

# Forward Transliteration of Dzongkha Text to Braille

*Tirthankar Dasgupta Manjira Sinha Anupam Basu*  
Indian Institute of Technology Kharagpur  
{iamtirthankar, manjira87, anupambas}@gmail.com

## ABSTRACT

In this paper we present an automatic Dzongkha text to Braille forward transliteration system. Dzongkha is the national language of Bhutan. The system is aimed at providing low cost efficient access mechanisms for blind people. It also addresses the problem of scarcity of having automatic Braille transliteration systems in language slime Dzongkha. The present system can be configured to take Dzongkha text document as input and based on some transliteration rules it generates the corresponding Braille output. We further extended the system to support an Audio QWERTY editor which allows a blind person to read and write Dzongkha texts or its equivalent Braille through a computer. The editor also contains Dzongkha voice feedbacks to further ease the use.

---

KEYWORDS: Braille, Forward Transliteration, Dzongkha, Audio QWERTY

---

## 1 Introduction

The Braille encoding system is one of the primary means of representing textual documents in a readable format for the Blind persons (Loaber, 1976; Weller and Klema, 1965; Basu et al., 1998; Lahiri et al., 2005). However, due to the scarcity of Braille compatible reading materials, blind people face difficulty in fulfilling necessities like education and employment. The unavailability of low-cost technical supports worsens the situation further. For them, the inability to communicate via the traditional writing system leads to complications at official places where the primary mode is still through writing. In order to bridge between the written text systems generally aimed at sighted persons and access mechanisms through which blind people can communicate with ease and efficiency, developments of systems such as automatic Braille transliteration and screen readers are highly needed.

Several works have been done on building automatic, bi-directional text to Braille transliteration system and speech enabled interfaces for the Blind community (Blenkhorn, 1995; Pennington and McCoy, 1998; Raman, 1996). However, all the present systems suffer from the central limitation of language dependency. A Braille transliteration system requires certain rules, based on which the system automatically transliterate a given text document into Braille. The rules are very specific to the input language. Therefore, system for a particular language cannot be extended to the others.

Dzongkha is the national language of Bhutan<sup>1</sup>. It has been found that more than 4% of the children in Bhutan suffer from early blindness. Therefore, it is essential to develop tools and technologies in such a country that will facilitate blind persons to access information from written text documents. The existing text on Braille transliteration systems cannot be directly applied to Dzongkha, mainly due to the following reasons:

1. Most of the systems are based on foreign languages like English, French, Germany, Spanish, Portuguese, and Swedish (winbraille, 2012; Duxbury, 2000; indexbraille, 2012; NFBTRANS, 2004). As Bhutanese scripts are quite different from these languages, separate rules are applied to transliterate from Dzongkha to Braille.
2. Foreign systems like, Duxbury (2004) and JAWS (2008) are costly and a large segment of the blind population is from poor economic background.

In order to overcome the limitations of the existing systems, we have developed a Dzongkha language text to Braille transliteration (DLBT) system. The system comprises of the following two key features:

- i. A generic framework for the transliteration of Dzongkha texts to Braille.
- ii. An audio QWERTY editor that provides pre-recorded voice feedback to any key-press operation performed by a blind user.

Apart from creating Braille encoded textbooks, such system will also assist blind people to create their own electronic text documents.

The rest of the paper is organized as follows: section 2 briefly discusses about the state of the art in text to Braille transliteration systems; in section 3 we have discussed about the Dzongkha

---

<sup>1</sup> [http://en.wikipedia.org/wiki/Languages\\_of\\_Bhutan](http://en.wikipedia.org/wiki/Languages_of_Bhutan)

Braille encoding system; in section 4 we have described the architecture of our proposed system and its the key components; we have discussed the Audio QWERTY editor in section 5 and finally concluded the paper in section 6.

## 2 Related Works

A number of texts to Braille transliteration tools are present for different languages. In English, systems like, Duxbury (2000), JAWS (2008), and WinBraille (2012) are popularly used. However, the problem of transliterating English text to Braille is relatively simpler than south-east Asian languages like, Indian or Bhutanese languages. It has been observed that due to the similarity in the scripts of these languages, rules to transliterate Dzongkha texts to Braille are very similar to that of the Indian language texts. The Sparsha Text to Braille Transliteration toolset (Lahiri et al., 2005; Dasgupta and Basu, 2009) is the only system encountered by us so far that can transliterate Indian language texts to Braille. The system provides some unique features like transliteration of mathematical symbols and tactile graphics to Braille. Apart from computer based transliterations of Braille, a Braille writing tutor system has been developed by Kalra et al. (2007). The prototype tutor system uses an e-slate device to capture a student’s action and tries to develop the Braille writing skills.

