

Hindi to Punjabi Transliteration using Phonetic and Orthographic Rules

Kamaljeet Kaur

Department of Computer Science and Engineering
Guru Nanak Dev Engineering College,
Ludhiana (Punjab) India

Parminder Singh, Ph.D

Associate Professor and Head,
Department of Computer Science and Engineering
Guru Nanak Dev Engineering College,
Ludhiana (Punjab) India

ABSTRACT

One of the important applications of Natural Language Processing is machine translation. Machine transliteration is an emerging and a very important research area in the field of machine translation. Translation systems translate message from source language to target language, keeping the exact meaning. While the transliteration system finds the same meaning word/sentence in another language, the transliteration helps us to pronounce them. This paper addresses the transliteration process from Hindi to Punjabi language using a rule based approach. Transliteration is more phonetic process, thus both source grapheme and its sound in the correct pronunciation of word have been considered for rule formation to achieve high performance and more accurate results. The system is tested on news domain text and achieved accuracy of 97%.

Keywords

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

1. INTRODUCTION

Transliteration is the conversion of a text from one script to another and thus a process of representing words from one language using the approximate phonetic or spelling equivalents of another language. Machine transliteration can play an important role in natural language processing applications, especially for handling proper nouns, technical terms and out of vocabulary words. 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. Hence, words and their underlying structures are supposed to change as per grammatical rules of target language, but the meaning must remain unchanged. Transliteration is different from translation. In other sense while translation tells the meaning of the word in another language, transliteration helps to pronounce them. Usually, transliteration is used to obtain phonetic equivalents in a target language for a given source language word. Transliteration is meant to preserve the sounds of the syllables in words. Transliteration is helpful in situations where one does not know the script of a language but knows to speak and understand the language nevertheless. This paper presents the rule based approach adopted to develop a Hindi to Punjabi transliteration system that can transliterate Hindi text into equivalent Punjabi text. Transliteration is more phonetic than orthographic so transliteration occurs on the basis of pronunciation. Thus we have considered both orthographic and phonetic information.

The remainder of this paper is organized as follows. Section 2 describes the related work done in machine transliteration. Section 3 introduces about the approach followed for transliteration. We described basic character to character mapping and rules for transliteration in sections 4. Experimental Results are discussed in section 5. Finally, we have concluded it in section 6.

2. RELATED WORK

Significant work in the field of machine transliteration has been done for Indian as well as foreign languages. Three approaches for transliteration being used are: Grapheme based models; Phoneme based models and Hybrid models [10]. 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 [9]. Lee and Choi (1998) have presented a statistical transliteration model for the transliteration from English to Korean. They have compared two SMT based methods i.e. direct method and pivot method. After comparison they have proposed a hybrid method that is more effective for transliteration [11]. 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 [13]. Malik (2006) has developed Punjabi Machine Transliteration System (PMT) that is used to transliterate words from Shahmukhi script to Gurmukhi script. The Punjabi Machine Transliteration System uses transliteration rules (character mappings and dependency rules) for transliteration of Shahmukhi words into Gurmukhi [12]. 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 [14]. 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. The system has been tested on various domains like proper names, city names, country names, river names etc. and gives accuracy of about 98% [4]. 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. There are some characters in Hindi for which no rule applies. For generating these characters, the Soundex technique has been employed. Soundex is a phonetic matching technique [6]. 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% [8]. Deep and Goyal (2011) have proposed a transliteration system that addresses the process of forward transliteration of person names from Punjabi to English by set of character mapping rules using grapheme based method and achieved accuracy of 93.22%. The system is accurate for the Punjabi words but not for the foreign words [2]. Josan and Kaur (2011) have developed a statistical model that is used for transliterating Punjabi text into Hindi text. The statistical method has achieved 87.72% accuracy rate [5]. Dhore

