A Linguistic Analysis of Misspelled Words in Urdu

Iram Sabir, Ph.D. in Linguistics

Abstract

The purpose of this article is to investigate the written Urdu errors by Urdu medium students in Auraiya. This paper aims to identify the written errors of Urdu as a second language in Uttar Pradesh, India. Researcher has provided quantitative validation for the qualitative analysis. This study is limited in scope in as much as Urdu has a writing system which is different from other languages such as English. The students in this study had difficulties in writing Urdu. Sometimes they insert or omit consonant and vowel in the word. For example: آبے پاش /a:be: pash/ (آب پاش /a:b-e-pash/) (Pot by which water was given to plants), آخیرت /a:xi:rat/ (آخرت /a:xrat/) (The other world after death), آگن /a:gan/, آنگَن /ā:gan/ (Court-yard) and sometimes they change the arrangement of letters in a word. Every letter is correctly written but with inaccurate placement. For example: لفط /luft/ as لطف /lutf/ (Enjoy), and اب ریشم /abri:sham/ as ابریشم /abre:sham/ (Raw silk).

Keywords: Orthographic errors, word boundaries, Insertion, Omission, Permutation, Substitution.

Introduction

Language teaching is a complex set of activities all aiming to language learning. A good language teaching theory always aims at meeting all necessities of learners in the best possible ways. Error analysis is the best tool for describing and explaining errors made by students. It helps to know the sources of difficulties of the learners. It also draws the main features of Urdu writing system, which consists of letters, associated forms and diacritical marks.

Urdu language is spoken in India as well as in Pakistan. Urdu is also among the languages whose writing system is diverse from that of English. Urdu has many irregularities in its spellings for a variety of reasons such as it has many homophonous letters and many words

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu
56

exhibit a variance between sound and spelling. In brief, for the most part, the speech sounds in Urdu are represented with more than one letter and vice versa.

Error Analysis

Error analysis is a type of linguistic analysis that focuses on the errors that learners make. Systematically analyzing errors made by learners makes it possible to determine areas that need reinforcement in teaching (Corder, 1974). Researchers are interested in errors because they are believed to contain valuable information on the strategies that people use to acquire a language (Richards, 1974; Taylor, 1975; Dulay and Burt, 1972). Furthermore, according to Lennon (1991), an error is "a linguistic form or combination of forms which in the same context and under similar conditions of production would, in all likelihood, not be produced by the speakers' native speakers' counterparts". Committing the error has always taken place along with teaching and learning processes and has always been used as a device of power and a teaching strategy.

Methodology

The analysis conducted is an attempt to classify the error patterns of Urdu medium students. All the participants are students of the reputed Urdu medium school from class VI to class VIII in Auraiya. A data base of 250 words / spelling errors was collected from hand written detritus of students from A.K.M. Inter College, Auraiya, UP, India. There are some examples of errors found in the Urdu words and ways to analyze them. These Urdu words were dictated to the students in the class room. The sample of study consisted of 45 students with 250 words selected. Then the samples were collected, analyzed and the findings were arrived at.

The essentials of spelling errors in Urdu writing are analyzed. Such analysis is of great help to teachers in recognizing error patterns that keep recurring. There are five kinds of spelling errors found in the data, viz., substitution (replacement of the word), insertion, omission, permutation and the word boundaries (splitting and merging).

Errors of substitution and insertion form the bulk of total spelling errors (79.13%). The effect of sociolinguistic setting or background of learners is also treated here since in language learning social variables play a crucial role in characterizing the language quality and learning efficiency of students.

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   57

The analysis is based upon the comprehensively collected data from 45 students of the reputed Urdu medium school of Auraiya. It is observed that spelling errors are governed by the writing system. An attempt is made to identify problem areas of Urdu spelling.

