Framework to Evaluate the Severity of Spelling Mistake during Answer Evaluation Process

Amit S. Walters
PhD Research Fellow
Department of CS & IT
SHIATS, Allahabad

Rajendra K. Isaac
Professor, Faculty of
Engineering & Technology
SHIATS, Allahabad

ABSTRACT
In this paper titled "Framework to evaluate the severity of spelling mistake during answer evaluation process", discusses a framework for various types of spelling errors occurs during writing a sentence and the severity of the mistake. The examples included in this paper are based on the analysis of writing samples collected for the answer evaluation and other researchers discussed on their papers. Procedures used in the framework are based on the algorithm for matching the word with list of known word and reassemble word for best match word. Framework is designed to analyze and evaluate the severity of mistake.

Keywords
Spelling errors, Causes of spelling errors, Type of spelling error, severity of spelling error.

1. INTRODUCTION
To spell a word involves different mental processes such as the correct association of spoken sounds (phonemes) and written symbols (graphemes). The correct association of sound (phonemes) and writing (graphemes) involves semantic, phonological, morphological and orthographic skills. Cook (1999) suggested in the Fig. 1.1reflecting the proportion of spelling mistakes of L2 learners with different L1, based on spelling literature, such as Brooks, Gorman & Kendall (1993) for L1 children and Bebout (1985) for L2 adults.

It can be observed with the help of proportion chart Fig.1 given by cook (1999) that 59% of mistakes are omissions and additions, with the higher proportion of omissions and additions of vowels than consonants at 37% and 22% respectively. The second highest is substitution with only 30%, where there is no big difference between vowels and consonants (18% and 12% respectively). Transposition (reversal) and sound based spelling mistakes are much lower and account to 5% and 6% respectively.

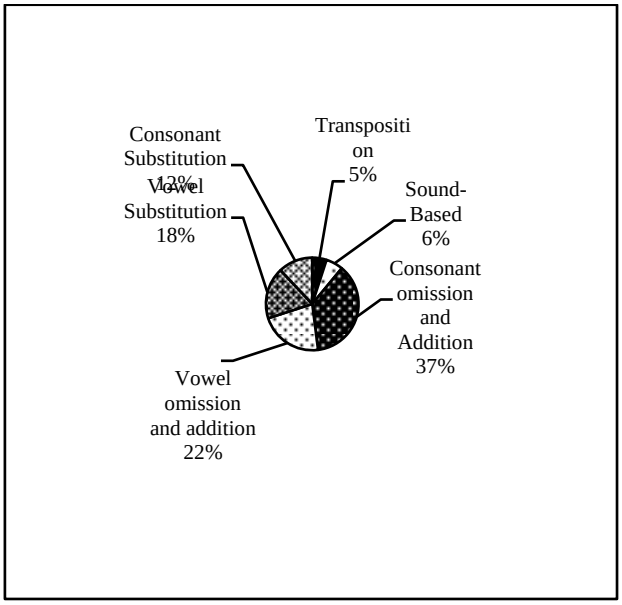

Fig.1 Proportions of spelling mistakes (Cook: 1999)

Further, relationship between pronunciation of word and writing correct spelling Sailaja (2009) observes that a number of English words (often mispronounced) are merely heard, thus it is common to see misspellings across the country, but all the mistakes can't be evaluated on the same scale as there are various resinous of spelling mistakes. Framework is designed to analyze and evaluate the severity of mistake.

2. LITERATURE REVIEW
On reviewing the past research literature it was observed that it is difficult to put all spelling errors into neatly defined categories. Ott (2007) has discussed that misspelling analysis is time-consuming and often reliant on judgment and it is not exact science–for different types of errors different explanations can be given. The error can be categorized effectively if we know the cause of errors. Krishnamurthy (1978) and Sailaja (2009) have talked about the errors of spelling pronunciation of English word in Indian, which in turn influences the spellings of the learners. Many words like "lovely", "dedicated", "irregular", "nicely" , etc. are very common to hear people, sometimes, don't ever look up the spellings of the words they hear in their environment. Other common types of spelling errors are the large number of homophonous words in the English language with entirely different spelling. There are many other known and unknown reasons for the occurrence of spelling errors in English. We discuss some them in our framework with the category of spelling errors.

3. FRAMEWORK

The proposed framework Fig. 2 analyses the given word mainly by two methods.

3.1 Method-1

It is assumed that the given word is currently spelled and tries to find match from the word list with +/- first and last latter Misspelled word 'telephon' has 8 latter it finds 9 latter (+1) word 'telephone' in the database.

