

Roman Urdu and its Effects on Standard Urdu

Naeem Abbas

naeemabbasrajpoot@gmail.com

Tahira Khan

tahira779@gmail.com

Iqra Girls College Sargodha

Zoya Khalid

zoyakhalid4513@gmail.com

Corresponding Author: Naeem Abbas naeemabbasrajpoot@gmail.com

Received: 15-01-2026

Revised: 02-02-2026

Accepted: 17-02-2026

Published: 05-03-2026

ABSTRACT

This study examines the impact of Roman Urdu on standard Urdu through a quantitative, corpus-based analysis of spelling variations. Data collected from 50 Urdu-speaking participants reveal significant orthographic inconsistency due to the lack of standardized rules in Roman Urdu. The findings show that while certain common spellings dominate, multiple variations coexist for the same words, reflecting phonetic interpretation and individual preference. The study also highlights that gender has minimal influence on spelling choices. Results indicate that Roman Urdu facilitates fast and convenient digital communication but may weaken proficiency in standard Urdu script. The study concludes that a balanced approach is necessary, promoting Roman Urdu in informal contexts while preserving standard Urdu in formal and academic settings.

Keywords: Roman Urdu, Standard Urdu, Orthographic Variation, Digital Communication, Language Standardization

INTRODUCTION

Roman Urdu refers to writing Urdu language by using English alphabets instead of Perso - Arabic script. It has transformed a limited, informal variety of language into a broadly used and influential variety in a digital discourse. It has vital influence on the standardization and pedagogical usage of standard Urdu. Roman Urdu dominated Perso- Arabic script in online settings due to ease of English keyboard and the absence of native script support in older mobile technology. Roman Urdu has shortage of fixed orthography which results in higher variation of spellings. Roman Urdu is leading to a perceived shortage in the fluency of reading and writing the conventional Urdu script among youth. The rare use of a standard Urdu lead to the difficulty of text mining tasks. With the increasing growth of WhatsApp, YouTube and other online messaging services users commonly pursue writing Urdu in the Roman script instead of conventional Nastaliq script. This shift is because of technological convenience, speed of typing and the extensive exposure to English alphabets. The other reason of predominantly used of roman Urdu is that digital devices are actually designed for English keyboard and it is easy to type English alphabets rather than Urdu script. The standard Urdu follows formal and conventional spellings, rules and script norms while Roman Urdu does not follow conventional rules, this creates a confusion weakening learner accuracy in formal Urdu writing.

It also believed that it is a natural transformation of language. Languages constantly change due to its dynamic nature and social situations, but it also threatening the Urdu script in formal settings. Therefore a moderate approach is needed. Roman Urdu should be used in informal digital communication settings while standard Urdu should be used in formal and academic settings.

Roman Urdu Definition

According to David crystal (2001) “Internet communication encourages script innovation and flexible spelling systems which explains the rise of forms like Roman Urdu”. Roman Urdu is the practice of writing Urdu language using English alphabets especially at the social media. It is commonly used in south Asian countries for digital communication. According to Paolillo (2007), "Roman Urdu is the transliteration of Urdu into the Roman alphabet used in computer mediated communication. One of reason of its use is due to limitations of keyboard in modern digital devices especially for the Nastaliq script. Roman Urdu has non –standard language because its spellings are not fixed, users write it based on their phonetic perception. According to Rahman (1996) “Language choice and script use in Pakistan are influenced by social, educational technological factors”

Why Roman Urdu is used by people attached with social media?

Mostly digital devices are commonly designed for English keyboard and it is easy to type English alphabets rather than Urdu script. It also provides flexibility to people who cannot read Urdu script properly can use Roman Urdu, Roman Urdu is commonly used on social media devices for creating clear tone. Roman Urdu is also widely used because of extensive use of code switching at social media.

One significant reason is the simplicity of typing. In modern digital devices, Nastaliq script and fonts are not easily accessible, so users mostly selected Roman alphabets. Roman Urdu is widely used because it can easily be typed on mobile phones and computers. “Online writings often reflects non –standard spellings and innovative orthographies as part of digital identity construction” (Androutsopoulos, 2011)

Romanized Urdu does not have standardized rules and spellings which makes it flexible, adaptable and creative. In post –colonial countries where educations and institutions are influenced by English. Mixing Urdu with English shows their bilingual nature. Roman Urdu allows for faster on –the –go typing rather than the Urdu script which makes it ideal for the widely use. According to Kachru (1985): “The use of English script for Asian languages reflects sociolinguistics adaptation and bilingual practices in postcolonial societies”

Many younger and diaspora speakers can speak Urdu fluently but cannot read traditional Urdu script, Roman Urdu allows them to express themselves digitally by using roman script. Companies also use Roman Urdu in their advertisement, as it is easily understood by rural and urban people. “Computer mediated communication promotes informal, rapid and speech like writing styles which supports the use of Romanized form” (Baron, 2008)

Does Roman Urdu inflect standard Urdu in bad way?

Many scholars claimed that wide use of Roman Urdu reduces familiarity with Nastaliq script, which is the traditional writing system of Urdu. Urdu is written in the Nastaliq script. Critics argue replacing it with the Latin alphabet serves the link to centuries of literary heritage and cultural identity. According to Gunther Kress (2010) "The changes in writing systems affect how people learn and process language". According

to Safder et al. (2020) “The lack of standardization in Roman Urdu poses significant challenges for digital information retrieval and language analysis”.

Roman Urdu does not have fixed orthography so the spellings may vary. This makes learner of standard Urdu confused about the original spellings. Roman Urdu may violate the grammar rule due to faster typing. Roman Urdu users mix English and Urdu which is natural in bilingual societies but may weakens the formal Urdu writing. Its extensive use especially in formal situations harm the standard Urdu.

