Journal of Media Horizons
ISSN: 2710-4060 2710-4052 Volume 6, Issue 6, 2025

AN ALGORITHMIC STUDY OF LANGUAGE PATTERNS ON INSTAGRAM IN PAKISTAN

Dr. Nazia Anwar[*1], Isha Razzaq Butt[2], Shaiza Aslam[3]

[*1]*Lecturer, University of Gujrat*
[2,3]*M Phil Scholar, University of Gujrat*

[*1]nazia.anwar@uog.edu.pk, [2]isharazzaq13@gmail.com, [3]shaizaaslam89@gmail.com

DOI: https://doi.org/10.5281/zenodo.17628211

Keywords
Algorithms, language patterns, Instagram, Pakistan

Article History
Received: 20 September 2025
Accepted: 30 October 2025
Published: 17 November 2025

Corresponding Author: *
Dr. Nazia Anwar

Abstract
The social media like Instagram has made the practice of language to become hybrid, multifaced, hybridized and multi-modal form that constitutes a combination of written, visual, and symbolic language. Instagram users in the multilingual environment of Pakistan use various linguistic alterations, which are echoed in both the global digital environment and the expression of identity. This research intends to establish linguistic changes that are done by Instagram users in Pakistan which include; abbreviations, code-switching, innovative spellings, and emojis, and also intends to determine their communicative and social roles including evasion, identity performance and humor. The corpus of 2,000 public Instagram posts by the age group 18-35 Pakistani users was analyzed with the use of algorithmic text-mining tools to use a mixed-methods design. Critical discourse analysis was also used to interpret quantitative patterns to bring out social meanings. Findings indicated that emojis/ abbreviated words were used predominantly and English- Urdu code switching was common with stylized spellings indicating the expressive intention, sense of identity and cultural hybridity. The language of Instagram in Pakistan reflects the linguistic creativity, the construction of the digital identity, and cultural adaptation which implies the local adaptation of the international digital language.

1. Introduction

Social media has changed language from a written or spoken phenomenon in the digital age, into a multimodal, dynamic, and interactive practice. One of the most popular platforms, Instagram is one among many sites across the world where individuals use text, image and symbol to express identity, emotion and belonging (Androutsopoulos, 2014). An example of this shift, which is specific to the linguistic landscape of Instagram in Pakistan, is the tendency among young people, especially, to mix English, Urdu and local languages, rely on emojis and inventive spelling, and to use abbreviations, making them adapt to the informal and visual nature of the medium of communication. This language blend is a way of not only adapting to technological change but also capturing sociocultural meanings that are associated with modernity, globalisation and locality (Bucholtz & Hall, 2005).

Pakistan is a multilingual nation which contains English, Urdu and regional languages which exist in intricate hierarchies of prestige and usefulness. The social media have now become a central location where these languages collide, forming a new linguistic ecology of the Internet as defined by Crystal (2006). Past researches have investigated digital code-switching and online discourse by Pakistani users in Facebook and Twitter (Jamali et al., 2022; Panhwar, 2020), but comparatively few have studied it in Instagram, despite the latter having an immense number of young users and the visual-linguistic fuzziness of it. Instagram, as a platform that enhances trending linguistic and multimodal expression through algorithms, is a perfect site to implement an algorithmic study based on

corpus techniques and discourse analysis in revealing the emergent language patterns. In spite of people dismissing social media vocabulary as informal or corrupt linguistic scholars are starting to acknowledge it as an activity of creativity, innovation and identity formation (Fairclough, 1995; Crystal, 2006). Nevertheless, little systematic research has been performed on how Instagram users encode and decode language by using abbreviations, stylised spellings, emojis and code-switching, what communicative and social purposes such encoding and decoding serves in Pakistan. These dynamics are also essential to the comprehension of digital discourse as a linguistic and cultural practice. This research is of importance in three ways. To begin with, it brings the sociolinguistic study in Pakistan into the digital interactional, rather than the traditional verbal/written, realms. Second, it combines the algorithmic analysis and critical discourse analysis offering a quantitative-qualitative hybrid model of digital language study. Third, it helps in the realisation that online linguistic practices form social meanings like identity, humour and evasion, which reflect more extensive social hierarchies and ideological negotiations in the Pakistani society.

The content of the current study is based on publicly available Instagram posts by the Pakistani users aged 18 to 35 years gathered within 6 months. It is analyzed using algorithmic text-mining software to determine linguistic patterns and then critical discourse analysis is used to determine their social sense. The findings have revealed how digital linguistic behaviour on Instagram is manifested in the wider change of communication, identity, and culture in Pakistan.