et al. (2012) have focused on the specific problem of machine transliteration of Hindi to English and Marathi to English using a phonetic based direct approach. They have developed a rule based phonetic model using linguistic approach. Total 15,244 named entities are tested and the accuracy of the proposed system is 74.14% [3]. Joshi et al. (2013) have proposed system that can do transliteration from Roman script to Devanagari script. Statistical machine learning approach was used for transliteration. The syllable theory has been used as basis for transliteration [7]. Vijaya et al. (2009) have developed English to Tamil Transliteration system and named it WEKA. They have demonstrated a transliteration model based on multi class classification approach for English to Tamil transliteration. It is a Rule based system. The transliteration model produced an exact transliteration in Tamil from English words with an accuracy of 84.82% [15]. Bhalla et al. (2013) have proposed rule based transliteration scheme for English to Punjabi. Some rules have been constructed for syllabification which is the process to extract or separate the syllable from the words. In this probabilities are calculated for name entities (proper names and location). The proposed approach has attained accuracy of 88.19% [1].

3. METHODOLOGY

The proposed system for Hindi to Punjabi transliteration follows rule based approach. Some rules are made depending upon the pattern of the sentences that can be given as an input. These rules are applied to the input sentence and the sentence gets transliterated according to that matching rule.

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. That is, Hindi character(s) is mapped to Punjabi character(s) via the spoken form associated with the word.

Phonetic transliteration also works according to the transliteration rules. First of all Hindi words are segmented into Hindi transliteration units. Figure 1 shows the word ‘कमलजीत’ in source language is segmented into source language transliterated units ‘क’, ‘म’, ‘ल’, ‘ज’, ‘ी’ and ‘त’.

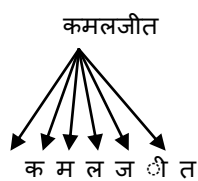


Figure 1. Segmentation of Source Language Word

After the segmentation of source word, transliteration rules are applied to various character combinations. First of all each character of Hindi language is mapped to the character of Punjabi language according to the matching rules. Direct mapping is applied to rest of the characters. Surrounding characters are also considered for more appropriate transliteration. Figure 2 shows transliteration of Hindi transliterated units to Punjabi.

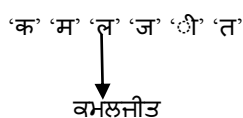


Figure 2. Transliteration of Source Language Transliterated Units

4. DESIGN AND IMPLEMENTATION

4.1 Hindi-Punjabi Character to Character Mapping Rules

Because Hindi and Punjabi are related languages due to lots of similarities in syntax and vocabulary, the sound of most of the letters is same in both scripts. Table 1 show that most of the letters have one to one correspondence, so this table also forms the base for direct mapping [4].

Vowels in Hindi also have one to one correspondence with Punjabi vowel symbols. The vowel diacritics are used to indicate that a vowel is attached to a consonant. The full vowel form is used when the vowel occurs alone, at the beginning of a word, or after another vowel. In other words, the full vowel form is used whenever there is no consonant for the vowel to attach to. The corresponding vowel mapping in Hindi to Punjabi is shown in Table 2 and 3.

Table 1. Hindi-Punjabi Consonant Mapping

क	ख	ग	घ	ङ
क्	ਖ	ਗ	ਘ	ਙ
च	छ	ज	झ	ञ
च्	ਛ	ਜ	ਝ	ਞ
ट	ठ	ड	ढ	ण
ट्	ਠ	ਡ	ਢ	ਣ
त	थ	द	ध	न
त्	ਥ	ਦ	ਧ	ਨ
प	फ	ब	भ	म
प्	ਫ	ਬ	ਭ	ਮ
य	र	ल	व	इ
ਯ	ਰ	ਲ	ਵ	ੜ
श	ख	ग	ज	फ़
ਸ਼	ਖ	ਗ	ਜ	ਫ਼
ळ	स	ह		
ਲ਼	ਸ	ਹ		

Table 2. Hindi-Punjabi Full Vowels Mapping

अ	आ	इ	ई	उ
ਅ	ਆ	ਇ	ਈ	ਉ
ऊ	ए	ऐ	ओ	औ
ਊ	ਏ	ਐ	ਓ	ਔ

Table 3. Hindi-Punjabi Vowel Diacritics Mapping

ा	ि	ी	ु	ू
ा	ि	ी	ु	ू
े	ै	ो	ौ	
े	ै	े	ै	

The major problem in machine transliteration system is ambiguity. Direct transliteration or direct character-to-character mapping does not provide any solution for ambiguity. There is need to develop rules for such types of problems and for more accurate transliteration.