Zareen Sharf (2019) argues that analysis techniques and lexical resources for Roman Urdu have significant limitations due to the complete lack of a standard orthography, which severely affects the accuracy of any linguistic processing. Roman Urdu influences orthography informally but it is not inherently bad, problem arise when learner rely extensively on Roman Urdu and use it in formal situations and academic texts .Roman Urdu significantly affecting standard Urdu in various ways especially in digital communication. The Roman Urdu has not standardized orthography this results several spellings based on the user that how they hear or type it which lower the quality of conventional, standardized Urdu. Using Roman Urdu results in English like pronunciation which blocks the native like pronunciation.

Digital writing often encourages informal spelling practices, which may reduce attention to standard orthography in formal contexts.” (Baron, 2008) Roman Urdu also creates challenges for NLP tasks because it is non- standard and making machine learning models less effective to understand the language. Roman Urdu is also considered the assisted suicide for language because it separate the phonetic beauty and rich literary script from language and transformed it into simplified Latin script.

Javed T. and Bhatti, M. (2015) they concludes that Urdu script is the best script for representing the Urdu language. Roman script was found inadequate for accurately capturing Urdu unique phonetics. Frequent use of Roman Urdu decreases the reading and writing practice in traditional Urdu script which is reducing the exposure to Nastaliq script. According to Rabia Khan (2025) “Roman Urdu is a slow cultural poison, destroying the ability of next generation to properly read and write in Urdu.”

RESEARCH METHODOLOGY

Research Design

The present study uses a quantitative corpus-based research design to explore orthographic variation in Roman Urdu. The pivotal aim of the study is to recognize and evaluate different spellings of selected Urdu words when written in Roman script. The research is descriptive in nature. It systematically explores different patterns in participants’ responses.

Participants

There are 50 participants that were approached through convenience sampling. All the participants were Urdu speakers with different educational background. This variation allows for a broader representation of spelling variation in Roman Urdu.

Data Collection Method

Elicitation technique was used to collect data. Respondents were asked to write few sentences (کیا حال ہے؟) (روزہ کیسا گزر رہا ہے؟ دنیا میں مختلف زبانیں بولی جاتی ہیں، اردو ایک شائستہ اور معزز زبان ہے۔) in Roman Urdu. Through this technique, it was ensured that data was naturally produced.

Corpus Development

All the data were filed into a self-constructed corpus of Roman Urdu text. Later on, it was ordered for analysis. The spelling variation of Urdu words were identified within the corpus.

Data Analysis Procedure

Orthographic variation analysis was used for the analysis of the corpus. The analysis was done step by step:

- Urdu words written in Roman script was extracted
- Spelling differences of the same word was identified
- On the basis of orthographic differences words were categorized
- Frequencies of each word was calculated

Descriptive statistics, including frequency counts and percentages, were used to present the variation patterns.

ANALYSIS

In this section the researcher is going to describe the spelling variation used by the respondents to write Roman Urdu sentence. The sentences in Urdu language are;

کیا حال ہے؟ روزہ کیسا گزر رہا ہے؟ دنیا میں مختلف زبانیں بولی جاتی ہیں، اردو ایک شائستہ اور معزز زبان ہے۔

Description of the Analysis

The sentence in Urdu language is; **کیا حال ہے؟**

Male Respondents

Description of spelling in Roman Urdu of word **کیا**:

Word	Frequency	Percentage
Kia	13	54%
Kya	5	20%
Kiya	3	12%
Kay	1	4%
Kaya	2	8%

The majority of male respondents (13 out of 24) used the spelling (KIA). It is 54% Of the male respondents, (5 out of 24) male respondents (20%) used the spelling (kya), the (3 out of 24) male respondents (12%) use the word (kiya), the (2 out of 24) male respondents (8%) use the word (kaya), the (1 out of 24) male respondents (4%) use the word (kay).

Description of spelling in Roman Urdu of word (حال):

Hal	11	45%
Hall	7	29%
Haal	6	25%

The majority of male respondents (11 out of 24) used the spelling (hal). It is 45% of the male respondents, (7 out of 24) male respondents (29%) used the spelling (hall), the (6 out of 24) male respondents (25%) used the spelling (haal).

Description of spelling in Roman Urdu of word (ہا):

Ha	14	58%
Hai	6	25%
Hy	2	8%
Hain	1	4%
Hay	1	4%

The majority of male respondents (14 out of 24) used the spelling (ha). It is 58% of the male respondents, the (6 out of 24) male respondents (25%) used the spelling (hai), the (2 out of 24) male respondents (8%) used the spelling (hy), the (1 out of 24) male respondents (4%) used the spelling (hain), the (1 out of 24) male respondents (4%) used the spelling (hay).

Female respondents

Description of spelling in Roman Urdu of word (کیا):

Word	Frequency	percentage
Kia	16	61%
Kya	5	19%
Kaya	2	7%
Kiya	2	7%

Kyaa	1	3%
------	---	----

The majority of female respondents (16 out of 26) used the spelling (KIA). It is 61% of the female respondents, (5 out of 26) female respondents (19%) used the spelling (kya), the (2 out of 26) female respondents (7%) use the spelling (kaya), the (2 out of 26) female respondents (7%) use the spelling (kiya), the (1 out of 26) female respondents (3%) use the spelling (kyaa).

Description of spelling in Roman Urdu of word (حال):

Hal	14	53%
Hall	7	26%
Haal	5	19%

The majority of female respondents (14 out of 26) used the spelling (hal). It is 53% of the female respondents, the (7 out of 26) female respondents (26%) use the spelling (hall), the (5 out of 26) female respondents (19%) used the spelling (haal).

Description of spelling in Roman Urdu of word (ہائی):

Hai	10	38%
Ha	9	34%
Hy	4	15%
Hn	1	3%
Hae	1	3%
Hay	1	3%

The majority of female respondents (10 out of 26) used the spelling (hai). It is 38% of the female respondents, the (9 out of 26) female respondents (34%) used the spelling (ha), the (4 out of 26) female respondents (15%) used the spelling (hy), the (1 out of 26) female respondents (3%) used the spelling (hn), the (1 out of 26) female respondents (3%) used the spelling (hae), the (1 out of 26) female respondents (3%) used the spelling (hay).

