English to Punjabi Transliteration using Orthographic and Phonetic Information

Kamaljeet Kaur
Guru Nanak Dev Engineering College
Ludhiana-141006, Punjab
meetk.89@gmail.com

Parminder Singh
Guru Nanak Dev Engineering College
Ludhiana-141006, Punjab
parminder2u@gmail.com

Abstract

Machine transliteration is an emerging and a very important research area in the field of machine translation. While the translation system finds the same meaning word/sentence in another language, the transliteration helps us to pronounce them. This paper describes the process of transliteration from English to Punjabi language using a rule based approach. Both source grapheme and phonetic information of words have been considered for rule formation to achieve high performance and more accurate result. Phonetic information proved vital for correct transliteration as well as for ambiguous words. The system is tested on news domain text of more than 10,000 words and achieved accuracy of 95%.

Keywords

Machine translation, transliteration, natural language processing, transliteration rules.

1 Introduction

Transliteration is the conversion of a text from one script to another. It is the process of representing words from one language using the approximate phonetic or spelling equivalents of another language. Machine translation (MT) is the process that takes a message in a source language and transforms it into a target language, keeping the exact meaning. Transliteration is meant to preserve the sounds of the syllables in words. This paper presents the development of English to Punjabi transliteration system that can transliterate English text into equivalent Punjabi text.

The remainder of this paper is organized as follows. Section 2 describes the related work done in machine transliteration. We have described basic character to character mapping and rules for transliteration in section 3. Performance evaluation is discussed in section 4. Finally, we have concluded it in section 5.

2 Related Work

Significant work in the field of machine transliteration has been done for Indian as well as for foreign languages. Three approaches for transliteration being used are: Grapheme based models; Phoneme based models and Hybrid models. Knight and Graehl (1998) have developed a phoneme based statistical model using finite state transducer that implements transformation rules to do backward transliteration. They have proposed method for automatic backward transliteration that can be used for transliteration of words from Japanese back to English. Oh and Choi (2002) have proposed English to Korean transliteration system using pronunciation and contextual rules. They have used phonetic information such as phoneme and its context as well as orthography. Malik (2006) has developed Punjabi Machine Transliteration System that is used to transliterate words from Shahmukhi script to Gurmukhi script using transliteration rules. Saini and Lehal (2008) have proposed a corpus based transliteration system for Shahmukhi script to Punjabi language. The transliteration system has been tested on a small set of poetry, article and story. The average transliteration accuracy of 91.37% has been claimed. Goyal and Lehal (2009) have proposed system in which Hindi words are transliterated into Punjabi words. They have implemented complex rules for accurate transliteration between Hindi-Punjabi language pair. Josan and Lehal (2010) have presented a novel approach to improve Punjabi to Hindi transliteration by combining a basic character to character mapping approach with rule based and Soundex based enhancements. Quite a reasonable improvement can be achieved by small amount of dependency or contextual rules.

D S Sharma, R Sangal and J D Pawar. Proc. of the 11th Intl. Conference on Natural Language Processing, pages 315–320, Goa, India. December 2014.

Kaur and Josan (2011) have proposed a system that addresses the issue of statistical machine transliteration from English to Punjabi using MOSES that is a statistical machine transliteration tool. After applying transliteration rules average accuracy of this transliteration system comes out to be 63.31%. Deep and Goyal (2011) have proposed a transliteration system that addresses the problem of forward transliteration of person names from Punjabi to English by set of character mapping rules. The proposed technique achieved accuracy of 93.22%. Josan and Kaur (2011) have developed a statistical model that is used for transliterating the Punjabi text into Hindi text. The proposed system has claimed 87.72% accuracy rate. Dhore et al. (2012) have focused on the specific problem of machine transliteration of Hindi to English and Marathi to English which are previously less studied language pairs using a phonetic based direct approach. Bhalla et al. (2013) have proposed rule based transliteration scheme for English to Punjabi. Some rules have constructed for syllabification. In this probabilities are calculated for name entities (proper names and location). The proposed approach has attained accuracy of 88.19%. Joshi et al. (2013) have proposed system that can do transliteration from Roman script to Devanagari script using statistical machine learning approach.