4.2 Rules for Ambiguity Resolution

The critical problem with transliteration system is ambiguity that refers to more than one transliteration of source word. Ambiguity is explained as the problem that a source language word can have more than one possible transliteration. Hindi words may have more than one representation based on the perception of the transliterator. For example Hindi word 'प्यास' can be represented in Punjabi as 'ਪਿਆਸ', 'ਪੀਆਸ', 'ਪਿਯਾਸ', 'ਪਯਾਸ'. Only first representation, when again transliterated back into Hindi, converts to the correct representation. Rules for resolution of ambiguity are based on the character sequences. There is need to refer to the surrounding characters for correct transliteration. Most of the ambiguous words are resulted due to the presence of Hindi consonant 'य'. The following rules are proposed to handle this problem.

Rule 1: if 'य' is at first position then it is transliterated to 'ਯ'.

For example, यादव→ਯਾਦਵ

Rule 2: if 'यो'/'ये'/'या'/'यी' are at last position then they are transliterated to 'ਓ'/'ਏ'/'ਆ'/'ਈ' respectively. For example, कोरीयो/ कारवायी→ਕੋਰੀਓ/ ਕਾਰਵਾਈ

Rule 3: if 'या' is preceded by half character then 'या' is transliterated to 'ਿ'+ਆ'. For example, प्यास→ਪਿਆਸ

Rule 4: if 'य' is followed by 'ी'/'ा'/'े'/'ै' then 'य' is transliterated to 'ਇ'. For example, सोसायटी/ सहायता/ दायरे/ डायरेक्टर→ਸੋਸਾਇਟੀ/ ਸਹਾਇਤਾ/ ਦਾਇਰੇ/ ਡਾਇਰੈਕਟਰ

Rule 5: if 'य' is followed by 'ि' then 'य' is transliterated to 'ਈ'. For example, रासायनिक→ਰਾਸਾਈਨਿਕ

Rule 6: if 'या' is preceded by 'ा' then 'या' is transliterated to 'ਇ'+ਆ'. For example, लगाया→ਲਗਾਇਆ

Rule 7: if 'या' is preceded by 'ि'/'ी' then 'ि'/'ी' and 'या' are transliterated to 'ੀ'+ਆ'. For example, रीया→ਰੀਆ

Rule 8: if 'यू' is preceded by half character then 'यू' is transliterated to 'ਿ'+ਊ'. For example, न्यू→ਨਿਊ

Rule 9: if 'यु' is preceded by half character then 'यु' is transliterated to 'ਿ'+ਊ'. For example, व्यु→ਵਿਊ

Rule 10: if 'ण' is preceded by 'य' then 'य' is transliterated to 'ਇ'. For example, नरायण→ਨਰਾਇਣ

Rule 11: if 'यो' is preceded by half character then 'यो' is transliterated to 'ਿ'+ਓ'. For example, कयोर→ਕਿਓਰ

Rule 12: if second and last character is 'या' and previous is consonant then 'या' is transliterated to 'ਿ'+ਆ'. For example, गया→ਗਿਆ

Rule 13: if there is 'य', previous character is 'ा'/'ो' and second previous character is consonant then 'य' is transliterated to 'ਇ'. For example, पायल/कोयला→ਪਾਇਲ/ਕੋਇਲਾ

Rule 14: if there is 'य', previous character is 'ि' and second previous character is half character then 'य' is transliterated to 'ਅ'. For example, प्रियंका→ਪ੍ਰਿਅੰਕਾ

Rule 15: if 'य' is preceded by half character then 'य' is transliterated to 'ਅ'. For example, व्यस्त→ਵਿਅਸਤ

4.3 Rules for Middle Vowel Character in Word

In Hindi vowels are mapped to different characters based on their position in the word. The surrounding characters also affect their mapping in target language. Examples of such vowels are 'ि', 'े', 'ै' and 'ं'. Hindi vowel 'ि' is not always mapped to 'ਿ'. Sometimes it is mapped to 'ੀ' according to its position and surrounding characters in a given word.