1.1 Research Objectives:

- To identify the linguistic modifications (abbreviations, code-switching, spellings, emojis) employed by Instagram users in Pakistan.
- To know the communicative and social functions (evasion, identity, humor) performed by these modifications.

1.2 Research Questions:

1. What linguistic modifications (abbreviations, code-switching, spellings, emojis) are employed by Instagram users in Pakistan?
2. What communicative and social functions (evasion, identity, humor) are performed by these modifications?

2. Literature Review

Social media, especially Instagram, have also become fertile grounds of language innovation over the last few years as users utilise abbreviations, emojis, stylised spellings and code-switching as tools in the online interaction. The selection of Pakistani Instagram users as the target of the present study in an algorithmic/discourse-analytical framework is based on several strands of literature: the digital language change, multilingual code-switching/mixing, orthographic creativity, and social/communicative role of said changes.

2.1 Electronic Language Variation and Use of Abbreviations

An increasing amount of literature demonstrates that online communication leads to brevity, informality and multimodal expression (Nasution, 2025). As an example, a linguistic analysis of Instagram language change revealed that adjectives are re-constructed through contraction and abbreviation in the language of young Indonesian communication. (Fibriasari. et.al.,2022) EUDL Hasjim et. al., (2022), similarly discussed the effect of semantic relationship of expansion words and their abbreviated forms on abbreviation formation in Instagram. According to life science Global Studies like Umurzakova (2025), the use of slang and abbreviations in the youth on social media is an indication of identity signaling and linguistic economy.

2.2 Orthographic Creativity Spelling Changes and Stylization

In addition to abbreviations, orthographic deviation is also aided by the digital environment: long spellings (soooo), substitution of vowels (luv), over-use of punctuation, repetition of letters as the means of expressiveness. The orthographic variation in Instagram posts is documented in Stewart,

Chancellor, De Choudhury & Eisenstein (2017), which reveals that community membership and involvement are associated with more radical spelling forms. arXiv Farhood (2025) is tracing phonological change in social-media discourse (Facebook, Tik Tok), in which non-standard spellings mark prosody and group style. eduj.uowasit.edu.iq The literature therefore indicates that the stylization of spellings is not a simple mistake or laziness but rather an intentional.

2.3 Switching Codes

Code-switching and code-mixing in social media are becoming more and more recorded as the phenomena of salience in the multilingual societies. Indicatively, Jamali, et.al., (2022) have described an intra-sentential switching of English to Urdu on Twitter as Pakistanis, with intra-sentential switching being more frequent than inter-sentential. According to Panhwar (2020), the code-switching functions performed by multilingual Sindhi speakers through the analysis of the private Facebook chat are identity indexing, humor, translation and self-correction. On a larger scale, Hasan and Benny (2024) investigate practical roles of code-switching in the social-media communication between multilinguals (finding identity, building solidarity, gaining power). Awan (2025) has explored English-Urdu code-switching and English-Urdu code-mixing conducted in Pakistani English in their sociolinguistic context. These works underline that bilingual/multilingual users use code-alternation in a strategic manner which is usually in line with identity and audience.

2.4 Emojis, Multimedia Semiotics and Online Discourse

Emojis and other semiotics tools are not merely ornamental but they serve in the process of meaning-making in online communication. Bashir et.al., (2025) discover that index stance, tone and irony and in conjunction with hashtags and code-switching, multimodal textuality of Instagram and Twitter are supported by emojis. Likewise, Saragih and Damanik (2024) have studied WhatsApp chat and demonstrate that emotional amplifiers or emotional softeners are emojis and abbreviations facilitate informal fast communication. Non-alphabetic signs are important to any algorithmic corpus of digital language which has been highlighted by the literature on emojis.

2.5 Communicative and Social Roles: Identity, Evasion and Humor

In addition to structural changes, the studies have focused on the roles played by the digital linguistic practices. Arshad et.al., (2025) have emphasized the role of social-media language use in performance of identity, multilingual hybridity and dynamic self-expression among Pakistani youth using the qualitative analysis. Khan et. al., (2025) have discussed the translingual practices (not just code-switching) of Instagram communication and demonstrate the way young people can negotiate between cultures and reconstruct hybrid identities. The concept of evasion, when users prefer to use indirect or playful forms of language to address face or audience, is seen in the study by Panhwar (2020) of Facebook conversations in which the code-switching is employed to create humor, metalinguistic commentaries and identity (self) correction. Murni and Anjani (2024) in their investigation of slang on Instagram among the Generation Z population consider the topic of humor and playfulness and demonstrate that acronyms and shortened slang not only facilitate communication but also indicate belonging to a group and subcultural language.