Researcher has encountered five (5) kinds of spelling errors in our data. The ratio to each pattern of error is given below in tables 1, 2, and 8. The total numbers of words are 250 that we have collected from our 45 respondents. Researcher has identified 145 (62 %) spelling errors out of 250 words which are also known as orthographic errors. After analyzing the data, researcher has found out that errors resulting from not knowing the rules of word-boundaries are approximately 10 (6.45%) in total out of 250 words. So the total number of error including word-boundaries are 155 (145/93.54% spelling errors and 10/6.45% word-boundaries errors).
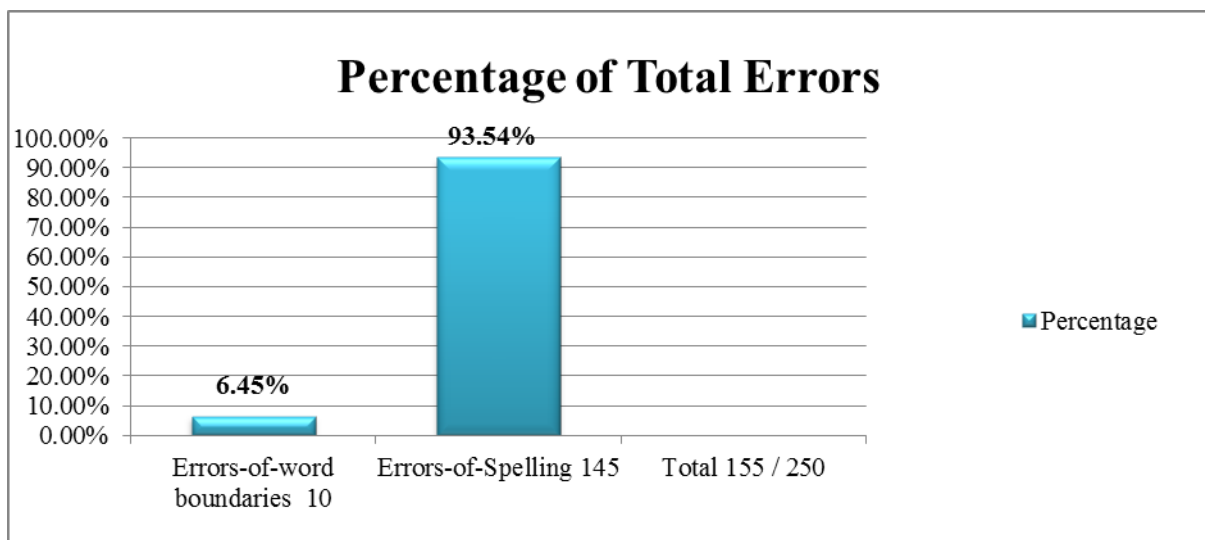
There are two types of orthographic errors:

1. Errors of spellings

2. Errors of word-boundaries.

Table .1: Orthographic Errors with their Percentage

Orthographic Errors	Number	Percentage
Errors-of-word boundaries	10	6.45%
Errors-of-Spelling	145	93.54%
Total	155/250	

Diagram.1: Orthographic Errors with their Percentage

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   58

Percentage of Total Errors

Errors-of-word boundaries 10 — 6.45%
Errors-of-Spelling 145 — 93.54%
Total 155 / 250

Table.2: Statistics of the Four Basic Types of Errors in the use of Letter (for Urdu language)

Spelling errors in the use of Letter	Number	Percentage
Substitution	65	44.82%
Insertion	50	34.48%
Omission	26	17.93%
Permutation	4	2.75%
Total	145/250	

