

A Corpus-Based Sociolinguistic Analysis of Code-mixing in Pakistani TikTok Comments

Ayesha Malik¹, Prof. Dr. Hafiz Ahmad Bilal², Dr. Naveed Nawaz Ahmad³,
Prof. Dr. Ijaz Asghar⁴

¹ MPhil English (Linguistics), Department of English Language and Linguistics, University of Sargodha, Email: ayeshamalek1219@gmail.com

² Higher Education Department, Punjab, Email: ahmadbilal.uos@gmail.com

³ Assistant Professor, Department of English Language and Linguistics, University of Sargodha, Corresponding Author's Email: naveed.nawaz@uos.edu.pk

⁴ Chairperson, Department of English Language and Linguistics, University of Sargodha, Email: ijaz.asghar@uos.edu.pk

DOI: <https://doi.org/10.70670/sra.v3i4.1301>

Abstract

This paper examines code-mixing in English-Urdu in comments on Pakistani Tik Tok through the corpus based sociolinguistic method. Since the emergence of social media, bilingual language practices have been practiced in digital communication, which may indicate, both, creativity and social identity. The existence of 1000 comments on TikTok were gathered by boiling down 20 popular TikTok videos and examined through quantitative and qualitative approaches. The quantitative analysis was used to test the frequency, type and place of English insertions in Urdu comments. It was found that word-level code-mixing was the most common (65% of cases), then there was phrase-level mixing (25%), and sentence-level mixing (10%). The thematic analysis qualitatively involved the presence of English words that were mostly used at the beginning or within the middle of the text as emphasizing and stylistic markers: amazing, funny, and love. The qualitative thematic analysis identified four key functions of code-mixing: identity construction, humor and stylistic expression, social interaction and solidarity, and creative digital vernaculars. The users who are younger especially those who are aged between 16 and 22 showed stronger code-mix frequencies, and there were gender variations in functional use. Males mainly used code-mixing as a means of humor and assertiveness whereas the females used it to reveal their feelings and to share with their peers which proved that code-mixing is contextually based and socially motivated. The paper identifies TikTok as a source of linguistic innovation and engagement in communities demonstrating that English-Urdu code-mixing has a purposeful social and communicative role. This study adds into the knowledge of bilingual practices of digital communication and motivates the idea of conducting further research on the use of language in the scenarios of online social media.

Keywords: Code-mixing, TikTok, English-Urdu, Sociolinguistics, Bilingualism, Digital Communication.

Introduction

Background of the Study

Language is a dynamic process which is an indication of social identity, culture and communication practice. Multiple languages are actively used by two or more people, which is why it is called code-mixing or code-switching (Zafar, 2024). The concomitant existence of both Urdu and English in Pakistan has created a fertile linguistic household where English Urdu code-mixing is becoming a widespread occurrence especially in young generations and in cities. The

introduction of social media platforms, including TikTok, has exposed new grounds of interaction where people have the freedom to mix languages to come up with hybrid forms of communication (Ghafoor, 2025). Such informal and interactive texts as comments in TikTok can offer researchers a rare corpus to study sociolinguistic behavior by analyzing not only the lexical and syntactic forms but also social meanings and identity answer construction behind language mixing. Corpus-based method enables researchers to systematically gather, organize, and analyze extensive volumes of textual data, and patterns in the use of language that might not be immediately evident in smaller qualitative samples are observable (Salman et al., 2023). The research conducted in Pakistani online settings has demonstrated that age, gender, education, and worldwide media exposure are the factors that affect code-mixing (Awan et al., 2025). By using a corpus-based sociolinguistic approach and examining TikTok comments, one can get a clearer picture of the modern trends of the English-Urdu hybridization, social purposes of code-mixing, and the construction of the online identity.

Problem Statement

Although the idea of code-mixing in Pakistani digital media is increasingly becoming common, the research on TikTok comments as a linguistic corpus. The focus of most current research has been on spoken discourse, on print media, or other social media platforms, such as Twitter and Instagram (Ghafoor, 2025). This creates a knowledge gap concerning the ways English Urdu code-mixing is appearing within the short, highly interactive, and performative setting of TikTok. In the absence of empirical evidence of this platform, sociolinguistic understanding of the current digital communication in Pakistan is not complete. Moreover, the studies on corpus are necessary to measure the trends of the use of both languages and discuss the sociocultural reasons of such linguistic behavior, especially among the youth using digital communities.

