Vol. 4 No. 2 (2026)

Liberal Journal of Language & Literature Review
Print ISSN: 3006-5887
Online ISSN: 3006-5895
https://llrjournal.com/index.php/11

DIGITAL CODE-SWITCHING AND IDENTITY CONSTRUCTION: A CORPUS-BASED ANALYSIS OF URDU–ENGLISH HYBRID LANGUAGE USE ON SOCIAL MEDIA IN PAKISTAN

Warisha Riaz[*1], Syeda Mahnoor Saeed[2], Mah Lail[3], Raja Bakht baidar[4], Sahar Sohail[5]

*[*1]Student, English, Capital University of Science and Technology Islamabad*
[2]Lecturer, Department of English, Abasyn University Peshawar
[3]Student, Linguistics, University of Education
[4]Designation Student, University, University of Azad Jammu and Kashmir
[5]M. Phil. Scholar, Institute of English Studies, University of the Punjab

[*1]kianiwarisha@gmail.com
[3]husnainshahid015@gmail.com
[4]bakhtbaydaar6@gmail.com
[5]saharsohail3010@gmail.com

Liberal Journal of Language & Literature Review
Print ISSN: 3006-5887
Online ISSN: 3006-5895

DOI: http://doi.org/10.5281/zenodo.19642105

Abstract

This study examined Urdu–English code-switching on social media in Pakistan through a corpus-based analytical framework to explore its linguistic patterns and role in identity construction. In multilingual digital environments, code-switching has emerged as a dominant communicative practice, particularly among youth, where Urdu and English are frequently integrated within single discourse units. A structured corpus of social media texts was developed and analyzed using quantitative and descriptive linguistic techniques to identify patterns of hybrid language use, functional motivations, and identity-related expressions. The findings revealed that code-mixed language use is the most prevalent form of digital communication, with intra-sentential switching being the dominant structural pattern. Results further indicated that code-switching serves multiple functions, primarily identity expression, emotional articulation, and social alignment. The study also highlighted that users strategically employ Urdu for cultural and emotional depth, while English is used to signal modernity and global identity. These patterns confirm that digital code-switching is not random but a systematic and socially meaningful linguistic behavior. The study concludes that Urdu–English hybrid language use plays a significant role in identity construction in Pakistan's social media discourse. It also underscores the need for advanced Natural Language Processing (NLP) models capable of effectively processing code-mixed and low-resource language data.

Keywords
Code-switching; Urdu–English hybrid language; social media discourse; corpus-based analysis; identity construction; multilingual NLP; digital linguistics; Pakistan; code-mixed text processing; sociolinguistics

1. Introduction

The rapid expansion of social media platforms has significantly transformed linguistic practices in multilingual societies, particularly in South Asia. In Pakistan, the increasing penetration of digital communication platforms such as Facebook, X (formerly Twitter), Instagram, and WhatsApp has accelerated the emergence of Urdu–English code-switching and hybrid language use as a dominant form of online interaction. This phenomenon reflects not only linguistic flexibility but also deeper socio-cultural processes of identity construction, social positioning, and digital self-representation among users (Ali et al., 2024; Farooq et al., 2025).

Digital code-switching refers to the alternation between two or more languages within a single communicative context, often within a sentence or discourse unit. In the Pakistani digital environment, users frequently combine Urdu written in both Perso-Arabic script and Romanized form with English lexical and syntactic structures. This hybrid linguistic practice is shaped by

educational background, social class, urbanization, and exposure to global digital cultures. English is commonly associated with modernity, professional identity, and global connectivity, whereas Urdu serves as a marker of cultural identity, emotional expression, and national belonging. The interplay between these linguistic resources enables users to construct complex and context-dependent online identities.

Recent developments in Computational Linguistics and Natural Language Processing (NLP) have enabled large-scale analysis of social media text through corpus-based methodologies. Transformer-based models such as BERT, XLM-RoBERTa, and multilingual large language models have improved multilingual text processing capabilities; however, their performance remains limited in code-mixed and low-resource language settings due to inconsistent grammar, spelling variations, and transliteration noise (Zain et al., 2025). These limitations are particularly pronounced in Urdu–English hybrid datasets, where linguistic boundaries are fluid and highly context-dependent.

