



Error Analysis: A Study of Errors made by Pakistani 'O-Level' Students in Writing Urdu

Asif Ali

PhD Scholar, English Linguistics, University of Central Punjab, Lahore, Pakistan
asifalinaseer@gmail.com, & L1F24PHDL0008@ucp.edu.pk

Dr. Nabeel Ahmed Nabeel

Associate Professor, Department of Urdu, Division of Islamic and Oriental Learning,
University of Education, Lower Mall Campus, Lahore, Pakistan
nabeel.ahmed@ue.edu.pk

Shaheera Saeed

PhD Scholar, English Linguistics, University of Central Punjab, Lahore, Pakistan
L1F25PHDL0001@ucp.edu.pk & shaheerasaeed00@gmail.com

Abstract:

Every academic year, thousands of Pakistani students appear in the Cambridge Assessment International Education Examination of Ordinary level and take the examination of Urdu as a compulsory subject. Urdu is spoken and learned at three levels in Pakistan. Firstly, it is the mother tongue of a minority of people. Secondly, approximately 6.8 % of the population speaks Urdu as their first language. Thirdly, almost 60% of people in Pakistan learn it as a second language. The main objective of this paper is to investigate the errors made in Pakistani 'O-level' students' texts in the light of the Error analysis approach of Corder (1967). Urdu is considered a challenging language owing to its cursive orthographic and largely borrowed vocabulary from Arabic, Persian, Hindi, and other native languages of the subcontinent. The paper adopts a qualitative research paradigm as the researcher aims to study students' errors in a natural linguistic setting. The study is limited to error analysis of 'O-level' students in writing Urdu only. The data was collected from the formative assessment papers of the sample of purposively selected 50, 'O-level' students of an English medium school in Lahore, Punjab, Pakistan. The study achieves its objective through a detailed analysis of the corpus of 313 errors made by the subjects of the study. The findings of the study revealed that 84.66 % of the investigated error corpus falls in the category of "single-letter errors". The findings also show that within the single letter category, the most frequent errors were "substitution errors" which is 58.49 % of the category. Moreover, the findings reflect that the visual and phonetic similarity of letters may have caused a comparatively high ratio of substitution errors in the single letter category. Finally, the study suggests some pedagogical implications for teachers to enhance students' ability to write error-free Urdu. Special focus on visually and phonetically similar letters, carefully designed practice exercises, differentiation of the functions of letters with their placement at the initial, middle and final stage of words, and regular feedback of errors may contribute to the improvement of the students in this regard.

Keywords: Urdu language, Orthographic errors, single letter errors, double letter errors

Introduction:

Urdu, the national language of Pakistan, is among the most widely spoken languages both in the subcontinent and globally. Estimates suggest there are around 11 million native Urdu speakers in Pakistan, with the global count surpassing 163 million speakers (Riaz, 2008; Iqbal et al., 2021). Urdu is spoken and learned at three levels in Pakistan. First, it serves as the mother tongue for a minority of people, with approximately 6.8% of the population speaking Urdu as their first language (Coleman, 2010). Second, nearly 60% of Pakistan's population learns Urdu as a second language (Dar et al., 2015). Finally, Urdu functions as a common medium of communication in Pakistan's multilingual society. It symbolizes unity and federation, connecting people across all provinces of the country (Rahman, 2006).

Urdu, with its extensive borrowed vocabulary from Persian, Arabic, Hindi, and other Asian languages, has a diverse orthographic system. Its writing system is heavily influenced by both Persian and Arabic, and it is written in a cursive, context-sensitive Perso-Arabic script from right to left. Urdu also exemplifies *digraphia*, a linguistic framework in which a language can be written using different scripts (Ahmad, 2011).

In Punjab, Pakistan, government schools use Urdu as the main language for teaching. A survey showed that 60.47% of teachers feel comfortable teaching primary-level classes in Urdu (British Council survey). While most 'O-level' schools prefer English as their main language for instruction, seeing it as a symbol of status and influence, they also include Urdu in a bilingual teaching approach. Other regional languages, however, receive little attention, as the focus remains mostly on Urdu and English (Rahman, 1998, cited in Mohanty, 2010).

Significance of the study:

Language and more specifically writing is one of the most effective sources of students' performance assessment in the field of education. It is also considered a powerful means of communication in any language (Graham, S., & Perin, D. 2007). Like any other language learners, errors in writing Urdu are common to 'O-level' students in Pakistan. The core purpose of this paper is to explore orthographic errors made by Pakistani 'O-Level' students while attempting their papers in the Urdu language. Nonetheless, Urdu is the national language, still, students commit certain errors because of unawareness of the linguistic rules. This study is significant as it is being conducted on error corpus in a natural setting and the major error categories are identified. It will highlight commonly repeated errors and will provide a detailed analysis of them. The study will also suggest some pedagogical strategies to enhance awareness and ways of improvement in the future for error-free writing.