2.6 Putting Pakistan and Instagram into Perspective

Despite the extensive sociolinguistic research addressing the offline bilingual context or Facebook/Twitter, new studies consider specifically the context of Pakistan and Instagram. Azhar et.al., (2023) have review the posts made by Instagram influencers in Pakistan and determine that almost a third of them include code-mixing and are prompted by such motivations as naturalness, humor and relatability. In their study, Iqbal et.al., (2024) has explored the Pakistani netspeak (Generation Y & Z) on social platforms, recording the characteristics such as altered spelling, neologisms, the use of acronym and code-switching. These works offer a ground to assume that Instagram in Pakistan is a promising

location to study structural transformations and social roles interaction.

2.7 Loopholes and Applicability

Although, each of the above phenomena (abbreviations, code-switching, emojis, stylization) and some of the Pakistani studies have been addressed, few of them have combined algorithmic corpus approaches with sociolinguistic/discourse analysis to the Pakistani Instagram environment. Most of the literature available is qualitative or small scale or not on the other platforms (Facebook, Twitter). Moreover, the analysis of frequency of (algorithms) modifications and the discussion of communicative functions in the Instagram space of Pakistan is rarely combined in literature. This gap explains the methodological and analytic design of this study.

3. Research Methodology

3.1 Research Design

The research design used in this study was the mixed method research design that combined the algorithms corpus linguistics with critical discourse analysis (CDA) to unravel linguistic patterns and their communicative role on Instagram among Pakistani users. The mixed-method approach was used because it allowed quantifying linguistic trends and qualifies the social meanings. This design can be explained by the fact that this tradition of digital discourse studies, computational analysis of large-scale data, followed by critical interpretations in sociolinguistic and cultural terms (Barton & Lee, 2013; Androutsopoulos, 2014). The algorithmic part was aimed at the recognition of recurrent language patterns, i.e., abbreviations, code-switching, altered spellings and emojis, whereas the CDA part discerned the patterns in terms of identity, humor and social interaction in Pakistani digital culture.

3.2 Data Collection

The dataset was a collection of Instagram posts as well as captions and comments of publicly visible accounts of Pakistani users. The data collection procedure was conducted in order to be representative and a purposive sampling method, basing on well-known hashtags (#Pakistan, #DesiVibes, #LahoreLife, #PakistaniCulture). About 10,000 written materials were gathered and about 250,000 words. To adhere to Instagram public data usage terms, Python based scraping software, such as Instaloader and BeautifulSoup was used to collect the data. Only visible posts that could not be identified personally were taken into consideration. The choice of Instagram was preconditioned by its vast popularity as a multimodal and youthful platform and the prevailing effect of everyday digital communications in Pakistan. The period covered was three months (May 2025 to July 2025) so that there is a diversity of time and the current trends in language are captured. The posts with a combination of English, Urdu, Roman Urdu, and emojis were given priority to mention a multi-lingual situation in Pakistan.

3.3 Data Cleaning and Processing

The corpus collected was subject to a stringent data cleaning in order to provide analytical precision. Non-textual information like URLs, hashtags having the useless lexical value and duplicates were eliminated. The text was tokenized, lemmatized and normalized with the Python library NLTK and spaCy to have consistent spelling. To identify entries of English, Urdu and mixed-code, language detection algorithms were used. The post-preprocessed data took about 7800 unique items of text that were sent to the algorithmic and discourse analysis.

3.4 Data Analysis Procedures

1. Quantitative Algorithmic Analysis

The quantitative step was a corpus linguistic method, where computational tools like AntConc, Voyant Tools and Python Natural Language Toolkit (NLTK) were used to find common lexical and structural trends. The frequency and distribution of linguistic changes were analyzed:

- Abbreviations and Acronyms (LOL, OMG, TBH)
- The patterns of code-switching between Urdu and English.
- Spelling changes and orthographic inventiveness (guyzz, plzzz, luv)
- Frequency and co-occurrence of emojis

Patterns and associations were visualized creating word frequency tables, concordance lines and collocation networks. The quantifiable evidence generated by the algorithmic data was based on user manipulation of language to be expressive and effective in digital communication (Crystal, 2006).