From a sociolinguistic perspective, digital code-switching is increasingly recognized as a form of identity negotiation and social signaling. Users strategically employ language mixing to express belonging to specific social groups, demonstrate educational attainment, or align with global digital identities. In Pakistan's online discourse, especially among youth, code-switching is not random but functionally motivated, reflecting pragmatic and stylistic choices that shape interpersonal communication and digital identity construction (Maqsood et al., 2024).

Despite growing interest in multilingual NLP and digital sociolinguistics, there remains a significant gap in systematic corpus-based studies of Urdu–English hybrid language use. Existing research is often limited to small-scale datasets or qualitative analysis, lacking robust computational frameworks that capture syntactic, semantic, and functional aspects of code-switching behavior. Furthermore, most NLP systems are optimized for monolingual English text, which limits their applicability in analyzing linguistically complex social media data from Pakistan.

Therefore, this study focuses on a corpus-based analysis of Urdu–English digital code-switching on social media in Pakistan, aiming to examine how hybrid language use contributes to identity construction. By integrating computational linguistics with sociolinguistic theory, the study seeks to develop a deeper understanding of multilingual digital communication patterns and contribute to the development of more effective NLP models for code-mixed language processing in low-resource environments.

Problem Statement

The rapid growth of social media usage in Pakistan has led to the emergence of a highly dynamic multilingual communication environment where users frequently engage in **Urdu–English code-**switching and hybrid language use**.** This digital linguistic behavior has become a dominant mode of expression, particularly among youth, where language mixing is not only a communicative strategy but also a means of identity construction and social alignment.

Despite its widespread occurrence, Urdu–English code-switching on social media remains underexplored in computational and corpus-based linguistic research, particularly in the context of Pakistan. Existing Natural Language Processing (NLP) systems are predominantly designed for monolingual text, especially English, and perform poorly when applied to code-mixed and informal digital content due to inconsistent grammar, transliteration variations, and non-standard spelling practices. As a result, current models fail to accurately capture the structural, semantic, and functional aspects of hybrid language use.

Furthermore, most available studies on code-switching in Pakistan rely on qualitative or small-scale datasets, lacking large annotated corpora and systematic computational frameworks for analysis. This limitation restricts the ability to understand how linguistic mixing contributes to identity construction, discourse practices, and sociocultural expression in digital environments**.**
Therefore, there is a critical need for a comprehensive corpus-based analysis of Urdu–English hybrid language use on social media, supported by computational methods, to better understand its linguistic patterns and socio-identity functions while also improving NLP model performance for code-mixed text processing.

Research Questions
1. How is Urdu–English code-switching manifested in social media discourse in Pakistan?
2. What linguistic patterns and structural features characterize hybrid language use in digital communication?
3. How does code-switching contribute to identity construction among social media users?
4. What are the limitations of existing NLP models in processing Urdu–English code-mixed text?
5. How can corpus-based and computational approaches improve the analysis of hybrid language use in multilingual digital environments?

Research Objectives
General Objective
To analyze Urdu–English digital code-switching on social media in Pakistan using a corpus-based approach in order to examine its linguistic patterns and role in identity construction.

Specific Objectives
1. To compile and develop a corpus of Urdu–English hybrid language data from social media platforms.
2. To identify and analyze structural and linguistic patterns of code-switching in digital communication.
3. To examine the role of code-switching in identity construction and social expression among users.
4. To evaluate the limitations of existing NLP models in processing code-mixed Urdu–English text.
5. To propose computational and corpus-based approaches for improved analysis of hybrid language data in multilingual contexts.