## 3 The Dzongkha Braille Encoding System

The Dzongkha Braille system is the standard technique of representing Dzongkha texts into Braille (Tibetan, 2004; Dzongkha, 2005). The system uses 6 dot cells to represent each Dzongkha character. The combination of these 6 dots can represent a total of  $2^6-1$  i.e., 63 different Braille characters. However, as described later, Dzongkha has more than 63 characters. This issue is handled by assigning multiple Braille characters for a single source language character. This is illustrated in the third row of table 1. Although Braille encoding for all languages share the same Braille font, different languages have different Braille representation rule, thus a single Braille character may be interpreted differently in the context of different languages. Table 1 contains the mapping of same Braille codes to different characters of three different languages: English, Hindi, and Dzongkha.

Braille	English	Hindi	Dzongkha
k	k	क	ཀ
⠠	.	त	མ
" R	" R	ऋ	ར

TABLE 1- Mapping of Braille codes to characters of English, Hindi and Bangla

The Dzongkha script, which is also known as the Bhutanese script, is used to write Dzongkha. The Dzongkha script has 30 consonants and 4 basic vowels. The 30 consonants can occur in three different positions: twice in nominal position and once in orthographic subjoined position (Bhutan, 2004; Tibetan, 2009; Dzongkha, 2005). This leads to a total of  $30*3=90$  consonants, 4

vowels and 1 extra character, i.e., 95 alphabets in Dzongkha. Therefore, Dzongkha scripts can be written either from top to bottom or from left to right (Bhutan, 2004) (refer to figure 1(b)).

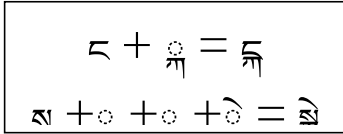


FIGURE 1(A) - Construction of Dzongkha conjugates

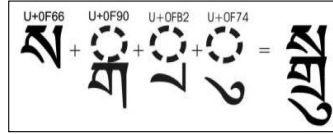


FIGURE 1(B) - Dzongkha characters arranged from top to bottom

As in other languages, two or more consonants can combine to form a conjugate Dzongkha character. However, these consonants are not separated by any special symbols (as in the case of Indian languages where the consonants within the conjugate are separated by a special symbol called halant). The consonant clustering in Dzongkha takes place between the consonant at the nominal position and consonant at the orthographic subjoined position. This is illustrated in Fig. 1(a). It can be observed from table 2 below that the conjugate characters, as constructed by clustering of consonants and vowels, may have an entirely different visual representation. However, the corresponding transliterated Braille is represented by a sequence of Braille cells for each of the characters.

Rule	Example	Braille
CC	ཅ + འ = ཅའ	@+K
CCV	ཅ + འ + འ = ཅའའ	gK o
CCC	ས + འ + འ = སའའ	@S@TR
CCCV	ས + འ + འ + འ = སའའའ	@S@TR9

TABLE 2- Conjugate construction rules with examples taken from Dzongkha language (C=Consonant, V=Vowel)

Unlike any Indian language texts, there exists no inter-word separation in Dzongkha. Each of the Dzongkha words are composed of single or multiple syllables. The syllables are separated by a special symbol called *tsheg*. Each of the syllables contains a root letter, and may additionally contain a prefix, suffix, vowel and post-suffix. Figure 2 illustrates this phenomenon with an example.

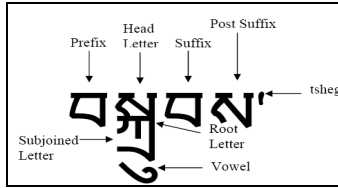


FIGURE 2- Construction of a Dzongkha syllable

#### 4 The System Architecture

The architecture of our proposed Dzongkha text to Braille transliteration (DLBT) system is shown in figure 3. The diagram presents the essential components of the system. The details are discussed in the subsequent sections.