Diagram.2: Spelling Errors in the use of letter with their   Percentage

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   59

Spelling Errors in the use of Letter
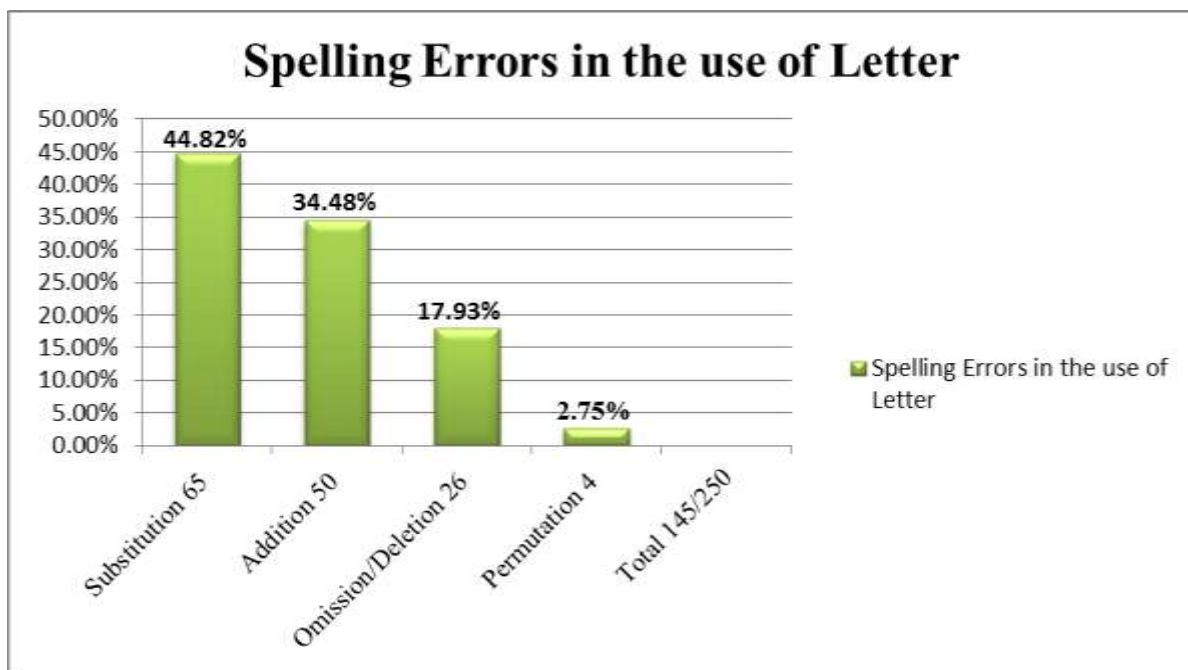
1. Errors of spellings

1.1. *Substitution*

A substitution error occurs when one letter has been written in place of another letter. Substitution errors are caused mainly by homophonic character which primarily represents the same sound. In fact, there is a lot of confusion about the use of homophones. Learners often mistake one homophone for another. These errors are frequently more complex, since homophones are pronounced identically; consequently, they are misspelled mostly. Students get confused with these words. They find it difficult to write them correctly.

Generally the students do not know the actual spelling of the word and they try to write it with possible spelling, guessing it with similar pronunciation of the letters that sound, to some extent, similar to the actual word. The homophone or similar sounding letters are commonly substituted in this type of error (e.g. ز for ط, ت for ض, ز/ذ for ص, س for ق, ک for /ذ for ظ, 'ه' /hah/ for 'ح' /hey/ or 'ت' /tey/ for 'ط' /toye/. e.g. "آکسیجن" (Oxygen: */a:ksi:jan/*) replaced by "عاقسیجن" and "تاریک" (darkness: /tari:k/) replaced by "طاریک".

Researcher performed the analysis of spelling errors of substitution of letters for similar sounding letters in various positions of words in writing of students of VI to VIII class. Researcher has found out spelling errors of substitution of single letter that is substituted for a letter which is similar in sound. Number of errors found in this category is 65 which come to 44.82%. Some such letters are as: ق ,ک, ع , غ ,ظ, ط , ض,ص , ح and ه

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu 60

Table 3 shows the list of all the Urdu letters that fumble according to their similar sound pronunciation.

Table. 3: Similarly pronounced letters in Urdu

	Similar Sound Character			S.N.
	ع	آ	ا	1
		ط	ت	2
	ص	س	ث	3
	ہ	ھ	ح	4
		ک	خ	5
ظ	ض	ز	ذ	6
		ک	ق	7
		غ	گ	8

1.2. *Insertion*

It covers spelling errors caused by insertion of letters to a word. These errors are caused by lack of knowledge with the result words are misspelled phonetically. Learners have a particular propensity to spell words according to the way they sound. Therefore, they make a false analysis of words. As a result, they misspell them.