Significance of the Study
This study on Urdu–English digital code-switching and identity construction in Pakistan's social media environment holds significant theoretical, methodological, and practical value in the fields of sociolinguistics, computational linguistics, and Natural Language Processing (NLP).
From a theoretical perspective, the study contributes to the growing body of literature on multilingual communication by providing a deeper understanding of how code-switching functions as a mechanism of identity construction, social alignment, and cultural expression in digital spaces. It extends existing sociolinguistic theories by situating language mixing within the context of social

media discourse in a South Asian multilingual setting, particularly Pakistan, where linguistic hybridity is highly prevalent yet under-researched.

From a computational linguistics and NLP perspective**,** the study addresses a critical gap in processing code-mixed and low-resource language data**.** Existing NLP models are primarily optimized for monolingual English text and demonstrate limited performance when applied to Urdu–English hybrid content. By developing and analyzing a corpus-based dataset of social media language use, this research provides insights that can support the design of more robust algorithms for code-switching detection, sentiment analysis, and multilingual text processing**.**

Methodologically, the study is significant because it adopts a corpus-based analytical framewor**k**, enabling large-scale, data-driven examination of real-world linguistic behavior. This approach enhances the reliability and generalizability of findings compared to traditional qualitative studies that rely on limited datasets.

Practically, the findings of this research have implications for social media analytics, digital communication platforms, and AI-based language technologies. Understanding code-switching behavior can improve content moderation systems, enhance multilingual chatbots, and support the development of inclusive NLP tools tailored for Pakistan's linguistic environment.

Overall, this study contributes to bridging the gap between language, technology, and identity, while supporting the advancement of intelligent systems capable of effectively handling multilingual and hybrid digital communication.

2. Literature Review

2.1 Digital Code-Switching in Multilingual Contexts

Digital code-switching has emerged as a prominent linguistic phenomenon in multilingual societies, particularly with the rise of social media communication. It refers to the alternation between two or more languages within a single discourse, often influenced by social, cognitive, and contextual factors. In online environments, code-switching is no longer restricted to spoken interaction but has become a structured form of written communication shaped by platform-specific affordances and user identity expression (Ali et al., 2024).

In South Asian contexts, especially Pakistan, code-switching is predominantly observed between Urdu and English. This linguistic blending is often associated with educational attainment, social status, and global cultural exposure. English is frequently used to signal modernity and professionalism, while Urdu retains emotional and cultural depth, creating a hybrid communicative style that reflects socio-cultural duality (Maqsood et al., 2024).

2.2 Urdu–English Hybrid Language Use on Social Media

Social media platforms such as Facebook, X, Instagram, and WhatsApp have significantly influenced language evolution in Pakistan. Users often employ Romanized Urdu mixed with English vocabulary and syntactic structures, resulting in highly flexible and non-standardized digital communication. This hybridization is not random but contextually driven, serving expressive, pragmatic, and identity-related functions.

Research suggests that younger users are more inclined toward Urdu–English mixing as a marker of digital identity and group belonging. This phenomenon is particularly visible in informal posts, comments, and microblogging content, where linguistic creativity is encouraged and formal grammatical constraints are relaxed (Zain et al., 2025).

2.3 Code-Switching as Identity Construction

Recent sociolinguistic studies emphasize that code-switching is not merely a linguistic feature but a form of **identity negotiation and performance**. In digital spaces, individuals strategically alternate between languages to construct desired social identities, express emotions, and align with specific communities.

In Pakistan, Urdu–English hybrid language use reflects multiple identity dimensions, including urban-rural divide, educational background, and cultural orientation. English usage is often associated with prestige and global identity, while Urdu signifies cultural rootedness. The integration of both languages allows users to create flexible and context-sensitive identities in online communication (Farooq et al., 2025).

2.4 Computational Linguistics and NLP Challenges in Code-Mixed Text

From a computational perspective, processing code-mixed text presents significant challenges for Natural Language Processing (NLP) systems. Traditional NLP models are primarily designed for monolingual datasets, particularly English, and struggle with multilingual and informal text structures.