3 Design and Implementation

3.1 Approach Followed

The proposed system for English to Punjabi transliteration follows rule based approach. A machine transliteration model should reflect the dynamic transliteration behaviors in order to produce the correct transliterations thus we have considered both source grapheme and phonetic information to achieve high performance and more accurate result. The input text goes through various stages, as shown in Figure 1, in order to get transliterated into equivalent Punjabi text.

Preprocessing module identifies language based on Unicode of the words. Segmentation module involves segmentation of source string into transliteration units of source language. Transliteration module transliterates English text to equivalent Punjabi text based on the rules framed. Surrounding characters in a given word are also considered for resolving ambiguities and for more appropriate result.

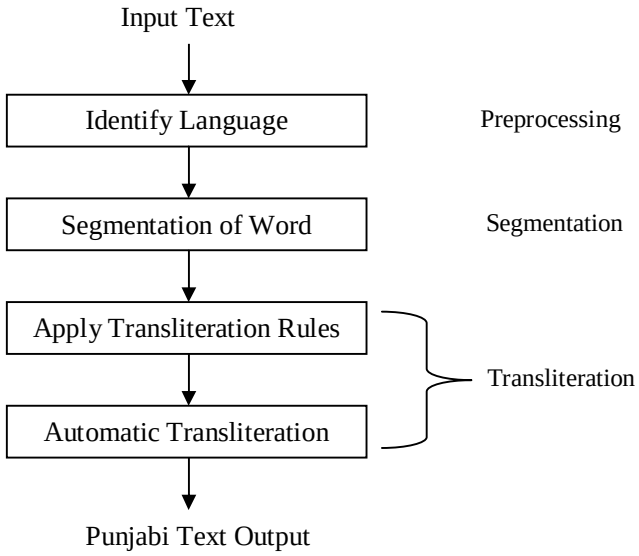

Figure 1. Rule based approach for Transliteration

3.2 English-Punjabi Transliteration Rules

Due to the differences in syntax and vocabulary of both English and Punjabi language, there is need to construct some hard rules. These rules are explained and illustrated in the following subsections.

3.2.1 English-Punjabi Character to Character Mapping

Direct character to character mapping of consonants and vowels for English-Punjabi language pair is shown in Tables 1, 2 and 3 (Deep and Goyal, 2011).

k ਕ	kh ਖ	g ਗ	gh ਘ	ng ਙ
ch ਚ	chh ਛ	j ਜ	jh ਝ	yan ਞ
t ਟ	th ਠ	d ਡ	dh ਢ	n ਣ
t ਤ	th ਥ	d ਦ	dh ਧ	n ਨ
p ਪ	ph ਫ	b ਬ	bh ਭ	m ਮ
y ਯ	r ਰ	l ਲ	v,w ਵ	rh,r ੜ
sh ਸ਼	khh ਖ਼	ghh ਗ਼	z ਜ਼	f ਫ਼
lla ਲ਼	s ਸ	h ਹ		

Table 1. English-Punjabi Consonant Mapping

a ਅ	a, aa ਆ	i, ya ਇ	i ਈ	u ਉ
u ਊ	e, a ਏ	ai ਐ	o ਓ	au ਔ

Table 2. Independent Vowels Mapping

a, aa ਾ	i ਿ	i, ee ੀ	u ੁ	u, oo ੂ
e ੇ	ai, ay ੈ	o ੋ	o, au ੌ	

Table 3. Dependent Vowels Mapping

3.2.2 First Vowel Character in Word

Only direct mapping is not sufficient for transliteration, which may lead to very low accuracy of the system. The vowels in English are mapped to different characters in Punjabi depending on their position in the given word. Following rules are proposed for vowels when they occur at first position in the word:

Rule 1: if first character 'a' is followed by consonant then it is transliterated to 'ਅ'. For example, aman→ਅਮਨ

Rule 2: if 'an' combination is at first position then it is transliterated to 'ਅੰ'. For example, ankit→ਅੰਕਿਤ

Rule 3: if 'ea' combination is at first position then it is transliterated to 'ਈ'. For example, eagle→ਈਗਲ

Rule 4: if 'ei' combination is at first position then it is transliterated to 'ਏ'. For example, eight→ਏਟ

Rule 5: if 'in' combination is at first position then it is transliterated to 'ਇੰ'. For example, india→ਇੰਡੀਆ

