatically flowed with a preset constant speed;

(ii) Speed and direction of the flowing text were controlled by subject;

(iii) A display unit on which the subject put his fingers was moved along the text line, similar to reading a Braille book, although with the display unit being slid on the desk, while, on the contrary, fingers are slid across the Braille book in conventional reading.

The results showed that the finger reading ability was very low in the case of passive reading, as in (i) (similar to the reading of a light news board on which the text was displayed by constant speed), and would not be useful. Results of the second method showed considerably better performance, although the subject could not easily find the desired position in the section or in the chapter when skipping several lines. The best results were obtained with the third method. In case of the third method, it was observed that the subjects could read the texts smoothly with training of several minutes or about a quarter hour, and it was shown that the production of a useful paperless Braille book reader would be possible by the dynamic display method of vibration dots. We are now making studies of vibration modes and dot separation.

5. Braille Translation and Editing

Two Braille translation and editing systems have been produced; one is aimed mainly for use in Braille data keying, and the other is aimed for use in regular editing.

5-1. A Braille Translation System

A Braille translation system using microcomputer, shown in Fig.3 , is developed for use in keying of Braille texts and also for simple editing and error corrections. This system is especially designed for a non-professional computer operator, i.e. an amateur such as a teacher of the school for the blind, at a low cost using the parts of a microcomputer system for a hobbyist. The data from a key board is monitored on a TV picture tube and recorded on a compact audio cassette tape (Kansas City Standard) in 8 bit data. The CPU is Intel Co.'s 8080A, and the video adapter is the Video RAM TK-80BS from NEC.

Three kinds of programs for this system were developed as follows;

(1) A program for 8 dot Braille : This is an editor program for the new 8 dot and the conventional 6 dot data from just a Braille key board, in essence, and this program treats the data in a page block.
The title and the page number are set from the Braille key board, and the maximum number of characters in a line as well as the maximum number of lines in a page can also be changed using the Braille key board. When the number of characters in a line reaches the maximum, a warning tone notifies the operator of having reached the limitation,

and the insertion of another character in the same line is blocked. The division of the data into a page block is automatically made according to the maximum line number.

(ii) A program for 6 dot Braille : This program has functions of editor similar to the preceding program and auto new line function which creates new lines and new pages according to a given maximum number of characters in a line and a maximum number of lines in a page. The recording and the playback on/from cassette tape are made in a page block and the computer system can treat 10 pages of data in its memory at any one moment.

(iii) A program for JIS key board : This program has the translation functions of normal to Braille sentences and Braille to normal sentences. By the auto-translation functions, the operator, who has learned only the spacing rule of words in Japanese Braille, can work as a Braille translator with a microcomputer system of owned as a hobby, which has a JIS key board and a video adapter. Last summer, we asked for volunteers for this system from a microcomputer club, and a Braille correspondence course is now in progress.

As insurance against a decrease in the number of volunteers, year by year, this system will be of benefit in developing volunteers for the new method of translating with a microcomputer.

5-2. Editing System

For regular editings and error corrections of Braille data, a minicomputer system is developed with a Braille key board, a JIS key board, a Braille display unit (BDU), a CRT, a Braille sheet reader and so on, as shown in Fig.4 . By this system, a blind or sighted operator can edit the data in on-line with the BDU and the CRT and also in off-line, with a Braille sheet printer which prints out an ink dot Braille, and with a Braille sheet puncher which produces a conventional Braille sheet and/or a zinc plate for massproduction. The Braille sheet reader (BSR) is developed for data input of the Braille sheets already produced. This BSR is a reader for a single-sided Braille sheet of the 6 and 8 dot Braille cell.

Three kinds of programs which were necessary to produce the paperless Japanese-Japanese dictionary were developed: a program to convert the character string of an item into a number string, a program to rearrange the items into numercal order and a program to produce the index file which is important for random access to an item. Using these programs, the first trial production of a paperless Braille dictionary was created on August 23rd, 1978, which contains a part of a new Japanese-Japanese dictionary of about 120,000 Braille characters and 2,084 items. This dictionary will finally contain about 40,000 items and will be recorded on two floppy disks. In the near future, several kinds of dictionaries will be produced in the form of floppy disks and the concise English-Japanese dictionary will become four floppy disks of 170g as shown in Table 1.

6. Reading Aids for the Paperless Braille Book

Trial reading aids for a paperless Braille book were produced using a minicomputer. Further studies into display methods and communication methods between the user and the device are now being carried on.

A Braille display unit is constructed by combining the image display unit of tactile television RAB-III and a Braille key board on a sliding bed.
The four keys on the upper part of the key board are command key, and they are named "P", "N", "C" and "D" from left to right, respectively: "P" is used to display the previous page, "N" is used to display the next page, "C" is used as a command header key and "D" is used to excute a command according to the string keyed in by the Braille key board.

The mean time from "DO" key pressing to display is about 0.59 sec.; a computer produces a number string from an item string, searches and reads a few blocks of data in a floppy disc and rearranges and displays a page in the form such that a given item is arranged at the top of a page of 24 lines and its explanation follows.
If the given item is not in the dictionary, an error message of "No such item as xxxx." will be displayed, and the nearest string will be used as an item and one page will be reconstructed and displayed. If the objective item is not in the particular volume, an error message of "See vol. xxx." will be displayed. These new functions are not in conventional dictionaries, neither Braille nor original, and these are very beneficial to the user

of the dictionary. These functions have been realized only since the introduction of computer technology.

The reading aids will be constructed using a micro-computer, a floppy disk drive unit, a cassette tape recorder, a Braille display unit, a Braille key board and a speaker functionally.

7. Conclusion

In the area of paperless Braille systems, several devices using a compact audio cassette recently have become available for use as a memo note and a book reader. However, the devices which have a random access function will be more convenient if the problems of weight and cost are solved.

According to our studies, a mean access time per item in a conventional Braille book was between 30 sec. to one minute. If we would be permitted to estimate the total figure of improvement by the simple multiplication of the values of each improvement, the total figure of improvement of a English-Japanese dictionary becomes about 12,000,000: an improvement in speed being 100, an improvement in weight being 292 and an improvement in volume being 416 as shown in Table 1. With such revolutionary improvement as this, things would become possible which were impossible with conventional systems, such as an increase in educational opportunities and fuller social participation for the blind.

A new study into how to transmit Braille data by telephone line has been started this year in our laboratory.

With wide use of the paperless Braille system in the near future, a blind person will be able to get the latest Braille book in his home over the telephone.

The studies mentioned above have been partially carried out in conjunction with a staff of Nippon Light House.

References

(1) O. Sueda: Construction of Character-Pattern Recognition Aids for the Blind (I), Technical Reports of IECE of Japan, MBE-74-22, 55-61, 1974.

(2) O. Sueda: Development System RAB-3A for the Tactile Television for the Blind, Technical Reports of IECE of Japan, Vol.77, 210, 45-48, 1978.

(3) O. Sueda: Applications of Mini- and Micro-Computer for the New Braille System for the Blind, JAACE Technical Reports, A-78-09-3, 7-8, 1978.

(4) O. Sueda: A Proposal for the Standardization of the Digital Braille Recording on Compact Audio Cassette, Proc. of Meeting on the Unification of Standards for Digitally Encoded Braille, Appendix VI, UNESCO Headquaters, Paris, France, 2-3 May, 1978.

0th dot ○ ○ 7th dot
1st dot ● ● 4th dot
2nd dot ● ○ 5th dot
3rd dot ● ● 6th dot

0th dot	Bit 7 (MSB)
1st dot	Bit 6
2nd dot	Bit 5
3rd dot	Bit 4
7th dot	Bit 3
4th dot	Bit 2
5th dot	Bit 1
6th dot	Bit 0 (LSB)

Fig. 1. The direct mapping of an 8 dot Braille cell.

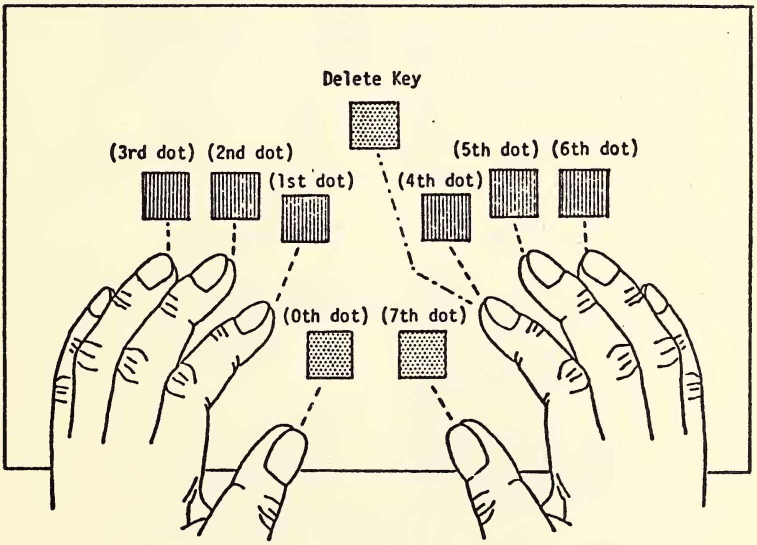


Fig. 2. A basic key configuration of the two hand Braille key board for both 6 and 8 dot Braille system.

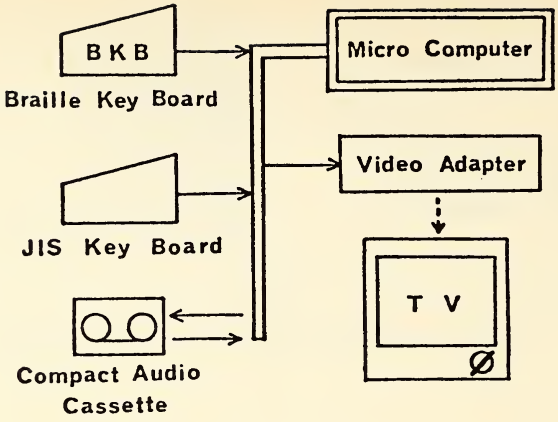


Fig.3 A block diagram of the Braille translation device of a microcomputer; one of two key boards is used for the key input device.

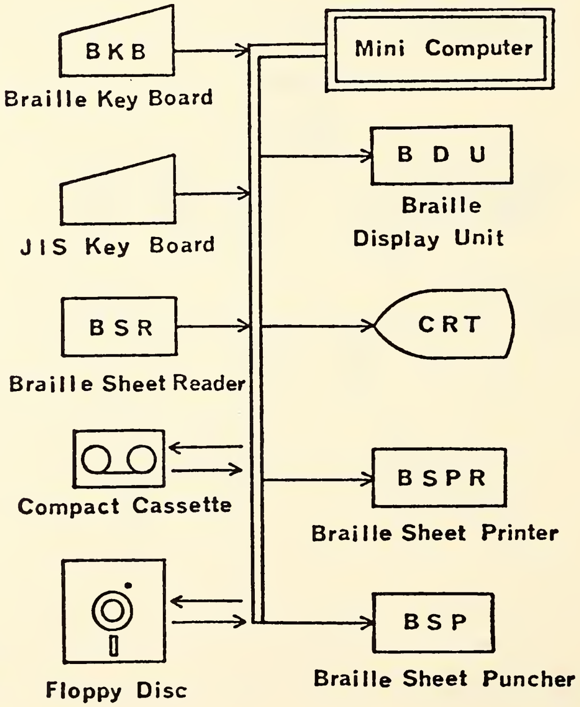


Fig.4 A block diagram of the Braille editing and error correction system using minicomputer for regular use in a printing house.

Table 1. Comparison of three English-Japanese dictionary.

	Original	Conventional Braille	Paperless Braille (Floppy disk)
Total weight	300g	49,700g	(170g)
Number of volumes	1	71	(4 sheets)
Total volume	340 cm^3	175,000 cm^3	(420 cm^3)
Number of characters	(?)	(4.7×10^6)	(4.8×10^6)

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Cenference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

BRAILLE FROM A WORDPROCESSING SYSTEM

by
J.A.H. Snelders
and
H. A. Spanjersberg

Delft University
Dept. of Electrical Engineering
PO Box 5031
2600 Delft
Netherlands

Braille from a wordprocessing system

by J.A.H. Snelders and H.A. Spanjersberg
Delft University of Technology,
Dept. of Electrical Engineering.

0. Introduction

In administrative environments a new tool was introduced during the last five years by the appearance of the wordprocessor. Formerly typists were very busy correcting letters, minutes and other documents they had typed, and although some of them could rectify the errors in such a way that the corrections could not be perceived any more, in general the results did not gain in beauty and legibility, so that very often whole pages had to be retyped, by which new mistakes were introduced etc. Moreover, even when the document was typed without errors, the writer of that document very often wanted to make changes in the text reading the finite result. Then the secretary snatched at the glue-pot and a pair of scissors and she recomposed the concept and gave it to the typist again to type the document. Everyone who has experience in this area will remember wiping ladies and angry heads.

It seems a cunning salestalk, but these problems are nullified using a wordprocessor. For now a typist can type a document as formerly, may it be that detected mistakes can be corrected instantly without ink-eraser or correction liquid, but even afterwards corrections can be introduced. This feature does not imply, however, that all documents in the future are sent out without any mistake, for we spoke about detected errors; how many undetected errors will still leave?

Let us now direct our attention to the wordprocessor itself: the wordprocessor is an electronic instrument for text handling, composed of the following devices:

- a keyboard
- a special purpose computer
- an information carrier
- a printer
- optional a display

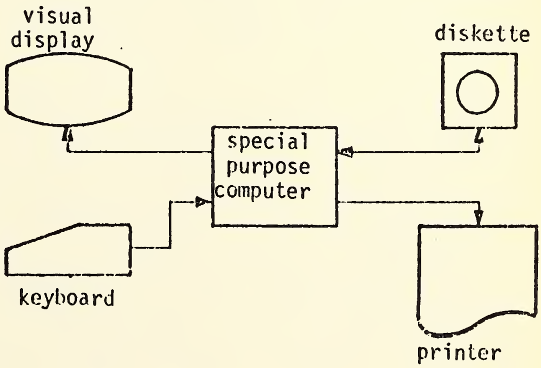


figure 1.

The typist puts in the text by pressing the keys of the keyboard, just as when operating a normal typewriter. By pressing a key an electric signal is sent out from the keyboard with a shape that is characteristic for the symbol on the key. The signals coming from the keyboard are handled by the special purpose computer; that signals can be divided into two categories: namely commands for the computer itself and text to be processed. The signals of the last category are saved up in the memory of the computer until a sufficient number of symbols has been reached, e.g. when a page is full. When this state has been reached, the saved characters are registrated onto the information carrier: usually a magnetic tape, a flexibele magnetic disk (floppy disk) or a magnetic cassette tape are used for that purpose.

The computer can retrieve at every moment any block of characters and bring it back into its own memory. Then characters, words and sentences can be added, removed, transposed etc. When the text is judged to be without mistakes, a print command is given by the operator by which the computer reads the characters from the information carrier and sends them to the printer together with control commands for the printer mechanism. It goes without saying that the operator needs a feed-back from the system about what he or she has keyed in: in general a visual representation of the text is given. The cheap category of wordprocessors realises that feature by the printer, usually being an electric typewriter, the more expensive wordprocessors at the contrary use to that end a visual display, at least it is offered as an optional. The benefits of this device are the quick response on the actions of the operator, the flexibility to scan several distinct areas of the text very rapidly and the legibility of the final result: at every moment the effect of all given commands, the last one included, can be seen.

It will be obvious that the number of features and the speed of execution grows with the amount of money one can spend for the system.

The computer is said to be a special purpose computer. In a number of cases that is the question indeed, because the functions were realized by special digital circuits, designed to perform that task and nothing else. But gradually microprocessors are introduced in this area too, so that the special purpose function has been realized by programs stored in inerasable memories (ROM's).

1. What have blinds to do with wordprocessors?

With the introduction in mind one should be inclined to ask what blinds have to do with wordprocessors, because these devices are fitted up as a tool for sighted people. In itself that conclusion is right, but we believe that with some extensions and modifications the wordprocessor can become a usefull tool for blind people too.

The first application we see is the production of copies in Braille of ink-printed documents. Fortunately blind people gradually get more responsible functions, such as municipal councillor, member of the board in any organization etc. To accomplish that function they have to read, as any one else, a large number of documents as minutes, reports etc. But how does the blind get the disposal of that documents in such a shape that he or she can read them? Let us at first trace how it is done now.

The document can be read by a sighted person and the spoken text can be registrated onto a magnetic tape, just like the talking books. An other way is to braille the documents with a hand driven braille machine. Sometimes there are more advanced features available to make a braille copy: computerized braille production. Usually that installations have been equipped with some information carrier. So when the text already has been "written" on an information carrier, what is the case with the use of wordprocessors, one should simply transfer the information carrier from the one device to the other and a braille printer could be started to print out the text in Braille. Although that should be an elegant and simple solution, it is a dream up to now. Every, or at least almost every manufacturer has his own information carrying device obeying his own conventions and instructions, which are secret because of patent rights, competition policies etc. But even now there are possibilities as will be pointed out in chapter 3.

Secondly, we can distinguish a slightly different usage of the wordprocessor with respect to the first application: when the wordprocessor is used to prepare the production of books and reviews in Braille. The difference concerns more the organizational aspects than the technical ones, so that we do not give attention to it further more.

At last the blind could become the operator of a wordprocessor, when there would be good facilities for the blind to read the text keyed in. For that purpose a tactile display is needed, representing a number of lines if possible.

2. Functional considerations

The most conspicuous difference between a normal printer and a braille printer is the format of the characters and by that the number of characters per line and the number of lines per page. A normal line printer has 80 or even 132 characters per line and about 60 lines per page, a braille printer on the other hand has up to 40 characters per line and about 30 lines per page, may it be that not only the format of characters contributes to this difference but also the dimensions of the sheets. Anyhow when a copy in Braille of an ink-printed document has to be made, the format of the text, proper for a normal printer, must be adapted to the format suitable to the braille printer in use.

This function introduces several problems. Let us, as an example, look at the following line:

"Each day the expedition department registers all books returned on a long list."

This line contains 79 characters and spaces. If that line should be reformatted for a 40 characters per line braille printer, the division should be made between the characters "s" and "t" of the word "registers". So after that "s" a break must be inserted followed by an "end of line"-command. But if we want to use a 32 characters per line printer, the division should be made between the "e" and the "n" in the word "department" and that is an illegal place to divide that word. A first approximation of the solution for that problem is to avoid this difficult division rules and shift the whole word that does not fit in the rest of the reformatted line to the next line.

An other problem we meet is the recombination of divided words: the first part is in the one line and the remainder of that word is in the next line. The break must be removed and the "end of line" symbol must be taken away. One exception of this rule can occur, anyhow in Dutch, if the last word part in the line belongs to a contraction of two words: then the break must be maintained and the "end of line" symbol must be replaced by the space character.

Shifting the remainder of a line to the next line, one can meet the problems of the paragraph bound: a new paragraph starts at a new line. So that next line must be tested on the occurence of the conventions for starting a new paragraph, e.g. two spaces at the beginning. If a new paragraph starts at that next line, a new line has to be inserted containing only the remainder.

The difference in page sizes has consequences for the page numbers too. If a braille page is full, usually the ink-print page has some characters left. That characters must be transferred to the next braille page. It is a nice feature to place as well the page number of the ink-print copy as the page number of the braille copy in the first line of the braille page, giving each number its own position.

Transferring the remainder of a page to the next one, the paragraph bound can be overstepped too. Moreover the problem can arise that at the next ink-print page a new chapter begins.

Besides the reformatting of the pages, the translation of the characters give some difficulties. At first there is the conversion from one ink-print character (e.g. a capital) into two braille characters with several exceptions to that rule, but the opposite appears too. All symbols with an accent sign are keyed in as two different characters, although in Braille they are represented as one character.

The problems mentioned up to now are all known in computerized braille production and can be solved or avoided with familiar measurements. The usage of wordprocessors introduces some new difficulties. Some wordprocessors have the facility of the splitted page: in stead of a long line of 80 or more characters, two columns of shorter lines are used, just as dictionaries and the Bible have been printed. In Braille that facility can not be adopted because of the short line length, so that we have to split up the ink-print page into several braille pages, tracing the left-hand column at first and thereafter the right-hand one. Depending on the column concerned a "l" or a "r" must be added to the page number of the ink-print original.

If in the original page some words have been underlined, in the braille text an emphasis character must preceed that words. It depends on the type of wordprocessor how the underlining is executed. Some wordprocessors give the same facility as an electric typewriter, that is to say that the word can be typed, the cursor must be shifted back to the first character of the word and then the underline-key must be pressed by which repetitively underline-characters are produced until that key is released. With the conversion of the ink-print text to the braille text all these back-space characters and the underline characters must be removed and in front of the first character of the underlined word an emphasis character must be inserted and a reset character must be added to the word after the last character underlined.

Other types of wordprocessors have a deviating method of underlining: before the word to be underlined is typed, a control key must be locked and after the last character that key must be unlocked. Although this method meets to the typist the receipt for brailling, the processor does not generate a special character when the control key is pressed, but derives from the next characters the command to underline. Depending on the type of printer the processor generates a series of characters suited for the printer to perform the printing of an underlined word. It is possible that the stream of characters is the same as described for the other method.

3. Technical considerations

Braille embossers are in general quite a lot slower than ink-printers; the fastest braille embossers produce 3 lines per second, that is 120 characters per second, while line printers reach 20 lines per second and more, that is 1600 characters per second and more. Usually however powerfull wordprocessors are equipped with printers of 5 lines per second, that is 400 characters per second, if the line length is 80 characters. Starting from these data, for our subject an ink-printer can be said to be 3 â 4 times faster than a braille embosser. That implies that anything must be done to meet the difference in speed.

The solution of that problem depends on an other aspect: there are namely several ways in which the braille embosser can be connected to the wordprocessor.

The first way is connecting the embosser at the line printer plug. To realize this manner there are several possibilities. Some wordprocessors have two printer plugs. In that case the embosser can be connected to the wordprocessor besides the line printer and to the wordprocessor the embosser must look like a slow ink-printer. Supposing the processor controls the printer by the use of flags or interrupts, the embosser interface sets that flags or interrupt request signals at the moment a new line of characters is needed. Wordprocessors at that level of performance usually have software features by which during the printing of a text at the one printer, the second printer can be used for an other job. In fig. 2 such a configuration is outlined.

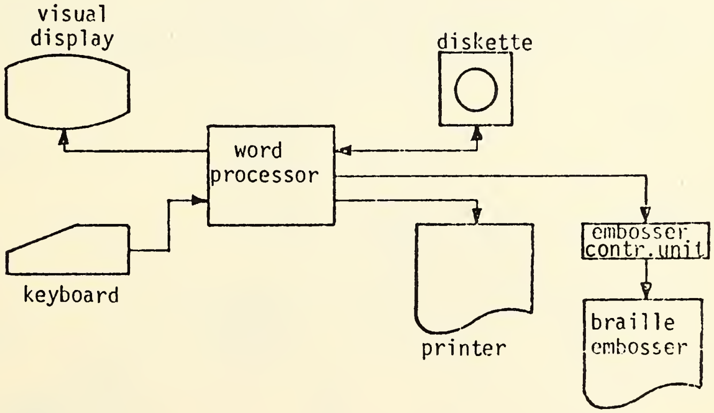


figure 2.