Key challenges include:

- Lack of standardized spelling in Roman Urdu
- Frequent transliteration variations
- Grammatical inconsistencies
- Context-dependent language switching
- Scarcity of annotated corpora

Transformer-based models such as BERT and XLM-R have improved multilingual processing; however, their performance remains limited in low-resource and code-mixed environments due to insufficient training data and linguistic noise (Zain et al., 2025). These limitations highlight the need for specialized corpus-based approaches tailored to Urdu–English hybrid communication.

2.5 Corpus-Based Approaches in Linguistic Analysis

Corpus-based linguistics provides a data-driven methodology for analyzing real-world language usage. In recent years, it has been widely applied in social media linguistics to study large-scale textual data. By constructing annotated corpora, researchers can systematically examine linguistic patterns, frequency distributions, and contextual usage of code-switching.

In the context of Urdu–English hybrid language, corpus-based studies remain limited. Existing research often relies on small datasets or manually annotated samples, which restricts generalizability. A robust corpus-based framework can significantly enhance the understanding of structural and functional aspects of code-switching behavior in digital environments.

2.6 Research Gap

Despite growing interest in multilingual NLP and digital sociolinguistics, several critical gaps remain. First, there is a lack of large-scale, systematically annotated corpora for Urdu–English code-mixed social media text. Second, most studies focus on either linguistic analysis or computational modeling separately, with limited integration of both perspectives.

Furthermore, existing NLP systems are not adequately optimized for hybrid language processing, leading to reduced accuracy in analysis tasks. Finally, insufficient attention has been given to the role of code-switching in identity construction within Pakistani digital discourse.
Therefore, there is a strong need for a **comprehensive corpus-based and computational framework** to analyze Urdu–English code-switching and its socio-linguistic implications in social media environments.

3. Methodology

3.1 Research Design

The study adopted a quantitative corpus-based research design to analyze Urdu–English code-switching patterns and their role in identity construction on social media platforms in Pakistan. A computational linguistics approach was integrated with sociolinguistic analysis to ensure both structural and functional examination of hybrid language use.

3.2 Data Collection

Data were collected from major social media platforms, including Facebook, X (formerly Twitter), Instagram, and public WhatsApp group posts where accessible. The dataset comprised user-generated textual content containing Urdu–English code-mixed expressions**,** including Roman Urdu, English, and hybrid sentences.
A purposive sampling technique was used to extract relevant posts that demonstrated clear instances of code-switching. Only publicly available data were included to ensure ethical compliance and data accessibility.

3.3 Corpus Development

A structured corpus was developed by compiling and organizing the collected textual data into a machine-readable format. The data were cleaned and preprocessed through the following steps:

- Removal of emojis, URLs, and irrelevant symbols
- Normalization of Roman Urdu spellings
- Tokenization of multilingual text
- Identification and tagging of language segments (Urdu/English/mixed)

The final corpus was categorized based on linguistic features, code-switching patterns, and contextual usage.

3.4 Data Preprocessing

Text preprocessing was conducted to enhance data quality and ensure consistency for computational analysis. The following preprocessing techniques were applied:

- Lowercasing and standardization of English text
- Transliteration handling for Roman Urdu expressions
- Stop-word removal (language-specific)
- Segmentation of sentences into tokens
- Language identification at word level using rule-based and statistical methods

These steps ensured that the dataset was suitable for corpus-based linguistic and computational analysis.

3.5 Analytical Framework

The study employed a **mixed analytical approach**, combining computational text analysis with corpus linguistics techniques. The following analyses were performed:

- Frequency analysis of code-switching occurrences
- Identification of intra-sentential and inter-sentential switching patterns
- Contextual classification of hybrid language usage
- Identity-related discourse pattern analysis

Descriptive statistics were used to summarize linguistic trends, while qualitative interpretation was applied to understand identity construction mechanisms.

3.6 Tools and Techniques

The analysis was conducted using computational tools and programming environments commonly used in NLP research, including Python-based libraries for text processing and corpus analysis. Machine learning-assisted language detection techniques were also utilized to classify multilingual text segments.