Rule 6: if 'oi' combination is at first position then it is transliterated to 'ਓ'+'ਇ'. For example, oil→ਓਇਲ

Rule 7: if first character 'u' is followed by 'double consonant' then 'u' is transliterated to 'ਉ'+'ੱ'. For example, utter→ਉੱਤਰ

3.2.3 Last Vowel Character in Word

Vowels are mapped to different characters when they occur at last position. In order to achieve more accurate result there is need to refer to previous characters because their pronunciation is also affected by the consonants or vowels, they follow. The following rules are proposed to handle last vowel character in a word:

Rule 1: if 'ia' combination occurs at last then 'a' is transliterated to 'ਆ'. For example, sonia→ਸੋਨੀਆ

Rule 2: if last character 'a' is preceded by consonant then 'a' is transliterated to 'ਾ'. For example, samrala→ਸਮਰਾਲਾ

Rule 3: if 'nu' combination occurs at last then 'u' is transliterated to 'ੰ' + 'ੂ'. For example, sonu→ਸੋਨੂੰ

Rule 4: if 'oa' combination occurs at last then it is transliterated to 'ੋ'+ 'ਆ'. For example, goa→ਗੋਆ

Rule 5: if last character is 'e' and second last character is 'i' then 'e' is usually skipped in that case. For example, like→ਲਾਈਕ

3.2.4 CVCC Pattern (Double Consonants)

In phonetics, gemination or consonant elongation happens when a spoken consonant is pronounced for an audibly longer period of time than a short consonant. In transliteration, double consonants in English are mapped with Punjabi gemination symbol 'Addak'. The two geminates 'mm' and 'nn' are written with 'Tippi' in many cases because they represent nasalized sound (Gill and Gleason, 1986). Proposed rules for gemination are described in Table 4.

Word Combinations	English Word	Punjabi Mapping	Punjabi Word
'a' + double consonant	bhatt	'ੱ'	ਭੱਟ
'e' + double consonant	dress	'ੈ' + 'ੱ'	ਡਰੈੱਸ
'i' + double consonant	gill	'ਿ' + 'ੱ'	ਗਿੱਲ
'o' + double consonant	boss	'ੋ'	ਬੋਸ
'u' + double consonant	full	'ੁ' + 'ੱ'	ਫੁੱਲ

Table 4. Rules for Gemination

3.2.5 CVVC Pattern (Long Vowels)

When two vowels are adjacent, they form a vowel combination. There are two vowels between consonants that create either one or two sounds. For example, word 'hair' and 'poor'

creates two sounds and word ‘four’ creates one sound. Table 5 shows mapping of various vowel combinations.

‘a’ Vowel Combinations (aa, ae, ai, ao, au)	**‘o’ Vowel Combinations (oa, oe, oi, oo, ou)**
‘ao’ → ‘ਾ’ + ‘ਓ’	‘oa’ → ‘ੋ’
(Sarao → ਸਰਾਓ)	(Road → ਰੋਡ)
‘aun’ → ‘ੌ’ + ‘ਂ’	‘ook’/’ood’ → ‘ੁ’ + ‘ੱ’
(Launch → ਲੌਂਚ)	(Took → ਟੁੱਕ)
‘au’ → ‘ੌ’	‘oo’ → ‘ੂ’
(Author → ਔਥਰ)	(Stool → ਸਟੂਲ)
‘aa’ → ‘ਾ’	‘oi’ → ‘ੋ’ + ‘ਇ’
(Taal → ਤਾਲ)	(Spoil → ਸਪੋਇਲ)
‘ai’ → ‘ੇ’	‘ou’ → ‘ੌ’
(Train → ਟਰੇਨ)	(Four → ਫੌਰ)
‘ae’ → ‘ਾ’ + ‘ਏ’	‘oe’ → ‘ੋ’ + ‘ਏ’
(Rae → ਰਾਏ)	(Poem → ਪੋਏਮ)
‘e’ Vowel Combinations (ea, ee, ei, eo, eu)	**‘u’Vowel Combinations (ua, ue, ui, uo)**
‘ei’ → ‘ੇ’	‘lue’/’rue’ at last → ‘ੂ’
(Veil → ਵੇਲ)	(True → ਟਰੂ)
‘ee’ → ‘ੀ’	‘ue’ → ‘ੀ’ + ‘ਊ’
(Mandeep → ਮਨਦੀਪ)	(Continue → ਕੋਂਟੀਨੀਊ)
‘eo’ →‘ੇ’ + ‘ਓ’	‘ua’ → ‘ਾ’
(Deol → ਦੇਓਲ)	(Guard → ਗਾਰਡ)
‘ea’ → ‘ੀ’	‘ui_e’ → ‘ਾ’ + ‘ਈ’
(Heat → ਹੀਟ)	(Guide → ਗਾਈਡ)
‘eu’ → ‘ੇ’ + ‘ਉ’	‘uit’/’uice’/ ‘uise’ → ‘ੂ’
(Heuristic→ ਹੀਉਰੀਸਟੀਕ)	(Fruit → ਫਰੂਟ)
‘i’ Vowel Combinations (ia, ie, io, iu)	
‘ia’ → ‘ੀ’ + ‘ਆ’	
(Pia → ਪੀਆ)	
‘io’ → ‘ੀ’ + ‘ਓ’	
(Jio → ਜੀਓ)	
‘ie’ at last → ‘ਾ’ + ‘ਈ’	
(Die → ਡਾਈ)	
‘ie’ → ‘ੀ’	
(Piece → ਪੀਸ)	
‘iu’ → ‘ੀ’ + ‘ਆ’	
(Celcius → ਸੈਲਸੀਅਸ)	