If we do not have a separate plug, a switch can be inserted between the wordprocessor and the printer. The embosser must be connected to the free pole of the switch. We have to take into bargain now that only one of the two printers can be controlled at the same time and that when the embosser is active the output speed is reduced. In fig. 3 this system is represented.

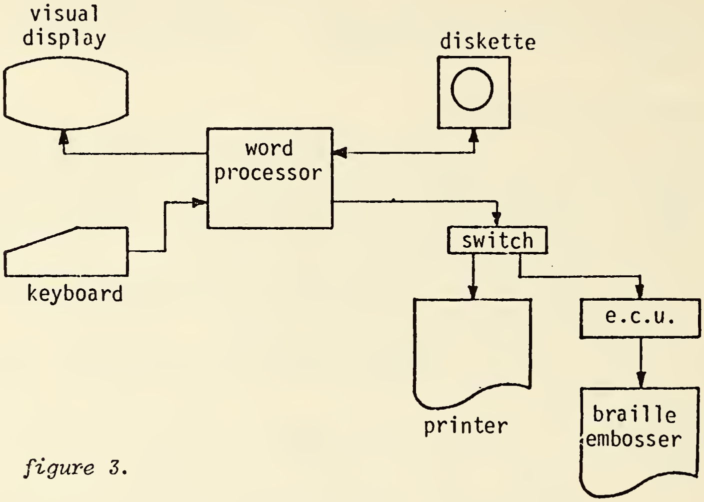


figure 3.

The drawback of the reduced speed can be nullified, even be converted to a benefit, by attaching a separate information carrier to the embosser control unit: the characters received from the wordprocessor are buffered into the memory of the control unit; as soon as the last character of a line has been received, the control unit requests for a new line. The control unit composes out of the buffered characters a braille page, which is registrated onto the cassette or the floppy disc. After a whole file has been transmitted the switch can be reset so that the ink-printer can be controlled by the wordprocessor and the embosser can be controlled by its own control unit, reading the braille characters from the information carrier. Fig. 4 depicts this method.

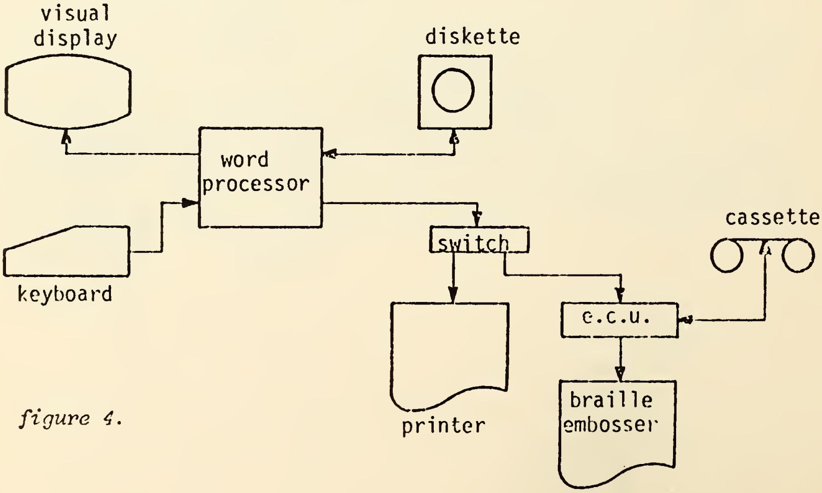


figure 4.

A slight change in this configuration can be made to increase the production if it is wanted that of any document an ink-pint copy must be made as well as a braille copy. Then the switch can be replaced by a tap: the printer is controlled and at the same time the embosser control unit takes the information from the data and control lines.

Now we have introduced a background memory, we can consider to connect the embosser control unit to the communications interface. If such an interface is available, the connection is quite a lot easier, because this interface will meet any C.C.I.T.T.-standard, usually V24. This is a two wire connection with a serial asynchronous information transmission at normalized signal speed, e.g. 1200 bits per second. An other advantage of this solution is that the same hardware can be used for several other types of wordprocessors. In fig.5 this system configuration is given.

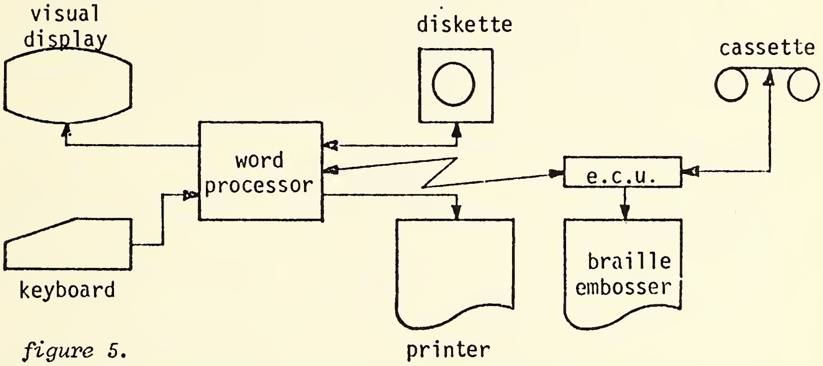


figure 5.

It will be clear that it closely depends on the money one has available, the performance one requires and the wordprocessor used, which of the solutions is selected.

The functional requirements as mentioned in chapter 2, and the forgoing technical considerations of speed adaption lead us to the conclusion that the embosser control unit must have some intelligence to do the job required. It is obvious that with the present day state of the technology a microprocessor is used to control the embosser.

If it is our starting-point that a blind should be able to operate the wordprocessor, the tactile display must replace the visual display. The microprocessor that controls the embosser can be used to control the tactile display too, as outlined in fig. 6

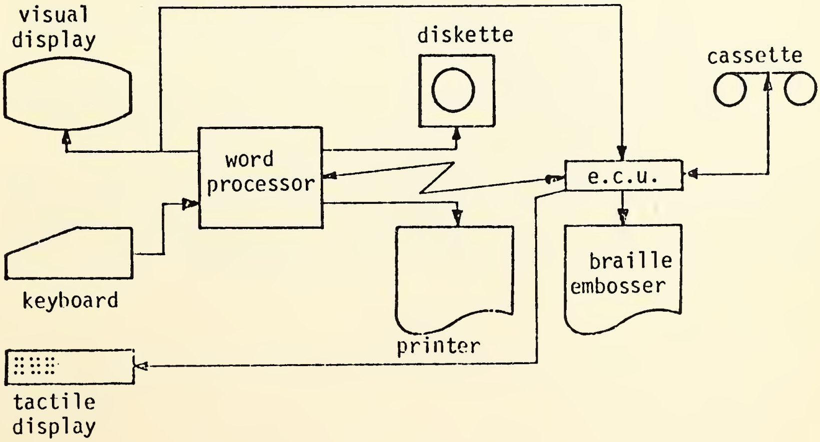


figure 6.

To that end we have to tap the connection between the wordprocessor and the visual display, what can easily be done because that is too a two wire connection, and lead one branch to the microprocessor. Here we have to do with almost the same functional requirements as when controlling the embosser, so that we can use nearly the same software. The memory of the microprocessor has to be large enough to contain a whole page of ink-printable text.

4. Integrated braille system

At the request of the "Vereniging Het Nederlandse Blindenwezen", a Dutch umbrella organization for unions of the blind, we at our laboratory made researches into the possibilities of connecting a braille embosser to a wordprocessor. Faced with a decision on which of the described solutions should be chosen, and especially what type of information carrier should be selected, we tried to include the realization into an integrated braille system we designed in behalf of the cooperating libraries for the blind in the Netherlands.

That system includes several modules, each with a specific function. To start with, there is the so-called braille input station. It is composed of the following parts as can be seen from fig. 7:

- a keyboard
- a mono audio cassette recorder
- a standard home t.v. set
- a microcomputer

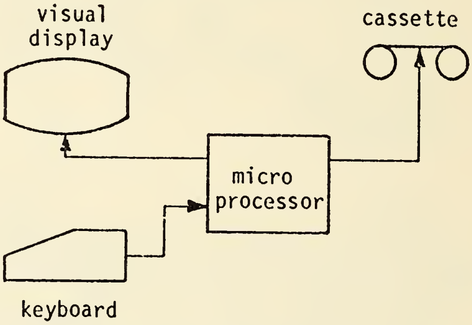


figure 7.

The operation can be compared with the working on a wordprocessor: the brailler puts in the text by using the keyboard and the characters are displayed at the t.v. screen. When the brailler should have made some errors, the brailler can use several commands to correct that mistakes. Coming to the judgement that the page is free of errors, the brailler presses a special key ("end of page") and then the station records that page on the cassette and the station prepares a new page input. It must be said that the software available now, only can handle text that is formatted in conformity with the page size, the so-called one-to-one input. But with a number of simple modifications in the software unformatted input can be allowed too.

Although the brailler having an input-station, has in principle all tools to produce an errorless tape, it may occur that some cassettes still contain some errors the brailler has overlooked. Or it may be that a cassette was prepared at the format for an embosser with a certain paper format but that another device must be used with an aberrant format. Still another possibility is that a cassette contains the text of a book (e.g. a schoolbook) that has to be updated to the newest edition. In all these cases the correction station can be used (fig.8).

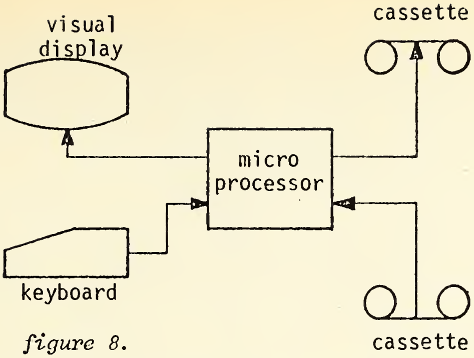


figure 8.

This station looks like the braille input station, but it is slightly different in that is has two recorders connected. The one recorder contains the cassette with the text to be edited, the other contains the cassette on to which the new text is recorded, or a copy of the old text depending on the need to edit a page. The operation of the correction station too differs slightly from the operation of the input station, because the correction station offers more features. It is enough to mention the three most important differences. The code of the input device is not necessary the same as the code of the output device. We have now implemented the ASCII-code as well as the Braille-MIT-code, but any code can be wanted. The operator has to determine what codes are to be used. A second difference is that it is possible to change the format. But changing the format generally implies that words have to be divided and that the parts of divided words have to be joined together, supposing the original text was formatted as "one-to-one". At last other devices than cassette recorders can be attached to the correction station, such as paper tape readers, paper tape punchers, floppy discs, etc. This feature supplies the possibility of changing the information carriers.

When the text is mature for multiplication one of the forms of distributing the information is braille embossing. The cassette coming from the input stations, whether or not edited by the correction station, are placed into a recorder to control a specific embosser or a zinc-plate machine to produce hard-copies in Braille (fig. 9).

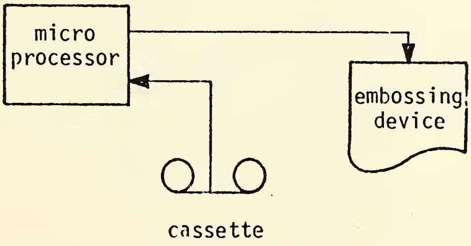


figure 9.

There are a large number of different braille embossers, varying from electrified handdriven embossers to highly specialized braille line printers. Because of this large number of different embossers we intend to make a teletype-compatible interface with several datacommunication speeds specified by program. With this interface we appeal to the designers of embossers to provide such an interface too.

Having the disposal of a tactile-display we can configure both the braille input station and the correction station as devices for visually handicapped professionals. Instead of the t.v. set a tacticle display must be attached and - as an option - a keyboard with seven keys, grouped just like the keys of a handdriven braille embosser, can replace the normal QWERTY-type keyboard. Then the input station can be used by blind people to write reports, articles etc. and the correction station offers the possibility to appoint blind correctors. Another application of the seven-keyboard is that other characters than the alphabet can be put in, such as musical notes, mathematical symbols etc. In that case the t.v. set control should have a second mode in which braille characters are generated instead of ink-print characters. When a tactile display is available other facilities come in sight. Instead of an output channel of any computer or processor, a large amount of text recorded on to an information carrier such as a cassette tape or a flexible disc can be used as the information source. The machine for interactive reading has been introduced now. The recorder can select a certain area of the whole text, for example a page of a book, and within that area a line or a number of lines he wants to read. To perform the selection the braille reading machine offers a number of commands. One can distinguish between two methods of reading, depending on the type of text to be read. The one method is reading a book for relaxation or study purposes. Then a page is read one line after another sequentially. The other method is used for looking up in dictionaries, encyclopedia and other reference books. Then an item, such as a cath-word, is given and the machine displays the lines containing that reference the one after another until the line wanted has been reached.

Looking at the long term, and stating already now a slowly growing offer in devices for speech synthesizing, we could think of a new output device: the speaking reader. We are convinced that it is too early now to anticipate at such devices, but it could be an important supplement of the devices described here, for only a small part of all blind people can read Braille.

Earlier we indicated already implicitly that there is interest in the possibility to process unformatted texts. This feature offers a number of advantages: reformatting is unnecessary because the text is formatted in the last stage before printing, the execution of correction commands is easier, the input speed is higher etc. But on the other hand now problems arise, because the last stage before printing must include a conversion from unformatted text into formatted text. That requires a fairly powerfull processor with a middle large back-ground storage.

5. Realization of the project

The introduction of microprocessors gave the opppertunity to think about a system for brailling based on microcomputers in such a way that not for each application a new device had to be designed, but that the hardware for each application was the same, or nearly the same, and that the systems only differed in the program they executed. Moreover, the important growth of the microcomputer market, caused a sharp fall in prices not only for the microcomputers but also for devices connected such as cassette recorders, floppy discs and memories not forgetting. We believe that the impact of this market on the further development of very cheap but nevertheless comprehensive microcomputers can hardly be overestimated. We tried to realize modularity in the system for braille input, for correction, for interactive reading, for interactive computer usage and for embossing. This modularity finds expression as well in the hardware as in the software.

The hardware consists of the Intel's 8080 microprocessor mounted on Eurocard-prints. For each device a peculiar interface print is designed which can be mounted into a backpanel. This backpanel is in fact the realization of the input/output bus of the microprocessor. According to the function of the unit programs can be supplied by inserting prints with ROM's (read only memories) and data storage can be enlarged by inserting RAM prints (random access memories, which can be written as well).

The basic idea we had was, to give all people that are brailling at home such a device. In the Netherlands there are a lot of volunteers that are brailling in their leisure time. Therefore we took cheap devices as t.v.sets and cassette recorders, which are available at home for almost everyone. The audio cassette recorder has a serious drawback in that the information recorded is not easily to retrieve, certainly not at random. A benefit of the audio cassette is that with the installations of the libraries the cassettes easily can be multiplied, what shows to full advantage when the reading machine can be distributed at large scale. The principle of registration on to the cassette is derived from the Kansas City Standard, which is based on a frequency shift keying method. Recording this standard a logical "0" is coded as 4 periods of a 1200 Hz sine wave and a logical "1" as 8 periods of a 2400 Hz sine wave. Bits are grouped into bytes and each byte must be preceeded by a starting bit and followed by two stop bits. Fig. 10 shows the typical signals for an arbitrary part of a byte.

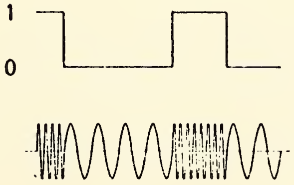


figure 10.

The Kansas City Standard normalizes a registration speed of 300 bits per second. We modified that standard in such a way that 1200 bits per second can be registrated, by coding a "1" as one period of the 2400 Hz sine wave and the "0" as one period of the 1200 Hz sine wave, as depicted in fig. 11.

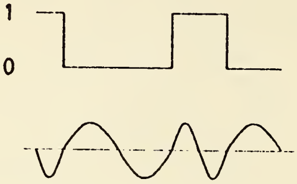


figure 11.

A braille page can be recorded now in 6.7 seconds, while using the Kansas City Standard it will take 37 seconds. With this modified registration a C60 cassette can contain about 440 pages of text.

6. Final remarks

Any technician who designed a system is proud of his results and propagates that as the final solution for all the problems in that area, and he has the perfect right to do so. Fortunately there are critical users that keep him with both feet on earth. We realize that and we hope that from both sides that competition will stay, and that this paper may contribute to an open discussion about the pros en cons of distributed systems for the blind.

It is necessary in our opinion that information is exchanged, as done at this congress. But we think that in the near future data must be exchanged too, and must be provided to perform that exchange.

With an eye to all developments up to day features and in the future, we believe that some standards are needed, such as codes, information carriers and their way of registration. Perhaps this conference could give a start to that discussion.

World Council for the Welfare of the Blind

Committee on Cultural Affairs

International Conference on

"Computerised Braille Production - Today and Tomorrow"

London, 30 May - 1 June 1979

COMPOSITORS TAPES
AND
LIBRARY OF CONGRESS BRAILLE PRODUCTION

by

Henry B. Paris, Jr.

and

Bruce E. Miller

The Library of Congress

Washington, DC 20542

1291 Taylor Street, N.W.

USA

COMPOSITOR TAPES
AND
LIBRARY OF CONGRESS BRAILLE PRODUCTION

by

Henry B. Paris, Jr. and Bruce Miller

Introduction

During the 1978 Braille Production Conference in Madrid, Dr. Van Vliet of the Netherlands described his successful application of compositor tapes in the braille production process. He discussed how he adapted the tapes to serve as an input to a computerized braille production system as an alternative to manual text entry. Compositor tapes are the storage media used by computers that compose print books. The tapes, either magnetic or paper, are used to drive typesetting machinery for printing a book.

Dr. Van Vliet's report stimulated the National Library Service for the Blind and Physically Handicapped (NLS) of the Library of Congress to re-examine the status of computer composition in the United States for braille application. In past years, similar investigations revealed inadequate sources of tapes and failure of the tapes to correspond to the desired final printed product, because the final editing process was not performed on the computer. This situation made program application impractical at that time. Since the summer of 1978, publishing industry representatives have been contacted through Associate Librarian of Congress Dr. Carol Nemeyer (former Executive Director of the Association of American Publishers) to learn more about the state of the art of photocomposition in the United States. This paper is a report of current findings and plans for the future.

Summary of Library of Congress Braille Production Requirements

NLS produces books for approximately 20,000 braille readers and distributes them via a network of forty braille-lending libraries. Some 350 braille book titles per year, and twenty weekly, monthly, or quarterly braille magazine titles are also produced each year. This results in a total annual expenditure of approximately $2 million. Additional funding supports purchase of quantities of books and magazines from other sources such as the Royal National Institute for the Blind and the Scottish Braille Press, which raises the total number to about 375 books and 32 magazines. An average of eighty copies per braille book and eight-hundred copies per issue of braille magazines is produced. A significant part of the average cost of a book title or magazine issue--summarized in Table I--is that portion required for preparing the master using manual entry methods. For books, mastering is estimated at $1,600 of the total $3,600 or 44%; for magazines, mastering costs are $900 out of $2,666 or 34%.

CURRENT ANNUAL BRAILLE COSTS

	Number of Titles Per Year	Number of Copies Per Title	Average Cost Per Copy Sub	Average Cost Per Title (Per Master)	Total Annual Budget
BOOKS	350	80	$45	$3600 44% ($1600)	$1,260,000
MAGAZINES	20	12978	$40	$2666 34% ($900)	$640,000

TABLE I

All braille is produced in the Grade 2 system without hypenation and justification. The average press braille book consists of four volumes of one-hundred pages each of interpoint braille and eighty copies of each are usually produced. NLS has no braille production facilities in-house and produces braille by contracting with braille printing houses. The American Printing House for the Blind (APH), which produces educational material as well as material for NLS, is the only source for press braille books. APH and three other printing houses share production of program magazines. Typically, a year's distribution of twenty magazine contracts is divided equally among APH, Clovernook Printing House for the Blind, National Braille Press (NBP), and Volunteer Services for the Blind as shown in Appendix I. Appendix II contains a summary of print composition methods currently used by publishers of magazines now produced in braille for the NLS program.

All braille printing facilities used by NLS employ computers to translate from print to braille and to edit and enter corrections. Appendix III contains a summary of these systems. Some have been described in previous papers at this conference. They differ in actual hardware and software, but perform the same basic functions. These functions are: manual entry of text onto the computer using a Video Display Terminal with varying levels of editing; conversion of print to braille using a system such as the Duxbury Translator; embossing a single braille proof copy using an LED 120 or similar peripheral; insertion of corrections; and output to either an automatic zinc plate embosser, a cassette-braille machine interface, or a single copy paper embosser.

Status of Photocomposition in the United States

Investigation began with discussions at Viking Press, a major publishing house, and Publishers Phototype, Inc. (PPI), a photocomposition bureau that services Viking. PPI cooperated in our study of its Atex photocomposition system. Atex systems are considered among the most advanced and are the most widely used throughout the United States publishing industry.

From these discussions and additional study it is apparent that the science of computer composition has expanded rapidly during the last few years. The basic functions of most systems in the United States are:

- "raw" unedited, basic story is input directly into the computer;
- subsequent editing is performed on the computer files through video display terminals;
- final edited tapes receive composition codes; and
- composition tapes are used to print the inkprint.

When a magazine is ready to print, the photocomposition data is transmitted from the editorial offices to the printer via telephone line, satellite, or microwave.

It is also concluded that:

- sufficient magazines are available in photocomposition tapes to make development of an NLS system worthwhile (Appendix IV lists magazines produced on Atex systems);
- editorial accuracy of the tapes is dependable; and
- these tapes can interface with the equipment of NLS braille producers.

Furthermore, publishing industry sources indicate that as many as 20 percent of the books published are printed from clean compositor tapes and that this sector may grow to as much as 80 percent within five years. If efficient ways to use magazine tapes are developed, the same system will be useable for producing books.

The key step in the process to interface compositor tapes with braille production equipment appears to involve:

- identification and removal or transformation of composition codes from tapes;
- conversion of some codes from unique local systems to Duxbury layout coding; and
- output of a medium acceptable by existing peripherals at each publishing house.

There is, however, very little standardization of composition codes within the publishing industry. This situation might result in the need for an interface program with the above features for each publisher.

Early this year Bill Raeder of NBP joined the NLS investigation. Independently he was already considering the feasibility of using compositors tapes. He specifically proposed to produce the National Geographic magazine for NLS using compositor tapes. Together, NLS and Mr. Raeder have embarked on a long-range program to develop a fully automatic method for converting a broad selection of available tapes to a format and media suitable for use by all NLS braille contractors. The first stage is development of software that converts Atex compositor tapes to a form suitable for use in the NBP braille production system. It is recognized that there may be some insoluable problems for using the computer alone, and that a man/machine interactive system may be the most practical outcome.

Project Definition: A two-phase project was undertaken by the Atex software group.

Phase 1.

The Specific--development of software to convert the text file of National Geographic magazine and US News & World Report to a format on magnetic tape to interface with the Duxbury Translator used by NBP. Software in this phase is to run on the Atex development computer (Modified DEC PDP-11) located in Bedford, Massachusetts. It will produce a modified version of the compositor tapes to run on the NBP computer and produce a braille magazine. During this phase, the degree of required human interaction will be determined by careful analysis of a record of changes made by the editors. The software module to be developed in this phase is part of a larger structure to be addressed in Phase 2. In addition to Dick Handverger, the key Atex software developer, Joe Sullivan of Duxbury and Jeff Matherly of NBP are working on the project.