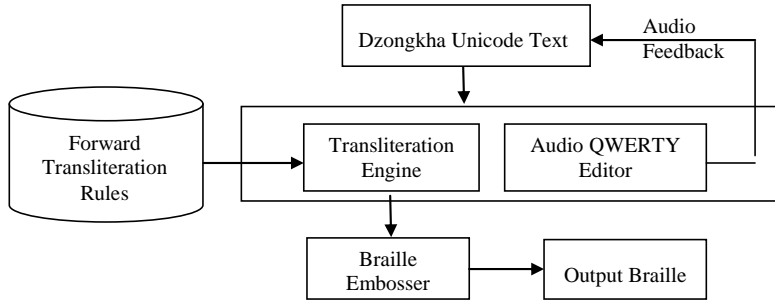


FIGURE 3 - The proposed system architecture

##### 4.1 The Input and Output of the System

The input to the system can either be a Unicode (Hudson, 2000) supported Dzongkha text document written in any popular word processor, or Dzongkha texts entered through a keyboard. Based on the user’s requirements, the system can transliterate the given text to Braille. The Braille output is then provided to the Blind user by printing the Braille texts on a large variety of commercially available Braille embossers (Taylor, 2000; Brailler). The current system has been evaluated on the Braille embossers like, Index Basic-S, Index Basic-D, Index 4X4 PRO, Romeo and Juliet Pro.

##### 4.2 Forward Transliteration

Based on the information discussed in section 3, we have constructed distinct forward transliteration rules for Dzongkha text to Braille. The transliteration engine consists of a *code table* which is nothing but a collection of the transliteration rules. An example of the code table structure is shown in figure 4.

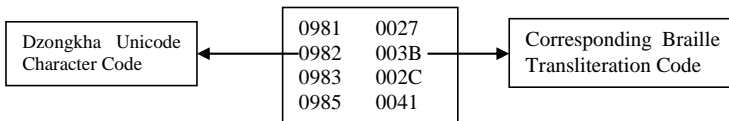


FIGURE 4- Code Table Structure of Forward Transliteration Engine.

Most of the transliteration rules are very similar to Indian language text to Braille transliteration (Lahiri et al., 2005; Dasgupta and Basu, 2009). However, there are few exceptional cases like, handling of some special conjugates in Dzongkha script, where the transliteration rules changes. For example, Braille representation of certain Dzongkha characters like “ra”, “la” and “sa” depends upon the occurrence of the character that follows it. This process is represented by the following rule in figure 5:

```

CONJ {PREFIX {U+0F62|U+0F63|U+0F66, U+0F90|U+0F92|U+0F94}}
      → PutBraille (53|59|57)

```

FIGURE 5- Context dependent text to Braille conversion rule in Dzongkha

The rule says, if the head letter of a Dzongkha conjugate begins with U+0F62, U+0F63 or U+0F66 and the root letter belongs to U+0F90, U+0F92, or U+0F94 then the Braille representation of the head characters will changed to “53”, “59” or “57”. Figure 6 illustrates the above rule with an example; figure 7 shows a screenshot of the working of the Dzongkha text to Braille transliteration system.

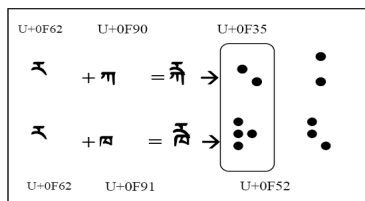


FIGURE 6- Illustration of the context dependent forward transliteration rules in Dzongkha

The Dzongkha Braille transliteration tool allows 38 characters per line and 25 lines per page. However, the system allows the user to change the configuration if needed. As mentioned above Dzongkha script does not allow any inter word space. However, the syllables and sentences are separated by the special symbols: *tsheg* and *she*. An interesting feature found in Dzongkha Braille formatting is that, each line of a Braille document must end with either *tsheg* or *she*. This issue is handled by an inbuilt auto formatting module of the transliteration system. The module first analyses the transliterated Braille output and puts 38 Braille characters per line at the preview window. If the last character of a line is not a *tsheg* or *she* then the module starts accumulating the previous characters into an array till it gets a *tsheg* or *she*. The array elements are then printed into the next line of the preview window. This results into the fact that after transliteration is over all the lines in the Braille output does not contain 38 characters.