Objectives of the study:

The primary aim of this study is to examine the different types of errors made by O-Level Pakistani students in their formative assessment papers for the Urdu language. Furthermore, the study seeks to analyze these errors, identifying the underlying reasons for their occurrence. Additionally, it will focus on strategies to enhance students' abilities to write in Urdu without errors.

Research Questions:

1. What types of errors are made by Pakistani O-Level students when writing in Urdu?
2. What is the frequency and percentage of each type of error identified?
3. What are the potential reasons for these errors?
4. What strategies can be employed to help students address and correct these errors?

Literature Review:

Error analysis is commonly defined as a systematic process for identifying, classifying, and interpreting errors produced by learners of a target language (Crystal, 2003; Brown, 2000). According to James (1998), it involves determining the incidence, nature, causes, and consequences of unsuccessful language use. The process entails collecting samples of learners' language, followed by the identification and categorization of errors based on their causes and nature, and finally evaluating the seriousness of these mistakes to facilitate their reduction (Keshavarz, 1999). Additionally, Ellis (1997) outlines four main stages of error analysis: identification, description, explanation, and evaluation. Error analysis provides thoughtful aspects of human insight to understand the acquisition process of a second language. It is a kind of linguistic analysis carried out in an organized manner to identify errors followed by descriptions and solutions to the reduction of errors. It also shed light on the linguistic interference in the occurrence of errors while learning a second language (Khan & Khan, 2016).

Committing mistakes and errors is a natural and essential part of the language acquisition process, whether involving a first language or a second language. Corder (1967) holds that learners of a second or foreign language often make errors as part of the language acquisition process. However, errors are not only confined to learning second or foreign languages; they also occur in the process of learning one's first language. Scholars and linguists have long studied these errors to assist teachers and provide valuable insights to improve learners' language skills and performance in their target languages. Numerous studies have been conducted to investigate the occurrence of errors and their causes throughout the language learning process (Richards, 1974; James, 1998; Norrish, 1983; Ellis, 1995; Ziahisseiny, 1999; Keshavarz, 2006; Darus, 2009).

Errors, as deviations from standard linguistic and orthographic rules, can be divided into two categories: "mistakes" and "errors." Mistakes are generally considered temporary lapses or slips in performance, often caused by factors such as fatigue, carelessness, or a momentary slip of the tongue. In these cases, learners typically have a solid understanding of the language, but external factors affect their performance (Brown, 2000; Corder, as cited in Richards, 1974). In contrast, errors reflect a learner's lack of competence in the language. They occur when a learner does not have a sufficient understanding of the language and cannot self-correct because they are unaware of the correct form (Gass & Selinker, 2008).

Researchers have identified several factors that may lead to mistakes during language acquisition. These include interference from the native language on the second language, lapses in attentiveness while speaking or producing the target language, and overgeneralization of linguistic rules. Additionally, limited understanding of complex structures, restricted exposure to morphological and

grammatical rules, and cognitive and psychological influences, such as anxiety, also contribute to errors (Brown, 2000; Ellis, 1995; Gass & Selinker, 2008; Richards, 1974). According to Stenson (1974), there are three main reasons for errors. Firstly, incompetency in the acquisition of the grammar of the target language. Secondly, the involvement of the internal and external factors in the teaching-learning situations of the target language. And, thirdly, errors took place due to normal human problems in the performance of language (Stenson, 1974).

Language learners often make errors across different language components, such as phonological, morphological, lexical, and syntactic aspects. For instance, phonological errors may occur when a learner cannot distinguish between the phonemes /p/ and /b/, leading to pronunciations like "pird" instead of "bird" and "brison" instead of "prison." Morphological errors include the incorrect use of morphemes, resulting in words like "mens," "sheeps," and "furnitures" (Touchie, 1986).

Urdu is among the languages with highly complex orthographic systems, containing 57 letters and diacritical marks (Hussain & Afzal, 2001). It shares orthographic traits with Arabic, Persian, and Hindi, making its script diverse and intricate. For instance, short vowels are represented by diacritics placed above or below consonants rather than as independent letters (Humayoun et al., 2006). Written in a modified Arabic script from right to left, Urdu inherits the complexities of Arabic, especially due to its use of Arabic loanwords (Hardie, 2003). A distinctive feature of Urdu letters is their multiple graphemic forms based on position within a word. For example, the letter *ain* [ع] changes form depending on whether it appears at the beginning (زت respect), middle (تعليم education), or end (اطلا information) of a word. Additionally, Urdu's cursive nature leads to numerous visually similar graphemes, distinguished only by their placement in a word or by the position of dots. For example, the words *kitni* [کتنی] (how much), *kehti* [کہتی] (she says), and *khaiti* [کھیتی] (crop) exhibit subtle yet crucial differences in their graphemic structures (Mirdehghan, 2010). Urdu is taught as a compulsory subject in all 'O-level' schools of Pakistan affiliated with the Cambridge Assessment International Examination (CAIE). CAIE offers two courses in Urdu; Urdu as the First language (code 3247) and Urdu as a second language (code 3248). The majority of 'O-level' students study Urdu as their first language in Pakistan. Whereas, Urdu as a second language because of its easy syllabus, is opted for by most overseas Pakistani students whose first language is English, Arabic, or any other language.