Rule 1: if 'ह' is preceded by 'े' then 'े' is transliterated to 'ਿ'. For example, सेहत→ਸਿਹਤ

Rule 2: if 'या' is preceded by 'ै' then 'ै' is transliterated to 'ਿ'. For example, तैयार→ਤਿਆਰ

Rule 3: if 'ह' is followed by 'ू' then 'ह' is transliterated to 'ਹਿ'. For example, महफूज→ਮਹਿਫੂਜ

Rule 4: if 'ि' is followed and preceded by 'ा' then 'ि' is transliterated to 'ੀ'. For example, जातिवाद→ਜਾਤੀਵਾਦ

Rule 5: if current and previous character is 'ि' then current 'ਿ' is transliterated to 'ੀ'. For example, नितिन→ਨਿਤੀਨ

Rule 6: if 'ि' is preceded by half character then 'ि' is transliterated to 'ੀ'. For example, अग्निपथ→ਅਗਨੀਪਥ

Rule 7: if ‘ि’ is at second last position and ‘या’ is at last then ‘ि’ is transliterated to ‘ी’. For example, एशिया→ऐशीआ

Rule 8: if ‘ि’ is at last then it is transliterated to ‘ी’. For example, स्थिति→सथिती

4.4 Rules for Half Words

Unicode does not encode half or subjoined letters for the Devanagari scripts. Unicode forms all consonant clusters (such as the “ल” in “पत्नी”) by inserting the character “Virama” or “Halant” (ँ) between the two relevant consonant letters. There is no special character present in Hindi for duplication of sound of a consonant. In Hindi sound duplication is represented by half character form consonants. But in Punjabi ‘Adhak’ is used to duplicate the sound of a consonant. The mapping of half character in target language is greatly depends on the next character combination.

Rule 1: if half ‘क’ is followed by ‘ख’ then it is transliterated to ‘ँ’. For example, मक्खी→मँकी

Rule 2: if half ‘ग’ is followed by ‘घ’ then it is transliterated to ‘ँ’. For example, मग्घ→मँघ

Rule 3: if half ‘च’ is followed by ‘छ’ then it is transliterated to ‘ँ’. For example, इच्छा→इँछा

Rule 4: if half ‘ज’ is followed by ‘झ’ then it is transliterated to ‘ँ’. For example, मज्झ→मँझ

Rule 5: if half ‘ट’ is followed by ‘ठ’ then it is transliterated to ‘ँ’. For example, मट्ठ→मँठ

Rule 6: if half ‘त’ is followed by ‘थ’ then it is transliterated to ‘ँ’. For example, पत्थर→पँथर

Rule 7: if half ‘द’ is followed by ‘ध’ then it is transliterated to ‘ँ’. For example, बुद्ध→बुँध

Rule 8: if half ‘प’ is followed by ‘फ’ then it is transliterated to ‘ँ’. For example, रप्फ→रँफ

Rule 9: if half ‘ब’ is followed by ‘भ’ then it is transliterated to ‘ँ’. For example, लब्भ→लँभ

Rule 10: if half character is followed by full character and both characters are same consonants then half character is transliterated to ‘ँ’. For example, मुहब्बत→मुहँबत

4.5 Rules for Nasalized Characters

Nasalization in Hindi is presented by anusvara, chandrabinu and some nasalized consonants, as shown in Figure 3. ‘Anusvara’ and ‘chandrabinu’ are written over the horizontal line of vowel or consonant it is attached to. Commonly used nasal character in Hindi is ‘Anusvara’ (ँ). In the Devanagari script, ‘Anusvara’ is represented with a dot (bindu) above the

letter e.g. ‘मँ’. Gurmukhi has two separate nasal characters, ‘Bindi’ (ँ) and ‘Tippi’ (ँ).

ँ ँ म न ण
Anusvara Chandrabindu

Figure 3. Hindi Nasalized Characters

Nasal character ‘Anusvara’ of Hindi is mapped to either ‘Bindi’ or ‘Tippi’ in Punjabi according to the matching rules.