Errors of insertion deal with the issue of the unnecessary use of و (*wao*), ی (*choti-ye*) ا (*alif*) and ں (*noon-ghunna*), (nasalization marker). These errors simply occur when a letter is mistakenly added by students to different positions of words, other than the ratio of errors made by VI and VII, class students is much higher than that of the students of class VIII.

Table. 4: Various Examples of insertion errors

Gloss	Wrong Word	Phonetic Transcription	Correct word	S.N.
Blister	آبلا / آب لا	*ablah*	آبلہ	1
Fire	آتیش	*a:tish*	آتش	2
Tear	آسوں	*ã:su:*	آ نسو	3
Officer	اوُفی سر/عوفیس	*a:fisar*	آفیسر	4

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   61

Total number of errors of insertion comes to 50 (34.48%) out of 250 words. This category of errors includes unnecessary use of certain graphemes. Generally errors of insertion are caused by vowels as well as consonants letter. Table 3 shows the list of all the Urdu letters

1.3. *Deletion /omission*

Researcher has evaluated the spelling errors of deletion by learners of all the three classes from VI to VIII. Errors of deletion have been analyzed. A deletion is when a letter has been omitted from a word. Sometimes a word may be correct except for one or two letters that are missing from the word. This happens when a learner fails to correctly analyze all the sounds within word.

Deletion/omission of letters representing vowels and consonants are particularly seen in the written text of the learners. Number of errors found in this category is 26 which come to 17.13% out of 250 words.

Researcher has discussed the errors caused by deletion of ح and ہ (*Choti-hey and Bari-hey*) in medial and final positions of the words. In Urdu the terminal ہ of the words is always pronounced as a long vowel /a: /, and the tables show the deletion of ہ and ح from the words.

Errors of omission of the letter ع (*ain*) have been examined. The deletion of ع usually occurs medially and finally. Researcher has analyzed 26 (17.93%) misspelled words. It deals with the deletion of diacritical marks, which represent short vowels. The deletion of following diacritical marks is treated viz; *tashdeed* {ٓ}, *do-zabar* {ﹰ}, *khara-alif* {ٰ}, *noon-e-ghunna* {ں} etc. Deletion of diacritical marks shows the misspellings that are caused by omission of *tashdeed*. It deals with the errors of omission of *khara- alif,* has been misspelled as a result of this deletion. In deletion/omission, researcher has described the errors of the diacritical marks *do-zabar*, and also errors caused by deletion of diacritical mark 'ں' (*noon-e-ghunna*) have been dealt with. The deletion represents nasalization of a vowel. It is concerned with the errors of deletion of 'ن' (noon) in the medial position of words.

Such analysis of errors of omission as done above is necessary to understand the difficulties that learners face while writing. It is noticed that most of the errors have been made only by the class VI students. These errors are decidedly larger in comparison with errors made by students *of* VII and VIII classes.

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu 62

Table. 5: Various Examples of deletion/omission errors

Gloss	Wrong Word	Phonetic Transcription	Correct word	S.N.
Willing	آمادا	*a:ma:dah*	آماده	1
Flame, heat	آچ/عاچ	*ã:c*	آنچ	2
Eye	آک	*ã:kh*	آنکھ	3
Gathering	اجتاما	*ijtama:á*	اجتماع	4
In the name of ALLAH	بی سمیلا/ بی مِلًّا	*bismilla:h*	بسم الله	5
Rulers	حوکم رانو	*ḥukmra:nõ:*	حُکمرانوں	6
Lid/Cover	ڈَقن	*ḍhakkan*	ڈھکّن	7
Prayers	سلوات/ صالاوات	*sala:t*	صلوٰت	8

1.4. *Permutation*

Researcher has also dealt with errors of permutation. These errors are caused by a change in the arrangement of a given number of elements in a group. One of the ways of organizing a group of things is called permutation. An error of permutation is perceived when all the letters of words are present but in a wrong order. This mirrors difficulty with sequential memory and shows learner's inability to see the picture of word in the mind. The result is a jumble of letters giving rise to meaninglessness. Total 04 errors are found for permutation which is 2.75% of total number of spelling errors out of 250 words. A permutation can be changed into another permutation by simply switching off two or more of the objects. One of the various ways in which we can combine or arrange a group of things is called permutation.