(<https://www.cambridgeinternational.org/programmes-and-qualifications/cambridge-o-level-urdu-first-language-3247/>)

Methodology:

This paper aims to investigate students' errors in Urdu writing, which are related to students' behavior towards learning a language and their performance. The qualitative research method is being applied in this research paper as it offers approaches to the identification, classification, description, explanation, and evaluation of the phenomenon. The possible error-free error analysis is to be done by applying the qualitative research method. It helps the researcher to investigate the phenomenon with an interpretive and naturalistic approach. It means that the qualitative research

method allows the researcher to study problems in their natural settings, making sense of the emerging pattern and interpretations. (Denzin & Lincoln, 2011).

Population and Sample:

The population of this study was a batch of 'O-level' final year students of a well-known English medium school in Lahore, Punjab, Pakistan. The population of the study, 350 students, studied Urdu for two years as their compulsory subject and as their first language. A total number of 50 students in the 'O-level' final year were chosen as a sample through random sampling.

Instrument and data collection:

The researcher used 50 written scripts as an instrument in this study. The participants were asked to answer three compulsory questions as part of their formative assessment of the Urdu language. The subjects were given a standardized time of 1 hour 30 minutes to complete the given task and they were not allowed to take any external help or to use any source during their assessments. One of the three given tasks was guided writing and the other two were free essay writing. The participants were given the following questions;

- A. Your younger brother is afraid of swimming and horse-riding. Explain to him via email the importance of sports and his participation. Include the following points in your email; healthy mind and body, positive activity/ exercise, and participation in competitions at the international level. [word limit 120]
- B. Write an essay on the topic of "Rural life is better than urban life". [word limit 200]
- C. Nasir Kazmi's poetry contains both feelings of hope and despair. Please comment with reference. [200-250 words]

The whole exercise was carried out fairly, the data collection was done with the prior permission of the management and bringing it to the knowledge of the participants.

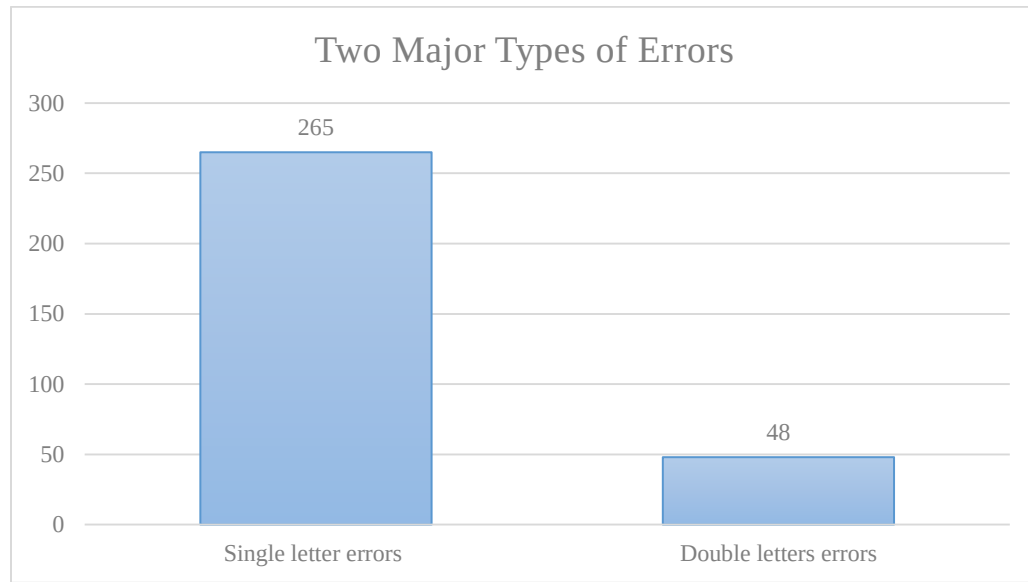
Data Analysis:

Data analysis is the next process after the successful collection of data. The process of data analysis involves the classification and interpretation of linguistic material to make meaning out of it (Flick, 2014). The participants' papers were read by the researcher at least three times meticulously. In the first step, all types of spelling errors were highlighted and identified in the submitted papers. A total number of 313 errors was found in the written scripts of the participants. The second step was to scrutinize the error corpus into predetermined categories; omission, insertion, substitution, transposition, and grapheme substitution (Corder, 1967; Burton et al. 2010). The above-mentioned categories were further divided into two major categories; first, single letter errors, and second, double letter errors. The frequency and percentage of single and double letter errors occurrence have shown in table 1.