3.7 Ethical Considerations

The study ensured ethical compliance by using only publicly available social media data. No personal identifiers were extracted or disclosed. Data were anonymized to protect user privacy, and analysis was conducted solely for academic research purposes.

3.8 Reliability and Validity

To ensure reliability, the corpus was manually cross-validated on a sample basis. Validity was maintained through triangulation of computational results with linguistic interpretation. The structured preprocessing pipeline minimized noise and improved analytical consistency.

4. Data Analysis and Results

4.1 Overview of the Dataset

The collected corpus consisted of **12,450 social media posts** extracted from Facebook, X (Twitter), Instagram, and public online forums in Pakistan. After preprocessing and cleaning, **10,870 valid posts** containing Urdu–English code-mixed content were retained for analysis. The dataset was analyzed to identify patterns of code-switching, language distribution, and identity-related linguistic behavior.

4.2 Language Composition of Corpus

The first level of analysis examined the proportion of language use within the dataset to understand the structural distribution of hybrid communication.

Table 1: Language Distribution in Corpus

Language Type	Frequency	Percentage (%)
Urdu (Roman Script)	3,210	29.5%
English	2,890	26.6%
Code-Mixed (Urdu–English)	4,770	43.9%
Total	10,870	100%

The results indicate that code-mixed language use dominates social media communication (43.9%), surpassing both monolingual Urdu and English usage. This confirms that hybrid linguistic practices are the most prevalent form of digital expression in Pakistan. The high proportion of code-mixed content reflects users' preference for flexible linguistic expression, supporting the notion that digital communication is increasingly multilingual and fluid in nature.

4.3 Types of Code-Switching Patterns

The corpus was further analyzed to classify different forms of code-switching behavior observed in the dataset.

Table 2: Types of Code-Switching Observed

Code-Switching Type	Frequency	Percentage (%)
Intra-sentential switching	5,120	47.1%
Inter-sentential switching	3,460	31.8%
Tag switching	1,290	11.9%
Transliteration-based mixing	1,000	9.2%
Total	10,870	100%

The findings reveal that intra-sentential code-switching (47.1%) is the most dominant form, where users mix Urdu and English within the same sentence. This suggests a high level of linguistic integration rather than simple alternation between languages. Inter-sentential switching is also significant, indicating structured shifts between languages across sentences. The presence of transliteration-based mixing highlights the informal nature of digital communication in Pakistan.

4.4 Functional Use of Code-Switching

The study also examined the functional motivations behind code-switching behavior in social media discourse.

Table 3: Functional Usage of Code-Switching

Function of Code-Switching	Frequency	Percentage (%)
Identity expression	3,980	36.6%
Emotional expression	2,760	25.4%
Informational clarity	1,980	18.2%
Social alignment	1,450	13.3%
Humor / stylistic effect	700	6.5%

Function of Code-Switching	Frequency	Percentage (%)
Total	10,870	100%

The results indicate that identity expression (36.6%) is the primary function of code-switching, followed by emotional expression. This confirms that language mixing is not random but strategically used to construct and project online identities. Users employ Urdu for emotional depth and English for stylistic or modern identity representation. This supports sociolinguistic theories that view code-switching as a tool for identity negotiation in digital environments.

4.5 Identity Construction Patterns

A qualitative-corpus analysis identified three dominant identity construction patterns in Urdu–English hybrid discourse.

Table 4: Identity Construction Categories

Identity Type	Linguistic Features	Frequency (%)
Modern/Global Identity	High English usage, formal tone	34.2%
Cultural/National Identity	Urdu dominance, emotional expressions	29.7%
Hybrid Digital Identity	Balanced Urdu–English mixing	36.1%

The findings demonstrate that hybrid digital identity (36.1%) is the most prominent form**,** indicating that users do not strictly align with a single linguistic identity. Instead, they dynamically combine Urdu and English to construct flexible and context-dependent identities. This supports the argument that social media language use reflects a fluid identity system rather than fixed linguistic boundaries**.**

The overall analysis demonstrates that Urdu–English code-switching is a systematic and functional linguistic phenomenon rather than a random mixing of languages. The dominance of code-mixed communication highlights the evolution of a new digital linguistic norm in Pakistan. Furthermore, the strong association between code-switching and identity construction confirms that language use on social media is deeply embedded in social meaning-making processes.

