# Moving Toward Greater Accuracy: Improvements and Challenges in Electronic print-to-Braille and Braille-to-Print Translation Since the Implementation of Unified English Braille

by Jennifer Dunnam

Manager of Braille Programs, National Federation of the Blind

Immediate Past Chair, Braille Authority of North America

## Abstract

Electronically-generated braille is a foundational tool for literacy for blind people in our digital age. Unified English Braille (UEB), with its clearly defined rules and symbol structures, provides, on its own or in combination with other specialized braille codes, a means to render both accurate braille representation of any non-image-based content that originated in print, and accurate translation from braille to print of material electronically typed in braille.

In the years since Unified English Braille became the official standard braille code in ICEB member countries (four years ago for the United States, and much longer ago for others), some strides have been made in improving the accuracy of electronic print-to-braille and braille-to-print translation. However, some significant difficulties remain, placing unnecessary limits on the reliability of braille as a means for communication and collaboration. This paper will explore the progress in improving accuracy of electronic braille since 2016 (including external, related technological developments); the details and impacts of some of the persistent but demonstrably correctable problems with electronic translation; and suggestions for bringing about improvements.

## Introduction

For many of us whose strong belief in braille as literacy is firmly grounded in direct observation and experience, the well-known challenges that seem to hinder the growth of the use of braille among those who are blind can sometimes seem intractable. The lack of access to quality braille education (for both teachers and students);lack of materials; unawareness of how to use available tools to produce quality braille; low expectations about braille's speed, efficacy, and adaptability; and other issues, continue to be combatted in ways large and small all around the world. Over the past few decades, talented and committed technology developers have worked tirelessly to bring about the hardware, software, and interfaces that enable braille to be used for communication in the digital age. While the high cost of braille technology has long been an impediment to its universal availability, concerted efforts on this front are beginning to change this equation. As the economic barriers begin to fall, it becomes more important than ever to increase awareness about and prioritization of solving another obstacle, which is, in some respects, actually the least intractable of them all—that is, increasing the accuracy and reliability of electronic braille translation.

At the 6th General Assembly of the International Council on English Braille in 2016, this author presented an exploration of electronic print-to-braille and braille-to-print translation—of the many ways it is accomplished and used, the types and examples of problems causing inaccuracy in electronic braille, and ideas for how to bring about improvement. This current paper seeks to provide an update to that paper[[1]](#footnote-1) by discussing the positive developments since 2016 and the challenges that remain, with a primary focus on UEB implementation in braille that is generated by computer screen readers or used in "on-demand" braille translations in online libraries.

## Notable Developments Since 2016

### Improvements in Liblouis Translation

The Liblouis open source software is used to generate braille in JAWS, NVDA, BrailleBlaster, ChromeVox, VoiceView, Bookshare BRF files, and other products.

Since 2016, there have been many improvements in the translation accuracy of the Liblouis software. As of this writing, its current version is 3.12. However, the screen readers and other products that rely on Liblouis generally do not deploy the latest version until much later than its release. At this writing, JAWS is using version 3.9, NVDA 3.10, and BrailleBack and Bookshare much older versions.

The improvements since 2016 include the correct display of a wider range of symbols, improvements to back translation, better application of mode handling and contraction rules, and the like. Some of the major issues yet to be addressed are discussed later in this paper, but the active ongoing development of the UEB translation in Liblouis is encouraging.

### Amendments to UEB Rules About Apostrophes and Quotation Marks

The Unicode character U+2019 is often used in print to represent both the apostrophe and the single closing quotation mark. This causes no issue in print, because the characters look alike. In braille, however, two very different braille symbols are used for these two characters, ('and ,0). For years, braille that is translated in real-time has been littered with single closing quotation marks where apostrophes should be—sometimes interfering with other rules and changing the formation of words entirely. Braille software developers have worked to improve the problems, with varying degrees of success. Braille producers have spent untold hours intervening manually or applying search and replace programming to ensure that apostrophes and quotes are brailled correctly.

The release of the revision of UEB §7.6[[2]](#footnote-2)in 2019 not only has made helpful changes in the rules for apostrophes and quotes, but it gives guidance directly to software developers to assist with the application of the new rules.

### Typeforms

For many years, most screen readers have attempted to indicate typeforms in braille by displaying dots 7 and or 8 in combination with each character of the text to which the typeform is applied, and using a one-cell character at one end of the refreshable braille display to indicate which attributes are being shown. This can be used to show color, strikeout, highlighting, and other typeforms that do not have specifically defined braille indicators. This method has the disadvantages of requiring active effort on the part of the reader to determine if a typeform is present, which typeform it is, and exactly where it starts and ends; it also can detract from the readability of the affected text.

At the end of 2016, NV Access released a version of its NVDA screen reader which renders the UEB typeform indicators for underline, italic, and bold correctly in braille when the speech settings are set to report the typeforms. Significantly, as of this writing, NVDA is the only screen reader that will display these indicators in real-time translation.

Superscripts and subscripts do not display with UEB indicators in any screen reader, although they are spoken with most. Braille typeform indicators cannot be typed in braille for translation to print.