Research Objectives

1. To identify and analyze the patterns of English-Urdu code-mixing in Pakistani TikTok comments
2. To examine the sociolinguistic factors influencing code-mixing in TikTok comments
3. To explore the social meanings and functions of code-mixing in the digital communication context of Pakistan

Research Questions

1. What are the prevalent patterns of English-Urdu code-mixing in Pakistani TikTok comments?
2. What sociolinguistic factors influence the use of code-mixing in TikTok comments?
3. How does code-mixing in TikTok comments reflect social identity, cultural trends, and online communication norms in Pakistan?

Significance of the Study

The proposed research is of value to the areas of sociolinguistics, digital communication, and corpus linguistics because it offers a solid empirical investigation of the notion of code-mixing between English and Urdu in a digital environment. Investigating the code-mixing patterns in TikTok posts can assist a researcher and educator to understand how language changes within a multilingual community, especially among the youth community using the social media. The results can also be used to inform digital media linguistics, emphasizing the role of online spaces in mediating language practices, identity, and cultural expression (Salman et al., 2023). Also, the corpus-based method offers methodological rigor and can be used as a basis of future research of the code-mixing in other digital settings in Pakistan and other multilingual settings.

Literature Review

Introduction

The literature review on the current state in the field of code-mixing, sociolinguistic analysis and

corpus-based methods in the digital environment, especially focusing on the English-Urdu interactions in Pakistan. It is basically divided into thematic sections, including (1) conceptualizations of both code-mixing and code-switching, (2) sociolinguistic factors that affect language mixing, (3) using corpus in linguistic research and (4) code-mixing in social media. The purpose is to place the current study in the context of the academic discourse, as well as to determine gaps in the research, in particular, regarding TikTok comments as a corpus to study socio linguistically. The two are code-mixing and code-switching which are closely related linguistic phenomena commonly researched in multilingual environments. Code-switching is a process of alternating between two languages on the sentence or discourse level (usually based on syntactic boundaries) and code-mixing is a process of inserting lexical items, phrases, or morphemes of one language into a different one (in a single utterance) (Asghar et al., 2021; Jabeen, 2018; Awan et al., 2025). The code-mixing between English and Urdu is common in the Pakistani context because of the colonial use of English, its official language in education, and the growing impact of global digital culture (Zafar, 2024). Research has revealed that code-mixing is not only a linguistic option but a social scheme that portrays identity, status and group affiliation. To illustrate, young urban speakers tend to incorporate English in Urdu to indicate modernity, education, or digital literacy (Channa et al., 2021). Sociolinguistic studies underscore the fact that contextual factors that affect language choice are age, sex, social status, education, and communication media. In Pakistan, Awan et al. (2025) discovered younger users tend to use English-Urdu code-mixing more in both oral and electronic communication, and especially in the informal context. On the same note, Jabeen (2018) emphasized that code-mixing in the media and in our daily speech has pragmatic and stylistic purposes, including humor, emphasis, or identity creation. Gender is also contributing with male and female users exhibiting various patterns in lexical borrowings and patterns of switching (Channa et al., 2021). These sociolinguistic variables have a significant role in examining the reasons and trends of code-mixing in social media where communication is extremely interactive and performative (Tariq et al., 2013).