Phase 2.

The General--development of the remaining software and hardware specifications for a system to convert any Atex-based compositor tapes to a suitable form for operating on any NLS contractors' computer systems. This phase will also identify problems associated with extending application to other than Atex-based compositor tapes.

Project Status

The project began with agreements from both National Geographic and US News & World Report to supply compositor tapes to the Library of Congress concurrently with the release of the magazine for printing. In the case of National Geographic, a monthly publication, this schedule means receiving the completed tape one to two months in advance of release of the magazine to newstands--a very comfortable braille production schedule. US News & World Report, however, represents an entirely different scheduling example. This weekly magazine of about one-hundred pages is transmitted from the compositor computer in Washington, D.C., on Friday nights, to printers in Connecticut, Illinois, and California, for weekend printing and newstand release on Monday mornings. Timely braille production of this type of publication is not possible using current braille production methods; however, compositor tape technology demonstrates many real advantages to the braille industry and makes possible production of weekly news magazines and similar publications.

The project development group began by pooling its experience and by defining the software requirements. As the project expanded, the group began working with the April issue of the National Geographic and identifying problems. The compositor tape received from the magazine was loaded onto the computer at Atex. Jeff Matherly, the NBP braillist who usually supervises manual entry of the magazine on the NBP computer, reviewed the tape with the aid of a powerful editing program. He then made it identical to tape from a manual full-text input operation. Due to its problem-solving nature, this process initially required a longer time than the manual entry process.

As Mr. Matherly completed sections of the magazine, Dick Handverger and Joe Sullivan compared the original compositor tape to the modification to look for new problems and test solutions to predicted ones. Examples of types of problems addressed to date appear in Appendix V.

As of the publication of this paper, the April National Geographic has been converted to braille and approximately 90 percent of the automatic software has been developed. Current indications are that use of compositor tapes will result in:

1) much shorter production time;
2) smaller number of errors; and
3) lower cost per magazine, hence greater numbers of magazines for the same budgeted dollars.

A period of production experience will be required before these advantages can be quantified.

Future Plans

Unless some radical, unsolvable difficulties arise, NLS will encourage implementation of systems to accept compositor tapes by all braille-production contrators. It is not possible to discuss anything but general concepts at this time; however, the system envisioned as a result of Phase 2 effort will contain:

1. a receiving terminal for an error-correcting communication system to permit electrical transfer of material from publishers to braille printing houses;

2. an edit system to facilitate both manual entry and proof-reading error correction;

3. a magnetic tape storage medium;

4. a computer capable of running the compositor tape conversion software and a print to standard-English braille conversion program;

5. a peripheral capable of producing high-speed single-sided braille proof copy;

6. a peripheral capable of producing high-speed interpoint zinc plates; and

7. an interface for cassette-braille reading machines.

The cost of implementing this system is estimated from $30,000 to $100,000 depending on the production facility, and will undoubtedly be shared among NLS, other United States government agencies, and non-profit braille-producers.

LIST OF APPENDICES

Appendix I - 1979 LC Braille Magazine Production Summary

Appendix II - Print Magazine Composition Summary

Appendix III - Summary of Braille Production Systems Employed by LC Contractors

Appendix IV - Magazines and Newspapers by Atex Photocomposition System

Appendix V - Interim Analysis - Atex/NBP Conversion Program

APPENDIX I

1979 LIBRARY OF CONGRESS BRAILLE MAGAZINE PRODUCTION SUMMARY

		CIRCULATION	FREQUENCY	PRODUCER	COPIES
1.	American Girl	400	Monthly	VSB	4800
2.	Better Homes and Gardens	1125	Monthly	CPH	13500
3.	Blind Data Processor	300	Bimonthly	APH	1800
4.	Boy's Life	275	Monthly	APH	3300
5.	Braille Variety News	1100	Bimonthly	CPH	6600
6.	Children's Digest	325	Monthly	VSB	3900
7.	Consumers Research	825	Monthly	VSB	9900
8.	Family Health	1150	Monthly	NBP	13800
9.	Fortune	300	Biweekly	APH	7800
10.	Galaxy	525	Monthly	CPH	6200
11.	Horizon	400	Bimonthly	CPH	2400
12.	Jack and Jill	285	Monthly	VSB	3420
13.	Journal of Rehabilitation	325	Bimonthly	APH	1950
14.	Ladies Home Journal	2025	Monthly	CPH	24300
15.	National Geographic	950	Monthly	NBP	11400
16.	New York Times L:T: Weekly	2000	Weekly	NBP	104000
17.	Playboy	1060	Monthly	CPH	12720
18.	Popular Mechanics	750	Monthly	CPH	9000
19.	Psychology Today	1015	Monthly	CPH	12180
20.	Seventeen	550	Monthly	CPH	6600
	TOTAL	15685			259,570
	AVERAGE	784			12978

APPENDIX II

PRINT MAGAZINE COMPOSITION SUMMARY

MAGAZINE	COMPOSITION METHOD		COMMENTS
	Computer	Other	
American Girl		Hot Metal	
Better Homes and Gardens	Photon		
Blind Data Processor			No Ink Print Version
Boys Life	Harris		
Braille Variety News			No Ink Print Version
Children's Digest		Hot Metal	
Consumers Research	ATF-4		
Family Health	Mergenthaler		
Fortune	"Time Life" IBM 360 & Videocomp		
Galaxy		Hot Metal	
Horizon	ATF-5		
Jack and Jill		Hot Metal	
Journal of Rehabilitation	ATF-4		

APPENDIX II

PRINT MAGAZINE COMPOSITION SUMMARY

MAGAZINE	COMPOSITION METHOD		COMMENTS
	Computer	Other	
Ladies Home Journal		Hot Metal	Planning Conversion to Computer Composition
National Geographic	Atex		
New York Times Large Type Weekly		Photo Process	
Playboy		Hot Metal	
Popular Mechanics		Hot Metal	Planning Conversion to Computer Composition
Psychology Today			
Seventeen		Gravure	

APPENDIX III

SUMMARY OF BRAILLE PRODUCTION SYSTEMS EMPLOYED BY LC CONTRACTORS

	INPUT DEVICE	DATA MANIPULATION TRANSLATION PROGRAMS	OUTPUT DEVICES
National Braille Press	2 Lear Ziegler ADM III video display	1-Sykes ComStor intelligent diskette drive MODEM TymShare Braille Translation	2 Triformation LED 120's 1 - Triformation PED 30
Volunteer Services for the Blind	3 DecScope video display	1-PDP 11/34 Braille Translation 2-disk drives RK-05	2 - Triformation LED 120 1 - Triformation PED 30
Clovernook Printing House	4 Digital VT 52 video display	1-PDP 11/34 Braille Translation (Duxbury program) 2-Disc Drives RK-05	1 - Triformation LED 120 1 - Triformation PED 30
American Printing House	16 video display diskette units	1-IBM 360/65 PUNCH CARDS	10 - Modified Stereograph Machines

APPENDIX IV

SOME MAGAZINES AND NEWSPAPERS COMPOSED FROM ATEX PHOTOCOMPOSITION SYSTEM

Boston Globe
Chicago Sun Times
Christian Science Monitor
Chronicle of Higher Education
Congressional Quarterly
Cosmopolitan
Dallas Times Herald
Economist
Electronic Design
Esquire
Federal Register
Forbes
Medical Economics
Minneapolis Star Tribune
National Geographic
Nations Business
Newsday
Newsweek
Readers Digest
Science Digest
Town and Country
U.S. News and World Report

Some publishing houses using Atex equipment in whole or in part:

Crown
Holt Rhinehart Winston
Knopff
Paulist Press
Time-Life Books
U. S. Government Printing Office
U. S. International Communication Agency
Viking Penguin

APPENDIX V

INTERIM ANALYSIS - ATEX/NBP CONVERSION PROGRAM

The conversion program strips unwanted coding, and flags elements of the text where recognizable direct conversion can take place from Atex to Duxbury systems. An experienced braille editor prescreens flagged elements and adjusts the input in accordance with known Duxbury performance. (All references are from English Braille, American Edition, 1959, rev. 1972)

1. Flagging (or extract) desirable and will be easily incorporated into next generation of program.

   Notes to National Geographic production staff.

   Footnote text was to be replaced (Sec. 22).

   Separation and termination lines should be placed.

2. Flagging of elements frequently occurring in inkprint tape which could be converted, but only with much more elaborate program.
   Italic signs were inserted.

   Footnotes occurred.

   Small capitals occurred (Sec. 2.f).

   A paragraph was introduced with a dropped initial and three words in full capitals.

3. Problems difficult to flag, to the extent that they can be flagged will greatly speed braille conversion.

   Acronyms were followed by a small "s", requiring capitalization of the "s" (in order to translate properly) and insertion of an apostrophe (Sec. 4.a).

   Proper names with short-form words (Sec. 47.b).

   Ordinal numbers required insertion of a letter; e.g. 3d, 2d should be 3rd, 2nd (Sec. 29 Exception).

Abbreviated years required insertion of an apostrophe; e.g. 50's should be '50's.*

Plural numbers, with "s" required insertion of apostrophe; e.g. '60s should be '60's (Sec. 4.a)

Proper name contractions may hinder pronunciation (Sec. 38.e).

Acronyms contained contractions (Sec. 27.2).

Ellipses must be checked for context (Sec. 7).

Natural pauses after the words: to, into, by, and, with, of, the, for and a (Sec. 37).

Insertion of comma may have been necessary after the first digit in four-digit numbers that are not years; e.g. 2000 pounds requires the insertion of a comma.*

Single letters followed by a closing parenthesis and not preceded by an opening parenthesis required deletion of the closing parenthesis and insertion of a letter sign before the letter (Sec. 12.a(2)).

Single letters followed by a period should be read to determine if the period is only ending a sentence. If so, a letter sign should be inserted before letter (Sec. 12.a(2)).

Occurrences of a, i and o as enumerators (as in an outline) are not followed by a period or enclosed by parentheses should have letter sign (plus sign) inserted.

Roman numerals V, X, L, C and M should not be preceded by a letter sign; letter sign should be blocked (Sec. 30).

A slash (oblique stroke) is used and should be replaced by a hyphen.*

---

* These are required by "Instruction Manual for Braille Transcribing", 1973 and/or "Transcriber's Guide to English Braille", 1974.

World Council for the Welfare of the Blind

Committee on Cultural Affairs

International Conference on

"Computerised Braille Production - Today and Tomorrow"

London, 30 May - 1 June 1979

THE NATIONAL GEOGRAPHIC MONTHLY MAGAZINE: FROM COMPOSITOR'S TAPES TO BRAILLE

by

William M. Raeder

National Braille Press
88 St. Stephen Street
Boston MA 02115
USA

World Council for the Welfare of the Blind
Committee on Cultural Affairs
Subcommittee on Computerized Braille Production
and Other Media
Chairman, Derrick W. Croisdale

International Conference on
"Computerized Braille Production--Today and Tomorrow"

Paper delivered June 1, 1979
by
William M. Raeder

"The National Geographic Monthly Magazine: From Compositor's Tape to Braille"

Thank you, Henry Paris, for sharing your precious time before this high-level conference to allow me to describe the National Geographic composition-file-to-braille project that we are now developing at the National Braille Press in Boston. The project is comprised of six steps, which I will describe, with a conclusion.

Step 1 was to select a publication worthy of the investment in conversion software. We selected the National Geographic for four reasons: (1) it is a publication desired in braille (1,000 copies of the braille edition are distributed monthly as part of the magazine program of the National Library Service for the Blind and Physically Handicapped of The Library of Congress); (2) the National Geographic Society has clean composition files available; (3) the publisher is very cooperative; and (4) it is a monthly periodical of significant size, approximately 180 braille pages, and thereby able to provide adequate return on investment in the conversion software.

Before proceeding to Step 2, I would like to comment on the word "clean" and on the importance of the publisher's cooperation. We have discovered that there are two definitions of the phrase "clean compositor's tapes." The first is that all the text material on the tape is found in the ink print and is faithful to it. The second is that all the material in the ink print is found on the tape and is faithful to it. Up to the time of the written version of this paper, we have found only one example, in the July issue, of "uncleanliness," where neither of these definitions was met; a sentence had been edited out of the text downstream of the file from which our tape was produced. With our first two tapes, for the April and May issues of this year, we found that the first, but not the second, definition held; that is, that there was copy in the print edition that was not on the tape. After verifying the braille file developed from the compositor's tape against the print edition, we learned that in each of these two issues, sixteen additional pages had to be typed into the computer. The missing material was mostly captions, map descriptions, and boxed articles.

We have noticed a marked improvement in the amount of material included on the tapes of the subsequent three issues which we have produced (June, July, and August). In the June issue, everything in the ink print was found on the tape and was in good order. This improvement is due to the superbly cooperative attitude of the National Geographic Society, their computer operator, and his supervisor. In the National Geographic's computer, each issue of of the magazine is broken up into many files derived from numerous departments; art, literary, photography, cartography, advertising, etc. The thoroughness with which the computer operator gathers these files and reads them onto a tape and the manner in which he and his supervisor have analyzed and met our needs have added significantly to the progress of our project.

Step 2 was to receive and mount our first tape, the April 1979 issue, and to make the conversion from the files provided by the publisher to an input file for our braille translation program. This conversion was done manually, using a computer text editor system, for the purpose of fully analyzing the job to be done by the as-yet-unwritten conversion program. It might not have been necessary to process so much text material manually in this way, but we did it because: (1) the April edition had to be produced; (2) we had the tape well in advance, providing ample time for learning this new process; and (3) we felt that carrying out this exercise, in a production mode, would give us a full appreciation of the complexities of the job.

Actually, there are two conversions to be made: the conversion from the computer environment and from the composition code of National Geographic to the computer environment and to the composition code of the National Braille Press. The first of these should be carried out in an Atex environment. Atex is the company which developed and distributes the very powerful and sophisticated publisher's photocomposition computer system used by the National Geographic Society and many other publishing and typesetting companies. Atex, Inc., is located near the National Braille Press and very generously offered the use of their home office computer installation and the assistance of one of their senior programmers so that our project manager could carry out the manual conversion of the April issue.

Having manually converted the April issue on the Atex system, we read the new files onto a tape for transport to the National Braille Press system, that is, to the commercial time-sharing company where we store our braille translation program and carry out all our computer braille translation. After mounting the tape on this PDP-11/70 system, we realized that further conversion was required; after running the file through the Duxbury translator and proofreading, we found that still further conversion was necessary and, as already mentioned, that additional text had to be typed in. This is how we produced the April issue for distribution, and we were pleased with the production quality.

Step 3 was to write the conversion program. This is being paid for by the National Library Service and is being written and revised by the same Atex programmer who assisted in the manual conversion. Even with the benefit of this program, the conversion should be made in the Atex environment. The Christian Science Publishing Society, which also uses the Atex system (and is only a five-minute walk from the National Braille Press), is contributing several hours of their computer and terminal time to us each month.

There are four conversion functions to be carried out by the program. The first is to strip out the substantial amount of composition commands not needed by the braille translation program. The second is to make one-to-one translations where possible from National Geographic commands to National Braille Press commands. The third is to make more complex translations from National Geographic commands to National Braille Press commands that require an algorithm within the conversion program. The translation of "20° C." to "dg#20.c" is an example (the period represents the braille capital sign). The fourth is to flag those places in the text where conversion requires human intervention, e.g., italicized words. When the program finds an italicized passage, it will make the necessary conversions to command the braille translator to italicize the passage, but will then flag the opening italic sign to alert the braillist-computer operator, who makes the determination whether italic is truly appropriate.

We are instituting four different flags for different kinds of braille problems. They are:

1. ^d indicates a dropped initial command which is followed by a fully capitalized paragraph leader.
2. ^i indicates the beginning of an italicized passage.
3. ^q indicates an opening inner quotation mark (closing inner quotation marks must be manually converted).
4. $$ indicates a miscellaneous problem area which may require human intervention for proper translation, e.g., spacing commands used for both letter and word spacing and characters with no braille code equivalent.

It becomes clear now that even with the benefit of the conversion program there is a considerable amount of human intervention required on the part of a braillist to produce high-quality braille.

Step 4 is to run several subsequent editions of the magazine through the new conversion program to discover its shortcomings and to draw up specifications for the first revision of the program. At the time of this writing, the specifications have been prepared but the program has not yet been revised.

Step 5 is to run several more issues of the magazine through the revised program and to prepare the final revision. Subsequent followup will be made to ensure that all the specifications for the final revision have been met.

Step 6 is to write two additional programs that will enable a more fully automated conversion process. Both of these will be interactive and multipass programs. The first will be a supplement to the main conversion program and is anticipated because the main conversion program is written to be a single-pass program and uses only a small buffer. We expect that this program will assist in doing such things as finding the beginnings and endings of articles and in the placement of footnotes.

The second of the two additional programs will be to help in handling braille idioms that require human intervention. We may ask this program to flag all whole- or part-word contractions that are both capitalized and used as part-word contractions to ensure that they are not used in a proper name; for example, "Friendly Fred Friendly was the noted TV producer who . . . ." Although in both instances the word "friendly" is spelled and capitalized identically, the "fr" contraction for "friend" would be used in the first word but not in the second because the second is a proper name. Other passes through the program might find natural pauses (e.g., "He quickly ran <u>into and</u> out of the building"), homographs (e.g., "said" is contracted, but "Port Said" is not because of the pronunciation difference), etc. With this program the braillist could, with one key stroke, instruct the program to make the proper braille translation. Keying "Y" says to the program, "Yes, make the contraction and show me the next potential braille idiom." "N" of course means, "No, do not make the contraction and show me the next idiom."

We have found that, with a small amount of training, input typists quickly learn to introduce many composition commands that a conversion program would not introduce in a straightforward process of converting from compositor's tapes. Fully automated braille from compositor's tapes therefore is inferior to the braille produced by input typists with only several months' experience. When working with compositor's tapes, consequently, we feel that it may be very helpful to have a program capable of flagging these braille idioms.

So far we are able to draw three major conclusions. The first is that the compositor's tape process is not simple; it is complex and expensive to develop. The conversion program is written specific to the composition code used by the publisher or typesetter who provides the tapes and, to a degree, specific to the publication. Even with clean tapes and a well-written conversion program, a considerable amount of human intervention on the part of the braillist is required to meet mass production quality standards. The need for proofreading is by no means eliminated; however, we expect a reduction in the amount of time required for proofreading when we further develop our proficiency with the system.

The second conclusion is that a considerable amount of production time can be saved by the use of clean compositor's tapes and a good conversion program. When manually stereotyped, each issue of the National Geographic requires between 55 and 75 hours hours of stereotypist's time for transcription and fabrication of zinc plates. If transcribed by an input typist and a computer (embossing proof pages on an LED-120 and zinc plates on a PED-30), 40 to 55 hours is required. (We have not actually reproduced a National Geographic this way, but we are able to extrapolate from other material produced by input typists.) Transcription of the April issue from compositor's tape, made without a conversion program and working with an unfamiliar process, required 88 hours of our braillist-computer operator's time. Transcription of the May and June issues from compositor's tapes, made with the conversion program, required 20 hours each. The August issue required 17 hours. The July, anniversary, issue was exceptionally long and complex.

With the benefit of the revised program, we anticipate that each issue will require less than 15 hours. With the further benefit of the two programs mentioned in Step 6 and the continuing aid of the powerful Atex editing system, we believe the time may be reduced to 6 to 8 hours, 3 hours of which is for platemaking.

The third conclusion is that delivery times can be much improved. This is for two reasons: the sharp reduction in the amount of stereotypist's or input typist's time for plate production, and the fact that computer composition files are of course prepared in advance of the inkprint copy and are available earlier than the ink print. (It is important to note that we must still await the advance copy of the ink print for verification.) Sometime next fall, in October or November, we look forward to having the National Geographic braille edition produced, enveloped, addressed, and packed ready for the mail waiting in the National Braille Press basement--waiting for a release date from the National Geographic Society because they, quite reasonably, do not want the braille edition distributed in advance of the inkprint edition!

Credits:

W. Jeffries Matherly, Project Manager
National Braille Press
Braillist/Computer Operator

Henry Paris
National Library Service for the
Blind and Physically Handicapped
The Library of Congress
For assistance in coordinating with Atex and for securing funding
for the conversion program.

John W. Seybold, President
John W. Seybold and Associates
For identifying the significance of Atex and introducing National Braille Press to Atex and for consulting on the project.

Thomas Hagan, President
Camex, Inc.
For consultation

J. M. Gill
Warwick Research Unit for the Blind
University of Warwick
For consultation

Richard Handverger, Senior Programmer
Atex, Inc.

Geoffrey T. McConnell, Systems Manager/Photographic Services
National Geographic Society

Lawrence F. Ludwig,Assistant Director of Photographic Services
National Geographic Society

Gail Samek, Electronic Editing Systems Manager
Christian Science Publishing Society

Harvey G. Bramson, Director of Technical Support/Publishing Systems
Atex, Inc.
For introduction to and coordination with Christian Science Publishing Society

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

TACTILE DISPLAYS:
THEIR CURRENT STATE AND A NEW APPROACH

by
P W F Coleman

Computer Unit
University of Warwick
Coventry, CV4 7AL
England

TACTILE DISPLAYS:
THEIR CURRENT STATE AND A NEW APPROACH

P W F Coleman

Technical Adviser
British Computer Association of the Blind

Computer Unit
University of Warwick
Coventry, CV4 7AL, England

ABSTRACT

This paper seeks to explore the types, functions and desired attributes of tactile displays and their current realisations. Much attention is given to design rationale and problems, and a method for classifying braille reading techniques is put forward. The paper looks at theoretical implementations of transient displays, before detailing the design on which I am working.

INTRODUCTION

For ten years now I have been giving a lot of attention to the subject of displays for the blind. The context for much of this work has been my project of an integrated information system for the blind (References 3 and 4); but I also maintain a watching brief on tactile computer terminals for the BCAB.

This paper confines itself to tactile displays. It reviews the current displays and their limitations against some of the desired attributes. It then reviews the design process, suggesting some necessary and urgent research, finishing with my own work.

I shall confine myself to tactile displays: it is true that speech output is, if anything, more important; but it is progressing quite nicely and, in any case, deserves at least one paper to itself.

First, then, let me define what I mean by displays, making some attempt to classify them, at least for the purpose of this paper.

REVIEW

In some respects display technology has moved little since 1969. Already then there were tape and page braillers; Sutherland was at work on his braille cell module, and Bliss on the Optacon whose display prefigured work on piezoelectric braille displays now taking place. All these devices are still very much with us though more or less refined (References 1 and 2).

Even the transient display was more than just a dream, with the work of Grunwald (Reference 1) and IBM. The latter appears to have been a one-off device, while the former has still to reach the marketplace. The sad thing is that, when these displays finally appear, they may use little that is not conventional technology.