Since 2016, a method was implemented in the Duxbury Braille Translator for mapping the transcriber-defined typeform indicators to any applicable typeforms in a document, so that the specified indicators will be applied when the document is imported from Word to Duxbury.[[3]](#footnote-3)

### Window-Eyes Discontinued

In 2017, sales and support of GW Micro's Window-Eyes screen reader were discontinued. The significance of this development lay in that, at that time, the UEB output of Window-Eyes was the most accurate of the PC-based screen readers.

### BrailleBlaster

Officially launched in 2017 but developed and tested for years prior, BrailleBlaster is a stand-alone translation program that the American Printing House for the Blind initially created to assist transcribers in working with publishers' files. In the years since, however, more file types have come to be supported, such as docx, html, and txt. This, along with its availability at no charge, has made BrailleBlaster a game changer for consumers and braille producers alike. For BrailleBlaster, the developers have implemented corrections to some of the problems in Liblouis discussed later in this paper. There are certainly some areas for growth; for example, some symbols such as shapes, which have defined braille symbols, are instead translated as their names enclosed in transcriber's note indicators. However, updates/improvements to BrailleBlaster are released regularly.

### Emoji in Braille

With the version of Apple's VoiceOver screen reader released in October 2017, the names of the emoji symbols began to be displayed in braille as lowercase words enclosed in transcriber's note indicators. This is a model for other software to follow. The emoji appear to be updated regularly as new ones are added. However, see below for discussion of emoji in Apple's optional Liblouis tables. Other screen readers display emoji as strings of characters not related to the names of the symbols (and therefore not readable in braille).

### Translation from contracted braille to print:

The ability to input using contracted braille became available in NVDA in 2017 and is now available in most all of the screen readers. It is usable but remains somewhat unreliable, particularly in screen readers with older versions of Liblouis. Voiceover's system UEB table generally yields the most accurate back translation, but, notably, it does not back translate the fraction symbols like (¼) as some other screen readers can.

## Significant Issues To Be Addressed

In most screen readers, errors exist that fall into the category of "not preferred" but do not cause ambiguity. In what follows, we will focus on errors that cause ambiguity or discrepancy.

### The Dash

The Em-dash — (U+2014) displays incorrectly even in the latest versions of Liblouis that have been deployed in screen readers. JAWS, NVDA, BrailleBack, VoiceView, and others are affected, as well as nearly every book in Bookshare's online library. This is a frequently-occurring character used in most cases where a "dash" is intended. In the UEB symbols list,[[4]](#footnote-4) this specific Unicode character is called "dash" and is defined in braille as dots 6, 36 (,-). Liblouis is displaying it as dots 5, 6, 36 (",-). The three cell symbol dots 5, 6, 36 is linked in the rulebook[[5]](#footnote-5)to the "long dash" (horizontal bar) ― U+2015. Liblouis displays u+2015 correctly. The en-dash – (U+2013) is being displayed by UEB using the symbol for dash. There is no definition for u+2013 in the rulebook.

Part of the reason for the persistence of this problem may lie in inconsistent representation of the em-dash in the UEB Rulebook itself. Excerpts from the Rulebook are used for creating some of the test files used for Liblouis.

The dash does display correctly in ScreenReader from Dolphin Computer Access, and in Apple's VoiceOver with the "system" UEB table selected. It also translates correctly in the Duxbury and BrailleBlaster programs.

### The Checkmark

The checkmark, (✓ U+2713) is singled out here because of the relatively recent adoption of the corresponding braille symbol (@%) in 2018. The latest deployed versions of Liblouis have not implemented this symbol. The only screen reader that displays it correctly is VoiceOver as long as the "system" UEB braille table is selected.

### A Numeric Mode Problem

In all Liblouis deployments as well as in the translator used by ScreenReader from Dolphin Computer Access, if a digit follows a comma or period that is immediately preceded by a letter or letters, the numeric indicator is incorrectly being placed before the comma or period rather than before the digit.[[6]](#footnote-6)For example, **No.9** displays as no#4i rather than no4#9. This issue also arises when reference numbers for footnotes and endnotes, (which are not shown as superscript in braille by the screen readers or Bookshare's translator), are adjacent to punctuation in a text. For example, **rules.23** displays as rules#4bc.

### Some Missing Grade 1 indicators

In the latest deployments of Liblouis, the grade 1 symbol indicator does not consistently precede letter sequences such as Bldg., LLP, and Blvd. The indicator is needed to show unambiguously that the "blind" and "little" shortforms are not being used. Note that this problem is properly corrected by applying a rule, not by fixing individual letter strings.[[7]](#footnote-7)Since this issue was reported some years ago using the examples "LLC", "BLT", and "BLVD", those strings have been corrected, but "LLP" and others still lack the indicator. To read **LLP** without a grade 1 indicator is not particularly confusing, but if someone becomes accustomed to reading it that way and then types ,,llp with six-keys for back translation, correctly-implemented translation rules will render it **LITTLEP**. Other words are similarly affected.