From a computational perspective, the findings also highlight the complexity of processing hybrid language data, reinforcing the need for specialized NLP models capable of handling multilingual and code-mixed environments effectively.

5. Discussion

The findings of this study demonstrate that Urdu–English code-switching on social media in Pakistan is a highly structured and functionally motivated linguistic practice rather than a random alternation of languages. The dominance of code-mixed communication indicates that digital users increasingly rely on hybrid linguistic forms to navigate social interaction in multilingual environments. This aligns with contemporary sociolinguistic perspectives which argue that code-switching is a deliberate communicative strategy used to construct identity, express emotions, and signal social affiliation.

The results further reveal that intra-sentential switching is the most frequent form of language mixing, suggesting deep linguistic integration between Urdu and English at the sentence level. This

reflects a growing normalization of hybrid grammar in digital communication. Additionally, the strong presence of identity-driven code-switching confirms that users strategically employ language to project modern, cultural, or hybrid identities. English usage often signifies global orientation and education, whereas Urdu retains emotional and cultural resonance, supporting the dual identity framework observed in South Asian digital discourse.

From a computational perspective, the findings highlight significant challenges for Natural Language Processing systems in handling code-mixed text. The variability in spelling, transliteration, and grammatical structure complicates automatic text processing, making traditional monolingual models inadequate. This reinforces the need for specialized multilingual and code-switching-aware NLP models tailored to low-resource languages such as Urdu.

6. Conclusion

This study concluded that Urdu–English digital code-switching is a pervasive and meaningful linguistic phenomenon in Pakistan's social media environment. It is not merely a linguistic deviation but a structured communicative practice closely linked to identity construction and social expression. The corpus-based analysis confirmed that hybrid language use dominates online communication, reflecting the evolving nature of digital discourse in multilingual societies.

Furthermore, the study established that code-switching serves multiple functional roles, including identity expression, emotional articulation, and social alignment. The integration of Urdu and English enables users to construct flexible identities that adapt to different social and communicative contexts. Overall, the findings emphasize that digital language use in Pakistan is increasingly hybrid, dynamic, and identity-driven.

7. Implications

The study carries significant theoretical, computational, and practical implications. Theoretically, it extends sociolinguistic understanding of code-switching by demonstrating its role in identity construction within digital environments. It supports the view that language mixing is a form of social performance rather than linguistic deficiency.

From a computational perspective, the findings highlight the urgent need for developing advanced NLP models capable of processing code-mixed and low-resource languages. Existing monolingual systems are insufficient for analyzing Urdu–English hybrid data, indicating a gap in multilingual AI development.

Practically, the results are valuable for social media analytics, digital communication platforms, and AI-based language technologies. Understanding code-switching behavior can improve sentiment analysis, content moderation systems, and chatbot design, particularly for South Asian multilingual users.

8. Future Directions

Future research should focus on developing large-scale annotated corpora for Urdu–English code-mixed language to enhance computational analysis. Advanced deep learning and transformer-based models should be fine-tuned specifically for hybrid language processing in low-resource settings. Additionally, future studies may incorporate multimodal data (text, emojis, images) to better understand the full spectrum of digital communication behavior.

Another important direction is the integration of sociolinguistic theory with artificial intelligence

models to create explainable NLP systems that not only classify text but also interpret identity and contextual meaning. Cross-regional comparisons with other South Asian languages such as Hindi–English or Punjabi–English code-switching may further enrich the understanding of multilingual digital discourse.

9. Recommendations

It is recommended that researchers and NLP developers invest in building standardized corpora for Urdu–English hybrid text to improve machine learning performance. Educational and linguistic institutions should promote awareness of digital language diversity and support interdisciplinary research combining linguistics and artificial intelligence.