DISPLAY TYPES

Definitions

For the purpose of this paper I make the following not too arbitrary definitions: a Display is any medium for showing information, from a book to a television set; it can be Static, Dynamic or Transient. In line with increasing usage, a Dynamic Display is one that shows a line (or page) at a time under the reader's control. An example might be a visual display unit showing a new page each time the carriage control key were depressed. The contents of a Static Display can never be altered, examples being a book or magazine on paper. This paper makes little further reference to static displays. Note that the definitions of Static and Dynamic are a little out of line with Reference 4.

I use the term Transient Display to mean one in which the information is continually changing. A scrolling vdu is an example of a transient page display.

Examples of Braille Displays

Among examples of hardcopy braille displays (virtually static) are the succession of paper tape braillers such as the Triformation BD3, and page braillers such as their LED range and Sagem's REM/TEM BR 8 terminals. Dynamic displays are at present usually solenoid-driven and include those used in the French, German and American information systems for the blind.

There remains one more and very important display type: the Graphical Display Unit (gdu). It allows pictures and other graphical material to be drawn on a screen, and either to be overwritten (Storage Tube) or dynamically altered (Refreshable). For neither form is there a tactile equivalent - what I shall call from now on a tgdu. The need is no less now than when I first postulated it in 2970 (Reference 1). Indeed it is now more urgent than ever.

DISPLAY ATTRIBUTES

General

Before going on to compare displays, we ought to define a set of properties for the ideal display; I shall call them Attributes. Firstly, then, Power Consumption, Weight and Space must all be the minimum consistent with the other desired attributes. Weight is particularly important because many of those who use these displays will be old or very young, or under medical advice not to lift heavy objects. Weight and power consumption are likely to be significant factors in at least determining which technology is not used in such devices (see below).

Noise is another factor of importance as many office environments are governed by strict noise regulations; in any case, even small amounts of noise can be a distraction to the blind user. The display I am describing later is a dynamic display virtually free from the traditional mechanical components, which should be quiet in operation, cheap to build, and reliable and easy to maintain when trouble does arise.

There is too the relation between noise and mechanical wear, and between wear and reliability. Since these displays are liable to be used in equipment upon which the blind user relies heavily, reliability coupled with rapid servicing is obviously of great significance.

Lines and Characters

There are a number of attributes encountered by the reader most directly; they concern Lines, Characters and Dots. At the highest level, the line should have a length natural to its use. Blind users might be content with a very short, 6-character display for a digital pocket notebook device, where size and weight must be minimal, and where the braille output is only needed for checking doubtful characters input. However this is quite unsuitable for a computer terminal: here 40 characters, half a card image, might be the ideal; this is not so long as to be unduly bulky, yet allowing the user a unit of display related to that of his seeing colleagues. None of these is ideal for reading a book: here something between 25 and 36 characters is more suitable. Current general-purpose displays have 12, 32 and 48 characters. All of these are static displays, providing the user with a discrete line at a time.

The user should, ideally, notice no difference between reading with his chosen display type and reading a book. One obvious difference in most displays is the absence of the next line: he cannot move to it while still scanning the end of the current one. Worse, on some displays there is a very noticeable waiting time before the next line appears. Half a second is probably the maximum tolerable time a user can wait. Even so this delay may slow a fast reader by as much as 25%; a refresh time nearer .1 seconds is nearer the ideal. (Similarly, in devices using tactile displays, system response times should be minimal, probably within three seconds for all functions - Reference 3).

Characters should be a standard size - Perkins size seems to have become a de facto standard, but research in this, as other, areas is needed to establish the ideal for various types of displays.

In the same way, Dot Size should be a standard, or the optimum determined by careful research. There may be a case for using giant (Jumbo) dots and characters: this would cause a minimal increase in machine size but none in the space occupied by the text: compare this with the consequent increase in conventional books in giant dot.

Dot Shape and Texture I suspect to be paramount. Dots should obviously be dome-shaped to pass easily under the fingers, but the lower part of the dot I am certain should be cylindrical so that there is clear separation between dots.

The phrases "I suspect", "I am certain", and so on are meant to point to the need for research talked about later. In this instance, reasonable questions would relate to the height of the cylinder and the type of dome - spherical, parabolic, etc.

Experience with both Thermoform and Solid Dot braille should have convinced all of us of the importance of texture.

Reaction by blind people to Thermoform suggests that the display surface should be smooth but not glossy: too rough a display would be hard on the fingertips and might give rise to ghost dots. Great care is therefore needed in the choice of material from which the display surface is made, particularly if it is a plastic.

Added to this is apparent texture resulting from vibration or movement, either of which could modify the texture as perceived by the user. Any apparent smoothing from motion might well be offset by the resulting friction on the skin and consequent irritation. Again, there is plenty of scope for research here.

Ease of Use

A great deal of thought must be given to how the display is going to be used and by whom. The best display design can be nullified if used as part of a device that is difficult or unnatural to use. Thus the positioning of the input keyboard and various controls is crucial; for example, controls such as line or character delete should not be part of the input keyboard though they might be located nearby. The nature of the keyboard and display will determine their relative positions, and in particular which is nearer the user. But either way, a lot of thought should be given to this problem - it can be awkward to reach over a deep keyboard to a braille display, but just as awkward to reach over a large surface to use the keyboard. Thus if the display is in front of the keyboard with respect to the user, every attempt must be made to minimise the space between keyboard and user. This may sound a trivial matter: however in reality a small problem can grow to a large one if constantly encountered.

Again, controls should be clearly labelled. If the display is a braille one, it surely makes more sense to use braille control markings than special shapes, as it assumes that the user is a braille reader; it is also sensible to have printed marks as well in case a sighted person (not least a service engineer) wishes to use them - this might well be the case in a mixed braille-speech system. I do take the argument that symbols rather than braille words make the system internationally acceptable; this would have more weight if there were an international standard for such symbols.

Control activation should elicit a rapid response, preferably instant, certainly within a second. With a display this means very rapid setup time (see above).

Controls should include adjustment to the display's speed, its ability to backtrack (to at least the beginning of the previous line), and instant stop-start (if possible just by lifting and lowering the fingers).

DISPLAY DESIGN

Just as there is no accepted standard for vdu's, so is there none for braille display units. However there has been a lot of research into print reading and its impact on the vdu. This has no parallel with the bdu. Often a display is postulated which is a rather thoughtless adaptation of equipment for the sighted: an example might be the addition of a single braille cell to either a braille keyboard or a normal vdu. A single-cell display might be fine for a pocket-sized device where weight and space must be minimal and the display is just for the occasional checking of input, but not for anything larger. For similar reasons a larger device carried over the shoulder might use only a 6- to 12-character display. 25-36 would be about right for a book reading device, 40-48 for a computer terminal. Thus the nature of the display is intimately tied up with its use.

The important thing is that design should be accompanied by consumer research, not based solely on the designer's hunches; nor solely on the opinions of his blind colleagues. This applies even if he himself is blind with a sound knowledge of braille.

In any research, the purpose of the proposed display should be clearly delineated - e.g., whether it is a pocket notebook or a computer terminal. Consumers should have ready access to a prototype both initially and after second thoughts about it. This will, of course, mean greater cost, but hopefully also a better product.

The designer should give careful thought to the need for computer simulation at the same time (or even before) testing mockups and prototypes. This can, for example, be used to check the correct functioning of units within the limits of their specifications.

Braille Reading Techniques

Many current systems seem to take little account of how blind people read and handle braille generally. Until recently there had been little systematic study of braille reading techniques, so that sighted designers, through no fault of their own, were working even more in the dark than their blind clients were said to be.

In an attempt to offset this, I have developed a (very tentative) classification of braille reading techniques (see Appendix I). The Appendix paints the picture for a right-handed braille reader; a similar set of types exist for left-handed people. With left-dominance, though, it is still the left hand that must do all the work of guiding and locating.

From this it can be seen that some read with only one finger, of one or both hands, or in various ways with more than one finger. Note the need with many readers to have the second line present at the same time as the first one. This means that no current display really suits them.

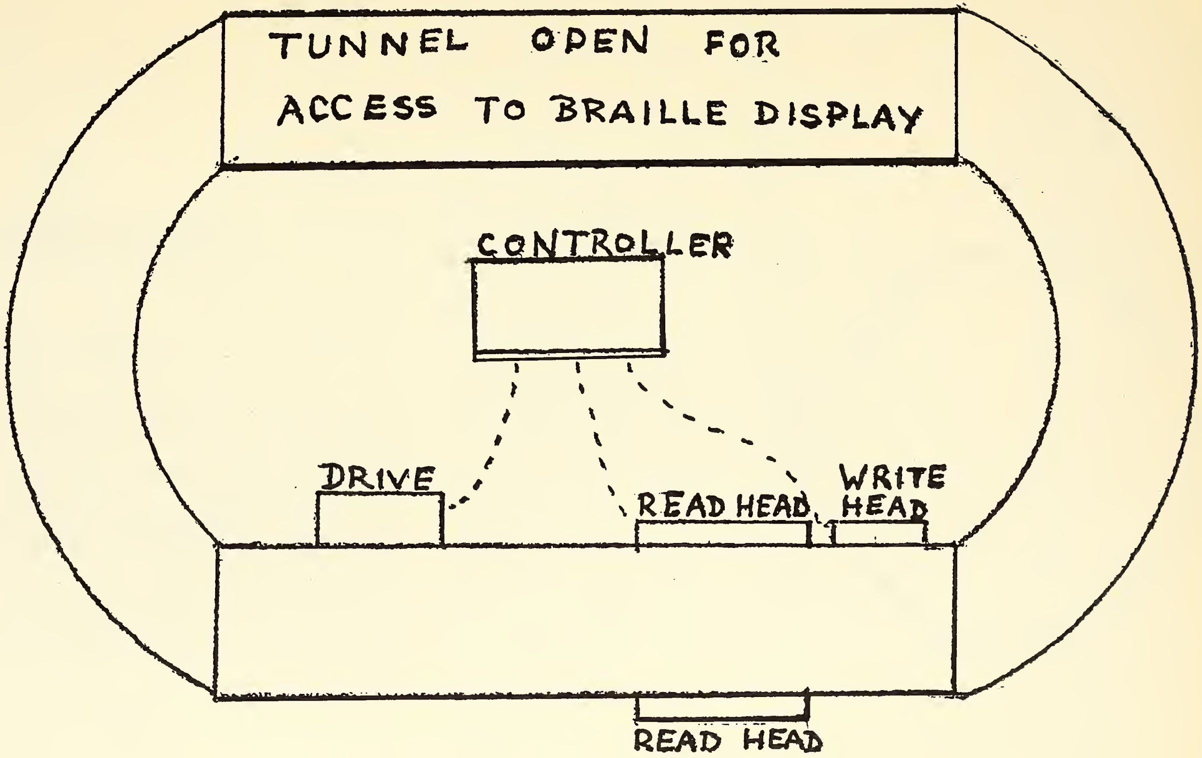


Figure 1: Layout of the Display Seen from Above

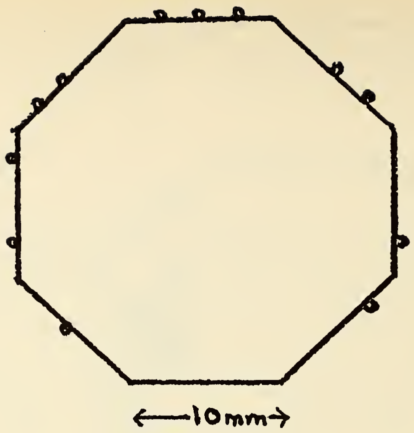


Figure 2a: Cross-Section of Octagonal Element

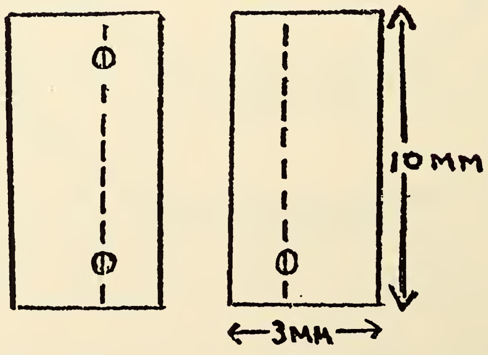


Figure 2b: Side View of two Elements Forming a Braille Cell

One possibility would be for a two-line display in which output went to alternate lines. The left hand would then move alternately to the first and second line as the next line to be read. This would, in all cases, add greatly to the cost, doubling the cost of the braille elements needed. A braille graphical display unit could, of course, adequately simulate a braille page for all kinds of readers, and this might be a neat if expensive answer. Clearly, though, the current generation of displays is somewhat defective.

A NEW APPROACH

Philosophy and Form of the Display

At this point I should like to elaborate on my own design for a display. It aims to fulfil all the criteria discussed above, in particular being light, compact, cheap, noiseless and reliable. It is a transient display similar in principle to that of Grunwald.

The display moves from right to left under the user's fingers and under his complete control. I envisage around 20-25 cells in the display allowing him to rescan misread characters.

The display consists of a series of octagonal elements forming a belt which runs through an octagonal tunnel (see Figure 1). The tunnel could be in a horizontal or vertical plane. Each element represents half a braille cell (Figure 2a), the eight faces containing between them the eight possible half-cell combinations of dots. The dots are slightly off centre, allowing the same element type to be used for both the left- and right-hand half of a cell (Figure 2b).

The elements would be of a hard polymer such as Dacron, the tunnel being of the same substance or one of similar hardness to minimise wear. Using this type of approach should minimise production costs: a mere 10,000 displays would need over a million half-cell elements, a very reasonable production run. If the tunnel were made in sections which were afterwards bonded together, this could be made to increase the production run for each type; this technique would in any case be necessary for the insertion into the tunnel of its elements and other components discussed below.

As shown in Figure 1, there would be a cut-out in the top of the tunnel to allow the reader access to the display. Very careful attention would have to be given to the nature of the surface of the element and its dots: it must not be so rough as to cause excess friction with the fingers. An additional problem might well be the buildup of perspiration on the surface, and some way might have to be found of removing this as the belt passed through the body of the display. Only experimentation would determine whether this was a problem.

Similar experiments would be needed to determine correct dot size and shape for this particular kind of display. I emphasise the need for such experimentation even if it results in a delay in the display's initial production. Current dynamic displays, while not ideal, are quite satisfactory for the time being, and allow us to be more thorough in developing the next generation of displays.

Movement and Control

Two stepper motors would provide movement. The first would move the display elements along the tunnel one half-cell at a time. One of the problems I have yet to solve is the type of contact between the motor and the belt. The obvious answer is a rubber wheel or belt pressing on the elements; however the wheel/belt would have to be rough to grip the surface satisfactorily. Although this might be arranged, it also introduces a potential weak link into the system due to the wear of the wheel/belt. I have looked at various alternatives but am not entirely satisfied with any of them.

One possibility is to have the write station described below itself move forward after aligning an element, then moving back to receive the next one. This may be fraught with problems. For one thing, some way of clamping the element temporarily into the station would have to be found, so that it moved forward with the station before being disengaged so that the latter can return without it; during its return the write station could pick up the next element.

Using this technique, a motor would not be needed to drive the belt; a solenoid would suffice, being loosely coupled to the write station - the station resting in a simple groove. However, this apparent simplification is offset by the need for some clamping mechanism in the station!

The other motor governs the orientation of the write station, and hence the element within it. The heart of the station, shown in Figure 3, is a wheel of thickness <= that of the half-cell element, which it receives into an octagonal cut-out in its centre. A gear connection with the second motor would enable this wheel to be rotated so as to position the current element within it, prior to its appearance in the display area. The wheel could be constructed with teeth around its circumference.

Thus, if this were also the drive station as discussed above, there would be two operations, which might also be run in parallel to speed up the process:
(a), rotate the current element to its new position;
(b), move it on one element width, thus pushing round the belt.

These two operations would be controlled by a central processor - a microprocessor or even a programmable logic array would be sufficient. This would have a number of inputs and outputs, described below. One output would control the write motor and another that for belt movement.

A read station is also envisaged: this would determine the position of its current element, to be stored by the controller against the time when it became the current element in the write station. It might be argued that this is unnecessary as, once set up, there is no way in which the element positions could change except in a way known to the controller. While this ought to be true, this "belt and braces" approach achieves two things, the second a happy byproduct of the first:
(a), it ensures beyond all reasonable doubt that the display output is 100% accurate even if an element is somehow disturbed;
(b), it means that volatile storage can be used in the controller, at least for most purposes.
Read-only memory would probably be needed only for the control program and the data for character-to-braille translation. This would cheapen the

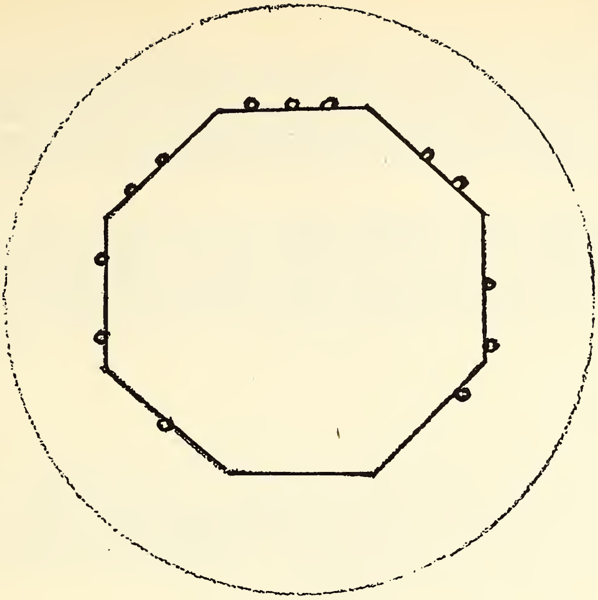


Figure 3: Cross-Section of Write Wheel
(gear teeth not shown)

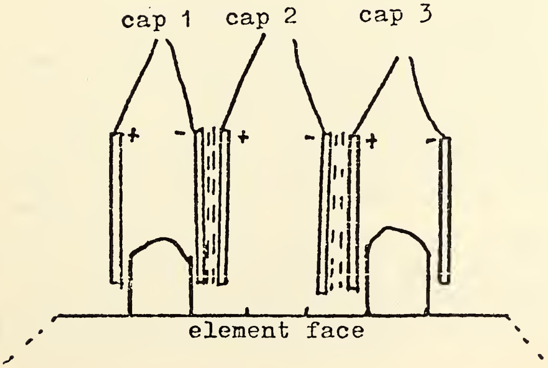


Figure 4: Capacitance Read Head for One Element Face

device a little.

Another input to the controller would indicate whether the user wanted the display to stop. This might be done directly by the activation of a switch, or indirectly by sensing whether his hands were in the reading position. Another input, set by the user, would determine the speed of motion of the display, the controller translating this into the operational speeds of the motors.

A third input would indicate the need for backtracking, in which case the controller might, for example, restart the current output buffer (but see below). For this reason, a cyclic buffer might be the most appropriate, able to store several times the display capacity. More complex functions would be under the control of the device using the display.

The Read Station

The read station could employ one of a number of techniques to detect the dots on the surface of an element. Depending on how a given technique were implemented, one head would be needed either for each face or for each dot position on a face - 8 or 24 heads.

The first is mechanical contact of some kind, e.g. spring-loaded wheels or rollers, one for each dot position. The problem here is the usual one of mechanical wear and near with consequent unreliability. (Research elsewhere indicates that this is not the best or easiest way to sense dots, at least on paper).

The second technique is optical sensing. Again the same research predicts problems with shadows, though at least in this case dot positions and sizes could be guaranteed. However, the light source could introduce another element of unreliability.

The third is ultrasound. I mention this only for completeness; in fact the distances involved are too small to make this a feasible proposition.

I suspect that the most likely method to succeed involves the use of capacitance. However I have not yet tested it out in any way. The idea is illustrated in Figure 4: two electrodes are positioned on either side of a face of an element, almost touching it, one pair for each dot position. These form the plates of a capacitor, its dielectric being the intervening air. This would be modified by the intrusion of a plastic dot, the modification being detected by appropriate circuitry. A simple dot or no-dot binary output would be given. One per dot position would be needed as, although this system would differentiate between different numbers of dots, it would not allow their positions to be determined.

For pragmatic reasons, I envisage the array of detectors spread over a number of cells as shown in Figure 1. It should be quite feasible to space the 24 heads over, say, six cells. The controller, aware of this spacing, would calculate the current orientation of a given half-cell from these binary inputs.

Interfacing

As already indicated, stop-start and speed controls could be part of the display. However, an extra input must be provided for buffer control from the device in which the display is used. It should allow for a buffer to be filled, and for a pointer to be moved by the device to any part of the buffer.

Thus the display should appear to its device as a simple, two-input device: a data input taking characters, and a control input taking a binary value representing the offset from the beginning of the buffer. Each new character would drive its prodecessors along the buffer as os usual.

CONCLUSION

This paper has sought to outline a new approach to the design of a transient display, deriving the rationale for it from a survey of current approaches. No detailed design work has yet commenced on this display, let alone implementation. It is my hope that someone else will take up the ideas I have discussed and developed them further.

ACKNOWLEDGEMENTS

I am grateful to C W Garland for suggesting to me a circular rather than a linear element, from which I finally derived the octagonal design. Also to the very many other people who have contributed to my thinking through discussion. Miss G B Warner helped greatly in preparing the diagrams, and in printing the final master from the computer; a special word of thanks goes to her.

My family deserve a mention for their patience with me when this work had sapped mine with them. I dedicate this paper to the glory of God from whom spring all my best thoughts.

REFERENCES

1. A Grunwald: A Braille Reading Machine. Science, vol. 154, 7th October 1966, pages 144-146.

2. P.W.F. Coleman: Computer Terminals For The Blind. Electronics & Power, March 1972.

3. P.W.F. Coleman: Microprocessors Aid The Blind. Microprocessor Workshop, University of Liverpool, September 1977.

4. P W F Coleman: Microprocessor Systems for the Blind - A Cautionary Tale. Proceedings of the Conference on Microprocessors in Automation and Communications September 1978; IERE Conference Proceedings No. 41.

APPENDIX I

POSSIBLE CRITERIA FOR CLASSIFYING BRAILLE READING TECHNIQUES

I have heard it said that there appear to be as many ways of reading braille as braille readers. This state of affairs is particularly unhelpful to those designing braille displays for equipment for the blind.

The following headings and diagrams, therefore, illustrate a set of possible classifications of braille reading techniques. I am particularly indebted to J Lorimer for pointing out that I had failed to allow specifically for right- or left-handed dominance in my initial classification. We therefore need two parallel classifications, one for left- and the other right-hand dominance. The right is given here.