"The word (Kia) is used 13 times by the male respondents and 16 times by the female respondents. Its cumulative frequency is 58%".

"The word (hal) is used 11 times by the male respondents and 14 times by the female respondents. Its cumulative frequency is 50%".

"The word (ha) is used 14 times by the male respondents and 9 times by the female respondents. Its cumulative frequency is 46%".

"The word (hai) is used 6 times by the male respondents and 10 times by the female respondents. Its cumulative frequency is 32%".

The sentence in Urdu language روزہ کیسا گزر رہا ہے؟

Description of spelling in Roman Urdu of word (روزہ):

Male respondents

Roza	23	95%
Razo	1	4%

The majority of male respondents (23 out of 24) used the spelling (Roza). It is 95% of the male respondents, the (1 out of 24) male respondents (4%) used the spelling (Razo). It seems that, this spelling is the result of autocorrect system. Otherwise it is impossible to think of such spellings.

Description of spelling in Roman Urdu of word (کیسا):

Kesa	9	37%
Kasa	7	29%
Kaisa	5	20%
Kaysa	1	4%
Kasy	1	4%
Kasay	1	4%

The majority of male respondents (9 out of 24) used the spelling (kesa). It is 37% of the male respondents. The (7 out of 24) male respondents (29%) used the spelling (kasa), the (5 out of 24) male respondents (20%) used the spelling (kaisa), the (1 out of 24) male respondents (4%) used the spelling (kaysa), the (1 out of 24) male respondents 4% used the spelling (kasy), the (1 out of 24) male respondents (4%) used the spelling (kasay).

Description of spelling in Roman Urdu of word (گزر):

Guzar	13	54%
Guzr	4	16%
Gozar	2	8%
Ghuzar	1	4%
Ghuzhar	1	4%

Gujar	1	4%
Guz	1	4%

The majority of male respondents (13 out of 24) used the spelling (Guzar). It is 54% of the male respondents. The (4 out of 24) male respondents (16%) used the spelling (Guzr), the (2 out of 24) male respondents (8%) used the spelling (Gozar), the (1 out of 24) male respondents (4%) used the spelling (Ghuzar), the (1 out of 24) male respondents (4%) used the spelling (Ghuzhar), the (1 out of 24) male respondents (4%) used the spelling (Gujar), the (1 out of 24) male respondents (4%) used the spelling (Guz). The spelling (Guz) is also the result of auto correction.

Description of spelling in Roman Urdu of word (جڑ):

Raha	19	79%
Rha	4	16%
Reha	1	4%

The majority of male respondents (19 out of 24) used the spelling (raha). It is 79% of male respondents. The (4 out of 24) male respondents (16%) used the spellings (rha), the (1 out of 24) male respondents (4%) used the spellings (Reha).

Description of spelling in Roman Urdu of word (ہا):

Ha	14	58%
Hai	7	29%
Hy	2	8%
Hain	1	4%

The majority of male respondents (14 out of 24) used the spellings (ha). It is the 58% of the male majority. The (7 out of 24) male respondents (29%) used the spellings (hai), the (2 out of 24) male respondents (8%) used the spellings (hy), the (1 out of 24) male respondents (4%) used the spellings (hain).

Female Respondents

Roza	24	92%
Rohza	1	3%
Raza	1	3%

The majority of female respondents (24 out of 26) used the spelling (Roza). It is 92% of the female respondents. The (1 out of 26) female respondents (3%) used the spelling (Rohza), the (1 out of 26) female respondents (3%) used the spelling (Raza). It is also the result of auto correction.

Description of spelling in Roman Urdu of word (کیسا):

Kesa	10	38%
Kasa	8	30%
Kaisa	3	12%
Kahsa	1	3%
Ksa	1	3%
Kiasa	1	3%
Kaysa	1	3%
Kase	1	3%

The majority of female respondents (10 out of 26) used the spelling (kesa). It is 38% of female respondents. The (8 out of 26) female respondents (30%) used the spelling (kasa), the (3 out of 26) female respondents 11% use the spelling (kaisa), the (1 out of 26) female respondents (3%) used the spelling (kahsa), the (1 out of 26) female respondents (3%) used the spelling (ksa), the (1 out of 26) female respondents (3%) used the spelling (kiasa), the (1 out of 26) female respondents (3%) used the spelling (kaysa), the (1 out of 26) female respondents (3%) used the spellings (kase).

Description of spelling in Roman Urdu of word (گزر):

Guzar	14	53%
Guzr	5	19%
Guzer	2	7%
Guhzar	1	3%
Ghuzar	1	3%
Gozar	1	3%
Gozr	1	3%
Guzzar	1	3%

The majority of female respondents (14 out of 26) used the spelling (Guzar). It is the 53% of the female respondents. The (5 out of 26) female respondents (19%) used the spelling (Guzr), the (2 out of 26) female respondents (7%) used the spelling (Guzer), the (1 out of 26) female respondents (3%) used the spelling (Guhzar), the (1 out of 26) female respondents (3%) used the spelling (Ghuzar), the (1 out of 26) female respondents (3%) used the spelling (Gozar), the (1 out of 26) female respondents (3%) used the spelling (Gozr), the (1 out of 26) female respondents (3%) used the spellings (Guzzar).

Description of spelling in Roman Urdu of word (راہ):

Raha	18	69%
Rha	4	15%
Rah	2	7%
Ra	2	7%

The majority of female respondents (18 out of 26) used the spelling (raha). It is 69% of the female respondents. The (4 out of 26) female respondents (15%) used the spellings (rha), the (2 out of 26) female respondents (7%) used the spellings (rah), the (2 out of 26) female respondents (7%) used the spellings (ra).