Corpus based studies offer a systematic approach to the study of language use, based on collection and computational analysis of large text collections (Salman et al., 2023). When applied within the framework of multilingualism, the corpus analysis enables the researcher to measure the incidences of the code-mixing phenomenon, define the common lexical or syntactic trends, and explore social roles of language alternation. In the past, corpora based on newspapers, online forums, and social media platforms (Instagram, Twitter) have been used to investigate the phenomenon of English-Urdu code-mixing in Pakistan (Ullah, 2023). The studies have shown that corpus techniques are useful in revealing the patterns of digital communication that provide replicable and empirical information to supplement traditional qualitative studies. The emergence of the social media has changed the linguistic landscapes; they offer plenty of information to observe the multilingual practices during real-time communication. Informal and interactive language use, which should be creative, takes place through digital platforms and typically incorporates both global and local languages. The study of Pakistani social media has recorded a significant amount of English-Pakistani mixing, including lexical borrowing and structural alternation, in posts on Instagram and comments in YouTube videos (Salman et al., 2023). TikTok is a rather short-form platform that creates a special scenario because of the conciseness of the text, multimedia, and the high-engagement rates, which promote the playful and theatrical use of code-mixing. This notwithstanding, very little scholarly interest is devoted to TikTok comments as a corpus, and little is known about how sociolinguistic factors work in this particular online setting.

Theoretical Framework

The hypothetical background gives the conceptual base on which the trends and reasons behind the English- Urdu mixes in Pakistani Tik Tok remarks can be understood. This paper is based on three main sociolinguistic theories Interactional Sociolinguistics Theory, Markedness Model, and Domain Theory proposed by Gumperz, Myers-Scotton and Fishman respectively. All of those

theories explain the functional, social, and situational features of the code-mixing process and guide the process of corpus-based study of the multilingual communication.

The Interactional Sociolinguistics Theory of Gumperz

John Gumperz (1982) came up with Interactional Sociolinguistics Theory, which focuses on the social context and the interaction cues in the interpretation of language use. Gumperz claims that code-switching and code-mixing do not occur randomly, but they are influenced by the situations of the conversation, the intent of the speaker, and social conventions. In online communications, e.g., in the comments in Tik Tok, code-mixing is actively used by users with the purpose to express meaning, humor, identity, or solidarity. This theory can be used to explain the way in which English-Urdu mixing is used to articulate the social relationships, group affiliation and negotiation during conversations among Pakistani speakers (Gumperz, 1982).

Markedness Model of Myers-Scotton.

The Markedness Model developed by Carol Myers-Scotton (1993) is a theory that assumes that speakers select a specific variant of language due to social norms and the urge to either adhere or to break the patterns of a norm. In code-mixing, the marked option is an expression of social meaning, e.g., prestige, modern or belonging to a group. When English words are used in Urdu sentences in the context of TikTok comments, it can be an indicator of higher education or cosmopolitanism or a need to align with the culture of the global youth. This theory is a guide towards examining the reasons behind the use of two languages by users and the social purposes that the change of languages serves in online communities (Myers et al., 1993).

Fishman's Domain Theory

Domain Theory of Joshua Fishman (1972) proposes language choices with regard to the domain where people communicate, family, education, workplace, or digital media. In every sphere, there are normative language and expectation rules. As a digital social space, Tik Tok promotes the use of informal, creative, and performative language, which frames the code-mixing use as not only tolerable but also significant in social terms. This theory underlines the exploration of the study with the domain of the platform forming patterns of English-Urdu integration in comments of users (Fishman, 1972).

Research Gap

Whereas many studies have been dedicated to researching code-mixing in spoken language, print media, and other social media platforms (Jabeen, 2018; Awan et al., 2025; Ghafoor, 2025), the literature on the topic is limited with references to Tik Tok comments. The interactive and the performative nature of Tik Tok offers distinct language innovation and identity expression possibilities unlike other platforms. Moreover, a corpus based quantitative analysis can offer more systematic and generalizable information which most studies have focused on smaller datasets or qualitative observations. This gap is the reason why the present study is necessary, as it aims at integrating both corpus-based quantitative and sociolinguistic approaches to the study of English to Urdu code-mixing in comments in Pakistani Tik Tok.

Research Methodology

Introduction

The systematized analysis of language alternation patterns and reasons is based on the sociolinguistic approach in which a corpus is used to analyze the information in a systematic manner. Methodology describes the research design, population, sampling method, corpus development, data gathering process, data analysis approaches and ethical issues. At the same time, it combines quantitative and qualitative methods to answer the research questions.