Table 5. Rules for CVVC Pattern

3.2.6 Silent Letters

A silent letter is a letter that, in a particular word, does not correspond to any sound in the word’s pronunciation. Some rules have been constructed for the most common silent letters- ck, wr, kn, mn, mb, mp, dg, gh, wh etc. Following are some English words examples with silent letters and their corresponding transliteration in Punjabi: Trick→ਟਰਿੱਕ, whole→ਹੋਲ, write→ਰਾਈਟ, know→ਨੋ, column→ਕੋਲਮ, company→ਕੰਪਨੀ etc.

3.2.7 CVCe Pattern

There are somewhat different rules for syllables with CVCe pattern. An ‘e’ is found at the end of the word. These words are transliterated according to the matching rules, as shown in Table 6.

Word Combinations	**English Word**	**Punjabi Mapping**	**Punjabi Word**
current_char is ‘a’ second_next_char is ‘e’	grace	‘ੇ’	ਗਰੇਸ
current_char is ‘i’ second_next_char is ‘e’	like	‘ਾ’ + ‘ਈ’	ਲਾਈਕ
current_char is ‘o’ second_next_char is ‘e’	vote	‘ੋ’	ਵੋਟ
current_char is ‘u’ second_next_char is ‘e’	tune	‘ੀ’ + ‘ਊ’	ਟੀਊਨ

Table 6. Rules for CVCe Pattern

3.2.8 Ambiguity Resolution

There are certain characters in English those correspond to more than one pronunciation. For example, the word ‘s’ can be pronounced as ‘ਸ’ or ‘ਜ’ as in case of ‘house’ or ‘resume’. Other ambiguous word is ‘g’. In order to resolve this problem following rules have been proposed:

Rule 1: if ‘g’ is followed by ‘en’/‘in’/‘ic’/‘im’ then ‘g’ is transliterated to ‘ਜ’. For example, general/ margin/ logic→ਜਨਰਲ/ ਮਾਰਜਨ/ ਲੋਜਿਕ

Rule 2: if there is ‘ange’/‘enge’ combination then ‘g’ is transliterated to ‘ਜ਼’. For example, orange/challenge→ਓਰੇਂਜ਼/ਚੈਲੇਂਜ਼

Rule 3: if 'g' is preceded by 'a'/'d' and followed by 'e' then 'g' is transliterated to 'ਜ'. For example, page/bridge→ਪੇਜ/ਬਰਿੱਜ

Rule 4: if 's' is preceded by 'ou'/'a' and followed by 'e' then 's' is transliterated to 'ਸ'. For example, mouse/base→ਮਾਊਸ/ਬੇਸ

Rule 5: otherwise if 's' is followed and preceded by vowel then it is transliterated to 'ਜ਼'. For example, visit/season→ਵਿਜ਼ੀਟ/ਸੀਜ਼ਨ

3.2.9 CVC Pattern (Short Vowels)

A short vowel word is any word that does not allow the vowel within it to generate that vowel's long vowel sound. Table 7 shows proposed rules for short vowels.