Right-hand, single-finger trace/vibrate,
retrace, locate:

```
  -------------\/\0/\/------------ 1

  -------------------------------- 2
```

Right-hand, single-finger smooth trace,
retrace, locate:

```
 -------------------0------------ 1

 -------------------------------- 2
```

Right-hand, single-finger trace and retrace,
left-hand locate:

```
| -----------------0-------------- 1
| 0
v -------------------------------- 2
```

Right-hand, multi-finger trace and retrace,
left-hand locate:

```
| ----------------000-------------- 1
| 000
v --------------------------------- 2
```

Right-hand & left-hand, single-finger
trace & retrace, left-hand locate:

```
---------0----------0------------ 1

--------------------------------- 2
```

Right-hand & left-hand, multi-finger
trace & retrace, left-hand locate:

```
-------000--------000------------ 1

--------------------------------- 2
```

Parallel right-hand/left-hand,
trace/retrace/locate:

```
--------------------------0000--- 1

                0000
--------------------------------- 2
''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''"
--------------------------0000--- 1

------0000----------------------- 2
''''''''''''''''''''''''''''''''''''''''''''''''''''''''''''"
--------------------------------- 1

------0000------0000------------- 2
```

ADDITIONAL PAPERS NOT FORMALLY PRESENTED AT THE CONFERENCE

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

A PROJECT FOR THE COMPUTERISED PRODUCTION OF BRAILLE IN SPAIN

by

Alfonso Recuero

Instituto Eduardo Tomoja
de la Construccion y del Cemento
Costillares Chamartin
Apartado 19002
Madrid 33

COMPUTERIZED BRAILLE PRODUCTION TODAY AND TOMORROW

London, May 30 to June 1, 1979

A PROJECT FOR THE COMPUTERIZED PRODUCTION OF BRAILLE IN SPAIN.

by ALFONSO RECUERO[1]

Since 1977 the author planned and tried to execute a project for the automatic production of Braille in Spanish. In March, 1979, the saving institution Caja de Ahorros y Monte de Piedad de Madrid decided to support the project giving financial aid to the ONCE (Spanish Organization of the Blind) to execute it. In spite of this approval, and due to bureaucracy, the funds have not yet been delivered, and because of this reason the real work has not yet begun. So it is only possible to speak now about the project, but not on the difficulties found or on the goals obtained.

The project consists of a series of chained phases, the first of which is considered as the fundamental one. This is the phase to be started immediately.

The main objective of this first phase is the automatic production of grade 1 Braille using as output device a Braille teletype, very probably the TEM-8-BR. Input data would be either keyed information or compositor tapes used by the printing houses in electronic text production.

The Instituto de Artes Gráficas Tajamar will also participate in the project. This center has designed a software package for electronic text edition. Its system is on work in some of the most importante printing houses in Spain. This Institute will modify the software package in order to fit the output to the embossing terminal chosen. They will also cooperate in the relations

---

(1) Instituto Eduardo Torroja de la Construcción y del Cemento -
Apartado 19.002 - Madrid-33 - Spain -

with the publishers who have the system in operation to allow us to use their data to produce Braille books.

The system will be supported by a computer PDP-11 with 32K words of core memory. It will be provided with two magnetic disks with a total capacity of 10 Megabytes, paper tape station, cathodic rays tube and a character printer besides the Braille embosser.

The package, in Assembler, allows multitask, so that every terminal may be making a different job. The system is a very evolutioned one. It produces a perfect hyphenation according to orthographic rules in Spanish and it permits a very varied range in text edition.

As soon as the system will be fully ready and the staff who will operate it is well trained, it will be delivered to the ONCE for its explotation. The period to get this first objective has been estimated in one year. The system would be productive from this moment on.

The objectives to be accomplished in future when this first step will be completed are:

- Production of paperless Braille, to be read by devices such as Digicassette, Braillechord, Braillex, Versabraille, etc, including in the recording keys for the eléctronic search of the items.
- Spanish grade 2 Braille production.
- Production of Braille text with a great number of copies, using Braille printing machines.

In parallel with the above mentioned objectives the system will be used - if required for many other functions in the ONCE, such as administrative work, control, and to facilitate to spanish blinds the access to the Informatics as a field of work in all its levels. Nevertheless all this parallel objectives, although they are intimately linked to this project, are outside its skope.

The autor thanks very much the World Council for the Welfare of the Blind for the organization of this conference, which allows him to know the problems found by those who have worked in these matters in different countries in the world. He hopes that this will facilitate his future work. He takes profit of this opportunity to offer his collaboration to all those who may want it. He also will welcome all the informations and cooperations which permit him to skip the difficulties that he will surely find in his work.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

SUMMARY OF PAPER ON FUTURE OF BRAILLE

by

Leslie L. Clark

Computer Centre for the Visually Impaired
Baruch College
University of New York

WCWB COMMITTEE ON CULTURAL AFFAIRS
INTERNATIONAL CONFERENCE ON COMPUTERISED BRAILLE PRODUCTION

For Interest

Summary of paper on Future of Braille

by Mr. Leslie L. Clark (formerly Director of Technological Information American Foundation for the Blind, New York)
Adjunct Professor,
Computer Center for the Visually Impaired,
Baruch College,
The City University of New York.

Although the population of blind persons is increasing, the use of braille as a reading, writing, and communication medium is declining. The reasons for this curious situation lie not only in benign neglect of the medium, but also in poor teaching, inadequate mix of materials needed, long ignored difficulties in the braille code, and lack of national debate on the issues involved. This paper tracks the major consequences of difficulties resulting from the above causes, gives brief descriptions of the problems involved, and suggests one or more solutions to each of the problems. As a system analysis of the production, deployment, and use of the braille medium, the paper is intended to provoke discussion about the problems treated, not to provide unique solutions of them.

The paper will be published in two venues:

1. *Braille Research Newsletter* No. 9 (in press)
2. The *Annals* of the New York Academy of Sciences (Section of Linguistics, 1978, Part II)

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

BRAILLE VIA THE SELECTRIC TYPEWRITER

by

Norman C. Loeber

IBM Corporation San Jose
California USA 95193

BRAILLE VIA THE SELECTRIC TYPEWRITER

Norman C. Loeber
International Business Machines Corporation
San Jose, California U.S.A. 95193

INTRODUCTION

An experimental system has been developed to provide a simple and direct method for producing a document in either braille or standard typewriter characters. This system directly benefits the blind and is an example of the potential of present day technology.

The primary goal was to make braille more readily available to the average user; that is, an every day working tool for the blind. It was not designed as a way of mass producing braille, but as a device that would create job opportunities for the blind and enable them to do the same work that sighted people do.

Experimental work on this system was made possible by a special grant from IBM to encourage professionals to broaden their perspective and explore devices and techniques outside of their normal job responsibilities.

This paper has two major purposes. First, to inform others of what has been accomplished and, second, to encourage a united effort in both the blind and sighted communities to produce a braille embossing capability for the average blind person striving to participate in today's sighted world. Much effort is needed to develop and standarize an acceptable universal code which all countries can use. This would reduce for all of us the development and production costs of a braille communicating translator.

As the blind are integrated into our schools and business communities, the need for braille copies of ordinary documents becomes more and more a necessity. Measuring the availability of such documents in days is unacceptable and, in some cases, useless. Availability in hours is also too slow. Immediate translation is required. The production of duplicate documents in typed characters or braille will open new job opportunities and broaden the perspective of the average blind individual.

The basic element of our system is the typewriter--the IBM Selectric (TM) typewriter. In recent years, the capabilities and versatility of this typewriter have been greatly expanded by word processing systems and the addition of electronic controls and memories. These devices can be and should be beneficial to the blind, both to the blind secretary/typist and to the professional.

IBM SELECTRIC TYPEWRITER

The typewriter used in this application has a standard keyboard for data entry, but its output versatility has been expanded. The document may be printed for use by a sighted operator or embossed in braille for the blind. The Selectric typewriter uses a removable typing element instead of type bars. Different elements with different type styles may be used on the same typewriter. Various foreign languages can be accommodated by simply changing the typing element. To accomplish the braille dot impressions, a print element was modified to include a braille cell with depressions for each of the six possible dots.

When braille output is required, the element with the braille cell is used along with the stylus assembly. Regular typing is accommodated by a standard element as chosen by the operator.

MAGNETIC CARD SYSTEMS

IBM has expanded the capability and versatility of its Selectric typewriter by incorporating it into a Magnetic Card System. This allows information entered at the keyboard to be stored for later printout or modification.

The Magnetic Card System is a word processing unit with a magnetic card recorder and reader. In addition to a temporary memory which may be altered or changed, it has a more permanent memory, namely the magnetic card. The magnetic card is the same size as the punched card used by IBM for many years. The magnetic card is coated with a recording material on one side. Information typed into the machine is recorded on the card in magnetic codes. Each card contains approximately 5,000 characters.

The card is read by the system and the information may be printed out as many times as required, or corrections and changes may be made, or the card may be recorded again. This magnetic card may be used with various models of Magnetic Card Systems to accommodate various job requirements. Some models can duplicate multiple cards for distribution to other users. With the incorporation of the braille device on the Selectric typewriter, information recorded on the magnetic card can be played back in either braille or typewriter characters from the same card.

Some of the Magnetic Card Systems have a communicating option which allows them to transmit the information by means of telephone lines to remote locations where a new magnetic card is recorded. This provides rapid distribution or transmission of information between offices or libraries.

BRAILLE EMBOSSER

One of the initial objectives of this device was to emboss on the front side of the paper so that the operator could read the braille embossing without removing the paper from the typewriter. The plan was to emboss the braille cells moving left to right across the page to allow reading in the normal left to right sequence. To accomplish this a special embosser mechanism was designed. The embosser is mounted in its own console located behind the typewriter. It includes a stylus assembly operated by six solenoids. The solenoids move the six stylus bars corresponding with the six dots of the standard braille cell pattern. The standard braille spacing configuration allows for 40 braille cells per line. Setting up a braille pattern to represent a desired character is accomplished by operating the solenoids which force the desired pins forward. This is followed by impacting the type element with its dot depressions against the protruding pins. Forming the paper over the protruding pins provides raised dots on the front of the paper.

TRANSLITERATION TABLES

A transliteration table converts the normal codes used in the Selectric typewriter to the specified braille code. These codes are stored in the read only memory. More than one table may be provided and customized to a particular braille code or language. For this prototype, braille is provided in direct single cell transliteration; that is, letter for letter representation. The transliteration occurs during the output function. With the machine set for braille output, information entered on the standard typewriter keyboard is translated into braille and embossed on the paper so the visually impaired operator can immediately read what has been typed.

Similarly, a card previously magnetically encoded with information, keyed in at the standard keyboard, can be read by the system and translated into braille with its embossed output. If the Magnetic Card System is equipped with a Communicating Feature, the device can be an interactive terminal. The computer output would be translated and printed out with the braille embosser. In this case, the system becomes a braille embossing terminal.

USERS

A device of this type can be beneficial to the blind in several ways. First, there is the secretary/typist who would actually be using the machine for transcribing various dictated letters or reports, or the individual who would type an original letter. Then, there are opportunities for the professional individual. A lawyer, for example, may not know how to type and he certainly does not want to take the time to type or braille his material. However, his sighted secretary could very easily run off a braille copy of correspondence or legal documents by simply feeding the magnetic card into the machine and having it emboss in braille the information recorded on the card. This is helpful in numerous business applications and provides the blind professional with a ready and immediate source of braille material and a ready source of information equivalent to that received by his sighted colleague.

Another use for such a system is the rapid and convenient distribution of limited amounts of information. Consider the library or school that wishes to send out a braille copy of an article or report. It can be typed and quickly converted into braille for the blind individual. In many cases, the article or report may have been typed and exists on a mag card or in a computer so that a second typing is not even necessary. Many organizations of and for the blind need to make a few copies of some document for distribution to their membership or a letter between officers and/or committees. This can be easily done with a magnetic card system equipped with a braille embosser.

APPLICATIONS WITH WORD PROCESSING SYSTEMS

Magnetic Card Systems equipped with embossers allow the visually impaired operator to enter and correct documents. If a typed copy is needed, it is only necessary to change to the type mode and set the machine to play back the magnetic card in conventional type. The visually impaired operator can then send the typed copy to a sighted individual. Conversely, a sighted person can type and record on the card. Letters, memorandums or reports can be corrected and updated and then, when a braille copy is needed for visually impaired associates, the embosser unit can be inserted. The Magnetic Card Systems using the transliterature table will automatically translate the document into the braille code and emboss the paper as required.

The Magnetic Card System, when equipped with a Communicating Feature, can be used as a braille embossing terminal by programmers or others. It can be used offline to record various programming codes or text material and then, after this has been organized and structured, it can be transmitted to a computer thus requiring a very limited amount of online connect time. Or it may be used online with the computer helping the blind operator by prompting or questioning incomplete responses.

One of the significant opportunities with the braille output device connected to a central computer is the interrelationship of various data bases. The many online bases that are currently available in various countries could be interrogated from remote locations and the output produced in braille. As an example, many of the laws for the state of New York are presently stored in a computer and are available through terminals.

Many visually impaired lawyers could benefit from the capability of interrogating this data base and receiving braille output directly in their office for their daily work. Or, they could receive from their secretaries a braille copy of various documents or briefs.

Another significant advantage of the Magnetic Card System is the portability of the recorded card. Mailing or transporting embossed braille is a difficult and awkward problem since it can be easily damaged unless carefully packed. A magnetic card, on the other hand, can be mailed in a plain envelope at low cost and with little damage anticipated.

Since magnetic cards can be created on one Magnetic Card System and played back on another, they may be distributed and printed out in different locations in braille or type as required. This opens up many opportunities in the business world, libraries or education areas for producing braille documents in small volume.

BRAILLE CODES

Our original work was accomplished with a simple translater to provide cell for cell translation. With the new large-scale integrated chips presently used on the newer computers, a much more sophisticated code could be handled by a terminal or word-processing system. It may be possible to provide grade 1-1/2 or perhaps even grade 2 braille from the original typewritten input. However, the development of such an LSI chip is very expensive and would require concurraece of a universal code between various countries to justify the large development expense.

The braille script of six dots, universal throughout the world, gives us an extraordinary basis of communicating between various languages. If we can properly identify a country's language to the braille cell, it might be possible to relate from the braille of one country to the braille of another country.

CONCLUSIONS

Several prototypes of a braille embossing terminal and embossing magnetic card system were built, tested and used by blind individuals in various job assignments. They have proven to be very helpful and beneficial. The direct using individuals have been able to undertake a greater variety of job responsibilities and to perform tasks that were previously impossible for them to do. Many others, though not actual users of the equipment, have benefited by the more timely information now available to them in braille.

BIBLIOGRAPHY

1. Braille Automated Newsletter, "Using Punched Cards for Automated Braille," Editors: J.M. Gill & L.L. Clark, Aug. 1976.
2. N.C. Loeber, "Proposed Braille Computer Terminal Offers Expanded World to the Blind," Fall Joint Comp. Conf., 1971.
3. IBM Corp., IBM Mag Card II Typewriter, Armonk, N.Y.
4. ACM SIGCAPH Newsletter, "Modified Standard Office Equipment for Braille Translation and Embossing," v. 21, Nov. 1976.
5. Think Magazine, "Breakthrough: Computer Braille for the Blind," W.D. Blankenship, July/August 1974.
6. American Foundation for the Blind Research Bulletin, "Interactive Terminal Use," v. 24, J. Boyle, W. Jacobs, N.C. Loeber, April 1974.

NORMAN LOEBER works in the Product Assurance group of the General Products Division of IBM. He has an extensive background of developing equipment and systems. As an inventor, he has contributed to storage technology, terminal, and printer developments. He was in charge of IBM's Braille Embossing research during the mid seventies.

He has done graduate work at several universities, and is a senior member of the Institute of Electrical and Electronics Engineers.

His interest in Braille dates back over thirty years, when as a volunteer for the Lutheran Braille Workers, he designed and built Braille producing equipment. His 1955 design of the Automated Braille Keyboard has been incorporated into several different systems and still serves as a model for current equipment. He has designed and built a number of high volume Braille embossing systems.

He has also been associated with the National Braille Association (having received their merit award for distinguished service), the Visually Impaired Data Processors International, the Visually Impaired Secretarial Transcribers Association and the Special Interest Group on Computers and the Physically Handicapped, a subgroup of the Association for Computing Machinery. He also is a member of the American Council for the Blind and is a past Director of the Lutheran Braille Workers Assciation, Inc.

* * *

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

THE ROSE BRAILLE DISPLAY READER

by
Stanley E. Rose
Joan B. Rose

Rose Associates
Falmouth, Massachusetts
USA

THE ROSE BRAILLE DISPLAY READER

A FULL-PAGE PAPERLESS BRAILLE DISPLAY DEVICE

by
Stanley E. Rose
Joan B. Rose

BACKGROUND

For the past three years ROSE ASSOCIATES, INC. of Falmouth, Massachusetts, USA has been developing a full page paperless braille display. This effort has been partly funded in the past year by the Department of Health, Education and Welfare, Office of Educations, Bureau of Education for the Handicapped.

The BEH project officer for this project is Dr. Allen Dittmann whose aid and suggestions are gratefully acknowledged in achieving our success to date.

Necessity being the mother of invention, it is the purpose of this paper to take the reader through ROSE ASSOCIATES'perception of the necessity for a full page braille display and how we arrived at the configuration we are now building in prototype form.

In order to understand the need for a full-page paperless braille display device and to understand the purpose it serves, it is useful to review the reading needs of blind persons.

Presentation of Information in Braille

In braille, the presentation of information is by the use of a "cell", which is made up of from one to six embossed dots, arranged in a two by 3 (three) matrix dots, more or less as required, as follows: ⠿ . By touching the cell and recognizing which of the 63 possible combinations of the dots in the cell, the blind reader identifies the braille character thus represented. There are various levels of braille. Grade One braille (used mainly for children and those learning braille, and also used for foreign language braille), is a form of notation which simply transcribes each inkprint letter into a corresponding braille symbol (cell). Because braille cells are uniform in size and are large, relatively speaking, Grade One braille is inefficient and uneconomical for general reading purposes and applications.

Grade Two braille is the common and customary formulation used for braille transcription and reading in English-speaking countries. Grade Two braille uses a kind of shorthand of "speedwriting" to reduce the number of braille cells needed to transcribe the inkprint to read. "Contractions" are used for this purpose. For example, the word "knowledge" is contracted to the symbol for the letter "k". The word "and" is contracted to a braille cell which contains dots numbered 1-2-3-4-6. The letters "st" when they appear together in an inkprint word are contracted to a single cell which contains dots numbered 3-4. The word "stand" can thus be contracted by using the cell (3-4) for "st" together with the cell (1-2-3-4-6) for "and". Because of the inherent idiosyncracies of the English language, the rules for using "contractions" are complex. Anyone who assumes responsibility for transcribing inkprint materials into Grade Two braille must be thoroughly versed in grammar, spelling and punctuation rules, in order to maximize the number of "contractions" which can be used to braille the materials at hand. The failure to use "contractions" and hyphenation of words is directly felt by the blind reader, not only because more pages of braille are needed to produce the materials and the books which are thus produced are heavier and bulkier as well as more costly for the reader, but also because the material will not read smoothly and its meaning may become unclear as a result.

Grade Three braille is an even more contracted form of braille, but at this time is not commonly used.

There is a special braille code for mathematical notations, where the materials are primarily mathematic rather than literary. There is also a special braille code for music notations. All use the standard braille "cell", but the code is different.

Historical development of Reading Devices for the Blind

At this point, it is useful to mention a little about the development of reading materials for the blind.

Originally, the only way that a blind person could obtain access to information contained in inkprint materials was to have a human reader recite from those inkprint materials. This, obviously, required that a sighted person be available and be willing to perform this service whenever and wherever needed. It also required that the sighted person have sufficient education and perceptivity to be able to recite from the inkprint materials in an informative, knowldegeable way. The inefficiencies and disadvantages of this method are self-evident.

MOON

The earliest substantive response to the needs of blind persons for a viable alternative to a human narrator was the development of "MOON", a system of embossing the Roman alphabet onto a page, using very large size letters. The blind reader was expected to feel each letter individually and recognize it, and to read by spelling out the words one letter at a time. This was terribly cumbersome for the reader, and the process of producing MOON was also slow and expensive. It freed the blind person from being tied to a human reader, but it tied him to a tediously slow system of "reading" which could not offer him much in the way of reading material even if he became adept at it. It also tied the blind reader to a system of symbol configurations--Roman letters--which are hard tactilely to identify and to read with any speed. The Braille system was a response to the deficiencies and inadequacies of MOON.

The braille system developed by Louis Braille gave the blind the ability to read with the same fluidity and speed as the sighted have when using their eyes. Empiric studies made by BRAILLE INC. of Falmouth, Massachusetts, a publisher of braille books, indicate that the average reading speed of a competent blind reader of Grade Two braille is about 2.5 minutes per braille page. Since a braille page usually contains 25 lines of braille, with about

40 cells of braille per line, this is equivalent to a reading speed of about 6 seconds per braille line.
This appears to be faster than the average narration time needed to recite meaningfully the same inkprint material.

Production of Braille Reading Materials

Unfortunately, while Louis Braille developed a viable system for the blind to read through tactual symbols, he lived in an era when science had not developed enough to make it possible to produce braille other than manually. Thus, even though the blind were freed from human narrators, they were still tied to human transcribers, who were not more available to do transcribing work than were the human narrators who were their predecessors. Even today, in the United States, the overwhelming bulk of braille for the blind is produced by human volunteer transcribers, using their own manual typewriters (braillers) and supplying their own paper, and making one copy at a time for blind beneficiaries of their charity. These volunteers work at home, mostly on a part-time, sporadic basis, so that there is no way of predicting how long or when their work-product will, if ever, be completed. In most instances, there is also no way of controlling what these volunteers choose to transcribe, so that they may devote their skills to books which are needed by none and be indifferent to the demand for other books for which there is a great demand.

The braille system was enhanced in this century by the development of the braille typewriter, the most commonly used of which is the well-known Perkins Brailler, made and sold by Perkins Institute for the Blind, Watertown, Massachusetts. This is a manually-operated desk-top type unit, actuated by pressing six dot-embossing keys in various combinations in order to type each braille cell. Recently, this machine has been updated to allow for electric powered operation by a human transcriber, in the same way as electric typewriters are powered. This type of operation tends to reduce errors caused by operator fatigue and also tends to make embossing equal in its appearance and size, and thus is an improvement over manually-powered braillers.

These electric braillers are only now being produced in relatively small quantities. Most braillists do not use them.

The standard work-product of braille transcribers is a paper page, which is a standard dimensions--11 inches by 11-1/2 inch page, which is the maximum size page that it can handle. A Perkins Brailler winds the blank paper onto its platen (roller) completely, and the paper is fed out of the typewriter after it has been embossed; and the machine is designed so that no more than an 11 inch page can be wound onto the platen. This has operated to establish a world-wide standard for a braille page. Shorter pages are possible, but are not used for anything except notes, special purposes and limited material where size is a factor, such as children's books.

Ordinarily, it takes about two (2) pages of braille to transcribe one (1) page of average inkprint literary materials. Therefore, in an inkprint book which is 100 pages long, it will probably take 200 pages of braille to transcribe it. However, the standards used in braille give rise to a number of functional characteristics, including serious limitations, which restrict what can be done to produce a braille book. For example, the very size and weight of the paper used for braille make it undesirable to include more than 100 braille pages in a volume. The braille embossing makes the book very thick, and the dimensions of the pages make the volume heavy, bulky and unwieldy. In addition, since a blind reader must usually read with one hand moving across the page (and often with the other marking the line being read at the moment), it is almost mandatory that the volume be bound in a way which permits the volume to be opened flat at all parts of the book. To bind more than 100 pages to meet this need is very difficult. It follows that ordinary braille books are produced in volumes of about 100 pages each.