Research Design

The research design was a mixed-methods one because it involves both quantitative analyses based on corpus and qualitative analysis based on sociolinguistic interpretation. Quantitative analysis was devoted to the frequency, types, and position of English insertion in the Urdu comments, and the qualitative analysis was devoted to the social and pragmatic role of the English insertion, including identity expression, humor, and social solidarity. The research design will be informed by the Interactional Sociolinguistics Theory by Gumperz, Markedness Model by Myers-Scotton and Domain Theory by Fishman where both structural and contextual aspects of the code-mixing are taken into account.

Population of the Study

The sample population consisted of TikTok users in Pakistan posting comments in Urdu and English or any of both languages. The research paper focused on the age group of users between the age of 16-30 years old, as they are the most active on TikTok. This group is quite a fruitful source of bilingual interactions which both expresses the globalization media influence and the local language.

Sampling Technique

The purposive sampling approach was used in order to make sure that the data was comprised of specifically the comments with the English-Urdu code-mixing. Pakistani creators with over 20 popular TikTok videos were picked using the above criteria of popularity, number of views and interest in the youth audience. Based on these videos, a thousand comments with bilingual contents were picked out to form a large and representative corpus. The purposive sampling method was appropriate to make sure that the data accurately reflected real-life examples of naturally occurred code-mixing, and at the same time, to keep irrelevant, monolingual or non-linguistic content out

Corpus Creation

The obtained remarks were gathered into an electronic corpus and given special identifiers. Metadata like the age of the user, gender (where necessary) and video context was captured. The samples have been filtered to eliminate duplicates and off-topic information that leaves 1000 useful comments. All the comments were coded by word-level, phrase-level or at the sentence-level code-mixing resulting in systematic quantitative and qualitative data analysis.

Data Collection Procedures

The TikTok comments that were publicly available were gathered manually. The original language context was maintained with the help of screenshots and the extraction of the text, and usernames and personal information were anonymized. Only genuine, naturally deriving comments were used to make sure the sociolinguistic analysis is valid.

Data Analysis Techniques

Frequency, types, and positions of English insertions were measured using quantitative analysis and Excel and SPSS were used to identify common lexical items. Social roles of code-mixing such as identity expression, humor, solidarity and digital creativity were interpreted by qualitative thematic analysis. The theoretical framework was tied to the analysis to be familiar with both interactional and social motivations of language alternation.

Ethical Considerations

There was adherence to ethical standards. The comments were limited to those that were publicly available and anonymized all personal identifiers. Only the sensitive or harmful contents were not included, and information was used only on academic purposes without violating the privacy and rights of online users.

Data Analysis and Results

Introduction

The case analysis of English-Urdu code-mixing in one thousand TikTok comments of Pakistan gives a close perspective on the way bilingual users interact with the linguistic resources in the digital environment. The broadened corpus facilitates the further analysis of the structural patterns, sociolinguistic motivation, and communicative functions presented in mixed-language exchanges. The remarks show the smooth interaction of English and Urdu under the influence of stylistic choice, identity expression, humor and community interaction. As an app that is based on the principle of immediacy, entertainment, and creativity, TikTok forms a very specific environment in which multilingual practices thrive. The analysis below depicts the quantitative patterns, sociolinguistic forces, thematic functions, and composite meanings of how and why code-mixing operates in this digital space.

Quantitative Analysis of Code-Mixing

The quantitative analysis of the one thousand comments shows that the word level of mixing is dominant, then the phrase level mixing, and lastly the sentence level mixing. There are six hundred and fifty word-level mixes in six hundred and fifty comments constituting sixty five percent of total dataset. English lexical items, adjectives, verbs, and evaluative expressions are expressive linguistic tools in such instances that enhance emotional or stylistic effect. Mixed phrases are found in two hundred and fifty comments, which is a quarter of the sample. Short English collocations are often used by users to support appreciation, stress or aesthetic value. The sentence-level mixing is not as high, and it is documented in one hundred comments or ten percent of the data. These only usually pertain to brief English phrases in Urdu discourse, which are employed to bring out humor or assertiveness or emphasis. The insertion of English words in the comments, also, shows intentionality: the majority of them are placed in the middle of the sentence, then initial and not so often final. The usage shows that English is an inbuilt expressive medium and not an addition.