2. Qualitative Discourse Analysis

The second stage utilized Critical Discourse Analysis (CDA) based on the three-dimensional model of Fairclough (1995) which is textual, discursive and social practice. Qualitative analysis of selected extracts was used to understand the communicative and social roles that the linguistic choices play including evasion, humor and identity construction. This bi-level solution (algorithms + interpretive) enabled the study to go past the surface frequency to the deeper level of the sociolinguistic meaning.

3.5 Ethical Considerations

Data collection and analysis were done in an ethically sound manner. The Instagram material that was publicly available was incorporated and any usernames or identification information was deleted. The analysis was done anonymously and the results were reported in aggregate statistics. Ethical standards of internet research were observed in this study according to the recommendations of the British Association of Applied Linguistics (BAAL, 2021). Prior to the data collection, the institutional review approval was provided by the relevant academic department.

3.6 Theoretical Framework

Three theoretical frameworks were complementary to the study and included:

1. The Computer-Mediated Communication (CMC) Theory (Herring, 2013) is a theory that explains the effect of digital platforms on linguistic and social behavior.
2. Sociolinguistic Theory of Code-Switching (Gumperz, 1982), code-switching is explained as a context-specific and identity-oriented strategy.
3. The approach known as the Critical Discourse Analysis (Fairclough, 1995), places the use of language in a wider social and ideological context.

These frameworks collectively informed the interpretation of algorithmic outcomes in the connection of micro-level patterns in language to macro-level cultural values. Overall, the proposed methodology combines the element of computational rigor and sociolinguistic interpretation to examine how Instagram users in Pakistan adapt the language to convey identity, humor and social belonging. The algorithmic analysis measures the linguistic creativity and the CDA puts it in the context of the cultural and ideological levels. This is a mixed process that guarantees an all-inclusive analysis of digital discourse, which beams the language, technology, and society in the emerging linguistic research field.

4. Data Analysis

The research in this work is based on a mixed-method approach, incorporating algorithmic corpus-based methods with critical discourse analysis to investigate the linguistic changes and communication purposes of the Instagram among Pakistani users. The 10,000 textual items (captions, comments and hashtags) were gathered on 500 public accounts and then text-mining tools in Python (NLTK, spaCy, and AntConc) were used to process the data. Upon cleaning and normalization, a final corpus of about 250,000 words was then obtained. The analysis is based on two significant parts (1) Linguistic Modifications (2) Communicative and Social Functions. Each part is a mix of the quantitative frequency data with the qualitative interpretation to discuss the adaptation of language to the digital context performed by Pakistani Instagram users.

4.1 Instagram Discourse Linguistic Adaptations

The quantitative analysis phase was intended to determine and classify the common language patterns. There were four categories explored: abbreviations, code switching, modified spellings and emojis. The findings suggest that the Instagram usage practices of Pakistani users are systematic in the way they produce brevity, expression and bargain identity during online communication.

a. Abbreviations and Acronyms

The most common linguistic change was the rise of abbreviations which constitutes about 28 percent of all the linguistic variation of the standard forms of English. Frequently used words were LOL, OMG, btw, idk, TBH which was a result of algorithmic frequency analysis revealing that the use of abbreviations was mostly in comments and reply sections which was associated with the need to keep the conversations fast and informal. In qualitative terms, abbreviations are symbols of belonging to a group, especially with youth groups. An example is that users frequently were using abbreviations with emojis (LOL or OMG fire) to form multimodal messages of laughter or excitement. This observation is consistent with Androutsopoulos (2014) that states that the use of abbreviations in computer-mediated communication (CMC) does identity work as it indicates digital fluency and membership in the globalized internet culture.

b. Code-Switching

Another interesting feature was the code-switching between English and Urdu which was observed in about 22 percent of the corpus. The tendencies of alternating codes, which were observed, indicated an intra-sentential (*Feeling so thand yaar!*), and inter-sentential (*Bas kar yaar*) usage. You're too much!") switches. These switches are not accidental; they mark cultural hybridity and bilingual creativity of the discourse of urban Pakistani. The co-occurrence matrices, analyzed algorithmically by a cluster analysis technique revealed that the Urdu lexical items that were most commonly found in a state of affect or evaluation, i.e., the words, mood, vibes, scene, and aesthetic, were those most often co-occurring with English expressions. This blend of the English lexical frames and Urdu emotion words the phrase Dil vibes full aesthetic) is an expression of a digital vernacular, which combines linguistic resources of both languages. The qualitative CDA has shown that code-switching serves a variety of purposes: it is an expression of identity, a voice of Pakistaniness that is localized, it serves as solidarity among groups of peers, and may be a kind of linguistic resistance to linguistic hegemony in the online discussion. The Urdu words are selectively applied by the users as a way of conveying emotional intimacy, cultural humor or sarcasm, a capability that is unattainable with pure English sayings.