Table 1: Two major types of errors

S. No	Type of Error	Frequency of error	Percentage
1.	Single letter errors	265	84.66
2.	Double letter errors	48	15.33
	Total	313	100

Hence a graph is also drawn for the presentation of errors data that is being analyzed.

Figure 1: Two major types of errors

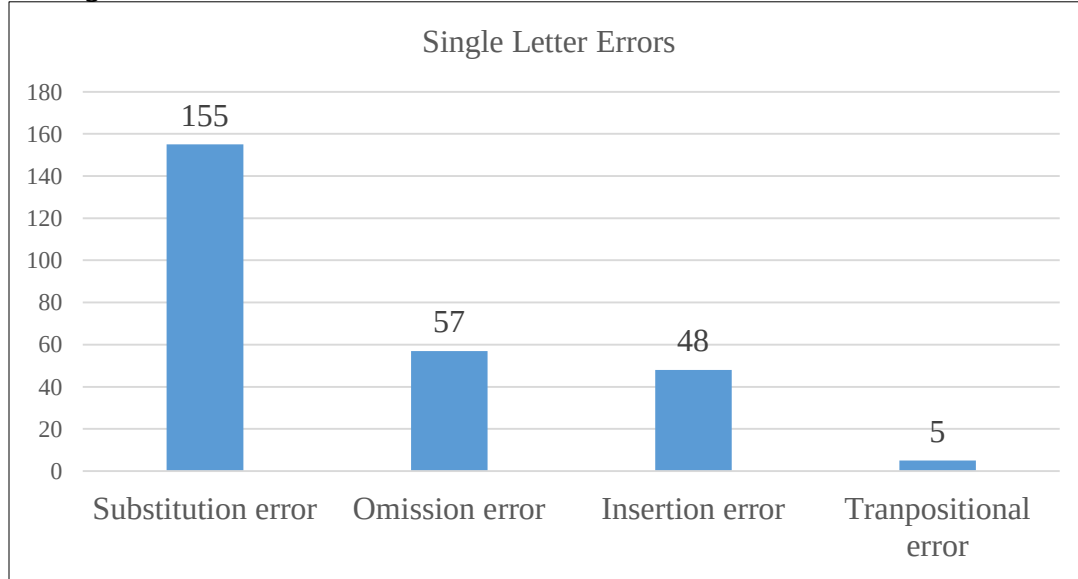
Result and Interpretation:

1. Single letter error:

Single letter error was the first category among three major types of errors. The participants committed 265 errors out of 313 which is 84.66% of the total errors. The main cause of the error was either substitution, insertion, omission, or transposition of a single letter within a word (Corder, 1967). A clear difference was marked in the participant's produced intended words compared to the standardized token of the language. The single-letter error type was further subcategorized into the following four types;

1. Single letter substitution: e.g. writing [آدی] instead of correct token [آدی].
2. Single letter omission/deletion: e.g. writing [متلق] instead of correct token [متعلق].
3. Single letter insertion: e.g. writing [تعالیم] instead of correct token [تعلیم].
4. Transposition of two adjacent letters: e.g. writing [استعمال] instead of correct token [استعمال].

The frequency of the occurrence of errors is shown in the figure 2.

Figure 2: Single Letter Errors

1.1.Substitution Error

A substitution error takes place when a letter is replaced with another incorrect letter in the formation of a word (Jobeen et al., 2015). Substitution errors are mainly caused by phonetic or visual similarity between the correct letter and letter produced by the subject of the study. Substitution error was the highest-rated error category in single-letter errors. As it is shown in graph: 2, the participants produced 155 errors of substitution out of 265 total single-letter errors, which is 58.49 % of this category. The highest frequency of substitution error found in the collected data is aligned with Cook (1999) who found substitution comparatively higher than the other three categories.

The students get confused while writing a word containing two letters either similar in pronunciation or shape. Taking the phonetic similarity first, For instance, one of the participants wrote a word /عادى/ man, in place of /آدى/ which is an incorrect token. The simple assumption behind this sort of error is the phonetic similarity between the two letters replacing each other. The initial letter of both correct /آدى/ and incorrect words/عادى/ is pronounced /ɑ:/ which sounds quite similar. So, a similar phoneme caused the problem and resulted in occur ace of error. A few more examples of phonetic similarity are given below in table 2.