Description of spelling in Roman Urdu of word (ہا):

Ha	10	38%
Hai	10	38%
Hy	3	11%
Hn	1	3%
Hae	1	3%
Hay	1	3%

The majority of female respondents (10 out of 26) used the spellings (ha). It is 38% of the female respondents. The (10 out of 26) female respondents (38%) used the spellings (hai), the (3 out of 26) female respondents (11%) used the spellings (hy), the (1 out of 26) female respondents (3%) used the spellings (HN), the (1 out of 26) female respondents (3%) used the spellings (hae), the (1 out of 26) female respondents (3%) used the spellings (hay).

"The word (Roza) is used 23 times by male respondents and 24 times by female respondents. Its cumulative frequency is 94%."

"The word (Kesa) is used 9 times by the male respondents and 10 times by the female respondents. Its cumulative frequency is 38%."

"The word (Guzar) is used 13 times by male respondents and 14 times by the female respondents. Its cumulative frequency is 54%."

"The word (raha) is used 19 times by the male respondents and 18 times by the female respondents. Its cumulative frequency is 74%."

"The word (ha) is used 14 times by the male respondents and 10 times by the female respondents. Its cumulative frequency is 48%."

"The word (hai) is used 7 times by the male respondents and 10 times by the female respondents. Its cumulative frequency is 34%."

The sentence in Urdu is:

دنیا میں مختلف زبانیں بولی جاتی ہیں۔

Description of spelling in Roman Urdu of word (دنیا):

Male respondents

Duniya	8	33%
Dunya	7	29%
Dunia	5	20%
Dynia	1	4%
Duna	1	4%
Duniyan	1	4%
Duniha	1	4%

The majority of male respondents (8 out of 24) used the spellings (Duniya). It is 33% of the male majority. The (7 out of 24) male respondents 29% used the spellings (Dunya), the (5 out of 24) male respondents 20% used the spellings (Dunia), the (1 out of 24) male respondents 4% used the spellings (Dynia), the (1 out of 24) male respondents 4% used the spellings (Duna), the (1 out of 24) male respondents (4%) used the spellings (Duniyan), the (1 out of 24) male respondents (4%) used the spelling (Duniha)

Description of spelling in Roman Urdu of word (میں):

Ma	13	54%
Main	3	12%
May	2	8%
Me	2	8%
Mein	2	8%
Myn	1	4%
Mai	1	4%

The majority of male respondents (13 out of 24) used the spellings (ma). It is the 54% of the male respondents. The (3 out of 24) male respondents (12%) used the spellings (main), the (2 out of 24) male

respondents (8%) used the spellings (may), the (2 out of 24) male respondents (8%) used the spellings (me), the (2 out of 24) male respondents (8%) used the spellings (mein), the (1 out of 24) male respondents (4%) used the spellings (myn), the (1 out of 24) male respondents (4%) used the spellings (mai).

Description of spelling in Roman Urdu of word (مختلف):

Mukhtalif	11	45%
Muktalif	6	25%
Mukhtlif	1	4%
Mukehtalif	1	4%
Mulktulaf	1	4%
Mutlif	1	4%
Mokhtalif	1	4%
Mukhtlaf	1	4%
Multif	1	4%

The majority of male respondents (11 out of 24) used this spelling (mukhtalif). It is the 45% of male respondents. The (6 out of 24) male respondents (25%) used the spelling (Muktalif), the (1 out of 24) male respondents (4%) used the spelling (Mukhtlif), The (1 out of 24) male respondents (4%) used the spelling (Mukehtalif), the (1 out of 24) male respondents (4%) used the spelling (Mulktulaf), The (1 out of 24) male respondents (4%) used the spelling (Mutlif), The (1 out of 24) male respondents (4%) used the spelling (Mokhtalif), The (1 out of 24) male respondents (4%) used the spellings (Mukhtlaf), the (1 out of 24) male respondents (4%) used the spellings (Multif).

Description of spelling in Roman Urdu of word (زبانیں):

Zabany	3	12%
Zubane	3	12%
Zubany	2	8%
Zubanain	2	8%
Zubanay	2	8%
Zubna	2	8%
Zabane	1	4%

Zabana	1	4%
Zubana	1	4%
Zabny	1	4%
Zabanai	1	4%
Zabunay	1	4%
Zoubana	1	4%
Zubanyn	1	4%
Zubanin	1	4%
Zubana	1	4%

The majority of male respondents (3 out of 24) used the spelling (Zabany). It is the 12% of male respondents. The (3 out of 24) male respondents (12%) used the spelling (Zubane). The (2 out of 24) male respondents (8%) used the spelling (Zubany). The (2 out of 24) male respondents (8%) used the spelling (Zubanain). The (2 out of 24) male respondents (8%) used the spelling (Zabunay). The (2 out of 24) male respondents (8%) used the spelling (Zubna). The (1 out of 24) male respondents (4%) used the spelling (Zabane). The (1 out of 24) male respondents (4%) used the spelling (Zabana). The (1 out of 24) male respondents (4%) used the spelling (Zubana). The (1 out of 24) male respondents (4%) used the spelling (zabny). The (1 out of 24) male respondents (4%) used the spelling (Zabanai). The (1 out of 24) male respondents (4%) used the spelling (Zabunay). The (1 out of 24) male respondents (4%) used the spelling (Zoubana). The (1 out of 24) male respondents (4%) used the spelling (Zubanyn). The (1 out of 24) male respondents (4%) used the spelling (Zubanin). The (1 out of 24) male respondents (4%) used the spelling (Zubana).

Description of spelling in Roman Urdu of word (بولى):

Boli	18	75%
Bole	3	12%
Bola	2	8%
Bolauy	1	4%

The majority of male respondents (18 out of 24) used the spelling (boli). It is the 75% of the male respondents. The (3 out of 24) male respondents (12%) used the spelling (bole), the (2 out of 24) male respondents (8%) used the spelling (bola), the (1 out of 24) male respondents (4%) used the spelling (Bolauy).