### Mathematics

In general, real-time braille translation support for mathematics/technical material remains far from complete although it has seen some improvements. There is now more robust braille support for mathML on the Web, although, for technical reasons, this seems generally to be true for Nemeth Code and not for UEB renderings of mathematical expressions. Items such as the UEB math tutorial from the American Printing House for the Blind[[8]](#footnote-8) rely on already-translated static braille font to render the examples and exercises in well-structured UEB for refreshable braille.

Amazon's Kindle hardware and applications have improved Math support as well, but many variables cause the experience of using math in these books to be uneven.

In Office 365 documents, MathType content will now display in Nemeth. Some, but not all, individual mathematical symbols will display in UEB in real-time translation in Microsoft Word documents.

With the improvements to Liblouis's back translation, it is now possible, in some screen readers when used in conjunction with a six-key braille device connected, to braille some individual mathematical symbols directly into Microsoft Word or other applications and have them translate to print. Supported symbols include fractions (⅞), shapes (△), arrows (↑), and many others (∠), (√).

### Computer Braille Code

Some situations remain where screen readers still unnecessarily display and require the use of Computer Braille Code (ASCII braille) within UEB.

### BrailleBack (Android)

For years, the screen reader for Android has lagged significantly behind others in its translation accuracy. The August 2019 update (BrailleBack 0.97 Beta) brought many improvements to general symbol accuracy. However, this application still appears to be using a version of Liblouis prior to 3.8, and therefore it contains many of the overarching problems that were corrected in versions 3.8 and later.

### VoiceOver's New Braille Table Options

In 2019, Apple introduced the option for a user to choose the Liblouis tables for braille display and input, rather than the default "system" tables. This new option in its current form is cause for concern, especially if it signals an intent by Apple to consider discontinuing maintenance of the "system" tables. The Liblouis implementation in Apple devices contains many global inaccuracies discussed above and not found in Apple's default UEB table.

Unlike with other screen readers using Liblouis, the names of emoji are displayed in VoiceOver with the Liblouis table invoked, but, ambiguously, the names are not enclosed in TN indicators but instead preceded and followed by dots 25.The use of dots 25 as an enclosure symbol has no connection with existing braille rules, can cause ambiguity, and violates principles of UEB. The decision to use a one-cell enclosure symbol could conceivably have been made in response to user request for an enclosure occupying fewer cells;if a need exists to address such a request, a different solution should be devised. However, it is important for developers to be aware that although UEB's TN indicators may be new to some adult braille users, they have an established, widely recognized purpose and meaning—to signal to the reader that the enclosed words do not appear in the print display but have been added to give information about a graphic representation.

### Bookshare

In addition to the other Liblouis issues previously mentioned, the braille-ready files made available by Bookshare have a unique issue not caused by Liblouis. When a space occurs in print before certain non-alphabetic, non-numeric symbols, that space is omitted in braille in Bookshare's BRFs. The affected symbols include but are not limited to brackets, ampersand, dollar sign and other currency, section symbol, etc. For example, the sentence "Goldman & Rollins raised $2,000 for the campaign" displays in braille as "Goldman& Rollins raised$2,000 for the campaign". The missing space can also cause other translation errors in adjacent words that can slow the recognition of these words. Note that this is specific to the conversion of the book to BRF and does not occur when a braille display is used to read the print-based formats of these books (such as the Web reader or with VoiceDream Reader). Other BRF converters do not yield such space omissions.

## Conclusion

There has been much continuing progress in the improvement of braille translation technology. Yet, inaccuracy in braille translation software for screen readers still creates ambiguity and outright misleading information for those reading the displayed braille, and unreliable output for those typing with the six braille keys. This remains a hindrance to full implementation of UEB, making the gain that we have sought by this tremendous international effort not yet fully realized. It adds unnecessary obstacles to the education of blind children, and to reliable reading and communication for all. To provide awareness, support, and guidance to the skilled technology developers in making all possible needed corrections to translation software must be a top priority as we continue to build toward increased literacy for those who are blind in the present and in the future.

Jennifer Dunnam

1. See 2016 General Assembly Papers (International Council on English Braille) http://www.iceb.org/gapapers16.html. [↑](#footnote-ref-1)
2. See "Update to the rules for quotation marks and apostrophes": http://www.iceb.org/ [↑](#footnote-ref-2)
3. Duxbury DBT 12.4 Help File, Typeform Indicators: https://www.duxburysystems.com/documentation/dbt12.4/working\_with\_word/word\_typeforms.htm [↑](#footnote-ref-3)
4. Rules of Unified English Braille, Appendix 3, page 314 [↑](#footnote-ref-4)
5. Rules of Unified English Braille, Appendix 3, page 308. [↑](#footnote-ref-5)
6. See UEB §6.4.1. [↑](#footnote-ref-6)
7. See UEB §10.9.5. Note that applying 10.9.5 requires applying 10.9.3(b and c), from which follows that any of the letters-sequences making up the mentioned shortforms and not followed by a vowel are read as a shortform. [↑](#footnote-ref-7)
8. UEB Math Tutorial (American Printing House for the Blind): https://uebmath.aphtech.org/ [↑](#footnote-ref-8)