This category is relatively un-controversial. The learners select the right forms and use them in the right context, but they arrange them in the wrong order, for instance, when the English word 'Sue' is written as 'Seu' or Urdu word as shown in the table:

Table. 6: Various Examples of permutation errors

Gloss	Wrong Word	Phonetic Transcription	Correct word	S.N.

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   63

A mistake	غل طی	*ɣalti:*	غلطی	**1**
Mystic	سوفی یانہ	*su:fiya:nah*	صوفیانہ	**2**
Honor	غے رت	*ɣairat*	غیرت	**3**
Enjoy	لفط	*Lutf*	لطف	**4**

2. Errors of word-boundaries

Researcher has dealt with the errors which result from not knowing the rules of word-boundaries. These are approximately 10 in total (i.e. 04%) out of 250 word samples. Researcher has found that the total number of errors including word-boundaries is 155 (145/93.54% spelling mistakes and 10/6.45% word-boundaries).

How a word is determined on boundaries on its both sides has been discussed here elaborately. It is revealed that learners are confused about tokenization. It is noted that the issue of tokenization in Urdu is rather more complex than in many other languages. Concept of space is conspicuous by its absence in Urdu language. It is added that issues of space insertion and word segmentation are very much tricky in Urdu and are very complicated too.

It has been found out that sometimes learners join two separate words by not leaving space between them and sometimes they break one single word into two words by leaving space in their middle.

Table. 7: Shows the various word-boundaries errors

Error	**Gloss**	**Wrong Word** (split/merge)	**Phonetic Transcription**	**Correct Word**
Substitution	Hospital	ہوسپٹل/ حوسپٹل	*haspata:l*	ہسپتال
Insertion	Exhibition	نوماعیش	*numa:ish*	نمائش
Omission	Touch word	ماشااللہ	*ma:sha: allah*	مساشاءاللہ
Permutation	Life	زن دگي	*zindagi*	زندگي

These are concerned with the violation of rules of word boundaries. Errors of split are 08 /10 i.e. 80%, while merger of the word boundaries is 02/10 i.e. 20%. It is seen that errors of split

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   64

are greater in number than the errors of merger. It is also observed that class VI children are more prone to make mistakes of split and merger compared to VII and VIII class students.

Table.8: Errors of Word boundaries with their Percentage

Errors-of-word boundaries	Number	Percentage
Split	8	80%
Merger	2	20%
Total	10/250	

Diagram.3: Errors of Word boundaries with their Percentage
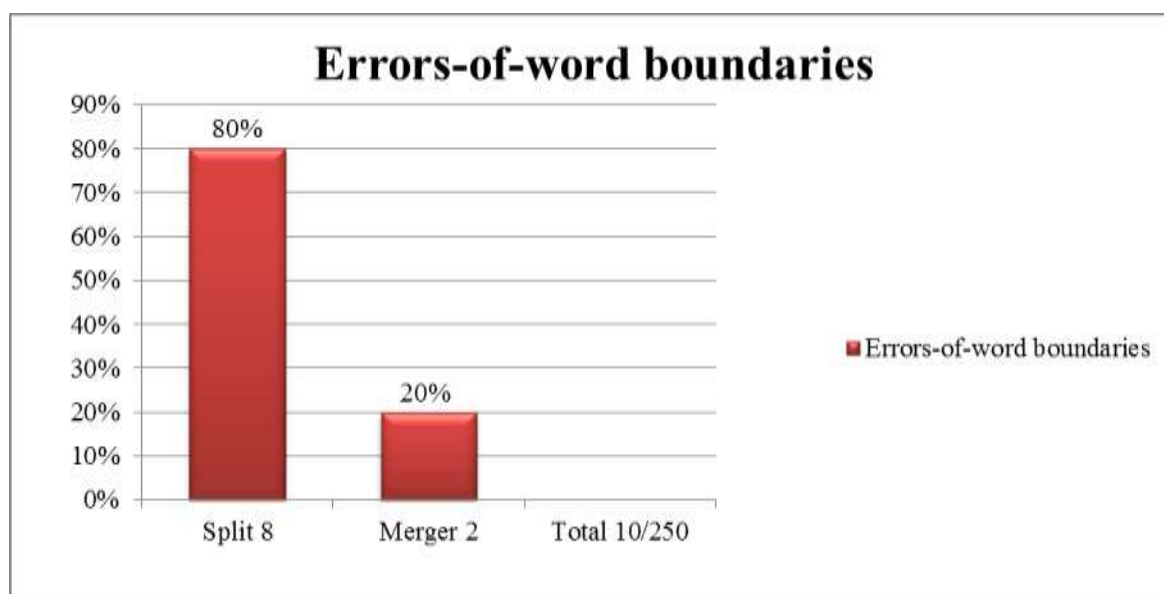
Conclusion