Table 4.1 Frequency of Code-Mixing Types in TikTok Comments

Code-Mixing Type	Frequency	Percentage
Word-level	650	65%
Phrase-level	250	25%
Sentence-level	100	10%
Total	1000	100%

Sociolinguistic Analysis

Demographic factors which influence the degree and form of English-Urdu code-mixing on TikTok and their presence in the corpus is shown by the sociolinguistic patterns. The most frequent mixing is exhibited by younger users, especially those who are between the ages of sixteen and twenty-two years, who usually include English phrases related to youth culture in the world, online humor, and tendencies. Their language attitude is a product of the impact of foreign entertainment services and the culture of interaction on Tik Tok. Users between twenty-three and thirty have been using it more often but in more organized and less experimental way. The disparity of genders is also revealed: male commentators often use English to convey humor, sarcasm, assertiveness, and performance of digital confidence, and female commentators mostly use English to convey praise, emotional attachment, beauty and supportive feedback. In all the demographic groups, English is an informational resource, which increases stylistic elegance, indicators of belonging to a group, and adds to the communicative depth of the platform.

Thematic Analysis

The thematic analysis of the one thousand comments shows that there are four major themes with several sub-themes that describe the social and communicative reasons behind code-mixing.

Along with the linguistic convenience, users combine English as a strategic option linked to identity, humor, emotionality, and digital creativeness. The themes indicate the practical richness of bilingual language in the virtual realm, as they describe how the users bargain to obtain the meaning, power of position, and relationship within the means of hybrid language.

Table 4.2 Themes and Sub-Themes of Code-Mixing in TikTok Comments

Main Themes	Sub-Themes	Description
Identity Construction	Modernity Signaling	Use of English to project globalized, fashionable, or educated identity.
	Digital Youth Identity	Aligning with online trends, memes, or global youth culture.
	Self-Presentation	Enhancing personal style or persona through selective English insertions.
Humor and Stylistic Expression	Comic Exaggeration	English used to intensify jokes, sarcasm, or playful commentary.
	Trend-Based Expressions	Use of viral English phrases (e.g., “literally,” “killed it”).
	Emphatic Styling	Creating rhythm, emphasis, or expressive flair.
Social Interaction and Solidarity	Complimenting & Appreciation	English phrases indicating praise or admiration.
	Agreement & Participation	Hybrid expressions used to join collective reactions.
	Relational Warmth	Use of English to convey positive interpersonal tone.
Creative Digital Vernaculars	Hybrid Slang	Blending English and Urdu into innovative expressions.
	Playful Linguistic Experimentation	Constructing new mixed forms characteristic of TikTok culture.
	Platform-Specific Creativity	Language adapted to the performative environment of TikTok.

The thematic results show that code-mixing is a complex communicative practice. The identity construction comes out clearly when the users use English terms to showcase modernity or cosmopolitan identity or conformity to digital culture. Another theme, humor and stylistic expression, echoes the theme of entertainment and expressive creativity of the platform. Most of the remarks use English in an attempt to build up on humor or create a stylistically lively response. The third theme is social interaction and solidarity, and in this case, code-mixing will enable the building of rapport and a sense of community and inter-user warmth. The last theme, creative digital vernaculars, shows how users create new forms of hybrid lingo that show the interaction of culture and platform. All these themes enable showing a conclusion that code-mixing is an intentionally used resource in the online discourse that is socially meaningful and creatively employed. Thematic and quantitative findings were integrated in a manner that allows comprehension of the findings and their related implications.

Synthesis of Quantitative and Thematic Results

The synthesis of quantitative and thematic results was done in a manner that enabled an understanding of the results and implications associated with the results. The synthesis of thematic insights and structural patterns shows that there is a consistent connection between the form and

the role of English-Urdu code-mixing in TikTok comments. The quantitative preponderance of the word level mixing correlates well with the thematic data that it is expressive in the immediate emotional response, humor and positioning of identity. Likewise, the sociolinguistic trends by age and gender mirror the observed thematic functions in the data, and they affirm that the English language is strategically implemented in order to achieve social meaning, stylistic correspondence, and online group membership. The thematic categories also support the purposefulness of mixing as demonstrating that English is not placed as an ad hoc insertion, it remains an instrument that users deploy in the creation of stylistic intensity, cultural signaling and relational interaction. The combined results present the way linguistic structure and social purpose collaborate in order to build a dynamic bilingual environment on TikTok.