Short Vowel	English Word	Punjabi Mapping	Punjabi Word
'a'	fan, scan	'ੈ'	ਫੈਨ, ਸਕੈਨ
'e'	bed, pen	'ੈ' + 'ੱ'	ਬੈੱਡ, ਪੈੱਨ
'i'	big, swim	'ਿ' + 'ੱ'	ਬਿੱਗ, ਸਵਿੱਮ
'o'	dog, stop	'ੋ'	ਡੋਗ, ਸਟੋਪ
'u'	cut, bus	'ੱ'	ਕੱਟ, ਬੱਸ

Table 7. Rules for Short Vowels

3.2.10 Unreduced Vowel (Schwa + Consonant)

English language has certain words with unstressed syllables that are often pronounced with full vowels like care→ਕੇਅਰ, fire→ਫਾਇਰ, clear→ਕਲੀਅਰ, affair→ਅੱਫੇਅਰ, real→ਰੀਅਲ etc.

3.2.11 Schwa Reduction

The main problem in machine transliteration system is mainly due to the existence of schwa vowel that is sometimes considered and sometimes not. Figure 2 illustrates the schwa sound deletion and retention process in general.

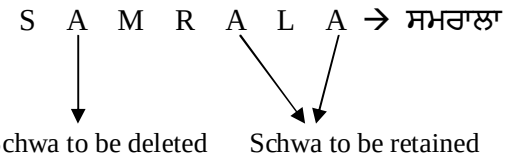

Figure 2. Process of Schwa Deletion

To find out when to retain schwa sound is not a trivial task. Some rules have been constructed to handle schwa sound:

Rule 1: if 3rd_last_char and last_char is 'a' then both are transliterated to 'ਾ'. For example, batala/samrala→ਬਟਾਲਾ/ ਸਮਰਾਲਾ

Rule 2: if 2nd_last_char is 'a' and 4th_last_char is vowel then skip 'a'. For example, sonam→ਸੋਨਮ

Rule 3: if 'a' is followed by 'rt'/ 'ph'/ 'th'/ 'rk'/ 'sk'/ 'st' then it is transliterated to 'ਾ'. For example, part/ graph/ dark/ flask/ last→ਪਾਰਟ/ਗਰਾਫ/ਡਾਰਕ/ਫਲਾਸਕ/ਲਾਸਟ

Rule 4: if 'a' is followed and preceded by same consonant then 'a' is transliterated to 'ਾ'. For example, brar→ਬਰਾੜ

4 Performance Evaluation

The evaluation of machine transliteration system can be found using Word Accuracy Rate that is used to capture the performance at word level. The word accuracy rate of proposed system is found to be 95% as shown in Table 8.

Language	Transliteration Units			Accuracy (%)
	Total	Wrong	Right	
English	10,000	495	9505	95%

Table 8. Accuracy of Transliteration System

We have manually evaluated the proposed transliteration system on news data published on a web site and the output is shown in Table 9.

Input Text
chief minister parkash singh badal had approved the agriculture power tube wells policy stipulating guide lines for the release of electric connections for agricultural pump sets in the state the punjab state electricity regulatory commission had mandated the state government to decide the number of agricultural pump connections to be released each year in the state citing reasons of ecology inadequate ground water potential declining water table among others well before the process for parliamentary elections was set in motion 593 farmers in the abohar electricity division were issued notices to deposit whopping amounts ranging between 3 lakh to get tube well connections only six of them have paid the demand notices during the recent visit of the chief minister at bahadurkhera village small farmers charan singh bhupinder singh and tejinder singh minced no words in telling him that they feel cheated since sharomani akaali dal leaders had given them an impression

