

# A iRhyme Rephraser

## Team PlayList :

Rishabh Sharma-[[rishabhr@usc.edu](mailto:rishabhr@usc.edu)][7628846249]

Akanksha Singh-[[akanksha@usc.edu](mailto:akanksha@usc.edu)][5250661320]

Tanvi Juneja-[[tjuneja@usc.edu](mailto:tjuneja@usc.edu)][8355373007]

Manas Mahanta-[[mmahanta@usc.edu](mailto:mmahanta@usc.edu)][7523347437]

## Motivation:-

We introduce “A iRhyme Rephraser” as a powerful tool for the users to experience the original Hindi poems in English while preserving the poetic effect. Many services exist that assist the users in knowing the meaning of the poem but doesn't tackle the major issue of paraphrasing and maintaining the rhyme, semantics and syntactic structure of the translated text. Research papers that address a major subset of this problem includes CRhyme and Hindi-English Translation system-IIT Bombay. CRhyme transforms English input sentences to its corresponding poem, by producing syntactically correct stanzas, but does so with altered meaning. English, being spoken worldwide and myriad collections of poems being in Hindi, but only limited to the Indian subcontinent, makes the poem translation task interesting. It is a challenging task to translate Hindi poem lines to its corresponding English text using google translate API as poems do not follow the standard SVO structure. Another uphill task is to impose rhyming constraints on the translated text and reorder phrases without affecting the semantics and coherence.

## Data Sources:-

The sources for Hindi poems includes various sites like geeta-kavita.com, lyricsindia.net, 10lyrics.com and many more which we will crawl and extract poems. We also employ CMU rhyming dictionary that consist of a list of all English words each mapped to its corresponding pronunciation.

## Procedure:-

We perform rule based Part of Speech Tagging for Hindi words and translate Hindi sentences to its respective English Words based on Phrase Based Statistical Machine Translation (PBSMT), used for chunking phrases and Google Translate API. We then divide each English sentence into sub phrases based on Rule Based approach preserving the syntax in each sub phrase. Some similar examples can be “He is walking hastily | towards his house” or “Johny is coming | by bus”. A rule of splitting for the above examples might be to split between the words if the current word is preposition given that the previous word is a verb or an adverb.

Consider the translated English text that is not rhyming,

“He was **going through** a bad state  
For she is mine **forever I** am fighting”

By rule based approach, we identified 4 phrases “Through a bad phase”, “he was going” in the 1<sup>st</sup> sentence. “I am fighting” and “for she is mine forever” in the 2<sup>nd</sup> sentence. The last words of each of the phrases are significant i.e. “going”, “state” in the first sentence “fighting” and “forever” in the 2<sup>nd</sup> sentence. Now, we enlist the possible meanings of each word in a {1: n} relationship fashion while preserving the **POS property** for each of the mapped words.

Going: {moving, advancing, progressing, passing, proceeding},

State: {stage, period, chapter, episode, part, step, point, time, juncture, aspect, days, phase},

Fighting: {quarreling, bickering, squabbling, differing, disagreeing, brawling, contending, opposing}  
forever:{always, incessantly, continually, limitless,}.

“**phase**” sounds like “FEY1 Z” in the CMU Pronunciation dictionary and “**always**” as “AO1 L W EY2 Z” and these 2 words have the same suffix.

The best rhyming word pairs are found by employing the CMU Pronunciation dictionary and swapping appropriate phrases, we get

“He was **going through** a bad **phase**  
**I** am fighting for she is mine **always** “

Evaluation Plan:

- Firstly, we will show that our phrase based input feeding into Google translate API is better than line-by-line translation for poetic lines. This is mainly because Google Translate API can translate Hindi sentences in SOV format to English but poem does not follow the SOV paradigm as a result of which the translation result is error prone.
- For showing the efficacy of our poem generator, we will use the table given below as meter to evaluate the coherence and the poetic effect of the translated poems.

R-Meter	Meaning
0	syntactically incorrect ,rhyme and meaning not preserved
1	syntactically correct , rhyme and meaning not preserved
2	syntactically incorrect, rhyming is preserved, meaning not preserved
3	syntactically correct, rhyming is preserved, remote relationship
4	syntactically correct, rhyming is preserved, meaningful

**WORD\_COUNT:-650**

**Division of Labor:-**

Part of Speech Tagging in Hindi: Akanksha Singh,Tanvi Juneja

Hindi to English Translation using PBSMT and Google Translate API: Manas Mahanta,Rishabh Sharma

Phrase based Chunking of English sentences by syntactic parsing: Tanvi, Juneja,Rishabh Sharma

Evaluation of generated English poems: Manas Mahanta,Akanksha Singh

**References:-**

cRhyme: A Computer Engine for Generating Rhyming Sentences:

<http://nlp.stanford.edu/courses/cs224n/2013/reports/shotan.pdf>

The IIT Bombay Hindi⇔English Translation System at WMT 2014:

<https://www.cse.iitb.ac.in/~pb/papers/wmt14-eng-hin.pdf>

Mining Hindi-English Transliteration Pairs from Online Hindi Lyrics:

[http://www.lrec-conf.org/proceedings/lrec2012/pdf/365\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/365_Paper.pdf)