Summary

The examination of one thousand TikTok comments can give a deeper insight into English-Urdu code-mixing being a structured, goal-oriented, and social meaning, linguistic activity. Word-level mixing is the most widespread and it can be seen as the expressive immediacy of online communication. Sociolinguistic habits that are age and gender correlated depict how various demographic groups use English to convey their identity, humor and relational meaning. The thematic analysis identifies four main functions, namely identity construction, humor and stylistic expression, social interaction and solidarity, and creative digital vernaculars which are reinforced by subtle sub-themes. All the results together point to the conclusion that TikTok creates an exciting bilingual environment where users negotiate the meaning, develop identity, and engage in the digital culture utilizing hybrid languages.

Discussion

The analysis showed that the most common ones were word level code-mixing, phrase level code-mixing and sentence level code-mixing respectively. This is in line with the reduced, informal and performative qualities of TikTok comments, where the emphasis is on the concise and expressive influence. The English lexical terms like funny, amazing and love were widely used as a way of underlining feelings or reactions and this evidence shows the strategic use of English as a symbol of style portrayal. Mixed phrases and sentences were less common but was used to express humor, sarcasm or subtle meanings, and is an expression of how bilingual speakers can be flexible in the online world. The sociolinguistic analysis showed that age and gender had a role to play in code-mixing behavior. Younger users (particularly, those aged 16-22) tended to combine languages more, as they have a higher level of digital literacy and exposure to international media. Gender trends indicated that males mostly used code-mixing to create humor and assertiveness whereas females used it to display feelings, support and compliments. These results confirm the Markedness Model offered by Myers-Scotton, which focuses on the fact that speakers engage in social decisions when using a language to mark themselves or correspond to the group expectations. Thematic analysis found four key roles of code-mixing, namely, identity construction, humor and stylistic expression, social interaction and solidarity, and creative digital vernaculars. The construction was done with English insertions to create a modern and cosmopolitan identity, which portrays the aspirations of the users and their involvement in the youth culture around the world. Code-mixing proved to be useful in improving playfulness and interest to the audience through humor and stylistic expression. There was also socialization and unity that hybrid language provides a common ground of communication that strengthens community ties. The engage the creative digital vernaculars put emphasis on the innovation and flexibility of bilingual communication on line whereby users change language to fit the platform-specific norms. These results also affirm the Interactional Sociolinguistics Theory by Gumperz because the options of users were determined by the conversational context and platform norms. The Domain Theory defined by Fishman was manifested in the social domain of TikTok which promoted the mixing of informal and playful languages and Myers-Scotton framework described the social factors that affected marked

selections of English in Urdu comments. All of these theories combined can be used to explain the functional and social meaning of code-mixing in online communication.

Recommendations

Based on the results, educators, linguists and researchers of digital communication are advised to appreciate the essence of code-mixing as a normal sociolinguistic phenomenon. Teachers are able to incorporate bilingual practices awareness in education programs to improve knowledge among students about language variation in the digital world. The social media platforms may take features into consideration that enable them to express in multi lingualize, as a natural way of communication among the users. It is suggested that research on cross-platform comparisons of code- mixing be undertaken to learn more about the differences between Tik Tok, Instagram, and other social media platforms. Additionally, the awareness programs may reiterate the positive aspect of code-mixing to promote creativity, cultural interaction and language identity among the young people.

Research recommendations in the future.

The scope of this study can be extended in a number of ways in future research. First, bigger corpora could be examined in various social media to draw a comparison between the trends of using codes across various digital domains. Second, longitudinal research might be employed to examine the dynamics of code-mixing as a phenomenon over time to capture the shifting linguistic norms on the one hand and cultural trends on the other hand. Third, studies can examine some regional and socio-economic differences in code-mixing to gain insights into the role of education, urbanization and media exposure. Fourth, the perceptions and the understanding of code-mixed messages by the audience could be tested through the experiments to determine its effectiveness in communication. Fifth, interdisciplinary studies may be conducted as to the overlap between linguistics and digital media and sociocultural identity to learn more about the roles and implications of online bilingualism.