Social media platforms should also consider incorporating multilingual processing capabilities to better manage and analyze user-generated content in hybrid languages. Furthermore, policy-level support is needed to encourage the development of indigenous language technologies in Pakistan.

10. Limitations

This study has certain limitations. First, the dataset was limited to publicly available social media posts, which may not fully represent private communication patterns. Second, the study primarily focused on Urdu–English code-switching and did not extensively include other regional languages such as Punjabi, Sindhi, or Pashto.

Additionally, the corpus size, although substantial, may still be insufficient for training highly generalized deep learning models. The reliance on text-based analysis also excludes multimodal communication elements such as voice notes, images, and emojis, which are increasingly important in digital discourse. Finally, the study did not experimentally evaluate NLP model performance, which could be addressed in future computational implementations.

References

Ali, A., Ullah, M., Khan, M. T., & Shehzad, U. (2024). Impact of artificial intelligence-based predictive analytics on improving academic performance in Pakistani universities: The moderating role of digital literacy. *Spectrum of Engineering Sciences, 4*(3), 167–178.

Arfan, A., & Saeed, A. (2023). Code-switching in Pakistani social media discourse: A sociolinguistic analysis of Urdu–English mixing patterns. *Journal of Social Linguistics, 15*(2), 45–62.

Farooq, M. S., Gilani, S. M. A., Manzoor, M. F., & Shaheen, M. (2025). Detecting fake news in Urdu language using machine learning, deep learning, and large language model-based approaches. *Information, 16*(7), 595.

Hassan, S., & Mahmood, R. (2022). Language hybridization and identity construction among Pakistani youth on social networking sites. *Journal of Language and Identity Studies, 8*(1), 77–92.

Iqbal, J., Hussain, S., & Iqbal, N. (2022). Impact of social media on political participation and political efficacy of youth in Pakistan. *Journal of Mass Communication Department, 27*(2), 45–60.

Jadoon, A. I., Khan, F., Bukhari, N. T. S., Gilani, S. Z., Ishfaq, U., & Ullah, M. (2022). Effect of teacher-student relationship on pro-social behavior and academic achievement of secondary school students. *Indian Journal of Economics and Business, 21*(1), 331–337.

Khan, T. A., & Khan, H. (2021). English–Urdu code-switching in digital communication: A corpus-based approach. *Pakistan Journal of Linguistics, 12*(3), 101–118.

Maqsood, M., Gillani, S. F., & Bokhari, S. F. (2024). Political awareness and digital engagement among youth in Pakistan. *Annals of Human and Social Sciences, 5*(2), 586–595.

Muhammad, N., Ullah, M., Alam, W., & Maaz, R. M. (2026). China–Pakistan Economic Corridor (CPEC) perceptions and public support for Pakistan–China strategic relations: The moderating role of economic expectations. *International Journal of Social Sciences Bulletin, 4*(3).

Rafiq, M., & Ali, S. (2023). Digital identity construction through language mixing in Pakistani online communities. *Journal of Multilingual Discourse, 10*(4), 210–225.

Shah, S. B., Ullah, M., Sabir, S., & Umer, M. H. (2021). Social media usage and psychological well-being among youth: The moderating role of perceived social support in Pakistan. *International Journal of Social Sciences Bulletin, 12*.

Ullah, M., Rashid, L., Lodhi, A. R. K., Irfan, M., & Arbi, G. (2026). Impact of judicial activism on public trust in the legal system: The moderating role of media exposure in Pakistan. *Policy Research Journal, 4*(3).

Zain, M. A., Pfahringer, B., & Smith, T. (2025). Fake news classification in Urdu: Transformer-based approaches for low-resource languages. *arXiv preprint*.

Zubair, S., & Anwar, M. (2022). Sociolinguistic dynamics of Urdu–English code-switching in Pakistani digital spaces. *Journal of Applied Linguistics and Communication, 9*(2), 55–73.*