Description of spelling in Roman Urdu of word جاتی

Jati	17	70%
Gati	4	16%
Jate	1	4%
Jatay	1	4%
Jadhi	1	4%

The majority of male respondents (17 out of 24) used the spelling (jati). It is the 71% of the male respondents. The (4 out of 24) male respondents (17%) used the spelling (Gati), the (1 out of 24) male respondents (4%) used the spelling (Jate), the (1 out of 24) male respondents (4%) used the spelling (Jadhi). The spelling (Jadhi) is the result of auto correct.

Description of spelling in Roman Urdu of word ہا

Ha	9	37%
Hain	9	37%
Hai	3	12%
Han	2	8%
Hy	1	4%

The majority of male respondents (9 out of 24) used the spelling (ha). It is the 37% of the male respondents. The (9 out of 24) male respondents (37%) used the spelling (hain), the (3 out of 24) male respondents (12%) used the spelling (hai), the (2 out of 24) male respondents (8%) used the spelling (Han), the (1 out of 24) male respondents (4%) used the spelling (hy).

Female Respondents

Description of spelling in Roman Urdu of word دنیا:

Duniya	13	50%
Dunya	5	19%
Dunia	4	15%
Dunian	1	3%
Donia	1	3%

Dinya	1	3%
Duniaya	1	3%

The majority of female respondents (13 out of 26) used the spellings (Duniya). It is the 50% of the female majority. The (5 out of 26) female respondents (19%) used the spellings (Dunya), the (4 out of 26) female respondents (15%) used the spellings (Dunia), the (1 out of 26) female respondents (3%) used the spellings (Dunian), the (1 out of 26) female respondents (3%) used the spellings (Donia), the (1 out of 26) female respondents (3%) used the spelling (Dinya), the (1 out of 26) female respondents (3%) used the spellings (Duniaya).

Description of spelling in Roman Urdu of word (میں):

Ma	7	26%
Main	6	23%
May	4	15%
Mein	3	11%
Man	1	3%
Maa	1	3%
My	1	3%
Mae	1	3%
Mn	1	3%
Me	1	3%

The majority of female respondents (7 out of 26) used the spellings (ma). It is the 26% of the female respondents. The (6 out of 26) female respondents (23%) used the spellings (main), the (4 out of 26) female respondents (15%) used the spellings (may), the (3 out of 26) female respondents (11%) used the spellings (mein), the (1 out of 26) female respondents (3%) used the spellings (man), the (1 out of 26) female respondents (3%) used the spellings (maa), the (1 out of 26) female respondents (3%) used the spellings (my), the (1 out of 26) female respondents (3%) used the spellings (mae), the (1 out of 26) female respondents (3%) used the spellings (mn), the (1 out of 26) female respondents (3%) used the spellings (me).

Description of spelling in Roman Urdu of word (مختلف):

Mukhtalif	7	26%
Muktalif	5	19%

Muktlif	4	15%
Mukhtlif	2	7%
Muktalf	1	3%
Muktaleef	1	3%
Mukhtalf	2	7%
Mukhtalif	1	3%
Muktlaf	1	3%
Muhaktalif	1	3%
Different	1	3%

The majority of Female respondents (7 out of 26) used the spelling (Mukhtalif). It is the 26% of the female respondents. The (5 out of 26) female respondents (19%) used the spelling (Muktalif). The (4 out of 26) female respondents (15%) used the spelling (Mukhtalif). The (2 out of 26) female respondents (7%) used the spelling (Mukhtlif). The (2 out of 26) female respondents (7%) used the spelling (Mukhtalf). The (1 out of 26) female respondents (3%) used the spelling (Muktalf). The (1 out of 26) female respondents (3%) used the spelling (Muktaleef). The (1 out of 26) female respondents (3%) used the spelling (Mukhtalif). The (1 out of 26) female respondents (3%) used the spelling (Muktlaf). The (1 out of 26) female respondents (3%) used the spelling (Muhaktalif). The (1 out of 26) female respondents (3%) used the spelling (different).

Description of spelling in Roman Urdu of word (زبانیں)

Zubanain	5	19%
Zubanay	4	15%
Zubany	4	15%
Zabany	4	15%
Zubana	1	3%
Zabainy	1	3%
Zabanay	1	3%
Zabaneiy	1	3%
Zabane	2	7%
Zubanya	1	3%

Zubainay	1	3%
Zabna	1	3%

The majority of female respondents (5 out of 26) used the spelling (Zubanain). It is the 19% of the female respondents. The (4 out of 26) female respondents (15%) used the spelling (Zubany), the (4 out of 26) female respondents (15%) used the spelling (Zabany), the (4 out of 26) female respondents (15%) used the spellings (Zubanay), the (2 out of 26) female respondents (7%) used the spelling (Zabane), the (1 out of 26) female respondents (3%) used the spellings (Zubana), the (1 out of 26) female respondents (3%) used the spellings (Zabainy), the (1 out of 26) female respondents (3%) used the spelling (Zabanay), the (1 out of 26) female respondents (3%) used the spelling (Zabaneiy), the (1 out of 26) female respondents (3%) used the spelling (Zubanya), the (1 out of 26) female respondents (3%) used the spelling (Zubainay), the (1 out of 26) female respondents (3%) used the spelling (Zabna).

Description of spelling in Roman Urdu of word بولی:

Boli	25	96%
bolie	1	3%

The majority of female respondents (25 out of 26) used the spelling (boli). It is the 96% of the female respondents. The (1 out of 26) female respondents (3%) used the spelling (bolie).

Description of spelling in Roman Urdu of word جاتی:

Jati	24	92%
Jaty	1	3%
Jatti	1	3%

The majority of female respondents (24 out of 26) used the spelling (jati). It is the 92% of the female respondents. The (1 out of 26) female respondents (3%) used the spelling (Jaty). The (1 out of 26) female respondents (3%) used the spelling (Jatti).