Table 2: Phonetic similarity in substitution error

S.No	Gloss	Intended word	Produced word	Substitution Error
1.	cleanliness	صفائی	سفائی	The initial letter /س/ is replaced with /ص/, both sound similar to the phoneme /s/
2.	agreed	متفق	متشک	The last letter /ق/ is replaced with /ک/, both

				sound similar to the phoneme /k/
3.	migration	ہجرت	حجرت	The initial letter /s/ is replaced with /ح/, both sound similar to the phoneme /h/
4.	issues	مسئلے	مشعلے	The middle letter /س/ is replaced with /ث/, both sound similar to the phoneme /s/
5.	cruel	ظالم	زالم	The initial letter /ظ/ is replaced with /ز/, both sound similar to the phoneme /z/

The second possible cause of substitution error is the visual similarity of two used letters. In such a case of error, a student simply substitutes a letter with an incorrect one because the shape of both letters seems similar. For instance, a participant intended to write the Urdu word /خارج/ and ended up writing /چارخ/, which is not any word of Urdu language in any sense. The initial letter /خ/ of correct word was changed with the incorrect phoneme /ج/ and that's how the error occurred. The initial phonemes /خ/ and /ج/ of both correct and wrong words belong to the same family of Urdu alphabets. So the apparent cause of confusion was the visual similarity between both phonemes. The change in the placement of a dot /-/ made the produced word incorrect and meaningless in the Urdu language. Few other examples of visually similar substitution errors are given below in table 3.

Table 3: Visual similarity in substitution error

S.No	Gloss	Intended word	Produced word	Substitution Error
1.	poetry	شاعری	ساعری	The initial letter /ش/ was replaced with /س/, both phonemes belong to the same class of Urdu alphabets and share the same shape.
2.	to break	ٹوڑنا	تورنا	The last letter /ڑ/ was replaced with /ر/, both phonemes belong to the same class of Urdu alphabets and share the same shape.
3.	pronunciation	تلفظ	تلظ	The last letter /ظ/ was replaced with /ط/, both phonemes belong to the same class of Urdu alphabets and share the same shape.
4.	need	ضرورت	صورت	The last letter /ض/ was replaced with /ص/, both phonemes belong to the same class of Urdu alphabets and share the same shape.
5.	bones	ہڈیوں	ہڈہوں	The initial form of the middle letter /ی/ was

replaced with /ز/, and the initial form of both phonemes seems similar in their position.

1.2.Omission/ Deletion Error

Omission or deletion error occurs when a linguistic item or phoneme that is required in the correct formation of a word is omitted (Sabir, 2013). The casual behavior of the learner or failure to analytical recognition of all sounds within a word are the possible causes of such type of errors. The collected data presents that omission error was the second-highest in the single letter error category. The researcher identified 57 errors of omission against 265 total single-letter errors, which is 21.50 % of this category, as shown in graph 2.

For instance, one of the research participants intended to produce the Urdu word [متعلق]/about/ and he ended up with an incorrect lexical item [متلق]. Only one phoneme, the initial form of /ع/ is omitted by the student which resulted in the wrong lexical item. It is noted that the omission phoneme /ع/ usually occurs in the middle and final position of a word. Two other phonemes, *noon* with inside dot/ن/, and *noon* without inside dot, *noon-e-ghunna* /ں/ are found omitted in the middle and at the final position of the word respectively. It occurs due to the nasalization sound of these two phonemes. For instance, in an incorrect lexical item /ماگوں/, the phoneme *noon* with inside dot/ن/ has been omitted which resulted in an omission error. The correct token of Urdu language is /ماگوں/it means to requesting of a thing. The example of phoneme *noon* without an inside dot, *noon-e-ghunna* /ں/ is found missing at the final position of words. The omission of *noon* without an inside dot, *noon-e-ghunna* /ں/ occurred in /چیڑے/, /دونو/, and /کیو/ whereas the correct words were /دونوں/both/, /چیزیں/things, and /کیوں/why. A few more examples of omission errors other than the above-mentioned phonemes are given below in Table 4.

Table 4: Omission/ Deletion Error

S.No	Gloss	Intended word	Produced word	Omission/ Deletion Error
1.	pollution	آلودگی	الودگی	A diacritical mark/َ/ placed on top of the Urdu alphabet /ا/ e.g./آ/ is omitted in the initial letter.
2.	live together	ساتھ گزارنا	سات	A phoneme /تھ/ represents an aspirated sound /th/is omitted from the final position of /سات/.
3.	pleasant	خوشگوار	خشگوار	Urdu phoneme /و/ pronounced as a vowel in Urdu

				is omitted next to the initial position because of unstressed sound.
4.	destroyed	تباہ	تبا	The last letter /b/ was omitted as /b/always pronounced as the long vowel /a:/ in the Urdu language at the final position.
	branches	شاخوں	شخوں	Urdu phoneme /h/ is omitted in the position next to the initial letter. The letter /h/ is included in the vowels of Urdu language.