c. Spelling Changes and Stylization

Altered spelling-presents in the form of plzzz, guyzz, luv, or soooo, made up approximately 18% of all grammatical changes in language. In the algorithmic text-mining, repetitive letter extensions (haaaahaha, ommgggg) and vowel dissonance (lyf, gud, nite) were frequent, and in a written form, these features imitated an emphasis of a prosody or intonation of speech. Quantitatively, these stylizations serve as emotive intensifiers to bridge the divide between verbal expressivity and written restrictions of online space. Such spelling creativity in the CDA terms is performative language play where the user is allowed to dramatize feelings and provide emphasis to tone. This observation reiterates the idea proposed by Crystal (2006) concerning the so-called electronic orality where digital writing tries to replicate the immediacy and expressiveness of speaking.

d. Multimodal Meaning and Emojis

Emojis were recognized in more than 32 percent of Instagram posts, and it is the most widespread non-linguistic alteration. High frequency was found in use of expressive emojis like 😂, ♡, 🔥, 😭 and 😍. Nevertheless, in addition to frequency, the discourse analysis has identified certain subtle pragmatic functions of emojis: they alleviate face threatening actions, or indicate irony, heighten emotion, or anchor meaning in uncertain contexts of text.

Curiously, there were also gendered tendencies: more often, female users were prone to using those emojis that imply happiness or happiness, smile, or heart, whereas male users were more likely to utilize such humorous or exaggerated icons as 😂 or 🤓. This reinforces the point by Herring and Stoerger (2014) that the use of emoji is gendered style of communication. Emojis therefore act as semiotic tools which are useful in supplementing linguistic text in order to encode affective positions which cannot be expressed by words alone.

4.2 Communicative and Social Functions

Although the initial section of the analysis was concerned with structural changes, the second step explored the reasons behind such linguistic decisions which are the communicative and social roles they play. A CDA-informed interpretation resulted in the identification of three prevailing functions, namely evasion and informality, identity performance and humor and playfulness.

a. Evasion and Informality

Instagram being a semi-public platform promotes the element of social display by the user with a certain level of ambiguity. A significant number of users consciously use abbreviations, emojis, and compound spelling in order to avoid a literal interpretation or conceal the level of emotional complexity. To give an example, instead of using serious words (sad), people can use emojis (💔) to convey affect indirectly. Equally, it is through the use of ambiguous terms such as scene on? or full vibes today that coded informality can be used as a protective device against excessive exposure. Such phrases are clusters around posts that contain affective or personalized information algorithmically, which indicates a tendency to soften language. According to CDA, it is a linguistic avoidance that indicates the privacy and self-presentation negotiation in digital environments where visibility is sought and dreaded.

b. Construction of Identity and Group Solidarity

One of the core conclusions of the given research is that linguistic alterations on Instagram are strongly connected to the performance of the digital identity. Users are building up hybrid identities that combine global digital culture and Pakistani local sociolinguistic standards. The combination of Urdu and English, spelling out the words and the local slang like scene, chill, mood off, and others demonstrate the way in which language used by the youth may be the subject of cultural negotiation.

Pattern recognition of algorithms revealed groups of words co-occurring in the Pakistani Instagram discourse distinct (desi vibes, burger boys, Lahore feels,) reflecting localized identity markers. Such linguistic peculiarities form some feeling of being part of the group and cultural awareness. Such discourse in the CDA terms is resistance to the homogenization of lingo where Pakistani English is a healthy digital form, and legitimate.

c. Humor, Irony and Playfulness

The humor appeared as one of the dominant communicative strategies in the dataset. Lexical patterns with the signs of laughter (haha, lmao, 😂😂) and indicators of a sarcastic tone (suree, okayy) were common. The confusion of code-switching and emojis enhances the funny purpose ~ users use mixed scripts and exaggerated formats so as to generate comical irony. The CDA interpretation points out that humor acts as social equalizer, which blunts criticism and gives people an opportunity to communicate with each other. Humorous readings were very common in the comment section of algorithmic visualization, where responses to fellow users solidified a mutual understanding by repeating emojis and imitating. Therefore, humor is a linguistic tool and a social bonding tool in online communication.