Notwithstanding the size, weight, bulk, and limitations on the number of pages per volume of braille books, they have represented the greatest source of freedom and benefit for blind readers to date. Empiric studies by BRAILLE INC. have also indicated that blind readers prefer to read braille more than they like to listen to narrated material on records or tapes. The problem of making braille books available to the blind is related to the attitudes of those who have tried to deal with the needs of the blind from a sighted person's point of view. They have worked to invent means to solve the problems of the sighted, instead of dealing with the problems of the blind.

Within the subject of conventional braille transcribing, press braille should be mentioned. Press braille is a form of paper braille which is produced using zinc plates and modified printing presses. The zinc plates are stamped out on a stereotype machine, so-called, which is operated by a human braillist or can be driven by a magnetic tape containing braille data. The value of press braille is that it permits embossing on both sides of a single sheet of paper, through interpointing of the dots, and so permits more braille to be bound in a single volume. However, even though press braille thus saves space, it is not practical to use it in a majority of instances because the cost of plate making and setting up the presses can only be justified where there is to be a long run on the press. If there is no immediate use for all of the copies made, there is a cost for the materials, and storage, and the printing itself that will not be reimbursed for some time. In most cases, there is no immediate demand for multiple copies of a single book in quantities which are economical for producing press braille. Children's basic braille text books, religious books, and similar materials are an exception.

For making single copies of braille books, a practical current solution has been to make thermoformed copies of original braille pages, one page at a time, on PVC sheets in a machine that has been developed for that purpose. The resulting book is essentially a plastic copy of the original materials, and is equally bulky, unwieldy, heavy and expensive. In addition, for some blind readers, something in the PVC causes a loss of tactile sensitivity after a limited reading period, so that the person's reading capability is reduced.

Recent Approaches for Production of Materials for the Blind

Virtually all current approaches to the problem of giving blind persons access to inkprint books have fallen into one of two categories, neither of which offers anything new to blind persons. One approach has been to produce Talking Books, socalled. This refers to the program sponsored by the U.S. Library of Congress, in which special phonographs are loaned out free-of-charge to blind persons, who can thereafter borrow from a regional Talking Book library various prerecorded phonograph records to be played on the borrowed machine. In Massachusetts, for example, in 1976 (according to the Massachusetts Commission for the Blind) there were 7456 Talking Book Machines (phonographs) in use by blind persons in the State. A more recent approach has been to produce prerecorded magnetic tapes, in standard tape cassettes, for the same purpose. Again, in Massachusetts, for example, in 1976 these were in short supply; and only 2,200 cassette books were in use.

What apparently has not been recognized in regard to the socalled Talking Book Program or the cassette tape program is that they actually are returning the blind person to the era before MOON and before braille was invented. They are placing the blind person in the position of being recited to by a human reader, with all of the inefficiency and disadvantages of that method. The blind person is forced to subordinate himself to the limitations of speed of speech, inflections, phrasing and timing, as well as the actual or nonexistent understanding of the subject, of the narrator. And though modern sound recordings have eliminated the utility factors whereby the blind listener could ask a human sighted reader to skip backward and forward quickly through the inkprint materials to refresh or reinforce what was being read.

Moreover, even though a phonograph or a tape recorder-player can be equipped with the means whereby a blind person can operate the device rapidly backwards or forwards for the purpose of refreshing and reinforcing the data being heard, these do not give the blind reader an easy way without disturbing their reading process to return to their place where they were primarily listening.

And although devices now exist to enable sound recordings to be played at a higher speed while retaining listening intelligibility, these devices do not solve the problems connected with searching for other data on the same recording and then returning. Furthermore, the more difficult the subject matter, the less likely it will be that higher speed playbacks will be useful.

It is believed that overwhelming evidence supports the proposition that a voice recording is not the intellectual or emotional equivalent of a "readable book". A book permits the reader, wether blind or sighted, to allow his own personal imagination and associative thought processes to function freely while he reads. The text of a book is converted into a highly personal experience and point of view, depending on how the reader's mind is struck by and with the material. A sound recording demands the full attention of the listener, for the most part, and thus limits his ability to apply his own imagery and values to the data being presented. For that reason, as well as others, the reading of a book ordinarily provides more personal satisfaction than any narration of the text of that book. The main exception to this is where the material is primarily intended for oral presentation, such as a theatrical play. Empiric studies by BRAILLE INC. indicate that blind persons prefer to read braille rather than to listen to voice recordings of books.

Another main approach to giving blind persons access to inkprint materials has been to dispense with braille and to return to a form of MOON with a device called the Optacon. The Optacon is a table-top device which combines an optical scanner of inkprint with an electromechanical device in the form of a black box in which the blind person places his hand. Inside the box, the blind person is given a raised copy of the optically scanned material to feel in the form of a vibrating surface of about 1.5 square inches. In other words, the Optacon presents the inkprint, one letter at a time, in enlarged tactual form for the blind person to "read". In other words, MOON! The major difference between an Optacon and MOON is that science has solved the sighted person's problem in producing raised Roman letters for the blind. There is, of course, an appropriate place for an Optacon in the lives of many blind persons, particularly those who have need to read inkprint materials of very limited nature and in varied forms, such as labels or titles or brief instructions. To braille these would be inefficient, while an Optacon may well be the best device to date for reading limited inkprint materials. For general reading and study, however, the Optacon, for all of its technical novelty, basically produces a relatively unwieldy and slow reading form similar to that which was abandoned long ago by the blind in favor of braille.

Use of Computer Aids for the Blind

Up to this point, nothing has been said about computer-generated braille. A number of systems have been developed already which allow for braille books to be produced by the use of computers. These utilize, usually, some form of inkprint input into a computer which is programmed to translate the inkprint data into Grade Two braille, with an output in the form of a punched or magnetic tape which is coded to control a high-speed braille embossing machine. The embossed material is delivered in the form of standard paper-braille pages. In other words, computer-generated braille eliminates or largely reduces the need to use human braille transcribers. It does virtually nothing to change the final braille book into a product that represents an improvement for the blind reader. The size, bulk and weight of the paper braille book is not different simply because it is computer-generated. It is as unwieldy as ever, and still as difficult and as expensive to store. For the blind user, there is no difference between press braille prepared stereotype plates, computer-generated paper-braille, and human volunteer-transcribed paper braille book.

One other form of computer-generated reading for the blind has appeared recently, which must be given mention. That is the system and device which combines optical scanning of inkprint materials with a means for producing a computer-generated voice output which recites the inkprint materials in synthesized human voice. While recognizing and respecting the scientific ingenuity which made such a device and system possible, it must be pointed out that it effectively is the equivalent of being read to by a sighted human reader, at best. It combines all of the deficiencies of human readers, and many of the disadvantages of voice recorded materials. Listening is not the same as reading, as has been pointed out. Computer-generated voice output readers at this time appear to have the same restrictive limitations as the Optacon, in that they are bulky and limit the blind listener to the location of the device.

It seems obvious that the above mentioned devices, while representing creativity and often spectacular inventiveness, do not resolve the real problems which actually face blind readers of braille, and which need resolution.

Paperless Braille Devices

An effort is now ongoing to deal with the needs of blind persons who wish to read braille materials which have been transcribed from inkprint materials by methods which do not offer the disadvantages and the restrictive conditions imposed by conventional braille books which contain embossed paper pages (or thermoformed plastic copies of the paper originals). Among the disadvantages and restrictive conditions which these methods seek to avoid are: Bulk, weight, unwieldiness, inordinate expense of individual books, expense of storage facilities for conventional braille books, and difficulty and expense of making copies of braille materials for general distribution. In other words, these methods are aimed at the problems of the blind consumer rather than of the producer of braille materials. To resolve those problems and to improve the utility of braille as an information delivery system for blind readers, efforts have been made to develop various paperless-braille display units operating from electronic bulk data storage sources.

The French ELINFA and the German BRAILLEX, as well as the Argonne Braille Machine are serious efforts to resolve that need. Each has its advantages. However, if any general characterization may be about them, it is that each is mostly a device for the user to make and store his own memos and notes in braille, so that they may later be recalled at will. While it is quite valuable, it does not deal with or resolve the problem faced by the average public school student and general reader, which is to find a way to carry with him and easily refer to the braille materials which have been supplied to him by others, which he must read and review from time to time.

It may be noted that the ELINFA offers only a twelve (12) braille cell "window", while the BRAILLEX can display up to thirty-two (32) cells at one time. These may in practice be quite restrictive in the same way that an inkprint book would be for a sighted student if each inkprint page contained only a single line of no more than twelve or thirty-two letters. A multi-line presentation is essential for many applications such as a chart, mathematical problems, a table, outlines, music, etc.

Approach taken by ROSE ASSOCIATES INC. to Display Braille Material

The question arises, then; What is the minimum number of lines on a display that would be most useful and still practical? Blind readers (like sighted readers) like to skim, skip or reread. Music, mathematics, and tabular data usually require many more than one or two simultaneously presented lines. ROSE ASSOCIATES INC. already has developed a technique and has reduced it to practice whereby it is feasible to display a full page (i.e., 1000 cells) or standard, conventional braille.

Furthermore, ROSE ASSOCIATES INC. at the outset, decided that the design guide would be to make the device as much like a braille book as possible. We opted to develop a convenient dependable reading machine that is as easy to use as a book. (Other features such as note taking would be in the form of an add-on plug-in braille keyboard.)

The device is named the ROSE BRAILLE DISPLAY READER, but is usually referred to as simply the ROSE READER. The ROSE READER will make it possible for the first time in history for local libraries to supply themselves with large braille collections at a moderate cost of acquisition and maintenance, for blind persons to own and maintain private libraries, and for blind school children to achieve near parity with sighted classmates by having available useful braille textbooks that are as easily read, stored, and transported as are the inkprint text. Combining the speed and skill of computer-generated braille with the practical and useful end-product provided by the ROSE RADER will make it possible for braille materials to be available to the blind as never before.

Operating Description of The ROSE READER

Figure 1 is an artist's sketch of the ROSE READER. At the date of this writing, the first prototype is not sufficiently assembled for a photograph, but a photographic slide will be available in time for the conference presentation.

To operate the display, the user inserts a magnetic tape cassette into the guide slot at the upper left edge of the display. This cassette stores the information contained in four hundred full pages of braille. The braille embossed label on the cassette will be readable even after the cassette is inserted; so that the blind reader can always ascertain which cassette he or she has in place.
At this point, the user has two options:

a. Go to the first page in the book, or
b. Randomly turn to any page in the book.

If the reader wishes to read the first recorded braille page, then he merely pushes the page-turner lever a one-half revolution away from himself. This erases any braille that may have been on the display and causes the first line of 40 braille cells to appear silently within two seconds. The user can then begin to read. While this first line is being read, the next line of characters will appear. The lines of characters appear sequentially, two or three times as fast as most blind readers can read, so that the lines of braille cells will be generated well ahead of the reader. The whole page will appear in an interval which depends on how many dots must be written. Once a page is generated

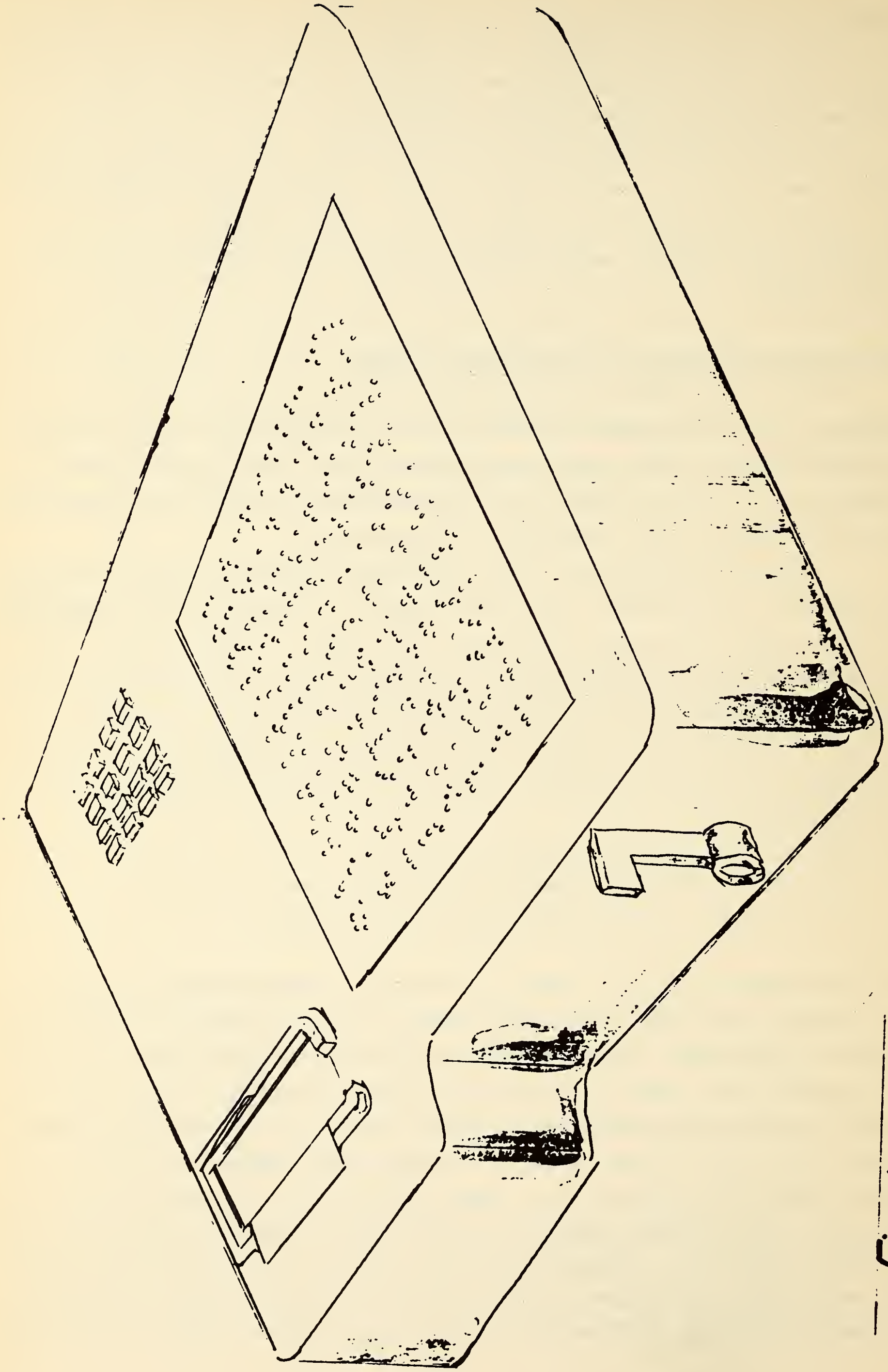


Fig 1.

it remains on the display without further use of power, until erased. The entire displayed page of braille, thus generated, consists of a maximum of 25 lines of 40 braille cells each whose dimensions and format correspond to the Perkins Braille Standard.

After the user has finished reading this page, the pageturner lever can again be operated which will instantaneously erase all of the displayed dots and automatically causes the next page in sequence to appear as above described. The identifying number of the page being displayed will appear in braille on the first line just as it would conventionally.

The reader may choose the other option to display a braille page other than the next sequential page. He or she will be able to do this by manually selecting the desired page number digit by digit from the keypad located at the upper right hand corner of the unit. After this selection, the reader again pushes the page-turner lever and, after a short searching time, the desired numbered page will be displayed. After each display is presented, the reader always has the option of automatically turning pages in sequential fashion or skipping forward or backward at random to any other chosen page. When the power is turned off, the last page displayed can be allowed to remain indefinitely, requiring no power to do so, or it can be erased and be left blank. If the book (on tape) that the user is reading has been transcribed from an inkprint edition, the user may wish to display a page identified by its inkprint page number. The keypad has provisions to allow the user to select an inkprint page number. This causes the corresponding braille page to automatically be displayed. This feature is especially useful for blind students who have braille editions of inkprint textbooks which their sighted classmates are using. In this manner, it can be seen that, for all practical purposes, the device appears to the user as if he or she were reading a braille book.

FUNCTIONAL DESCRIPTION OF THE ROSE READER

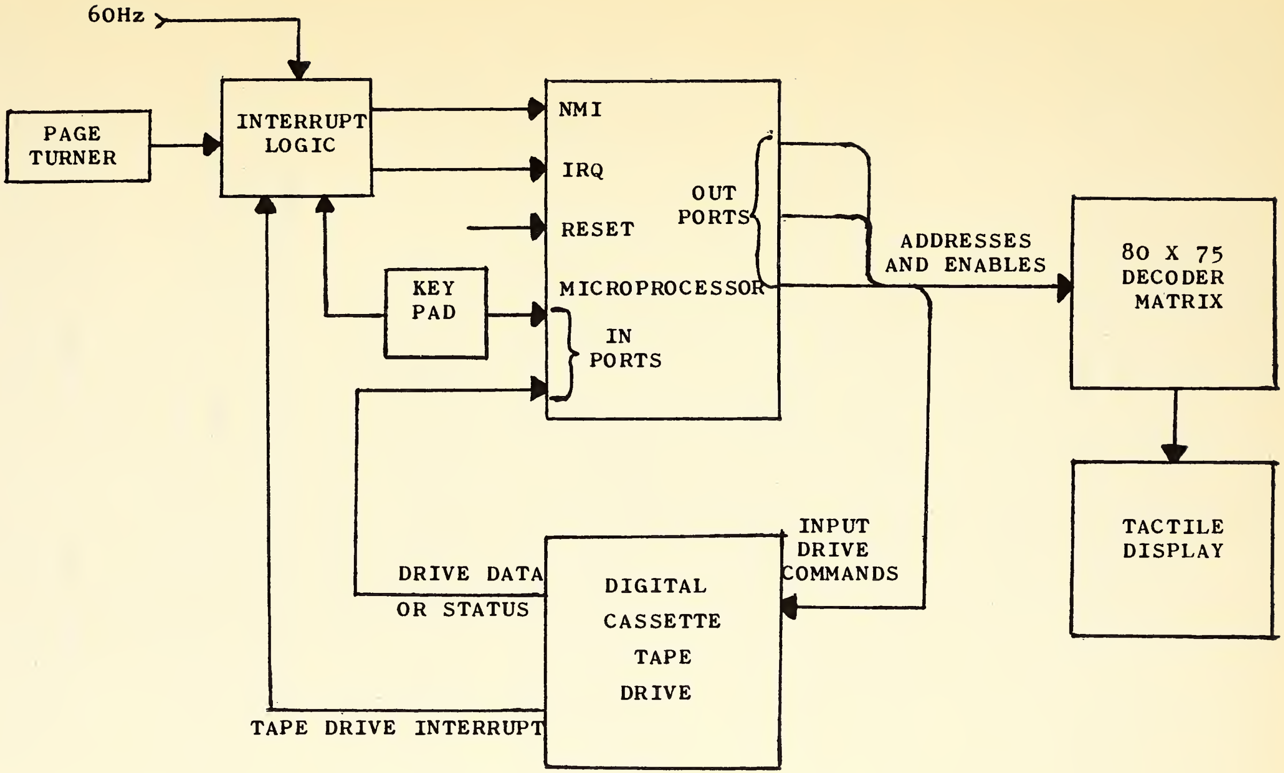
The Rose Reader has a microprocessor based logic which is used to control the tactile display and the tape cassette drive. It also processes commands from the page searching keypad and page turner-lever. The RAM in the microprocessor buffers the data from the tape and the ROM stores the control program. Figure 2 is a functional block diagram of the Rose Reader.

Tactile Display

The Tactile Display is a Rose Associates, Inc. proprietary mechanism which selectively "writes" dots on its surface based upon electrical address signals from the decoder matrix. As in a standard braille page the Tactile Display can "write" as many as 6000 dots on its surface. The design of this mechanism is such that the dot height and dot density conforms to the Perkin's Standard. Even at this dot density, the construction of the mechanism is such that it can be mass produced.

Decoder Matrix

The decoder matrix is an 80 x 75 switching matrix driven by 80 SCRs in the horizontal and 75 SCRs in the vertical. The gates of the SCRs are selectively driven by demultiplexers when appropriate addresses are generated by the microprocessor. The software which generates these addresses causes the dots to be written in a raster scan fashion. That is, the top row of required dots in the first line of braille characters appears first, then the second row of required dots in the first line of characters is "written" and finally the total first line of braille characters appears after the third raster scan. The raster scanning rate is such that it takes 4 microseconds to go from one dot position to the adjacent position. When the scan comes to the position where the dot is to be "written" it dwells in this position for 8.33 milliseconds. The microprocessor is synchronized to the 60 Hz power supply so that when the supply voltage is going positive the appropriate gates cause the SCRs in the matrix to conduct. Tis sends a pulse of energy to the display during the positive half of the cycle and a dot is "written."



ROSE READER BLOCK DIAGRAM

FIG 2.

During the negative 8.33 milliseconds half cycle no dots can be "written," but this time is used to scan to the next dot position to be "written." The process continues this way until 75 raster scans have been completed which finishes the pages. From this process, it can be seen that the time required to write a complete page depends essentially on the number of dots to be written.

Microprocessor

The microprocessor is designed around the Motorola 6800 CPU. It has 16 input data lines through two input ports. The output is through 24 latched data lines arranged in three output ports. In addition, it has one maskable interrupt, one non-maskable interrupt and provision for power-on reset.

The memory has 8192 bytes of ROM and 2084 bytes of RAM. About 800 bytes of the RAM is used to buffer the data coming from the tape of 8000 bits/second. Some of the remaining RAM is used for the page search routine. This leaves RAM for expansion when additional features will be added to the Rose Reader.

The stored program requires about 4K bytes of ROM leaving the remaining 4K bytes for planned new features of the Rose Reader.

The Digital Cassette Tape Drive

To minimize the skill and attention required for the user to operate the Rose Reader, we have selected a digital grade cassette drive rather than an audio cassette. The reliability and automatic features of the cassette drive we have chosen essentially requires no attention from the user.

The cassette drive design supports a tape data density of 800 bits/in. or 360K bytes per cassette. The uncorrected error rate is 10^{-7} which translates into no more than one dropped bit per cassette.

Without risking damage to the cassette or the tape, it can be rewound at 120 inches/sec. Page search can be conducted at 80 inches/sec. The data on the tape is read at 10 inches/sec., to allow a transfer data rate of 8000 bits/sec. into the RAM. The tape speeds are automatically selectable for rewind, search and read. The tape tension is kept constant without capstans or pinch rollers by means of two self-tracking d.c. motors in a phase-lock loop servo. The servo speed is controlled by a "clock" track prerecorded on the magnetic tape.

The cassette drive automatically signals the microprocessor when the following conditions are detected:

a. Cassette loaded
b. End of Tape, Beginning of Tape
c. No clock track (broken tape)
d. Interrecord gap (for page search).

As a result, all tape drive commands (rewind, high speed search, low speed read, etc.) can be generated through the microprocessor as a result of user operation, involving only the insertion or ejection of the cassette, operation of the page turning handle or page selection keys.

All data and command signals between the microprocessor and the cassette drive are conducted via a parallel interface into a bus and the non-maskable interrupt line.

Mechanical Design

The overall dimensions of the Rose Reader as shown in Figure 1 are 17 x 15 x 4 inches. It weighs about 15 lbs.