3.2 Method-2

In this method it takes every given word as scrambled word and finds all possible latter combination. It finds combination with +/- 1 latter, which can match with a word in the dictionary database. Misspelled word 'wather' has 6 latter framework finds all possible combination of 5 –7 latter words in the database. Further for the severity of mistake framework evaluates the word and assigns points.

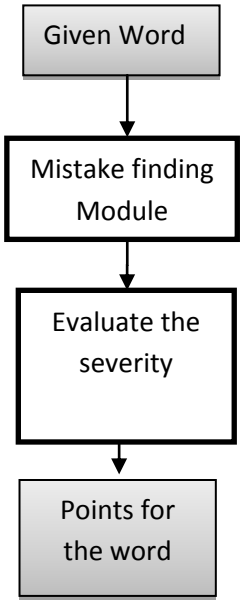

Fig. 2: Architecture of Framework.

4. ANALYSIS AND DISCUSSION

Following are the findings during the experiment.

- **Is the given word is in the dictionary database**: If the module of the framework finds the word in the dictionary database it returns the pointer to the structure with confidence level 100% on exact match.
- **Number of letters** omitted or added wrongly: If the module of the framework do not finds the word in the dictionary database it finds nearly matched word and returns the pointer to the structure with confidence level less than 100% depending on the number of mismatched letters and error level grater then 0.
- Location of wrongly added or omitted letters: If the module of the framework do not finds the word in the dictionary database it finds nearly matched word and returns the pointer to the structure with confidence level less than 100% depending on the location of the mismatched letter. Error level grater then 0.
- Type of letters omitted or added wrongly: If the modules of the framework do not finds the word in the dictionary database it finds nearly matched word and returns the pointer to the structure with confidence level less than 100% depending on the type of the mismatched letter (vowels / consonant). Error level grater then 0.

Table 1 : Value returned by Mistake finding Module

word	Confidence level	severity level	Number of wrong letter(s)	Best matched word	Padded word
schedule	100	0.00	0	schedule	schedule
schadule	87	0.12	1	schedule	sch-dule
schedul	85	0.75	1	schedule	schedul-
chedule	85	0.75	1	schedule	-chedule
representative	100	0.00	0	representative	representative
reprasentative	92	0.07	1	representative	repr-sentative
representativ	92	0.43	1	representative	representativ-
pepresentativ	92	0.43	1	representative	-epresentative
bread	100	0.00	0	bread	bread
bred	75	0.20	1	bread	bre-d
brea	75	1.20	1	bread	brea-
dread	80	1.20	1	bread	-read

On analyzing the data shown in the Table 1 it was observed that severity of the spelling error directly related with the word size. Smaller word which is easy to pronounce and memorize severity level is higher than the longer word, which is relatively, difficult to memorize.

5. CONCLUSION

Spell a word involves memorizing the proper phonetic pronunciation of the word. All spelling mistakes cannot be considered same. Severity level of the mistake depends on the size and complexity of the phonetic pronunciation of the word. Values for analysis in Table 4.1 shows, that all mistakes may not be taken equal; it can be observed that there is only one error in different sized word the confidence level and severity level is different. The framework for the evaluation of the spelling mistake works in the same format. Further importance of the word in the sentence should be considered and more methods can be added for batter confidence level.

6. REFERENCES

[1] Al-Jarf, R. (2008) - Listening Spelling Strategies in EFL Arab College retrieved 19 January 2010

[2] Bebout, L. -'An error analysis of misspellings made by learners of English as a first and as a second language'.Journal of Psycholinguistic Research (1985) 14/6: 569-593.

[3] Bowen, T. (2008) - Teaching English spellinghttp://www.onestopenglish.com/section.asp?docid=146381 retrieved 10 January 2010

[4] Brooks, G., Gorman, T. & Kendall, L. - Spelling It Out: the spelling abilities of 11- and 15-year-olds. Slough:NFER (1993)

[5] Kharma, N. Hajaj, A. - Errors in English among Arabic Speakers:Analysis and Remedy, Longman Group UKLimited (1989)

[6] Cook, V. (1999) – Teaching Spelling http://homepage.ntlworld.com/vivian.c/Writings/Papers/ TeachingSpelling..htm retrieved 15 January 2010

[7] Thompson-Panos, K. Thomas-Ruzic, M. – The Least You Should Know About Arabic Implications for the ESLWriting Instructor, TESOL QUARTERLY V, Vol. 17, No. 4, December 1983

[8] Radi, A. – Difficulties of Arabs Learning English, Language Centre FORUM, Spring 2001