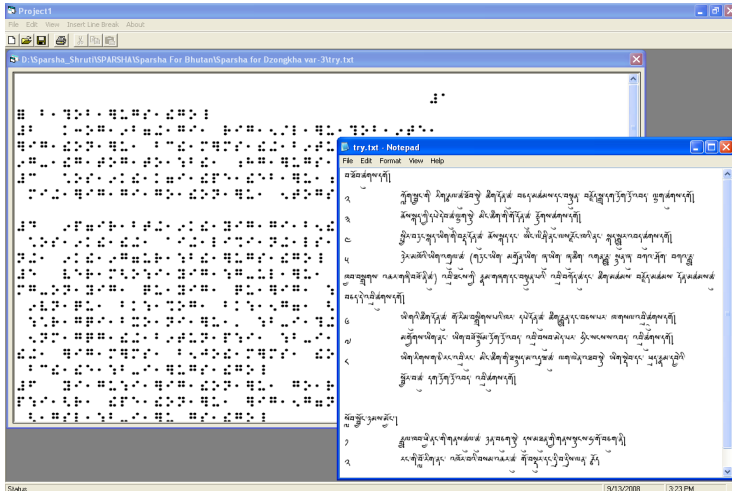


FIGURE 7: WORKING OF THE DZONGKHA TEXT TO BRAILLE TRANSLITERATION SYSTEM

## 5 Dzongkha Text and Braille Audio QWERTY Editor

The primary goal of the audio QWERTY editor is to allow Blind people to create Dzongkha Text/Braille documents (see figure 8). This requires an interface for accepting the regional language text entered through the keyboard and performing different operations on it, like formatting, printing and saving the text. The creation of a new editor interface was not warranted as it would put additional burden on the user to learn the new system. Hence, we have chosen to use some already existing standard editor with the required capabilities. Our investigations have proved that Microsoft Word can be configured to accept text in regional languages including Dzongkha. Although the Audio QWERTY editor plug-in can be integrated to any other Unicode enabled text editors like, Notepad, and WordPad, the reasons for choosing Microsoft Word are:

1. Support for Unicode—ensures the use of multi-language text documents
2. Rendering of Fonts —uses proper rendering engines for correct rendering of regional language fonts including glyph shaping and repositioning (Rolfe, 2001).
3. Well documented object model – the editor exposes a comprehensive set of objects for interacting with it, this simplifies the task of programming.
4. Ease of Use – the existing popularity of this editor predicts a low learning curve for the proposed system. Further, the editor provides a large number of keyboard-shortcuts.

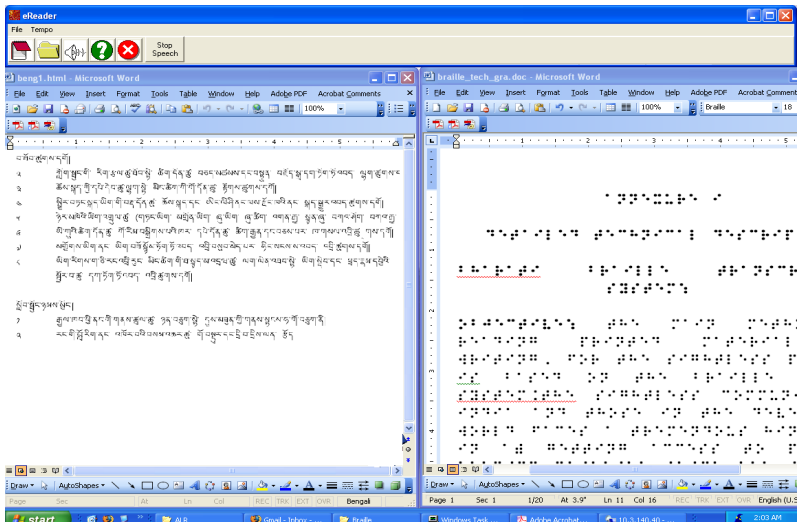


FIGURE 8- The Dzongkha Braille Audio QWERTY Editor. For any key press operation, the Blind user will get audio based feedback.

The audio QWERTY editor is integrated with the pre-recorded Dzongkha alphabet voices that make it different from other commercially available text editor; as a result, each of the keyboard operation performed through this editor is followed by a voice feedback in Dzongkha. This enhances its ease and efficiency for a Blind person to read and write Braille texts.