1.3.Insertion Error

Insertion errors are caused by the insertion of the unwanted letter at the initial, middle and final position within a word (Jobeen et al., 2015). The confusion of learners' minds regarding the accurate pronunciation of words and careless behavior while writing is the main causes of the occurrence of insertion errors. Insertion errors were found at the third-highest ratio in the collected data by the participants of the research. 48 errors of single letter insertion were identified by the researcher out of 265 total single-letter errors, which is 18.11 % of this category, as shown in figure 2.

The data depicts the unnecessary insertion of certain Urdu letters, e.g. /ا/, /د/, /ی/, /ع/, and /و/ repeatedly. For instance, the Urdu word [تنگ]/tight/ is produced incorrectly /تانگ/, so the insertion of a phoneme /a/ has been made right after the initial letter. Another participant who was supposed to write [موبیٹی] ended up [موبیٹی] and he mistakenly added /ع/ next to the initial letter /م/. A few more examples of insertion errors are shown in table 5.

Table 5: Insertion Error

S.No	Gloss	Intended word	Produced word	Insertion Error
1.	area	علاقہ	اعلاقہ	An unnecessary insertion of /ا/ has been made in the initial position of the word.
2.	advice	مشورہ	مشہورہ	A phoneme /h/ is mistakenly added in the middle position of the word.
3.	confession	تسلیم	تسلیم	Urdu phoneme /ع/ is inserted in the middle position of the word.
4.	old	پرانے	پورانے	The insertion of the Urdu phoneme /و/ has been made in the position next to the initial letter.

5.	new	نئی	نئی	Urdu phoneme /ی/ is added in the middle position of the word.
----	-----	-----	-----	---

1.4. Transposition Error

A transposition error occurs when two adjacent letters swap their places within a lexical item. The disordering of two side-by-side letters constitutes an incorrect word that was supposed to be correct if this disarrangement has not taken place (Sabir, 2013). Transposition errors are also attributed to the careless behavior of students while writing (Carney.2012). The participants committed 5 transposition errors against the total number of 265, which is 1.88% of the category of single-letter errors. The statistics are shown in graph 1. The ways a change in two letters caused an error could be seen in the error of [دکھی] which was supposed to be written [دکھی] (sad). In the above example, two phonemes /د/ and /ک/ were altered in their places. The phoneme /ک/ was supposed to be used first, followed by /د/ to constitute an accurate lexical token of the Urdu language. Another aspect of transposition error is that no additional letter is used rather misplacement of already existing letters causes an error. A few more examples of transposition errors are shown in table 6.

Table 6: Transposition Error

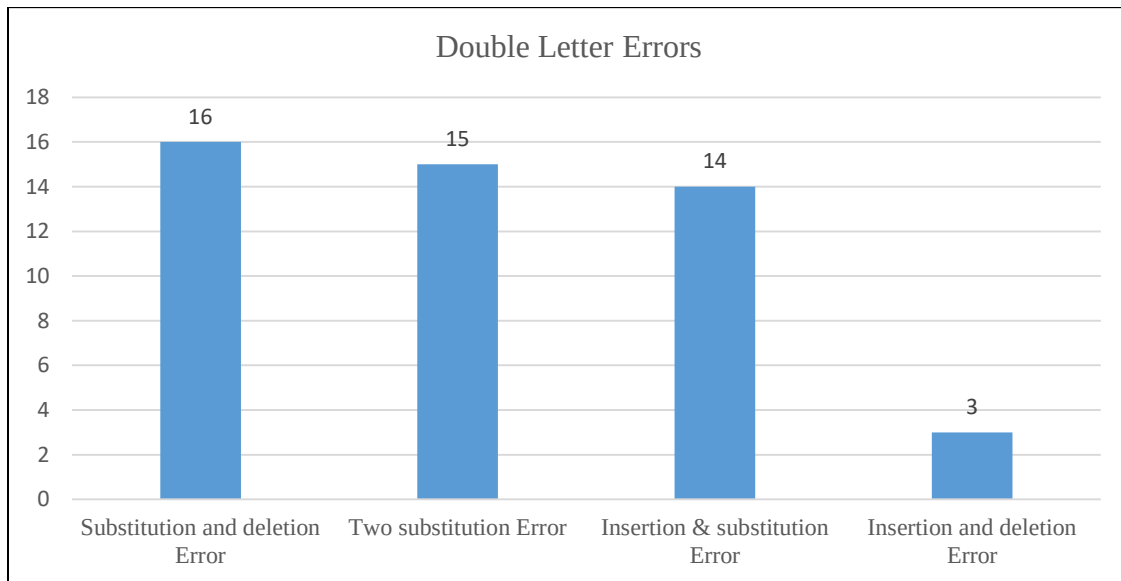
S.No	Gloss	Intended word	Produced word	Insertion Error
1.	love	محبتیں	محبتیں	Letter /ی/ is swapped with /ت/ which is supposed to be placed before /ی/.
2.	use	استعمال	استعمال	Letter /م/ is placed before /ع/, which is supposed to be written after /م/.
3.	muscles	پٹھے	پٹھے	Letter /ٹ/ is changed with /ھ/ which is supposed to be placed after /ٹ/.
4.	poetry	شاعری	شعاری	Letter /ع/ is placed before the letter /ا/, it was supposed to be written before the letter /ر/.