As a result of some preliminary user tests, we have incorporated into the mechanical design a "resilient" feel to the individual dots as the user moves his or her fingers over them. We have found that the very "stiff" dots tended to desensitize the users fingers. The rest of the reading plane is textured to provide a comfortable sensation even if the users hands are perspiring.

Guides are provided to ensure that a blind user will have no trouble inserting the tape cassette. An ejection lever is provided to easily remove the cassette.

The reading area of the tactile display is subdivided into 40 column subassemblies each occupying two vertical columns of 75 dots. These subassemblies are plug-in units and can easily be replaced if repair should be required. The mechanical boundary between these vertical dot column assemblies is undetectable to the user. For all intents and purposes, the reading surface has the tactile appearance of a page of braille.

TAPE CASSETTE RECORDING FORMAT

The tape cassette conforms to the ASNI/ECMA standards including a standard Bi-Ø-Level recording format at a density of 800 bpi. For use in the microprocessor, the data is converted to the NRZ format.

One track contains a prerecorded clock of 800 clocks per inch. This clock track and an appropriate one-shot circuit, provide accurate dual speed control and constant tape tension.

Each braille page constitutes a record. The record is started with an appropriate preamble containing a synch. byte and information concerning the number of bytes in the record (usually 750 bytes for a full page). The data is <u>not</u> recorded in mapped braille character format, but rather each 10 bytes conforms to a line of dots across the page for the raster scan. This simplifies the software and increases the amount of data that can be stored on each tape.

Each record is followed by an interrecord gap of about 0.5 inches. During the high speed page search, these interrecord gaps are detected and counted to keep track of the braille page.

DEVELOPMENT AND PRODUCTION STATUS

As of this writing, Rose Associates, Inc. has completed most of the hardware and software development and the construction of the first full prototype is nearly completed. Prototype functional testing will begin in early summer of 1979.

Simultaneously with the prototype development production, tooling has been designed and at this writing some of it is being fabricated. Five such prototypes in all will be produced by the end of the summer of 1979 for field testing and environmental testing during the fall.

The results of the testing will be factored into the final design and the production tooling will be complete by the spring of 1980 at which time Rose Readers would be manufactured for sale.

FUTURE PLANS

The design philosophy in the prototype and first production Rose Readers has emphasized high reliability, good maintainability, ease of operation and yet low cost producibility. The use of a microprocessor and a digital grade cassette drive will help to assure the reliability and ease of operation.

Designing the unit with reliable plug-in modules will help assure maintainability. At this point in the program, the design goals for reliability, maintainability and ease of operation have increased the unit cost as would be expected. As we gain experience from field and environmental testing, it is hoped that some of these measures will prove to be unnecessary and the cost can be reduced in the production models.

Of course, during the manufacturing engineering phase of the program, other cost reduction measures will be pursued. Also, as sales increase and economies of scale can be realized, Rose Associates expects to be able to lower the price of the Rose Reader.

The Rose Reader has been designed basically as a reading machine as has been stated before. As a result, some "nice to have" features have not been offered as part of its initial design. Instead, it has been our philosophy to anticipate some of these desirable features and provide for them as plug-in ancillary units. Provisions for new features were made when we selected a microprocessor whose memory space is not completely occupied.

Mechanically, the design and modular construction of the tactile display allows the Rose Reader to be expanded to a 42 cell line by plugging in additional dot column subassemblies. Also, provisions have been made for plugging in add-on units such as those described below.

NOTE TAKING DEVICE

The cassette drive now has the ability to both write as well as read cassette tapes. A unit with keys corresponding to the Perkins Brailler will be offered to be plugged in so that a user can type a page of notes or a letter on the surface and then record the page on the tape cassette. Four hundred such pages will be able to be recorded on one cassette.

MASTER/SLAVE DISPLAYS

For classroom work a Rose Reader acting as a Master will be able to drive a number of remote displays. Remote displays, not needing cassette drives or microprocessors, will be much less costly than the master. A variation of this multi-display concept will provide a two-page display to a user for instance with braille text on one display and graphics on the other.

LIBRARY REFERENCE

The Rose Reader will be able to plug into a larger bulk storage device such as a floppy disk drive providing encyclopedic information to blind researchers.

COMPUTER AND TELEPHONE TERMINALS

An interface device will be able to be provided to plug into the Rose Reader and specified computers, to make the overall system a computer terminal. By also combining it with the device mentioned in a) above, it will be able to be made into an interactive terminal. By using telephone audio couplers, this braille terminal can be located anywhere with respect to the computer allowing blind users to have a timeshare computer terminal.

BLIND-DEAF USERS

One obvious outcome of the above will be to permit deaf-blind users to communicate with each other via the telephone and Rose Readers in combination.

CONCLUSIONS

In summary, in this paper we have attempted to make the point that it is necessary for the blind to be able to gather and assimulate information tactually just as the sighted read inkprint. Although other techniques for transmitting information are useful such as synthesized speech or raised Roman letters from inkprint scanning, they do not take the place of the "written" word. Braille is useful and will always be needed. It can be made even more useful if it is put in the form most convenient to the blind user.

We at ROSE ASSOCIATES INC. feel that a major step will be taken in this direction when the data is stored on magnetic tape and displayed as a full page with the features similar to a book of braille.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

DOTTRAN-A CLASS OF FORMAL BRAILLE SYSTEMS

by
Peter Duran

ARTS Associates, Inc.
Boston, USA

DOTTRAN
A CLASS OF FORMAL BRAILLE SYSTEMS

Peter Duran, President

ARTS Associates, Inc.

INTRODUCTION

Braille is a system of sixty-four tactile characters. The Braille characters represent a few inkprint characters such as upper- and lower-case letters, digits, punctuation marks, and so forth.

The basic unit of the Braille system is the "cell". The cell is a rectangular array of six dots, arranged as three rows and two columns. By omitting one or more dots from the cell, a Braille character is formed. There are, therefore, sixty-four Braille characters. (The space is counted as a character.)

This paper only discusses Grade I and Grade II Braille, American style. The principal reference is English Braille, American Edition, 1959-1972. The "Typical and Problem" Word List refers to the list of transcribed words in this reference.

The main thesis of this paper is that the Braille system, as presently constituted, is a confused collection of dogma! The Braille system is too hard to learn, read, and program. These contentions are amply supported by the fact that the number of Braille readers, in the United States, is rapidly decreasing. This is especially true of readers who require the transcription of technical materials.

Brailled materials are vitally needed. Using present production methods, they are not readily available to Braille consumers.

Analysis reveals that the Braille system is irreparable by minor alterations here and there. Review of the past several revisions of the Braille system indicates that the committees involved have tinkered with the Braille system but have not significantly improved it or increased its utility.

The Braille system is analogous to a very old house with a fresh coat of paint; it is attractive from the outside and a shambles on the inside. Only periodic patching keeps it from falling down. The Braille system must be restored! It has to be rebuilt from its foundation up! The end result must be a sturdy structure having the same outward appearance as the old edifice, but without the blemishes and decay of the decades.

The Braille system can be discussed either as an "informal" system or as a "formal" system. As an informal system, the Braille system has certain purposes which give it utility. Typically, the methods for achieving those purposes are described in general terms, with many details left implicit or even vague. An informal presentation of the Braille system is embodied in the above mentioned reference. Like most informal presentations, this one suffers from vagueness, incompleteness, ambiguity, and inconsistency.

As a formal system, the Braille system has certain structural properties enabling it to represent inkprint materials. Like other formal systems, a formalization of the Braille system would possess abstractness (permitting it to be applicable to many natural languages without modification), preciseness (reducing the number of rules to a minimum), consistency (removing contradictions, intended or accidental), and completeness (leaving no doubt about any particular instance of a rule).

The merits of a formal approach to the Braille system are gauged by its "unexpected" triumphs. In what follows, a formal approach is presented based upon a very few general principles. During the presentation, emerging achievements are pointed out as justification for the approach.

DOTTRAN is a modest attempt to formalize the Braille system. There are many possible formalizations. The one selected has several principal goals. DOTTRAN increases the kinds of inkprint materials that are transcribable. DOTTRAN dispenses with a math code in the elementary cases. DOTTRAN entirely dispenses with a computer code. The rules of DOTTRAN are completely specified; thus dispensing with ambiguous cases. The rules of DOTTRAN are fewer in number, simpler, and more general than those of the Braille system.

INKPRINT CHARACTERS

One of the most frequent complaints about the Braille system is that it is very limited in the kinds of printed materials that it can adequately represent. This complaint is well-founded, as blind persons having to learn more than one Braille system will attest. This limitation has two general sources. Not enough inkprint characters are represented, and not enough Braille characters exist (there are only sixty-four).

The first source of the limitation is readily overcome. A larger set of inkprint characters is merely selected for representation. The selected set of inkprint characters can be arbitrarily large. For evolving a formalization of English Braille, it is selected to be the characters encompassed within the ASCII code--the American Standard Code for Information Interchange. There are one hundred and twenty-eight characters within ASCII; ninety-four are inkprint characters and the others have special computer-related uses.

ASCII is a natural choice for three principal reasons. First, huge amounts of material are being produced using this character set. Second, this character set contains all the characters typically represented in the Braille system. Third, this character set lends itself to computer Braille implementation.

THE CHARACTER CORRESPONDENCE

The task at hand is putting inkprint characters into correspondence with Braille characters. Since there are more inkprint characters than there are Braille characters, only some of the inkprint characters can be placed into correspondence with the available Braille characters. Which inkprint characters are put into correspondence with which Braille characters is a matter of personal preference. For example, the English Braille system and the French Braille system put the lower-case alphabet into different correspondences with Braille characters. The question, which correspondence is better, is a matter for committees to haggle over. With respect to a formal system of Braille, it does not matter which of the possible correspondences is actually employed. It is only required that the correspondence be well-defined as next described.

An inkprint character that is selected to correspond to one of the sixty-four Braille characters should correspond to just one Braille character--to at least one and to at most one. This one-to-oneness of the correspondence is one of the fundamental requirements of the formal Braille system.

The Braille system fails to guarantee uniqueness of this correspondence. For example, there are two Braille representations for the "double quote"; there are three Braille representations for the "apostrophe". It is never precisely specified under which circumstances each ought to be employed.

EXTENDING THE CHARACTER CORRESPONDENCE

In the Braille system, the twenty-six lower-case letters are put into correspondence with twenty-six distinct Braille characters. If this correspondence is extended in the obvious manner, only sixty-four ASCII characters are representable. This dilemma is resolved by putting inkprint characters into correspondence with well-defined sequences of Braille characters.

A sequence of one or more characters is called a "string". The best example of this practice is illustrated by the formation of capital letters. Each lower-case Braille letter is prefixed by a capital sign to form the corresponding upper-case Braille letter. Thus, by using one Braille character, twenty-six additional inkprint characters are represented.

Similarly, the letter sign is generalized. It prefixes each lower-case Braille letter, unless defaulted; that is, omitted in desirable circumstances.

In addition to the capital sign and the letter sign, a "control" sign is included. It prefixes the lower-case Braille letters to form the corresponding ASCII control characters. Similarly, the number sign prefixes the first ten lower-case Braille letters. to form the corresponding digits 1-9 and 0.

Characters such as the capital sign, letter sign, control sign, and number sign are called "descriptors". They prefix single Braille characters. The resulting Braille character strings can be placed into one-to-one correspondence with inkprint characters.

The notion of a descriptor receives its significance from its structural properties. These properties are briefly discussed next.

Descriptor Properties

From what has been said so far, it can be inferred that two Braille characters are required to represent most inkprint characters. Thus, a Braille character string is typically twice as long as the corresponding inkprint string. To reduce the number of Braille characters required to represent an inkprint string, the notion of "exportation" is introduced; that is, the moving of a prefixing character from in front of every character to merely in front of the entire string. This notion is a generalization of the use of the double capital sign. A descriptor is assigned a value of single or double. If neither single or double, it precedes each character of the Braille string. The alphabetic indicators of the Nemeth code are examples of non-exported descriptors. If single, it once precedes the entire Braille string. The number sign is an example of a single descriptor. If double, it twice precedes the entire Braille string. The capital sign is an example of a double descriptor.

DOTTRAN treats descriptors systematically. Each is assigned a particular class of inkprint characters to represent, and each is assigned a particular exportation value. This approach simplifies the rule of the Braille system for the sake of the reader and the programmer.

To permit even further abbreviation, each descriptor is assigned a class of "transitional" characters; that is, characters over which it is allowed to cross. For example, the double capital sign is permitted to cross over the hyphen. Thus, a compound word all in capitals has only the first word, rather than both words, preceded by a double descriptor.

Summary

The above brief outline of Grade I DOTTRAN indicates several things. A large set of inkprint characters are representable. A string of Braille characters is systematically constructible. The Braille character construction process is precise but not especially complicated.

By listing all the descriptors with their exportation values and their transitional classes, a general Grade I Braille system is specified. Exportation values and transitional classes can be chosen to meet special requirements. Thus, Grade I DOTTRAN is readily able to conform to other existing Braille systems for other natural languages.

CONTRACTIONS

Definition. A contraction is a string of characters substitutable for another string of characters. A substitution is typically only permitted under specific circumstances. These circumstances are discussed in detail below.

The use of contractions has numerous justifications. A few are mentioned next.

They reduce the amount of physical space and materials required for a "hard" copy of a transcription. A "hard" copy is one produced on some physical medium such as paper, plastic, and so forth.

They reduce the amount of perceptual data presented to the reader. For example, the word "character" is reduced from nine Braille characters to two Braille characters. The contraction is perceived as the word, after sufficient experience has accrued.

They enhance readability. Many contractions represent common letter groupings such as "ness". Thus, dividing a word is simplified for a reader.

ENGLISH BRAILLE

SUBSTITUTION CRITERIA

There are numerous criteria for deciding whether a particular contraction should, or should not, be substituted for a word or a word fragment. Unfortunately, these criteria are vague and rather arbitrary. A few examples shall suffice.

Criterion 1. A contraction is not substituted if it crosses a primary syllabic break within a word. The notion of "primary" syllabic break is never made explicit. Typically, appeal to a particular dictionary of a given edition is made. The Typical and Problem Word List violates this criterion in many instances. For example, the contraction for the two letters "ar" is always used, whether a primary syllabic break is crossed or not.

Criterion 2. A contraction may be used if it crosses a minor syllabic break within a word. The notion of "minor" syllabic break is never made clear. In the Typical and Problem Word List, the syllabic break in "erase" is major, and the syllabic break in "around" is minor. Again, "part" in the word "partial" is contracted but not in the word "partake". This type of arbitrary nonsense makes reading Braille and programming Braille tiresome!

Criterion 3. A contraction is used if it and the string it is being substituted for have the same sound or pronunciation. This criterion is a function of historical time, culture, dialect, and so forth. This criterion is too arbitrary to be reasonable!

Criterion 4. A contraction is used only in "familiar" words. What is "familiar" is merely a function of prior experience. It is quite unlikely that different Braille users would ever have the same experience with words. In fact, dozens of words in the Typical and Problem Word List occur virtually at the end of every reliable frequency count of American English.

DOTTRAN

SUBSTITUTION CRITERIA

The above criteria don't make much sense! They all suffer from vagueness, nonapplicability, arbitrariness, and appeal to authority (which changes from time to time).

The only thing that is constant is the words themselves; that is, the strings of inkprint characters to be represented. Taking this obvious fact as a principle, DOTTRAN uses only syntactic and positional notions for determining occurrences of contractions. That is, notions depending on the string of characters to be represented, and depending on the position of the contraction within the string or on a line. Some of these notions are briefly mentioned next.

String position: A contraction can occur at the beginning, in the middle, or at the end of a word. It can also represent the entire word. For example, the contraction for "con" is only permitted at the beginning; the contraction for "bb" is only permitted in the middle; the contraction for "ar" is permitted under all circumstances; and so forth.

Line position: A contraction is permitted at the beginning of a line or at the end of a line. This criterion governs the use of contractions when hyphenation occurs. For example, the contraction for "con" is permitted at the beginning of a line but not at the end of it.

Syllable position: A contraction can be a whole syllable, cross a syllabic break, or be bounded by syllabic breaks. For example, the contraction for "con" is always a whole syllable; the contraction for "ar" can always cross a syllabic break; and the contraction for "upon" must always be bounded by syllabic breaks.

Summary

All contraction criteria are syntactic and positional. They are specified in terms of primitive notions such as position within a string, position on a line, position relative to syllabic breaks, and so forth. Based only on these criteria, DOTTRAN is able to produce "perfect" Braille. That is, DOTTRAN produces transcriptions of quality on a par with certified transcribers!

Does DOTTRAN do perfect English Grade II Braille? Since the current Braille system is so confused, this is difficult to determine. It is well-known that if the same manuscript is submitted to two certified transcribers, their transcriptions will, almost always, differ. The only thing that can be asserted is that ignoring the obscurities, inconsistencies, and ambiguities of English Braille, its rules and those of DOTTRAN produce the same results.

HYPHENATION

Contractions only reduce the total amount of "hard" copy by less than twenty percent. Thus, even a little space on a line or a Braille display is significant.

The primitive notions of DOTTRAN permit it to perform automatic hyphenation at the end of lines. This feature brings DOTTRAN closer to human transcribers than all other extant Braille translators. This feature secures a saving in space and promotes uniform transcription.

Perusal of a Brailled copy of this paper indicates that DOTTRAN does hyphenation as well as certified transcribers. Absolutely no interaction is required of the author to ensure accurate hyphenation!

TEXT PROCESSING

Up to this point, only string translation has been discussed. Text formatting, the arranging of the translated strings on a on a one-dimensional line, on a two-dimensional page, and in temporal or continuous sequence for refreshable Braille displays, is another major aspect of the Braille system.

The Braille format rules are even more numerous and confused than the rules governing string representation and contraction. No attempt has been made to model these rules as presented! There are many justifications for this decision. A few are discussed next.

Text Integrity

Many Braille format rules unnecessarily rearrange the inkprint text for Brailling. Thus, the blind person reading the transcription has no way to infer the inkprint layout. For example, special signs such as the percent sign are put before the number rather than after it, violating inkprint custom. Informal surveys of Braille readers overwhelmingly indicate that they resent this alteration of text.

To be sure, some format alteration is required. The Braille line is two thirds as long as the typical inkprint line. This necessitates some conversion. Similarly, the rearrangement of a table or chart may be required. Except for required format alterations, no vandalism should be committed on the inkprint text.

There are a multitude of possible formats. One kind is employed when transcribing literary texts such as a novel, and others are used for textbooks.

Typically, a translator is accompanied by a text formatting program. A particular format feature is selected by issuing a command indicating the action to be performed. For example, a "Centering" command precedes a line of text to be centered. Most formatters have a specific list of formatting features that can be used. The transcriber is typically unable to design a new format feature when needed.

The inability to design one's own formatting commands is a severe limitation. The transcriber has to compromise and settle for what is available.

DOTTRAN is initially written to run on the UNIX Time-Sharing system, developed and distributed by Western Electric Corporation, Inc., USA. This system has a "new" formatter with extraordinary capabilities. A transcriber is able to design virtually any style format for novel situations. Unlike other formatters, UNIX has a "language" in which a user can specify new format features. Once specified, a new format feature is readily useable on other occasions. The same formatter is used for printing text and for Brailling text. The formatter can create tables of contents and the like without any transcriber intervention. There is a whole host of pre- and postprocessors available; that is, extra programs for performing special tasks such as laying out tables automatically, creating cross references, and so forth.

DOTTRAN IMPLEMENTATIONS

At present, ARTS Associates, Inc., is distributing a DOTTRAN running under the UNIX Time-Sharing system. It is soon to be released running under other systems and as a "stand-alone" system.

In July 1979, ARTS Associates, Inc., will offer a transcription service using the UNIX version of DOTTRAN. This paper is available in Braille produced under this system.

DOTTRAN TRIUMPHS

The major differences between DOTTRAN and English Braille are summarized next. DOTTRAN is able to transcribe all possible ASCII strings; English Braille is unable to do so. DOTTRAN is able to handle elementary mathematics and computer generated text; English Braille is unable to cope with either. In fact, the plus and minus signs are excluded from English Braille.

The rules of DOTTRAN are few in number, precise, and easily stated; the rules of English Braille are many, vague, and verbose. The rules of DOTTRAN are based on factors inherent in the text to be transcribed; the rules of English Braille are based on historical accident, appeal to opinion and authority, and quasi-semantical notions.

DOTTRAN hyphenates uniformly and with consistency. Transcribers tend to vary in their hyphenation performance.

DOTTRAN preserves the inkprint format wherever possible. English Braille, with confusion, alters the inkprint format unnecessarily. (Using DOTTRAN, it would be possible to corrupt the inkprint format as demanded by current Braille practices by specifying the required format features. This practice, for the sake of the blind reader, is strongly discouraged!)

DOTTRAN surpasses all existing Braille translators with respect to design. DOTTRAN only depends on aspects of a string for purposes of translation. All other translators depend on "large" tables to govern translation. DOTTRAN's logical rules apply in all possible cases. Other translators fail in all instances not covered by their tables. Since there are always instances not included in a finite table, these translators are subject to errors. Due to DOTTRAN's elegant design, the contraction table presently contains less than two hundred and fifty entries, whereas other translators have close to one thousand. The translator itself requires five thousand bytes of memory (far less than all other high-quality translators).

We hope that this elegant formalization of the Braille code and DOTTRAN will encourage a world-wide rejuvenation of Braille.

World Council for the Welfare of the Blind
Committee on Cultural Affairs

International Conference on
"Computerised Braille Production - Today and Tomorrow"
London, 30 May - 1 June 1979

A BRIEF HISTORY OF BRAILLE IN JAPAN AND AN ACTUAL EXAMPLE OF ITS COMPUTERIZATION

by

Sadao Hasegawa

School for the Blind
Tsukuba University
Japan

A Brief History of Braille in Japan and An Actual Example of Its Computerization

by Sadao Hasegawa

School for the Blind Attached to
Tsukuba University

The official braille in Japan is Louis Braille's six-dot code. Braille was first actually read and written in Japan in 1887, at the public Blind and Deaf School (now the School for the Blind attached to Tsukuba Univ.) in Tokyo, utilizing the English alphabet and a braille board imported from England.

Three years later, in 1890, a braille code using Japanse phonetic symbols - cursive or angular scripts known as Kana - was invented which placed 48 basic Kana on a single braille code board. Kuraji Ishikawa, the code's inventor, used the braille figures as is. This code is presently used in Japan and, utilizing the English alphabet is able to express mathematics, physics and chemistry formulas in the same manner as European and American braille systems.

Kanji (Chinese-based characters or ideographs) make up about 30% of Japanese literature as commonly written. These characters (originally based on pictures) express word meanings in a visual manner which makes a composition more easily understood. Kanji do not, however, exist in official Japanese braille, which utilizes only the phonetic Kana. Because it is tactively difficult to understand Kanji through braille, an official braille for Kanji has not been invented. Today, when we want to mutually transcribe regular writing and braille by computer, a major obstacle lies in the fact that braille Kanji do not exist, making it essential that an easily read braille Kanji be invented.

Braille boards seem to have been manufactured in Japan around 1892-3, and the first Japanese zinc plate machine was manufactured in 1897, having been modeled after an American machine.

Presently, zinc plate machines are available in each school and institution for the blind, and braille typewriters are now widely used by braille users.