Description of spelling in Roman Urdu of word ہائ:

Hain	10	38%
Hai	6	23%
Ha	5	19%
Hy	2	7%
Hn	1	3%
Han	1	3%

Hay	1	3%
-----	---	----

The majority of female respondents (10 out of 26) used the spelling (hain). It is the 38% of the female respondents. The (6 out of 26) female respondents (23%) used the spelling (hai), the (5 out of 26) female respondents (19%) used the spelling (ha), the (2 out of 26) female respondents (7%) used the spelling (hy), the (1 out of 26) female respondents (3%) used the spelling (HN), the (1 out of 26) female respondents (3%) used the spelling (Han), the (1 out of 26) female respondent (3%) used the spelling (Hay).

"The word (Duniya) is used by 8 males and 13 females. Its cumulative frequency is 42%."

"The word (Dunya) is used by 7 males and 5 females. Its cumulative frequency is 24%."

"The word (Dunia) is used by 5 males and 4 females. Its cumulative frequency is 18%."

"The word (Mukhtalif) is used by 11 males and 7 females. Its cumulative frequency is 36%."

"The word (muktalif) is used by 6 males and 5 females. Its cumulative frequency is 22%."

"The word (Boli) is used by 18 males and 25 females. Its cumulative frequency is 86%."

The sentence in Urdu is:

"اردو ایک شانستہ اور معزز زبان ہے"

Description of spelling in Roman Urdu of word (اردو)

Male Respondents

Urdu	24	100%
------	----	------

The majority of male respondents (24 out of 24) used the spelling (Urdu). It is 100% of the male respondents.

Description of spelling in Roman Urdu of word ایک:

Aik	10	41%
Ik	8	33%
Ak	4	16%
Aikh	1	4%
Ek	1	4%

The majority of male respondents (10 out of 24) used the spelling (Aik). It is the 41% of the male respondents. The (8 out of 24) male respondents (33%) used the spelling (Ik), the (4 out of 24) male respondents (16%) used the spelling (AK), the (1 out of 24) male respondents (4%) used the spelling (Aikh), the (1 out of 24) male respondents (4%) used the spelling (ek).

Description of spelling in Roman Urdu of word شائستہ:

Shaista	13	54%
Shatsia	2	8%
Shayista	1	4%
Sasta	1	4%
Shahta	1	4%
Saista	1	4%
Shaisata	1	4%
Shaishta	1	4%
Shaysta	1	4%
Shatsia	1	4%
Shahista	1	4%
Shasta	1	4%

The majority of male respondents (13 out of 24) male respondent used the spelling (Shaista). It is the 55% of the male respondents, the (2 out of 24) male respondents (4%) used the spelling (Shatsia), the (1 out of 24) male respondents (4%) used the spelling (Shayista), the (1 out of 24) male respondents (4%) used the spelling (Sasta), the (1 out of 24) male respondents (4%) used the spelling (Shahta), the (1 out of 24) male respondents (4%) used the spelling (Saista), the (1 out of 24) male respondents (4%) used the spelling (Shaisata), the (1 out of 24) male respondents (4%) used the spelling (Shaishta), the (1 out of 24) male respondents (4%) used the spelling (Shaysta), the (1 out of 24) male respondents (4%) used the spelling (Shahista), the (1 out of 24) male respondents (4%) used the spelling (Shasta).

Description of spelling in Roman Urdu of word اور:

Or	17	70%
Aur	4	16%
Ore	1	4%
Auz	1	4%
Our	1	4%

The majority of male respondents (17 out of 24) used the spelling (or). It is the 70% of the male respondents. The (4 out of 24) male respondents (16%) used the spelling (Aur), the (1 out of 24) male respondents (4%)

used the spelling (ore), the (1 out of 24) male respondents (4%) used the spelling (Auz), the (1 out of 24) male respondents (4%) used the spelling (our).

Description of spelling in Roman Urdu of word معزز:

Muaziz	9	37%
Muhazib	2	8%
Moaziz	2	8%
Muazzaz	1	4%
Mauza	1	4%
Moaj	1	4%
Moizez	1	4%
Mozize	1	4%
Moziz	2	8%
Muiziz	1	4%
Muazziz	1	4%
Muazaz	1	4%
Muhaz	1	4%

The majority of male respondents (9 out of 24) used the spelling (Muaziz). It is the 37% of male respondents. The (2 out of 24) male respondents (8%) used the spelling (Moaziz), the (2 out of 24) male respondents (8%) used the spelling (Muhazib). This spelling (Muhazib) shows the misunderstanding or disability of the respondents as it is the spelling of another Urdu word (مہذب) the (2 out of 24) male respondents (8%) used the spelling (Moziz), the (1 out of 24) male respondents (4%) used the spelling (Muazzaz), the (1 out of 24) male respondents (4%) used the spelling (Mauza), the (1 out of 24) male respondents (4%) used the spelling (Moaj), the (1 out of 24) male respondents (4%) used the spelling (Moizez), the (1 out of 24) male respondents (4%) used the spelling (Mozize), the (1 out of 24) male respondents (4%) used the spelling (Muiziz), the (1 out of 24) male respondents (4%) used the spelling (Muazziz), the (1 out of 24) male respondents (4%) used the spelling (Muazaz), the (1 out of 24) male respondents (4%) used the spelling (Muhaz).

Description of spelling in Roman Urdu of word زبان:

Zuban	11	45%
Zaban	10	41%
Zabuna	1	4%
Zubain	1	4%

The majority of male respondents (11 out of 24) used the spelling (Zuban). It is the 45% of male respondents. The (10 out of 24) male respondents (41%) used the spelling (Zaban), the (1 out of 24) male respondents (4%) used the spelling (Zabuna), the (1 out of 24) male respondents (4%) used the spelling (Zubain).