during the recent elections that substantial subsidy would be made available whereas the corporation authorities had told them to bear the total cost of tube well connections
Output Text
ਚੀਫ ਮਿਨੀਸਟਰ ਪਰਕਾਸ਼ ਸਿੰਘ ਬਾਦਲ ਹੈਡ ਅੱਪਰੂਵਡ ਦ ਐਗਰੀਕੱਲਚਰ ਪਾਵਰ ਟੀਊਬ ਵੈੱਲਸ ਪੋਲਿਸੀ ਸਟੀਪੁਲੇਟਿੰਗ ਗਾਈਡ ਲਾਈਨਸ ਫੋਰ ਦ ਰੀਲੀਜ਼ ਆਫ ਇਲੈੱਕਟਰਿਕ ਕਨੈਕਸ਼ਨਜ਼ ਫੋਰ ਐਗਰੀਕੱਲਚਰਲ ਪੰਪ ਸੈੱਟਸ ਇਨ ਦ ਸਟੇਟ ਦ ਪੰਜਾਬ ਸਟੇਟ ਇਲੈੱਕਟਰੀਸਿਟੀ ਰੇਗੁਲੇਟੋਰੀ ਕੋਮਿੱਸ਼ਨ ਹੈਡ ਮਨਡੇਟਡ ਦ ਸਟੇਟ ਗਵਰਨਮੈਂਟ ਟੂ ਡੇਸਾਈਡ ਦ ਨੰਬਰ ਆਫ ਐਗਰੀਕੱਲਚਰਲ ਪੰਪ ਕਨੈਕਸ਼ਨਜ਼ ਟੂ ਬੀ ਰੀਲੀਜ਼ਡ ਈਚ ਈਅਰ ਇਨ ਦ ਸਟੇਟ ਸਾਈਟਿੰਗ ਰੀਜ਼ਨਸ ਆਫ ਏਕੋਲੋਜੀ ਇੰਨਏਡੁਕੁਏਟ ਗਰਾਊਂਡ ਵਾਟਰ ਪੋਟੈਂਸ਼ੀਅਲ ਡੇਕਲਾਈਨਿੰਗ ਵਾਟਰ ਟੇਬਲ ਅਮੋਂਗ ਅਦਰਸ ਵੈੱਲ ਬੀਫੋਰ ਦ ਪ੍ਰੋਸੈੱਸ ਫੋਰ ਪਰਲਿਆਮੈਂਟਰੀ ਇਲੈੱਕਸ਼ਨਜ਼ ਵਾਜ਼ ਸੈੱਟ ਇਨ ਮੋਸ਼ਨ 593 ਫਾਰਮਰਸ ਇਨ ਦ ਅਬੋਹਰ ਇਲੈੱਕਟਰੀਸਿਟੀ ਡੀਵੀਜ਼ਨ ਵਰ ਇੱਸ਼ੂਡ ਨੋਟੀਸਜ਼ ਟੂ ਡੀਪੋਜ਼ੀਟ ਵੋਪਿੰਗ ਅਮਾਊਂਟਸ ਰੇਂਜ਼ਿੰਗ ਬੇਟਵੀਨ 3 ਲਖ ਟੂ ਗੈੱਟ ਟੀਊਬ ਵੈੱਲ ਕਨੈਕਸ਼ਨਜ਼ ਓਨਲੀ ਸਿੱਕਸ ਆਫ ਦੈੱਮ ਹੈਵ ਪੇਡ ਦ ਡੇਮਾਂਡ ਨੋਟੀਸਜ਼ ਡੀਊਰਿੰਗ ਦ ਰੀਸੈਂਟ ਵਿਜ਼ੀਟ ਆਫ ਦ ਚੀਫ ਮਿਨੀਸਟਰ ਐਟ ਬਹਦੁਰਖੇਰਾ ਵਿੱਲੇਜ ਸਮਾਲ ਫਾਰਮਰਸ ਚਰਨ ਸਿੰਘ ਭੁਪਿੰਦਰ ਸਿੰਘ ਐਂਡ ਤੇਜਿੰਦਰ ਸਿੰਘ ਮਿੰਸਡ ਨੇ ਵਰਡਸ ਇਨ ਟੈੱਲਿੰਗ ਹਿੱਮ ਦੈਟ ਦੇ ਫੀਲ ਚੀਟਡ ਸਿੰਸ ਸ਼ਰੋਮਨੀ ਅਕਾਲੀ ਦਲ ਲੀਡਰਸ ਹੈਡ ਗਿਵਨ ਦੈੱਮ ਐਨ ਇੰਪ੍ਰੈੱਸ਼ਨ ਡੀਊਰਿੰਗ ਦ ਰੀਸੈਂਟ ਇਲੈੱਕਸ਼ਨਜ਼ ਦੈਟ ਸੁਬਸਟੈਂਸ਼ੀਅਲ ਸਬਸਿਡੀ ਵੁੱਡ ਬੀ ਮੇਡ ਅਵੇਲੇਬਲ ਵੇਅਰਐਜ਼ ਦ ਕੋਰਪੋਰੇਸ਼ਨ ਔਥੋਰੀਟੀਜ਼ ਹੈਡ ਟੋਲਡ ਦੈੱਮ ਟੂ ਬੀਅਰ ਦ ਟੋਟਲ ਕੋਸਟ ਆਫ ਟੀਊਬ ਵੈੱਲ ਕਨੈਕਸ਼ਨਜ਼