Rule 1: if ‘ँ’ is not preceded by ‘ै’/‘े’/‘ा’/‘ी’/‘ो’/‘ौ’ then ‘ँ’ is transliterated to ‘ँ’. For example, मंत्री→मँत्री

Rule 2: if ‘ँ’ is at last position in a given word then it is transliterated to ‘ँ’. For example, दायरों→दायँरें

Rule 3: if ‘ँ’ is preceded by ‘ै’/‘े’/‘ा’/‘ी’/‘ो’/‘ौ’ then ‘ँ’ is transliterated to ‘ँ’. For example, कांग्रेस→काँगरेस

Rule 4: if ‘ँ’ is preceded by ‘ु’/‘ू’/‘consonant’ then ‘ँ’ is transliterated to ‘ँ’. For example, मुकुंद→मुकुँद

In some words containing long vowels (e.g. ‘आ’ and ‘ऊ’), the nasalization dot is accompanied by a small moon (ँ). This sign is called ‘Chandrabinu’. It usually means that the previous vowel is nasalized. It is not a part of standard Hindi. ‘Chandrabinu’ is also mapped to either Punjabi ‘Bindi’ or ‘Tippi’ according to following matching rules.

Rule 1: if ‘ँ’ is preceded by ‘ु’/‘ू’/‘consonant’ then ‘ँ’ is transliterated to ‘ँ’ otherwise it is transliterated to ‘ँ’. For example, हूँ→हुँ

Rule 2: if ‘ँ’ is at last then it is transliterated to ‘ँ’. For example, कहाँ→कहाँ

Commonly used nasalization consonants in Hindi are ‘म’, ‘न’ and ‘ण’. When written with vowels or half character form then they are either mapped to ‘Bindi’ or ‘Tippi’ in Punjabi. These characters are transliterated according to following matching rules.

Rule 1: if half ‘म’/ half ‘न’ are preceded by ‘ा’/‘ो’/‘ौ’/‘े’/‘ै’/‘ी’ then half ‘म’/ half ‘न’ are transliterated to ‘ँ’. For example, फ्रान्स→फँरंस

Rule 2: if there is half ‘म’/ half ‘न’/ half ‘ण’ then it is transliterated to ‘ँ’. For example, पन्ना/पण्डित→पँना/पँडित

4.6 Rules for Common Conjunct Consonants

Use of conjunct consonants is limited in Punjabi. Three consonants are used in Punjabi as conjuncts. The three conjunct consonants are called *paireen* ‘haahaa’, ‘raaraa’ and ‘vaavaa’, as shown in Table 4. Conjuncts are created using a Devanagari

like 'Halant'. Thus, if you wish to type 'पू' you would enter ('प'), Halant ('ँ') and finally ('र').

Table 4. Gurmukhi Conjunct Consonants

Hindi Word	Punjabi Conjunct Consonants	Punjabi Mapping	Punjabi Word
बुल्हड़	਼	ਲ+੍+ਹ	ਬੁਲੁੜ
स्वर्ग	ँ	ਸ+੍+ਵ	ਸ੍ਵਰਗ
प्रेम	ँ	ਪ+੍+ਰ	ਪ੍ਰੇਮ

There are conjunct consonants in Hindi that have no direct mapping in Punjabi. These letters will never be mapped in Hindi to Punjabi transliteration using a direct mapping approach. They are mapped according to the rules. A conjunct consonant is a combination of two or more consonants which are pronounced together without the pronunciation of the inherent 'a' vowel between them.

The groups of conjunct consonants in Hindi are forms that are quite distinct from the consonants they are formed from. Table 5 shows common conjuncts used in Hindi and their corresponding Punjabi mapping.

Table 5. Hindi Conjunct Consonants Mapping

Hindi Characters to form Conjunct	Hindi Conjunct Consonants	Corresponding Punjabi Characters
क+्+ष	क्ष	ਕ+ਸ਼
ज+्+ञ	ज्ञ	ਗਿ+ਅ
क+्+त	क्त	ਕ+ਤ
त+्+र	त्र	ਤ+ਰ
द+्+व	द्व	ਦ+ਵ
श+्+र	श्र	ਸ਼+ਰ
द+्+द	द्	ੱ+ਦ
द+्+ध	द्ध	ੱ+ਧ

5. PERFORMANCE EVALUATION

The evaluation of machine transliteration system can be found using Word Accuracy Rate that is the percentage of correct transliteration from the total generated transliterations by the system. It is used to compute the performance at word level. The system is successfully tested on news domain of more than 10,000 words and other domains like names of persons, locations, institutes etc. Table 6 shows the result obtained. The word accuracy rate for Hindi to Punjabi transliteration is found to be 97%.