**Conclusion**

The Dzongkha text to Braille transliteration system is an attempt to develop low cost technology to assist blind people. The system will help a large segment of Blind population of Bhutan to access printed text book materials. The present version of the system can only perform the forward transliteration of Text to Braille. In future, we will try to incorporate the reverse transliteration approach where given a Braille document it can be transliterated back to its original Dzongkha font. Further, we will provide an online version of the system so that it can be accessed from different parts of the world.

**References**

A. Basu, S. Roy, P. Dutta and S. Banerjee, "A PC based multi-user Braille reading system for the blind libraries", IEEE Transactions on Rehabilitation Engineering, Vol. 6, No. 1, March 1998, pp.60—68

A. Lahiri, J. S. Chattopadhyay, A. Basu, "Sparsha: A comprehensive Indian language toolset for the blind". Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility. 2005.



T. Dasgupta, and A. Basu, "A speech enabled Indian language text to Braille transliteration system", Information and Communication Technologies and Development (ICTD), 2009, pp-201-211

P. Blenkhorn, "A system for converting Braille to print", IEEE Transactions on Rehabilitation Engineering, Vol. 3, No. 2, June 1995, pp. 215-221

HAL. Dolphin Computer Access, [www.dolphinuk.co.uk/products/hal.htm](http://www.dolphinuk.co.uk/products/hal.htm)

C.A. Pennington and K.F. McCoy, "Providing intelligent language feedback or augmentative communication users", Springer-Verlag, 1998.

T.V. Raman (1996). "Emacspeak – a speech interface". Proceedings of CHI96, April 1996

winbraille. (n.d.). Retrieved on September 14, 2012, from [www.braille.se/downloads/winbraille.htm](http://www.braille.se/downloads/winbraille.htm)

Duxbury.(n.d). In Duxbury Braille Translator, 2000. [www.duxburysystems.com](http://www.duxburysystems.com)

indexbraille. (n.d.). Retrieved on September 14, 2012, from [www.indexbraille.com](http://www.indexbraille.com)

NFBTRANS. National Federation of the Blind, 2004, <http://www.nfb.org/nfbtrans.htm>

MONTY, VisuAide. <http://www.visuaide.com/monty.html>

JAWS. 2008. In JAWS for Window. Freedom Scientific. [http://www.freedomscientific.com/fs\\_products/software\\_jaws.asp](http://www.freedomscientific.com/fs_products/software_jaws.asp)

N. Kalra, T. Lauwers, D. Dewey, T. Stepleton and M. B. Dias, "Iterative design of a Braille writing tutor to combat illiteracy", Proceedings of the 2nd IEEE/ACM International Conference on Information and Communication Technologies and Development, December, 2007.

A. Taylor. "Choosing your Braille embosser", Braille Monitor, October 2000. Available at [www.nfb.org/bm/bm01/bm0110/bm011007.htm](http://www.nfb.org/bm/bm01/bm0110/bm011007.htm)

Brailler. (n.d). Retrieved from [www.brailler.com/](http://www.brailler.com/)

J. Hudson for Microsoft Typography, "Windows glyph processing: an open type primer", November 2000,

<http://www.microsoft.com/typography/glyph%20processing/intro.mspx>

R. Rolfe, 2001, "What is an IME (Input Method Editor) and how do I use it?" <http://www.microsoft.com/globaldev/handson>

Bhutan. (2004). downloaded from <http://www.tcllab.org/events/uploads/pema-bhutan.pdf>

Tibetan. (2009). In The Online Encyclopedia of Writing Systems and Languages, Retrieved October 14, 2009, from <http://www.omniglot.com/writing/tibetan.htm>

Dzongkha (2005), Retrieved from <http://www.learn tibetan.net/grammar/alphabet.htm>

N.C. Loaber, "Modified standard office equipment for Braille translation and embossing", ACM SIGCAPH Computers and the Physically Handicapped, Issue 21, November 1976, Pages 9-19

Walter J. Weller, Virginia C. Klema, "A rapid Braille transliteration technique for certain IBM machine", Communications of the ACM, Volume 8 Issue 2, Feb. 1965, Pages 115-116