Table 9. Result of Proposed System

5 Conclusion and Future Scope

In this paper we have described transliteration system for English-Punjabi language pair using a rule based approach. There are certain syllable patters in English such as CVC, CVVC, CVCC, CVCe etc. Only a single rule is not applicable to all these patterns so different rules have been proposed for the transliteration of such syllable patterns. Also various rules have been proposed to handle ambiguous words, unstressed vowels and schwa sound. The proposed system is also applicable for transliteration of some person names and locations.

The existing work done in English to Punjabi transliteration emphasis on transliteration of named entities. In our work we have considered all possible English words for transliteration.

Still schwa sound deletion and multiple representation of same word need to be more explored. There are many English words that are spelled the same but pronounced differently. To resolve this issue, the future work can be based on referring to the context of the words in order to choose correct transliteration. We can say that accuracy of proposed system is depends on the rules framed so it can be furthered improved by adding new rules when identified.

References

Bhalla, D. and Joshi, N. (2013), "*Rule Based Transliteration Scheme For English To Punjabi*", International Journal on Natural Language Computing, Vol. 2, No. 2, pp. 67-73.

Deep, K. and Goyal, V. (2011), "*Development of a Punjabi to English Transliteration System*", International Journal of Computer Science and Communication, Vol. 2, No. 2, pp. 521-526.

Dhore, M., Dixit, S. and Dhore, R. (2012), "*Hindi and Marathi to English NE Transliteration Tool using Phonology and Stress Analysis*", in proceedings of 24th International Conference on Computational Linguistic, Mumbai, India, pp. 111-118.

Gill, H. and Gleason, H. (1986), "*A Reference Grammar of Punjabi*", Publication Bureau, Punjabi University, Patiala, India, pp. 38-42.

Goyal, V. and Lehal, G. (2009), "*Hindi-Punjabi Machine Transliteration System (For Machine Translation System)*", George Ronchi Foundation Journal, Italy, Vol. 64, No. 1, pp. 1-7.

Josan, G. and Kaur, J. (2011), "*Punjabi To Hindi Statistical Machine Transliteration*", International Journal of Information Technology and Knowledge Management, Vol. 4, No. 2, pp. 459-463.

Josan, G. and Lehal, G. (2010), "*A Punjabi to Hindi Machine Transliteration System*", Computational Linguistics and Chinese Language Processing, Vol. 15, No. 2, pp. 77-102.

Joshi, H., Bhatt, A. and Patel, H. (2013), "*Transliterated Search using Syllabification Approach*", Forum for Information Retrieval Evaluation, Delhi, India.

Kaur, J. and Josan, G. (2011), "*Statistical Approach to Transliteration from English to Punjabi*", International Journal on Computer Science and Engineering, Vol. 3, No. 4, pp. 1518-1527.

Knight, K. and Graehl, J. (1998), "*Machine transliteration*", in proceedings of the 35th annual meetings of the Association for Computational Linguistics, Madrin, Spain, pp. 128-135.

Malik, M. (2006), "*Punjabi Machine Transliteration*", in proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL, Sydney, pp. 1137-1144.

Oh, J. and Choi, K. (2002), "*An English-Korean Transliteration Model Using Pronunciation and Contextual Rules*", in proceedings of the 19th International Conference on Computational Linguistics, Taipei, Taiwan, pp. 758-764.

Saini, T. and Lehal, G. (2008), "*Shahmukhi to Gurmukhi Transliteration System: A Corpus based Approach*", Research in Computing Science (Mexico), Vol. 33, pp. 151-162.