Table 6. Result of Proposed System on News Domain

Input Text
पटना: राष्ट्रीय जनता दल के सुप्रीमो लालू यादव ने बिहार

के मुख्यमंत्री नीतीश कुमार के खिलाफ विवादित बयान देते हुए कहा कि मैंने ही नीतीश कुमार को पैदा किया था। लालू ने कहा कि मुझे पता होता कि नीतीश बबूल का पेड़ निकलेंगे तो मैं उसे उसी समय गर्म पानी से जला देता। गौरतलब है कि थोड़े दिन पहले बिहार के लोगों के डीएनए में ही जातिवाद का बयान देकर भारतीय जनता पार्टी (भाजपा) के पूर्व राष्ट्रीय अध्यक्ष नितिन गडकरी विवाद खड़ा कर दिया था। राज्य में चुनाव प्रचार करने आये गडकरी ने यहां भाजपा के प्रदेश कार्यालय में संवादाता सम्मेलन के दौरान कहा था कि बिहार के लोगों के डीएनए में जातिवाद है इसलिए यहां कुछ लोग जातिवाद की ज्यादा बात करते हैं।

Output Text

पटना: रासटरी जनता दल के सुप्रीमो लालू यादव ने बिहार के मुखमंडरी नीतीश कुमार के खिलाफ विवादित बयान देते हुए कहा कि मैंने ही नीतीश कुमार के पैदा कीया था। लालू ने कहा कि मुझे पता होता कि नीतीश बबूल का पेड़ निकलेंगे तो मैं उसे उसी समय गर्म पानी से जला देता। गौरतलब है कि थोड़े दिन पहले बिहार के लोगों के डीएनए में ही जातिवाद का बयान देकर भारतीय जनता पार्टी (भाजपा) के पूर्व राष्ट्रीय अध्यक्ष नितिन गडकरी विवाद खड़ा कर दिया था। राज्य में चुनाव प्रचार करने आये गडकरी ने यहां भाजपा के प्रदेश कार्यालय में संवादाता सम्मेलन के दौरान कहा था कि बिहार के लोगों के डीएनए में जातिवाद है इसलिए यहां कुछ लोग जातिवाद की ज्यादा बात करते हैं।

6. CONCLUSION

This paper describes the approach adopted to develop the Hindi to Punjabi transliteration system using a rule based approach. Besides direct character-to-character mapping, various rules have also been devised for transliteration in order to resolve various ambiguities and thus to obtain more accurate transliteration. Both orthographic and phonetic information of Hindi language is considered. Phonetic transliteration is followed to generate Punjabi equivalent of Hindi that works according to the transliteration rules.

Rules are also constructed according to the position of the vowels in the word. Surrounding characters in a given word are also considered for more accurate transliteration. In Hindi language, most of the words are written with half character form and conjunct consonants that have no direct mapping in Punjabi. These characters are mapped to equivalent Punjabi characters according to the proposed rules. The system is tested on the news domain and other domains and it is providing promising results.

7. REFERENCES

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] Joshi, H., Bhatt, A. and Patel, H. (2013), “*Transliterated Search using Syllabification Approach*”, Forum for Information Retrieval Evaluation, Delhi, India.
- [8] 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.
- [9] 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.
- [10] Kaur, K. and Singh, P. (2014), “*Review of Machine Transliteration Techniques*”, International Journal of Computer Applications, Vol. 107, No. 20, pp. 13-16.
- [11] Lee, J. and Choi, K. (1998), “*English to Korean Statistical transliteration for information retrieval*”, Computer Processing of Oriental Languages, Vol. 12, pp. 17-37.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] Vijaya, M.S., Ajith, V.P., Shivapratap, G. and Soman, K.P. (2009), “*English to Tamil Transliteration using WEKA*”, International Journal of Recent Trends in Engineering, Vol. 1, No. 1, pp. 498-500.