A study of joining the computer and braille was already begun in 1959 as the result of joint research by Tokyo Education University and Japan Telegram and Telephone Corporation. At that time, however, computers had not been fully developed and costing was prohibitively expensive, so this project did not continue for long, although it was later to have great influence.

Braille machine from this early stage are still kept at our school. The auto-zinc plate machine, manufactured at Kyoto Light House in 1970 was operated by computer paper tapes. This was utilized only experimentally, however, and was not actually used in book printing.

Although I have here reviewed the history of braille up to 1970, much research has continued since then.

From here on, I would like principally to discuss my research in terms of my thesis, which I presented to a Japanese academic institute. What is unique about it is that the thesis was written in braille and then converted to ordinary written Japanese characters by means of an automatic braille reading machine. The name of the academic society before which this thesis was read is the Society of Instrument and Control Engineers (SICE) and the date of presentation was Aug. 23, 1978.

Braille Data Processing Essential to the Study and Work of the Visually Handicapped

(by Sadao Hasegawa)

1. Preface

The greatest problems for the visually handicapped and blind are walking and reading and writing. If these could be resolved, most of the difficulties experienced by the visully handicapped would be removed. I am a visually handicapped person who uses braille extensively, and I would like to briefly report on the results my research partners and I obtained from braille data processing.

Because I am not a mechanical engineer, some of my explanations may be insufficient, for which I ask your patience. I would like to point out that this thesis was rewritten from braille by an automatic braille writing technique, with the braille manuscript read by an automatic reader and printed by a Kanji printer.[1]

2. Kana Braille System

The universally recognized 6 - dot braille system currently in use was invented by a visually handicapped Frenchman, Louis Braille, in 1825. This system employs convex symbol unit containing a maximum of 6 dots which are read by touch.

Japanese braille applies the 64 standard braille symbols to Kana (Japanese phonetic script) and writes by mixing these symbols with Arabic numerals and the English alphabet. Kana braille is very convenient for extensive use by the visually handicapped because, by following the dots with their fingertips, they can read almost all the letters by phonetics.

Ordinary Japanese writing, however, includes Kanji (Chinese-based characters or ideographs), which are very difficult to link with braille and ordinary symbols in a machine system; thus, the differences in the two writing systems become a great problem.

The best way to solve this would be to convert Kanji and Kana by means of software research[2] which is presently under way.

Because of the extreme difficulty of the foregoing, however, another method to solve this problem - the invention of new braille Kanji - is also being explored.

3. The Integrated Braille System (IBS)

In the Integrated Braille System (hereafter IBS) which includes Kanji, one ordinary character corresponds to one braille character, based on the "Japan Braille Inscription Method" which consists to a great extent of Kana, or phonetic symbol, braille. Under this system, Kana and numerals written in Kanji[3] are transcribed precisely into hiragana symbols and Arabic numerals; with additional codes placed before each, Kana and numerals written in Chinese characters are produced. Codes for the period, comma, question mark and exclamation mark remain unchanged.

The system is characterized by the fact that Kanji are interpreted by both on and kun readings.[4] This makes it much easier to read and write Kanji since their functional factors often accord with their phonetic pronunciation. People who have lost their sight early in life, are not familiar with the shape of ordinary characters, but they usually learn to read Kanji in two ways; the word "sea", for example, is read kai and umi. Therefore, with this system, it is possible for such people to form Kanji with their existing knowledge and easier for them to understand. When a Kanji cannot be structured with on and kun readings, a hieroglyphic, characteristic of the Kanji, is symbolized and used supplementally.

4. Braille Input Equipment

4.1. Braille Data Typewriter: This typewriter consists of six-dot braille keys and of function keys for the space and the line-shift. This equipment can record signals on either paper tape or magnetic cassette tape by monitoring the braille; some also possess a microcomputer and can modify signals.

4.2. Automatic Braille Reading Equipment: Originally developed by the Braille Cassette System Development Study Group, this equipment makes it possible to read braille symbols optically and to adapt its system to reproducing braille symbols and automatic writing. It reads one braille page in ten seconds.

5. Braille Output Equipment

The following devices have been developed and some are already in use: a braille printer which prints braille symbols in one-dot or one-cell units; a braille line printer which prints one braille line by dividing it into three (upper, middle, lower) levels; an automatic zinc plate type making machine for braille printing; braille ink printing type making machine; and braille display equipment which indicates braille symbols by protruding pins.

6. Automatic Writing Equipment

This system prints ordinary letters by using as input power braille signals which come from either a braille data typewriter or automatic braille reading equipment. In this case, by using the IBS, it is possible to distinguish three different categories of Japanese writing - Kanji, hiragana, and katakana. Input from automatic braille reading equipment makes it possible to produce printed symbols out of handwriting symbols. Since this system can specify the size of printed letters according to the type of equipment, it also can print enlarged letters for the weak-sighted.

7. Automatic Braille Translation

7.1. Fully automatic braille translation is a method which uses the signals from an optical character reader or kanji printer and computer phototypesetter directly for braille translation. This method is characterized by the fact that there is no need of input by manpower. With use of the IBS, complete braille translation is possible. It is important to note that Kanji can be converted into Kana symbols through this method.

7.2. Semi- Automatic Braille Translation: Assuming that braille is to be translated, this method permits input with Kana or Roman symbols in accord with the rules of braille Kana writing and spacing. Although use of this method necessitates that the inputter study braille rules and have some assistance in inputting, it is quite possible to make a complete braille translation of the Kana symbol braille system.

8. Braille Reproduction

This method enables reproduction of braille materials already signalized in accord with requested numbers. Methods of making braille signals include reading braille materials made by the visually handicapped themselves or volunteers through automatic braille reading equipment, and making braille signals over a braille data typewriter or through automatic braille translation. The printing method to be used is selected from braille output equipment according to the particular purpose involved.

9. Artificial Sound Application

Artificial sound can be of great assistance to the visually-handicapped in picking up information as well as braille letters. It is very convenient when one operates braille keys or ordinal English or kana symbol keys to be able to confirm them by sound. It is also possible to use a calculator if signals are indicated by sound alone or together with the braille symbols. We are now in the process of working with automatic reading by artificial sound.

10. Afterword

It is true that scientific technology will improve the capacity of the handicapped. I, myself, as a visually-handicapped person, have hopes and proposals regarding the development of technology and I would like to request the assistance and cooperation of specialists from different fields.

I would also like to take this opportunity to express my sincere gratitude to Prof. Kensuke Hasegawa of the Tokyo Institute of

Technology and Messrs. Takeshi Kambayashi, Yoshiaki Ishii and Kazuyoshi Adachi of the Japan Computer Center Co., Ltd., for their invaluable assistance in my writing of this thesis through automatic writing.

Foot notes:

1. Kanji Printer is able to print all kinds of letters such as an alphabet, numerals, Kana, Kanji and special symbols.

2. This software can convert ideographic symbols to phonetic symbols and vice versa.

3. Japanese phonetic writing is done in hiragana (a cursive script used to transcribe readings of native Japanese words also written in Chinese-based characters or ideographs called Kanji) and Katakana (an angular script used primarily to transcribe foreign words). Numerals are expressed either in Arabic numerals or Kanji. No such differences exist in Japanese braille symbols.

4. Chinese character (Kanji) may be given either of two readings, on or kun. On derives its original pronunciation from ancient Chinese. Kun is the native Japanese pronunciation given to a character. Since the reading of a Kanji depends on context, it is extremely difficult to read these Chinese-based characters by touch.

Reference: "Kanji Input and Braille Information Processing and Management of Information and Documentation by Braille", Sadao Hasegawa, Vol.21. No.1: 33/41 (1978)

WCWB COMMITTEE ON CULTURAL AFFAIRS

Sub-Committee on Computerised Braille Production and Other Media

INTERNATIONAL CONFERENCE ON COMPUTERISED BRAILLE PRODUCTION

30 May - 1 June 1979

PARTICIPANTS OBSERVERTS

*blind person ÷speaker #exhibitor

AUSTRALIA

÷ Mr. John W. Berryman	Manager, Braille Library & Computer Services, Royal NSW Institute for Deaf Blind Children, Box 4120 GPO, Sydney, NSW 2001.
Mrs. M. Fialides (Observer)	Manager, Tertiary Resource & Library Services, Royal Victorian Institute for the Blind, 557 St. Kilda Road, Melbourne 3004, Victoria.

BELGIUM

* Mrs. Irma Beugnies	Head of Transcription Centre, Les Amis des Aveugles, 37 rue de le Barriere, 7410 Ohlin en Mons.
* Mr. Narcisse Bouillon	Copyist, Les Amis des Aveugles.
Mrs. Micheline Haurez-Godart.	Directrice, Les Amis des Aveugles.

BRAZIL

* Mrs. Dorina de Gouvea Nowill	Chairman, WCWB Committee on Cultural Affairs. President, Fundacao para o Livro do Cego no Brasil, Rua Dr. Diogo de Faria, 558, 04037 Sao Paulo.

*blind person †speaker #exhibitor

Miss D. Nowill (guide)	
† Mr. Ricardo J. Machado	Marketing Analyst, IBM of Brazil, Rua do Ouvidor, 107, Caixa Postal 2106 - ZC.00, Rio de Janeiro-RJ-20.000.
Mr. P. Nigro	IBM of Brazil
CANADA	
Mr. Alan J. Davidson	Manager of Computer Services, Canadian National Institute for the Blind, 1929 Bayview Avenue, Toronto, Ontario M4G 3E8.
† Professor M. S. Doyle	Computer Science Department, University of Manitoba, Winnipeg, Manitoba R3T 2N2.
Mr. A. F. Filiatreault	Vice-President, Services Converto-Braille.
† Mr. P. A. Fortier	Department of French, University of Manitoba.
* Mr. R. G. Galarneau	President, Services Converto-Braille-Cypihot-Galarneau.
Mrs. Galarneau (guide)	
†*Mr. D. Keeping	Supervisor, Computer Blind Project, University of Manitoba.
Mr. A. Cheung (guide)	

*blind person †speaker #exhibitor

DENMARK

Mr. Pall Narvi	Statens Bibliotek of Trykkeri for Blinde, Ronnegade 1, DK-2100 Copenhagen.
Mr. Gunnar Reisler	Managing Director, Statens Bibliotek of Trykkeri for Blinde.

FRANCE

Bougé, Mme	RITM Braille, 23 rue du 8 Mai 1945, Areueil 94110.
Mr. Bernhard Causse	CERFIA (Cybernetique des Entreprises Reconnaissance des Formes Intelligence Artificielle), Université P. Sabatier, Toulouse.
Mrs. M. D. Desmerger	CNAM (Conservatoire National des Arts et Métiers), Laboratoire de Recherche pour la Reinsertion Professionnelle des Handicapés, 292 rue Saint Martin, 75141 Paris, Cedex 03.
* Mr. Jean Frontin	TOBIA, (Transcription par Ordinateur en Braille Integral ou Abrege), Université Paul Sabatier, Toulouse, 31077 Toulouse, Cedex.
Mr. G. Gouardères	CERFIA
† Mr. E. Grandjean	Université Scientifique et Medicale de Grenoble, BP 35X, 38041 Grenoble.
* Mr. Jacquin	Assistant Secretary General, Association Valentin Hauy, 5 rue Duroc, 75007 Paris.
Mrs. Jacquin (guide)	

*blind person ∔speaker #exhibitor

* Mr. A. L. Lequeux	CNAM, Paris.
∔ Mr. B. Mathieu	Université Scientifique et Medicale de Grenoble.
# Miss B. Rottier	SAGEM, Teletransmissions Department, 2 rue Ancelle, 92521 Neuilly-sur. Seine.
* Mr. P. Schneider-Maunoury	Secretary General, Association Valentin Hauy, Paris.
Mrs. Schneider-Maunoury (guide)	
# Dr. A. Tretiakoff	President, Société ELFINA, 3 bis rue le Corbusier, Silic 231, 94528 Rungis.
Mrs. O. Tretiakoff	Director, Société ELFINA.
∔ Miss M. Truquet	Director, TOBIA.

GERMAN DEMOCRATIC REPUBLIC

* Dr. S. Tschirner	Director, Deutsche Zentralbücherei für Blinde.
Mr. Horst Wandelt	International Relations Officer, Blinden-und-Sehschwachen-Verband der DDR, Poststr. 4/5, DDR-102 Berlin.

GERMAN FEDERAL REPUBLIC

* Dr. Karl Britz	Deutsche Blindenstudienanstalt, Am Schlag 8, D-355 Marburg.
Mrs. Britz (guide)	
Mr. Albrecht Drees	(BRAILLEX), Papenmeier Elektrotechnik, Postfach 1620, D-584 Schwerte 1.

*blind person †speaker #exhibitor

† Mr. B. Eickenscheidt	Münster University, Computing Centre, Einsteinstr. 60, D-4400 Münster.
# Mr. G. Einolf	Deutsche Blindenstudienanstalt.
† Mr. M. Harres	Forschungszentrum für Rehabilitation und Prävention, Bonhoefferstr. 1, 6900 Heidelberg.
Mr. H. W. Herzog	Infratest Gesundheitsforschung, Landsberger Str. 338, 8000 München 21.
Mr. J. Hertlein	Director, Deutsche Blindenstudienanstalt.
Dr. P. Janssen	Münster University, Computing Centre.
Mr. H. Kahler	Deutscher Blindenverband e.V., Bismarckallee 30, 5300 Bonn 2.
Dr. H. Kamp	Münster University, Computing Centre.
Mr. Franz Kutschera (Observer)	F. Kutschera & Co., Königswörther Str. 7, 3000 Hannover.
# Mr. Manfred Rommel	F. H. Papenmeier Elektro-Technik.
Mr. K. P. Schönherr	Ingenieurbüro für med. Technik, Schloss Solitude 3, 7000 Stuttgart.
† Dr. W. A. Slaby	Münster University, Computing Centre.
† Prof. Dr. J. Splett	Germanistisches Institut der Westfälischen Wilhelms-Universität, Johannisstr. 1-4, 4400 Münster.
# Mr. H. Trus	Deutsche Blindenstudienanstalt.

*blind person †speaker ‡exhibitor

Prof. Dr. H. Werner	Head, Computing Centre, University of Münster.
Mr. R. F. Witte	Deutsche Blindenstudienanstalt.
* Mr. Ferdinand Zekel	Manager, Deutscher Blindenverband E.V. Chairman, Technical Commission of WCWB European Regional Committee.
Mrs. Maria Zekel (guide)	

ISRAEL

Mr. E. Shapir	Ministry of Labour, Department for the Blind, Hacharouzim 10, Jerusalem.

JAPAN

† Mr. Osamu Sueda	Faculty of Engineering Science, Osaka University Toyonaka, Osaka 560.

THE NEDERLANDS

* Mr. H. Louis Jonkers (Observer - Friday only)	Dept. of Aerospace Engineering, Delft Univ., PO Box 5031, 2600 Delft.
Mr. Jan Ming	Vereniging het Nederlandse Blindenwezen, Kipstraat 54, 3011 Rotterdam.
Mr. L. van Noorden	Vereniging het Nederlandse Blindenwezen.
† Mr. J. A. Snelders	Delft University, Dept. of Electrical Engineering, PO Box 5031, 2600 Delft.
† Mr. H. A. Spanjersberg	Delft University of Technology.
†*Dr. M. Vliegenthart	Philips Research Laboratories, Dravikstraat 25, Geldorp.

*blind person †speaker ‡exhibitor

†*Dr. R. van Vliet	Nederlandse Blindenbibliotheek, Zichtenburglaan 260, 2544 La Hage.

NORWAY

† Mr. Knut Grimnes	Division Manager, Automatic Control Division, SINTEF (Foundation of Scientific & Industrial Research, University of Trondheim), Aud 48, N-7034 Trondheim.
Mr. Svein Saebøe	Norsk Data A.S., PO Box 4, Lindeberg Gaard, Oslo 10.

REPUBLIC OF SOUTH AFRICA

* Mr. E. J. Kruger	Director, S.A. Blind Workers Organisation, Braille Services, 102-8th Avenue, Mayfair, Johannesburg.
* Mrs. M. Kruger	Proofreader.
Mr. A. C. Zeelie Mrs. J. Zeelie	Senior braillist & Production Supervisor, Braille Services.

SAUDI ARABIA

* Sheikh Abdullah M. Al-Ghanim	President, Regional Bureau of The Middle East Committee for the Welfare of the Blind, Imam Abdul-Aziz bin Moh'd Street, PO Box 3465, Riyadh.
Mr. S. A. Al-Majid (guide)	

SPAIN

Mr. Luis Garcia Alos	Technical Adviser, Organizacion National de Ciegos de Espana (ONCE), Jose Ortega y Gasset 18, Madrid 6.

*blind person †speaker #exhibitor

* Mr. Jose Lorente	Head, Employment Department, ONCE.
Mr. Justo Martinez	Director, Printing House, Barcelona, c/o ONCE.
* Mr. Antonio Maso	Director, Printing House, Madrid, c/o ONCE.
Mr. Alfonso Recuero	Instituto Eduardo Tomoja de la Construccion y del Cemento, Costillares Chamartin, Apartado 19002, Madrid 33.
* Mr. Pedro Zurita	Head, International Relations Department, ONCE.

SWEDEN

* Mr. S. Becker Mrs. Becker (guide)	Synskadades Riksforbund (SRF) S-122 88 Enskede.
* Mr. Bengt Eriksson	System Analyst, c/o SRF.
† Mr. Barry Hampshire	Synskadades Riksforbund.
Mr. G. Hellström	Trask Datasystem Ab., Stockholmsvagen 34, 18274 Stocksund.
# Mr. B. E. Jacobsson	Trask Datasystem Ab.
Mrs. Krystyna Matisiak	National Swedish Board for Technical Development, Fack, 100 72 Stockholm.

UNITED KINGDOM

Mr. J. H. Adams	Manager, Scottish Braille Press, Craigmillar Park, Edinburgh EH16 5NB.

*blind person †speaker ‡exhibitor

Mr. G. Alner	British Computer Association for the Blind, 89 St. John's Road, Boxmoor, Hemel Hempstead, Herts.
Mr. D. Bell (Observer)	Director of Publications, Royal National Institute for the Blind (RNIB), 224 Great Portland Road, London W1N 6AA.
† Mr. G.F.S. Bennett	Computer Services Manager, RNIB.
* Mr. Eric T. Boulter	Director-General, Royal National Institute for Blind.
Mr. G. Bull	Production Director, RNIB.
Mr. S. Cheese	Assistant Braille Works Manager, RNIB.
†*Mr. P.W.F. Coleman	Information Officer, Computer Unit, University of Warwick, Coventry CV4 7AL.
Mr. D. W. Croisdale	Chairman, WCWB Sub-Committee on Computerised Braille Production and Other Media, c/o RNIB.
Professor J. L. Douce	University of Warwick Research Unit for the Blind.
Mr. C. W. Garland (Observer)	Technical Officer, RNIB.
Dr. John M. Gill	Warwick Research Unit for the Blind.
† Mr. R. J. Hagger	Braille Works Manager, RNIB.
† Mr. J. Humphreys	Warwick Research Unit for the Blind.

*blind person †speaker #exhibitor

* Mr. John Lorimer	Research Centre for the Education of the Visually Handicapped, University of Birmingham, 59 Selly Wick Road, Birmingham 29.
Mrs. Pam Lorimer (guide)	
* Mr. M. Milligan (Observer)	Association of Blind and Partially-Sighted Teachers and Students.
* Mr. R. Morrison	Deputy Manager, Scottish Braille Press.
* Mr. T. J. Parker	National Uniform Type Committee, 2 Tenterden Road, London N17 8BE.
* Mr. W.B.L. Poole	Chairman, National Uniform Type Committee, 97 New Bond Street, London W1.
Mr. E. J. Venn	Deputy Director-General, RNIB.
# Mr. Ed Vitu	Telesensory Systems Inc., Woodland Cottage, Dunstall Green, Chobham, Surrey, GU24 8HP.
* Mr. J. N. White	British Computer Association of the Blind, c/o The Open University, Walton Hall, Milton Keynes, MK12 6AA.
* Mr. John Wilson	National Uniform Type Committee, Civil Service Dept., Room 217, Gildengate House, Upper Green Lane, Norwich.

*blind person †speaker ‡exhibitor

UNITED STATES OF AMERICA

Mr. A. Ben-Dor	Telesensory Systems, 3408 Hillview Avenue, PO Box 10099, Palo Alto, California 94304.
Mr. G. Carbonneau	President, Triformations Systems, 3132 S.E. Jay Str., Stuart, Florida 33494.
Mr. R. Daily	Manager, Telesensory Systems.
† Professor Ian Davidson	Assistant Professor of Electrical and Computer Engineering, University of Cincinnati, Mail Loc. No. 30, 898 Rhodes Hall, Cincinnati, Ohio 45221.
† Mr. Martin Droege	Workshop Director, Clovernook Home & School for the Blind, 7000 Hamilton Avenue, Cincinnati, Ohio 45231.
†*Mr. Peter S. Duran	President, ARTS Associates Inc., Boston.
Mrs. M. Duran (guide)	
Mrs. W. N. Epstein	President, National Braille Association, 2470 SW 22 Avenue, Miami FLA 33145.
† Mr. Robert L. Haynes	Data Processing Manager, American Printing House for the Blind, 1839 Frankfort Avenue, Louisville, Kentucky 40206.
†*Dr. Kenneth R. Ingham	ASI Teleprocessing Inc., 123 Water Street, Watertown, MA 02172.

*blind person †speaker #exhibitor

Mrs. Ingham (guide)	
† Mr. Norman C. Loeber	146 N. Bonanza Avenue, Tuscon, Arizona, AZ 85710, USA.
Mrs. F. J. Mann	National Braille Association, 900 Lake Shore Drive 260 b, Chicago, Illinois 60411.
† Dr. Gerald W. Mundy	Executive Director, Clovernook Home and School for the Blind.
† Mr. Henry Paris	Chief, Materials Development Division, The Library of Congress, 1291 Taylor Street, NW, Washington DC 20542.
* Mr. W. M. Raeder (Observer)	Managing Director, National Braille Press, 88 St. Stephen Street, Boston MA 02115.
Mrs. Raeder (guide)	
Mr. L. G. Rooney	Director of Services, Lighthouse for the Blind in New Orleans, 123 State Street, New Orleans, Louisiana 70118.
Mr. H. Snider	President, Access for the Handicapped Inc., Suite 803, 1012 14th Street NW, Washington DC 20005.
Mr. J. E. Sullivan	Duxbury Systems Inc., 56 Main Street, Maynard, Massachusetts 01754.

*blind person †speaker #exhibitor

YUGOSLAVIA

* Mr. J. Duretic (Observer)	Director, Braille Printing House, Ustanicka 25 II, Postanski fah 807, 11001 Belgrade.
Mr. M. Maryanovic (Observer)	Union of the Blind of Yugoslavia.

Computerised Braille Production

This volume represents a survey of the state and trends of hard and soft ware in computerised Braille production. Braille – the writing blind person can feel with their fingers – represents a short hand based on logical as well as linguistic and semantic rules. Therefore its automatic production poses serious problems. Several approaches – depending on the languages involved – are described. Other papers review newly designed hardware and technical problems still to be solved.

ISBN 3-540-12057-2
ISBN 0-387-12057-2