From the discussion and studies given in this paper it can be concluded that in Urdu, learners seemed to lack phonetic spelling strategies. One of these is the frequent occurrence of substitution errors caused due to the form of similarity of the letters in Urdu alphabet and also due to similar pronunciation of various letters. Other kinds of error found in Urdu language are the omission of space character at the word boundaries. Based on this study of Urdu, it can be supposed that these results and error trends will also apply to other languages that use

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   65

writing systems similar to Perso-Arabic script. They might require modifications to provide for the script specific issues of spelling errors. Teachers should also need to try to teach the rules and conventions of writing more efficiently.

==

References

Corder, S.P. 1974. Error Analysis in Allen, J.L.P. and Corder, S.P. (eds); *Techniques in Applied Linguistics*. Oxford.Oxford University Press.   .

Dulay, H.C. & Burt, M.K. (1972). Goofing: An Indicator of Children's Second Language Learning Strategies. *Language Learning,* Vol. 22, Issue. 2, pp.235-245.

Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin.   . *Topics in Language Disorders*, *20*(3), 19-36.

Larsen, S., Hammill, D., & Moats, L. (1999). *Test of Written Spelling* (4th ed.; TWS-4). Austin, TX: Pro-Ed.

Lennon, P. 1991. "Error: some problems of definition and identification", in Applied Linguistic, vol. 12, num. 2, Oxford, pp. 180-195.

Masterson, J. J., & Mooney, R. (2006, February). *Criterion validity of the spelling performance evaluation for language and literacy (SPELL).* Poster presented at the Annual Pathways to Communication Conference. Missouri State University, Springfield, MO.

Naseem, T. And Hussain, S. *A Novel Approach for Ranking Spelling Mistakes in Urdu, to appear in Language Resources and Evaluation*, in June 2007

Naseem, T. And Hussain, S. (2007).*"Spelling Error Trends in Urdu". In the Proceedings of Conference on Language Technology (CLT07)*,University of Peshawar, Pakistan.

Richards, J.C. 1974. *A Non-contrastive Approach to Error Analysis*. In J. C. Richards (Ed), *Error Analysis:perspectives on Second Language Acquisition* (pp. 172-188). London: Longman.

Taylor, B. P. (1975). The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL. *Language Learning*, 25, pp. 73-107.

Wilkins, D.A. 1972. *Linguistics in Language Teaching*. London: ELBS.

==

.

Dr. Iram Sabir
Assistant Professor

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   66

Department of English
Al-Jouf University
Kingdom of Saudi Arabia
dr.iramameen@gmail.com

Address (INDIA):
C/o. Mr. Sabir Hussain (Teacher, Minto Circle)
H/no. 4/340, Alam Bagh,
Bhamola, Aligarh
Uttar Pradesh
India

Language in India www.languageinindia.com ISSN 1930-2940
13 : 3 March 2013
Iram Sabir, Ph.D. in Linguistics
A Linguistic Analysis of Misspelled Words in Urdu   67