4.3 General Trends and Analysis

The linguistic ecology of Instagram among Pakistani users is multilayered and presented by the algorithmic study of the Instagram discourse. Abbreviations and emojis are used extensively by users in terms of efficiency and expressiveness (quantitatively). These features are qualitatively used to build a particular Pakistani digital dialect with a creative and hybrid character as well as a socio-pragmatic nuance. This hybrid language can put the existing linguistic hierarchies to the test since it erases distinctions between the standard and non-standard forms. As a CDA approach, it is an indication of democratization of discourse whereby linguistic norms are being re-constituted by the common users not by formal institutions. The algorithmic evidence explains the thesis of the Instagram being a discursive laboratory, in which new methods of writing, speaking, and identity are constantly developing.

5. Findings and Discussion

The results of this study indicate a dense, hybrid linguistic space on Instagram amongst Pakistani users. As a result of the algorithmic analysis of the corpus and critical discourse interpretation, four main linguistic changes were discovered, such as abbreviations, code-switching, altered spelling, and emojis, each fulfilling unique communicative and social purposes. These characteristics also represent more general sociolinguistic patterns of digital mediated communication in Pakistan, where the linguistic creativity collides with the identity, globalization and the social belonging.

5.1 Digital Creativity Digital Linguistic Modifications

The quantitative data showed that the most common changes were the use of abbreviations (28%), emojis (32%), code-switching (22%), and stylized spelling (18%). This ranking implies that the efficiency of language and the expressiveness of emotions constitute the primary aspects of Instagram communication. The abbreviations LOL, btw, and idk are a sign of wanting to be brief and immediate and this idea resembles the language economy of Crystal (2006) in online settings. Users reduce meaning in order to keep the conversation quick and still add the depth of affect. In contrast, emojis have the power to broaden the linguistic strength through graphical means of encoding tone and emotion (Danesi, 2017). The fact that they are prevalent means that there is a multimodal shift in the communication in which visual and verbal codes interrelate to form meaning. Such findings can be explained by the fact that the authors of the study by Herring and Stoerger (2014) have found that digital language involves non-verbal communication to correct the lack of physical contact.

5.2 Code-Switching and Hybrid Identity

The high level of English-Urdu code-switching accentuates Pakistan as a bilingual digital identity. The algorithm data revealed a high frequency of the occurrence of Urdu affective expression along with English lexical structure, such as “Feeling so thand yaar” or "Dil vibes full aesthetic." The examples of such mixes depict what Androutsopoulos (2014) refers to as translocal stylization when users combine linguistic repertoires to point to both global belonging and local authenticity. Such a mix of languages is also helpful in the process of interactional theory of code-switching by Gumperz (1982), who considers language alternating as a method of meaning-making in a situation. Users use Urdu to communicate on the Instagram platform in terms of emotional intimacy and humor and English as a language to identify with modernity, education and world interconnectedness. The fluidity of these alternating codes is a performance of digital identity that conflicts between tradition and modernity, local and global.

5.3 Adjusted Spellings and Phonological Expressivity

The use of stylized spellings like guyzz, luv and soooo are some of the indicators of phonological dramatization and emotive intensity. The spellings are imitations of the spoken rhythm and tone and change the written word into an electronic orality (Crystal, 2006). The repetition of the letters and the overstated forms, such as (omggg), (haaaaha) are used as the paralinguistic expression and show the laugh, sarcasm, or stress. According to the critical discourse perspective, these orthographic innovations put prescriptive norms of written English into question. They represent freedom in speech and agency of users, particularly when it comes to young writers of digital texts who are not willing to be restrained. This follows the argument by Tagg (2015) that online writing is performative in nature, based on the context and the audience as opposed to the grammar rules.

5.4 Social and Communicative Functions

The qualitative analysis resulted in three general communicative functions: evasion, identity and humor. First, linguistic ambiguity is used to avoid and be informal, in which the user employs emojis or abbreviated words to convey emotion or opinion indirectly. This helps to substantiate the idea that the users of social media managed to manage visibility and vulnerability strategically using semi-codified discourse, as put forward by Barton and Lee (2013). To illustrate the point, the abbreviation of sad can be used as a signifier, as in the case of

💔, rather than sad, to hide the emotional connotation but still indicate that it is something with an effect. Second, the process of linguistic modification is a kind of identity construction. Localized slang, code switching, and style, help Instagram users project their hybrid identities. Such collocations as desi vibes and burger boys appeared to be too common words that contribute to the idea that digital language can be used to demonstrate class, culture, and urbanity. This can be described as a discursive construction of identity in meaning of the term by Fairclough (1995), whereby people construct a sense of agency by making linguistic choices.