Conclusion

In this paper, the author offers an in-depth examination of English Urdu code-mixing in Pakistani TikTok comments, which involves both quantitative and qualitative methods. It shows that code-mixing is not only a widespread, regular and socially significant process, but also that it fulfills a variety of roles: it is used, among other things, to construct identity, to joke, to socialize, or to express creativity. The results highlight the importance of digital platform as the area where linguistic novelty is created and social cultural identities are discussed. The research makes a contribution to the sociolinguistic theory by adopting the paradigms of Gumperz, Myers-Scotton and Fishman to digital communication to provide some understanding of how bilingual speakers negotiate between global and local languages. Finally, the paper also demonstrates how language is changing in the digital world and how the significance of code-mixing as a dynamic and socially situated phenomenon.

References

- Aijaz, R. (2019). Code-switching in digital communication: Pakistani youth on social media. *International Journal of Linguistics*, 11(4), 45–62. <https://www.ijlinguistics.com/article/view/2019-04-05>
- Asghar, I., Umar, A., Khatoon, H., Bilal, H. A., & Khan, M. T. (2021). Code-switching in an EFL classroom: A study of language choice. *Psychology and Education*, 58(1), 6663–6680.
- Awan, Z. N., & Khan, M. (2022). English-Urdu code-mixing in social media: Patterns and functions. *Journal of Sociolinguistic Studies*, 16(2), 89–105. <https://www.jss.edu.pk/articles/code-mixing>
- Awan, Z. N., Begum, U., & Khan, M. (2025). Code-switching and code-mixing in Pakistani English: A sociolinguistic analysis. *Social Science Review Archives*, 3(1), 45–61. <https://policyjournalofms.com/index.php/6/article/view/446>

- Bhatt, R. M. (2017). Social factors and bilingual language practices: A study of urban youth in South Asia. *Journal of Multilingual and Multicultural Development*, 38(6), 523–536. <https://doi.org/10.1080/01434632.2017.1285555>
- Fishman, J. A. (1972). *The sociology of language: An interdisciplinary social science approach to language in society*. Rowley, MA: Newbury House.
- Ghafoor, S. (2025). Digital vernaculars in Pakistan: A corpus-based study of Instagram posts. *Contemporary Journal of Social Science Review*, 12(1), 101–120. <https://contemporaryjournal.com/index.php/14/article/view/1486>
- Gumperz, J. J. (1982). *Discourse strategies*. Cambridge University Press.
- Jabeen, S. (2018). Code-mixing, code-switching, and borrowing in Urdu and Pakistani English in media and daily life conversations. *International Journal of Advanced Research*, 6(11), 45–60. <https://www.researchgate.net/publication/329708624>
- Myers-Scotton, C. (1993). *Social motivations for code-switching: Evidence from Africa*. Oxford University Press.
- Qureshi, I., & Ahmad, N. (2020). Sociolinguistic motivations for code-switching in online discourse. *Pakistan Journal of Social Sciences*, 40(2), 77–92. <https://pjss.org.pk/article/view/405>
- Rahman, T. (2002). *Language and politics in Pakistan*. Oxford University Press.
- Salman, M. U., Hanif, A., Shehata, S., & Nakov, P. (2023). *Detecting propaganda techniques in code-switched social media text (English ↔ Roman-Urdu)*. arXiv. <https://arxiv.org/abs/2305.14534>
- Tariq, A. R., Bilal, H. A., Abbas, N., & Mahmood, A. (2013). Functions of code-switching in bilingual classrooms. *Research on Humanities and Social Sciences*, 3(14), 29–34.
- Ullah, F. (2023). A study of code-switching patterns in Pakistani Urdu print media. *Poverty and Social Research Journal*, 5(1), 55–72. <https://poverty.com.pk/index.php/Journal/article/view/28>