2. Double Letter Errors

Double-letter errors is the second of the two major types of errors. The collected data determines that the majority of the produced two-letter errors were in the categories of substitution, omission/deletion, insertion, and transposition. The main difference in this type of error was that the errors were caused by making two mistakes of the subcategories within a single word. The participants

committed 48 double letter errors out of a total of 313 errors which is 15.33 % of overall errors. The frequency and percentage of the double letter errors have been shown in table 1, graph 1, and a further division of 48 double letter errors is illustrated in figure 3.

Figure 3: Double letter errors



The total of 48 double letter errors is further divided into the following four subcategories;

2.1.Substitution and deletion

The participants produced 16 errors in which they committed two mistakes one after another within a single word. In other words, the substitution and deletion of two different letters have been made in the same lexical item. For instance, a participant wanted to write [سَلَّام] (salary) but mistakenly ended up writing [سَلَّام]. The first mistake was the substitution of phoneme /س/ with /ل/ and the omission of the phone /ا/ was the second mistake, that's how double letters error occurred in this lexical item. 11 out of 16 errors were committed by making substitution first followed by deletion. In vice versa cases, 5 out 16 errors were in which deletion was made first followed by substitution of a letter. For example, writing [مَنْزِل] instead of the correct token [مَنْزِل] (safe). Here, in this case, the letter /ا/ was omitted, and then /ل/ was substituted with /ن/.

2.2.Double letter substitution

Double substitution error occurs when the substitution of two different letters takes place in one word. The data depicts that 15 errors of double substitutions were produced by the participants. For

example, writing [مذبت] instead of the correct token [مذبوط]. In this case, two substitutions have been made; in the first place, the letter /ض/ with /ذ/, and in the second place, the letter /ط/ with /ت/.

2.3. Insertion and substitution

The researcher found a total 14 errors involved insertion and substitution together in a single word. 9 of 14 were in which insertion was made first followed by substitution. For instance, writing [موتابق] instead of the correct token [مطابق]. Here, in this case, a letter /و/ was inserted unnecessarily after /م/, and then the letter /ت/ was substituted with /ط/. The remaining 5 errors were caused by substitution at first followed by the insertion of an unnecessary letter. For example, a participant wrote [زائج] in place of the correct word [زائج]. A phoneme /ز/ was substituted with /ض/ at the initial position and then the phoneme of /ی/ was inserted before the last letter of the word e.g. /ع/.

2.4. Insertion and deletion

A word containing the errors of insertion and deletion together is the fourth subtype of double letters errors. 3 errors of this category emerged in the identification of double letter errors. An example of this would be, writing [چنٹا] instead of the correct token [چنٹا]. In this case, a letter /و/ was added next to the initial letter, and then the phoneme /ن/ was omitted in the same word.

Conclusion:

The main objective of the study was to conduct an error analysis of the errors committed by Pakistani 'O-level' students in Urdu writing. The qualitative nature of the research methodology used in this study offered a wide spectrum of the identification, classification, description, analysis, and explanation of the selected error corpus. The findings of the study reveal that the sample of 50 'O-level' students committed 313 errors in their formative assessment papers. It is pertinent to mention that the identified 313 errors were divided into two major categories; single letter and double letters errors. A reasonable amount of difference has been found in the ratio of a single letter and double letters errors. The data determines that the ratio of the occurrence of single-letter errors was much higher with 84.66 % than double letters errors with 15.33%. The findings disclose that the frequency of substitution errors was also higher at 58.49 % as compared to the other three types of errors in the single letter category. It is found that in a multilingual society like Pakistan interference of mother and other languages, visual or phonetic similarity of alphabets, involvement of diacritical marks, lack of conceptual clarity regarding the shape and sound of letters, and carelessness of the learners caused occurrences of errors in Urdu writing. Therefore, it is recommended that Urdu language teachers should emphasize students' weak areas attributed to committing errors. Moreover, awareness of phonetic and visual similarity causes and regular practice of frequent errors would be helpful to overcome errors in Urdu writing. It is concluded that students without a substantial understanding of basic writing units and teachers' constant and meticulous guidance would not be able to write error-free Urdu.