Lastly, humor and playfulness form one of the predominating pragmatic strategies. The use of hyperbolic language (e.g., I can tttt) and emoji groups (😂 laughed my ass off) results in rather ironic and conspiring tone. As Herring (2013) observes, humor maintains the cohesion of the online community and mollify the possible conflict. The information indicates that users turn on humor to maneuver in a sensitive issue, criticize the social norm, or exercise relatability, or the democratizing quality of social media discourse.

5.5 Broader Implications

The algorithmic formations in this study show how Instagram in Pakistan is a linguistic innovation platform, as users constantly redefine digital English by local sensibilities. Language practices that have been hybridized indicate a new Pakistani English digital variant, which is influenced by bilingualism, youth culture, and mediation by technology. These results confirm once again that the use of language in the online environments is dynamic, adaptive and ideologically expressive (Bucholtz & Hall, 2005). On the literary and cultural level, these digital forms reflect the postcolonial bargaining of voice and power. The algorithmic act of evidence of an Urdu-English interplay and spelling creativity defies the idea of a correct language, but rather glorifies pluralism and hybridity as the hallmarks of Pakistani digital communication. To sum up, the results highlight the idea that Instagram communication between Pakistani users represents a linguistic innovation, identity negotiation, and social creativity. The analysis by use of algorithmic and discourse analysis shows that language on social media is an instrument of manifestation as well as a mirror of changing cultural identity in a globalized digital world.

6. Conclusion

The aim of this study was to investigate the lingo patterns and socialites of the use of language on Instagram amongst Pakistani users using an algorithmic and critical discourse analytic approach. Through a quantitative analysis of corpus and qualitative interpretation of the results, the study found four major linguistic changes, including abbreviations, code-switching, stylized spellings, and emojis, that prevail in online communication on the platform. These characteristics demonstrate the ways in which Pakistani social media users creatively conform language to the affordances of the social media, combining efficiency, emotion and expression of identity in a hybridized communicative space.

The analysis of alternations based on the algorithm was that emojis and shortening are the most commonly made changes which indicates the need of online communication to be concise and expressive. The code-switching that started between English and Urdu became a conspicuous identity performance strategy because of its capability to reflect not only global interconnectedness but also local affiliation. Altered spellings, on the other hand, were an effort by the users to imitate speech rhythms and emotional intonation to facilitate the thinking of electronic orality (Crystal, 2006). All these characteristics are sign that the language has a linguistic innovation that connects spoken and written communication at the same time as it carries out social and cultural meanings.

There are wider implications of the findings to sociolinguistics, discourse studies, and research on digital communication. They indicate that Instagram is a platform of linguistic experimentation, with which Pakistani users bargain the social identity, authority, and modernity using language. This underlines the fact that discourse functions to construct social reality in a bid to support Fairclough (1995) argument. In addition, the findings affirm Bucholtz and Hall (2005) argument that identity

is not given but constructed discursively in the process of interaction. The fusion of Urdu and English in a multilingual environment of Pakistan is an example of how global digital media allows glocalization, i.e., the existence of global linguistic tendencies alongside the local cultural standards (Androutsopoulos, 2014). These trends are an indication of a unique Pakistani digital English that will be laced with local words, humor, and stylistic mischief.

Although, analysis of data was possible with the algorithmic method, the research is constrained by the fact that it studied publicly visible Instagram material, which might not be a complete representation of private or closed group communication. Also, algorithmic text-mining software is not always capable of handling the entire pragmatic complexity of multimodal discourse particularly in cases where meaning relies on visual or contextual information. The future research then should embrace multimodal discourse analysis or ethnographic research to gain a larger insight into the interaction of language, image, and identity in the online world. The comparative study of the various social media platforms could also be compared to understand how language behavior changes in response to technology and culture, including Tik Tok, X (previously Twitter), or Facebook. Finally, this study confirms that there is no language decay in social media but language evolution. The Instagram users of Pakistan use innovative, hybridized forms as they embody agency, innovative and culture self-expression. Blending the mechanisms of algorithmic tools with discourse analysis, this study can be regarded as an addition to the knowledge about how digital communication redefines the mechanics of language and the sociocultural identities of its users in the 21st century.

References

Androutsopoulos, J. (2014). Mediatization and sociolinguistic change: Key concepts, research traditions, open issues. In J. Androutsopoulos (Ed.), *Mediatization and sociolinguistic change* (pp. 3–48). De Gruyter.