Description of spelling in Roman Urdu of word سمجھی:

Samji	12	50%
Samjhi	7	29%
Samje	1	4%
Sumgi	1	4%
Sajmu	1	4%
Samjhe	1	4%

The majority of male respondents (12 out of 24) used the spelling (Samji). It is the 50% of the male respondents. The (7 out of 24) male respondents (29%) used the spelling (Samjhi), the (1 out of 24) male respondents (4%) used the spelling (Samje), the (1 out of 24) male respondents (4%) used the spelling (Sumgi), the (1 out of 24) male respondents (4%) used the spelling (Sajmu), the (1 out of 24) male respondents (4%) used the spelling (Samjhe).

Description of spelling in Roman Urdu of word جاتی:

Jati	20	83%
Jatay	1	4%
Gati	3	12%

The majority of male respondents (20 out of 24) used the spelling (jati). It is the 83% of male respondents. The (3 out of 24) male respondents (12%) used the spelling (Gati), the (1 out of 24) male respondents (4%) used the spelling (Jatay).

Description of spelling in Roman Urdu of word :

Ha	14	58%
Hai	8	33%
Hain	1	4%
Hy	1	4%

The majority of male respondents (14 out of 24) used the spelling (ha). It is 58% of the male respondents. The (8 out of 24) male respondents (33%) used the spelling (hai), the (1 out of 24) male respondents (4%) used the spelling (hain), the (1 out of 24) male respondents (4%) used the spelling (hy).

Female Respondents

Description of spelling in Roman Urdu of word 

Urdu	26	100%
------	----	------

The majority of female respondents (26 out of 26) used the spelling (Urdu). It is 100% of the female respondents

Description of spelling in Roman Urdu of word :

Ak	10	38%
Aik	10	38%
ek	4	15%
Ik	2	7%

The majority of female respondents (10 out of 26) used the spelling (AK). It is the 38% of the female respondents. The (10 out of 26) female respondents (38%) used the spelling (Aik), the (4 out of 26) female respondents (15%) used the spelling (ek), the (2 out of 26) female respondents (7%) used the spelling (Ik).

Description of spelling in Roman Urdu of word :

Shaista	18	69%
Shahista	2	7%
Sahaista	1	3%
Sahseta	1	3%
Shayista	1	3%

Shaesta	1	3%
Shastha	1	3%
Shista	1	3%

The majority of female respondents (18 out of 26) used the spelling (Shaista). It is the 69% of the female respondents. The (2 out of 26) female respondents (7%) used the spelling (Shahista), the (1 out of 26) female respondents (3%) used the spelling (Sahaista), the (1 out of 26) female respondents (3%) used the spelling (Sahseta), the (1 out of 26) female respondents (3%) used the spelling (Shayista), the (1 out of 26) female respondents (3%) used the spelling (Shaesta), the (1 out of 26) female respondents (3%) used the spelling (Shastha), the (1 out of 26) female respondents (3%) used the spelling (Shista).

Description of spelling in Roman Urdu of word اور:

Or	16	61%
Aur	8	30%
R	1	3%

The majority of female respondents (16 out of 26) used the spelling (or). It is the 61% of the female respondents. The (8 out of 26) female respondents (30%) used the spelling (Aur), the (1 out of 26) female respondents (3%) used the spelling (r).

Description of spelling in Roman Urdu of word معزز:

Muaziz	13	50%
Muazaz	2	7%
Moazsis	1	3%
Mauzaz	1	3%
Muhziz	1	3%
Muziz	1	3%
Muazziz	1	3%
Mohaziz	1	3%
Moazzis	1	3%
Moaziz	1	3%
Muzzaiz	1	3%

Mozuzi	1	3%
--------	---	----

The majority of female respondents (13 out of 26) used the spelling (Muaziz). It is the 50% of the female respondents. The (2 out of 26) female respondents (7%) used the spelling (Muazaz), the (1 out of 26) female respondents (3%) used the spelling (Moazsis), the (1 out of 26) female respondents (3%) used the spelling (mauzaz), the (1 out of 26) female respondents (3%) used the spelling (Muhziz), the (1 out of 26) female respondents (3%) used the spelling (Muziz), the (1 out of 26) female respondents (3%) used the spelling (Muazziz), the (1 out of 26) female respondents (3%) used the spelling (Mohaziz), the (1 out of 26) female respondents (3%) used the spelling (Moazzis), the (1 out of 26) female respondents (3%) used the spelling (Moaziz), the (1 out of 26) female respondents (3%) used the spelling (Muzzaiz), the (1 out of 26) female respondents (3%) used the spelling (Mozuzi).

Description of spelling in Roman Urdu of word زبان:

Zuban	19	73%
Zaban	5	19%
Zubain	1	3%
Zoban	1	3%

The majority of female respondents (19 out of 26) used the spelling (Zuban). It is the 73% of female respondents. The (5 out of 26) male respondents (19%) used the spelling (Zaban), the (1 out of 26) female respondents (3%) used the spelling (Zubain), the (1 out of 26) male respondents (3%) used the spelling (Zoban).

Description of spelling in Roman Urdu of word سمجھی:

Samjhi	14	53%
Samji	7	26%
Samaji	1	3%
Samje	1	3%
Samjha	1	3%
Smji	1	3%
Samjha	1	3%

The majority of female respondents (14 out of 26) used the spelling (Samjhi). It is the 53% of the female respondents. The (7 out of 26) female respondents (26%) used the spelling (Samji), the (1 out of 26) female respondents (3%) used the spelling (Samaji), the (1 out of 26) female respondents (3%) used the spelling (Samje), the (1 out of 26) female respondents (3%) used the spelling (Samjha), the (1 out of 26) female respondents (3%) used the spelling (Samji), the (1 out of 26) female respondents (3%) used the spelling (Samjha).