References

- Ahmad, R. (2011). Urdu in Devanagari: Shifting orthographic practices and Muslim identity in Delhi. *Language in Society*, 40(3), 259-284.
- Brown, H. D. (2000). *Principles of language learning and teaching*. Longman, Inc.
- Burton, M., Davey, J., Lewis, M., Ritchie, L., & Brooks, G. (2010). *Progress for adult literacy learners*. London,, England: National Research and Development Centre for Adult Literacy and Numeracy.
- Carney, E. (2012). *A survey of English spelling*. Routledge.
- Coleman, H. (2010). Teaching and learning in Pakistan: The role of language in education: *The British Council*, 1-56.
- Cook, V. (1999). Teaching spelling. Retrieved February 10, 2013, from <http://homepage.ntlworld.com/vivian.c/ Writings/index.htm>
- Corder, S. P. (1967). The significance of learners' errors. In J.C. Richards (ed.) 1984 Error Analysis: Perspectives on second language acquisition, pp 19 – 27. London: Longman.
- Corder, S. P. (1974). Error Analysis. In J. P. B. Allen and S. P. Corder (eds.), *Techniques in Applied Linguistics* (3rd Ed). Oxford University Press.
- Crystal, D. (2003). *Dictionary of linguistics and phonetics* John Wiley & Sons., 30.
- Dar, M. Anwaar, H. Vihman, M. Keren-Portnoy, T. (2015). Developing an Urdu CDI for early language acquisition. *York Papers in Linguistics*, 14(2), 1-14.
- Denzin, N. K., & Lincoln, Y. S. (Eds.). (2011). *The Sage handbook of qualitative research*. sage.
- Darus,S and Subramanian, K.(2009).Error Analysis of the Written English Essay of Secondary School Students in Malasia: A Case Study *European Journal of Social Sciences*, 8 (3).
- Ellis, R. (1997). *Second language acquisition*: Oxford University Press.
- Ellis, R. (1997). *SLA research and language teaching*. Oxford University Press, 198 Madison Avenue, 10016-4314.
- Flick, U. (2014). Mapping the field. *The SAGE handbook of qualitative data analysis*, 1, 3-18.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Routledge.
- Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools*. A report to Carnegie Corporation of New York.
- Hardie, A. (2003). Developing a tagset for automated part-of-speech tagging in Urdu. In *Corpus Linguistics 2003*.
- Humayoun, M., Hammarström, H., & Ranta, A. (2022). Urdu morphology, orthography and lexicon extraction. *arXiv preprint arXiv:2204.03071*. (<https://www.cambridgeinternational.org/programmes-and-qualifications/cambridge-o-level-urdu-first-language-3247/>)
- Hussain, S., & Afzal, M. (2001, December). Urdu computing standards: Urdu zabta takhti (uzt)

- 1.01. In *Proceedings. IEEE International Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century*. (pp. 223-228). IEEE.
- Iqbal, M., Tahir, B., & Mehmood, M. A. (2021, May). CURE: Collection for urdu information retrieval evaluation and ranking. In *2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2)* (pp. 1-6). IEEE.
- James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. Harlow, Essex: Addison-Wesley Longman.
- Jobeen, A., Kazemian, B., & Shahbaz, M. (2015). The role of error analysis in teaching and learning of second and foreign language. *Education and Linguistics Research*, 1(2), 52-62.
- Keshavarz, M. H. (1999). Contrastive analysis and error analysis.
- Keshavarz, M. H. (2006). *Error analysis: A practical course for English students and teachers*. SAMT, Tehran.
- Khan, S. R., & Khan, M. R. (2016). *Error Analysis in English Writing.. International Multidisciplinary Research Journal*. www.rersearch-chronicler.com ISSN-2347-503X.
- Mirdehghan, M. (2010). Persian, Urdu, and Pashto: A comparative orthographic analysis. *Writing Systems Research*, 2(1), 9-23.
- Rahman, T., *Language and Politics in Pakistan*, Karachi, Oxford University Press, 1998, quoted in Mohanty, A. K., 'Language policy and practice in education: Negotiating the double divide in multilingual societies', in *Multilingual Education Works*, edited by K. Heugh and T. Skutnabb-Kangas, Orient Blackswan, 2010, pp. 164–175.
- Riaz, K. (2008, October). Baseline for Urdu IR evaluation. In *Proceedings of the 2nd ACM workshop on Improving non English web searching* (pp. 97-100).
- Richards, J. (1974). *Error Analysis and Second Language Strategies*. Longman.
- Sabir, I. (2013). A Linguistic Analysis of Misspelled Words in Urdu. *Language in India*, 13(3).
- Schumann, J. H., & Stenson, N. (1974). *New Frontiers in Second Language Learning*.
- Touchie, H. Y. (1986). Second language learning errors: Their types, causes, and treatment. *JALT journal*, 8(1), 75-80.
- Ziahosseiny, S. M. (1999). A contrastive analysis of Persian and English and error analysis. *Tehran: Nashr-e Vira*.