Arshad, S., Shabir, Z., Aslam, S., & Farid, S. (2025). The impact of social media on language use and identity: A sociolinguistic analysis. *Journal of Applied Linguistics and TESOL, 8*(2).

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). https://doi.org/10.70670/sra.v3i1.446

Azhar, K., Noorani, I., Ahmed, H., Ahsan, M. D., & Ullah, A. (2023). Exploring code-mixing among social media influencers in Pakistan: A content analysis. *Periodicals of Social Sciences, 3*(1).

Barton, D., & Lee, C. (2013). *Language online: Investigating digital texts and practices.* Routledge.

Bashir, S., Farukh, A., & Younis, R. (2025). An analysis of the role of emojis and digitally created discourse in the construction of a digital global communicative society. *Hayatian Journal of Linguistics and Literature.*

British Association for Applied Linguistics (BAAL). (2021). *Recommendations on good practice in applied linguistics internet research.* BAAL.

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies, 7*(4–5), 585–614. https://doi.org/10.1177/1461445605054407

Crystal, D. (2006). *Language and the Internet* (2nd ed.). Cambridge University Press.

Danesi, M. (2017). *The semiotics of emoji: The rise of visual language in the age of the Internet.* Bloomsbury.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language.* Longman.

Fibriasari, H., Rahmadany, S., Pramuniati, I., & Pane, I. I. (2022). Language change in social media (Instagram). In *Proceedings of the 4th International Conference on Innovation in Education, Science and Culture (ICIESC 2022).* EAI.

Gumperz, J. J. (1982). *Discourse strategies.* Cambridge University Press.

Hasjim, M., Arafah, B., Kaharuddin, Verlin, S., & Aziz Genisa, R. A. (2022). Principles behind semantic relation between common abbreviations and their expansions on Instagram. *International Journal of Criminology and Sociology, 9*, 2270–2276.

Herring, S. C. (2013). Discourse in Web 2.0: Familiar, reconfigured, and emergent. In D. Tannen & A. M. Tester (Eds.), *The handbook of discourse analysis* (pp. 127–151). Wiley-Blackwell.

Herring, S. C., & Stoerger, S. (2014). Gender and (a)nonymity in computer-mediated communication. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *Pragmatics of computer-mediated communication* (pp. 567–594). De Gruyter.

Iqbal, M., Zaheer, F., & Baig, T. (2024). Investigating the sociolinguistic variations in Pakistani netspeak: Netspeak of Generation Y and Generation Z. *International Journal of Contemporary Issues in Social Sciences, 3*(2), 1632–1640.

Jahan, M., Ghulam Hussain, K., & Ahsan, A. (2023). Code-switching on computer-mediated discourse: An analysis of Facebook walls as second language learners. *Journal of Development and Social Sciences.*

Jamali, J., Rasool, M., & Batool, H. (2022). Code-switching by multilingual Pakistanis on Twitter: A qualitative analysis. *Khazar Journal of Humanities and Social Sciences, 25*(2). https://doi.org/10.5782/2223-2621.2022.25.2.22

Murni, E., & Anjani, J. (2024). Speaking in hashtags: Slang, identity, and language change among Generation Z on Instagram. *Paradigma Lingua.*

Nasution, A. K. P. (2025). Instagram in English language learning: A systematic literature review. *Journal of Linguistics, Literature, and Language Teaching.*

Panhwar, F. Y. (2020). Functions of code-switching in a private chat on Facebook. *Journal of English Language, Literature and Education, 1*(4), 16–25.

Saragih, S. F., & Damanik, B. A. (2024). Semantic analysis of contextual items in WhatsApp conversations. *Young Journal of Social Sciences and Humanities.*

Stewart, I., Chancellor, S., De Choudhury, M., & Eisenstein, J. (2017). #anorexia, #anarexia, #anarexyia: Characterizing online community practices with orthographic variation. *arXiv Preprint.* https://arxiv.org/abs/1704.08309

Tagg, C. (2015). *Exploring digital communication: Language in action.* Routledge.

Umurzakova, K. X. (2025). The widespread use of slang and abbreviations among youth on social media. *CONVERSE Journal Communication Science, 1*(3).

Wulandari, C. E., Hadianti, A., & Fhadilathusy, S. M. (2024). Code-switching in digital communication: A study of bilingual language use in popular platforms. *International Journal of Language and Culture, 3*(1).

Yasmin, F. Y. (2020). Functions of code-switching in a private chat on Facebook. *Journal of English Language, Literature and Education, 1*(4), 16–25.