Description of spelling in Roman Urdu of word جاتی:

Jati	21	80%
Jahti	1	3%
Jti	1	3%
Jatti	1	3%
Jate	1	3%

The majority of female respondents (21 out of 26) used the spelling (jati). It is the 80% of female respondents. The (1 out of 26) female respondents (3%) used the spelling (Jahti), the (1 out of 26) female respondents (3%) used the spelling (Jti), the (1 out of 26) female respondents (3%) used the spelling (Jatti), the (1 out of 26) female respondents (3%) used the spelling (Jate).

Description of spelling in Roman Urdu of word های:

Ha	11	42%
Hai	9	34%
Hae	1	3%
Hay	1	3%
Hy	3	11%

The majority of female respondents (11 out of 26) used the spelling (ha). It is the 42% of the female respondents. The (9 out of 26) female respondents (34%) used the spelling (hai), the (3 out of 26) female respondents (11%) used the spelling (hy), the (1 out of 26) female respondents (3%) used the spelling (HN), the (1 out of 26) female respondents (3%) used the spelling (hae), the (1 out of 26) female respondents (3%) used the spelling (hay).

"All the males and female respondents used the same spelling of the word (Urdu). Its cumulative frequency is 100%."

The word Urdu written with English spelling is very common and it is seen by every student by the very first day. On the other hand, every student tries to write the spelling of other words on his/her own as a result all the words in Roman Urdu are different from one another. The Roman spelling of word Urdu are usual with all the students and we find no difference or mistake in the spellings of Urdu.

"The word (Shaista) is used by 13 males and 18 females. Its cumulative frequency is 62%."

"The word (Muaziz) is used by 9 males and 13 females. Its cumulative frequency is 45%."

"The word (Zuban) is used by 11 male and 19 female. Its cumulative frequency is 60%."

"The word (Zaban) is used by 10 males and 5 female. Its cumulative frequency is 30%."

"The word (Samjhi) is used by 7 male and 14 female. Its cumulative frequency is 42%."

"The word (Samji) is used by 12 males and 7 females. Its cumulative frequency is 38%."

CONCLUSION

This study was aimed at exploring orthographic differences in Roman Urdu through a quantitative, corpus-based analysis of elicited data from 50 Urdu-speaking respondents. The results strongly express that Roman Urdu is featured a high degree of spelling variation and the reason is the absence of standardized orthographic script. During the analysis of the sentences and lexical units different spellings were detected for the same Urdu word. It confirms that Roman Urdu functions as a flexible, user-driven writing system rather than a structured script.

The analysis highlights that there are a number of words that represent the dominant spellings such as (KIA, HAL, ROZA, RAHA, BOLI, JATI and URDU) yet a noticeable number of alternate forms coexist. The reasons of this variability include phonetic perception, individual preferences and the impact of English orthographic structure. The words with complex phonological structures, such as "mukhtalif" and "zubanain" exhibited a prominently higher variability whereas simple and commonly used words showed consistency in spellings.

The assessment based on gender highlights that gender does not play a pivotal role in establishing spelling selection. The difference is very minor as female participants expressed somewhat higher consistency in certain cases, while male participants showed greater variation in others.

The analysis also points out the routine use of reduced and simplified forms (ha, rha, ma). It exposes the impact of growing digital communication and the requirement of the typing competence. These results favor that statement made in the introduction that Roman Urdu is broadly used due to technological convenience, speed, and ease of typing on English based keyboards, especially in social media and informal contexts.

The study confirms the outlook of previous scholars that Roman Urdu is a natural outcome of linguistic adaptation in digital environments. But the anxiety is that it is damaging the standard nature of Urdu and flourishing the non-standard nature. The lack of proper spelling structure not only causes inconsistency but may also add to reduce competency in standard Urdu (Nastaliq script), especially among younger users. Moreover, the high frequency of variation is a challenge for natural language process (NLP), text mining, and computational analysis, because inconsistent spellings are a hurdle in precise data processing.

In short, Roman Urdu is a source of effective and easy medium for informal communication but its broad use without standard pattern may have insinuations for language learning, orthographic efficiency, and the preservation of traditional Urdu script. Consequently, a balanced approach is suggested: Roman Urdu can be efficiently used in informal digital context, while standard Urdu should be encouraged and sustained in formal, educational, and academic settings. Researchers can also explore standardization strategies, sociolinguistic impact, and the influence of Roman Urdu on literacy and language policy.

REFERENCES

- Androutsopoulos, J. (2011). From variation to heteroglossia in the study of computer-mediated discourse.
- Baron, N. S. (2008). *Always on: Language in an online mobile world*. Oxford University Press.
- Crystal, D. (2001). *Language and the Internet*. Cambridge University Press.
- Javed, T., & Bhatti, M. (2015). Roman Urdu usage in digital communication and its linguistic implications. *Journal of Social Sciences and Humanities*.
- Kachru, B. B. (1985). Standards, codification, and sociolinguistic realism.
- Khan, W. M., Farooq, M. U., & others. (2025). Impact of social media engagement on emotional intelligence and communication patterns (including digital language use in Urdu contexts). *RC Research Archive Journal*.
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Mahmood, Z., Safder, I., Nawab, R. M., Bukhari, F., Nawaz, R., Alfakeeh, A. S., Aijohani, N. R., & Hassan, S. U. (2020). Deep sentiments in Roman Urdu text using recurrent convolutional neural network model.
- Paolillo, J. C. (2007). How to write Urdu in Roman script: Language choice and variation in computer-mediated communication. In B. Danet & S. C. Herring (Eds.), *The multilingual Internet: Language, culture, and communication online* (pp. 303–328). Oxford University Press.
- Rahman, T. (1996). *Language and politics in Pakistan*. Oxford University Press.
- Sharf, Z. (2019). *Framework for performing discourse-based opinion mining on Roman Urdu datasets* (PhD thesis). Shaheed Zulfiqar Ali Bhutto Institute of Science & Technology, Karachi, Pakistan.