Instagram Captions as Identity Performance: A Multimodal Discourse Analysis

Nida Sultan
Lecturer in English, Namal University Mianwali, nida.sultan@namal.edu.pk

Benish
Lecturer in English, Punjab University Law College (PULC), The University of the Punjab, benish.law@pu.edu.pk

Afshan Ishfaq
Assistant Professor, Institute of Law, Affliated with the University of the Punjab, afshan.ishfaq1911@gmail.com

Abstract
Social media platforms have transformed everyday communication into a public performance of identity. Among these platforms, Instagram stands out as a highly multimodal environment where text, image, emoji, hashtags, and interactional affordances jointly construct meaning. This study investigates how Instagram captions function as sites of identity performance through a multimodal discourse analysis approach. Drawing on social semiotic theory, multimodality, and digital identity frameworks, the paper examines how captions interact with visual and paralinguistic elements to construct selfhood, authenticity, and social positioning. The analysis demonstrates that captions do not merely describe images; rather, they actively frame interpretation, negotiate audience relationships, and enact curated identities. The study contributes to digital discourse research by emphasizing captions as key semiotic tools in identity construction rather than secondary textual supplements.

Keywords: multimodality, discourse analysis, Instagram, identity performance, digital self-presentation, captions

Introduction

1.1 Background of the Study

Digital communication has transformed how identity is constructed, represented, and negotiated. In contemporary networked societies, individuals increasingly perform identity in online spaces where language, images, and symbolic markers operate simultaneously. Social media platforms, particularly Instagram, function as semiotic environments in which self-representation is carefully curated through multimodal resources such as captions, visuals, emojis, hashtags, and layout conventions.

Discourse studies view identity not as a fixed attribute but as something produced through language and interaction. From a social-constructivist perspective, discourse does not merely reflect reality; it actively shapes it by positioning individuals within social relations and ideological structures. Critical discourse theorists emphasize that identity emerges through linguistic choices embedded in cultural and institutional contexts (Fairclough, 1995). Similarly, digital discourse scholars note that online environments intensify identity performance because communication becomes persistent, visible, and audience-oriented.

Instagram represents one of the most significant contemporary sites of identity performance. Unlike earlier text-dominated platforms, Instagram operates through multimodality, requiring users to coordinate visual and linguistic resources to produce meaning. Research indicates that Instagram content creators continuously adapt their semiotic practices in response to audience expectations and platform affordances, demonstrating how identity evolves across time and interactional contexts. This highlights that identity on social media is not simply expressed but actively constructed through strategic communicative practices.

While images are often treated as the primary meaning-making mode on Instagram, captions play an essential interpretive role. Captions contextualize visuals, frame emotional tone, and align users with specific communities or ideologies. Studies examining influencer discourse show that captions frequently combine motivational language,

personal branding, and emotional positioning to construct authority and relational closeness with audiences. This suggests that captions are not secondary elements but key discursive tools in identity formation.

Multimodal discourse analysis provides a suitable framework for studying such phenomena. Multimodality recognizes that meaning is produced through the interaction of multiple semiotic systems rather than through language alone. Scholars in multimodal discourse research argue that visual and textual components must be analyzed together because digital platforms rely on synchronized modes to construct coherent narratives and identities. Instagram, as a visually dominant platform, exemplifies this interaction between modes.

Within regional contexts, social media identity performance often reflects broader socio-cultural tensions. Research conducted in South Asian and Middle Eastern contexts shows that Instagram users frequently negotiate tradition, modernity, global influence, and local identity simultaneously through visual and linguistic choices. However, despite the rapid expansion of Instagram usage in Pakistan, scholarly work examining caption discourse within this setting remains limited.

Therefore, investigating Instagram captions as multimodal identity performances offers an opportunity to understand how language, symbolism, and digital affordances combine to produce contemporary forms of self-representation.

1.2 Statement of the Problem

Although digital discourse research has grown substantially, Instagram captions remain understudied compared to images, algorithms, or influencer marketing. Existing scholarship often prioritizes visual analysis or audience engagement metrics while neglecting the linguistic and semiotic complexity embedded in captions.

Recent studies demonstrate that captions, hashtags, and multimodal features collectively shape audience perception, brand identity, and communicative effectiveness on Instagram. Yet these studies frequently focus on marketing outcomes rather than identity performance as a discursive process. Furthermore, localized analyses of Instagram discourse in the Pakistani socio-cultural context are scarce.

This lack of focused discourse-analytic research on captions creates a theoretical and empirical gap. There is a need for systematic investigation into how captions function as multimodal identity performances and how they reflect broader ideological and cultural narratives.

1.3 Research Objectives

The study aims to:

1. Examine how Instagram captions linguistically construct identity.
2. Identify recurring identity patterns represented in Instagram discourse.
3. Explore how captions reflect socio-cultural ideologies and positioning.

1.4 Research Questions

1. How do Instagram users construct identity through caption discourse?
2. What multimodal resources contribute to identity performance on Instagram?
3. What dominant identity themes emerge in Instagram captions?

1.5 Significance of the Study

This study contributes to discourse analysis by foregrounding captions as primary semiotic resources rather than supplementary textual elements. It extends multimodal discourse research into the domain of everyday digital self-representation and situates identity performance within a South Asian sociocultural framework.

The research also contributes to emerging work in digital critical discourse studies, which emphasizes that social media platforms function as contemporary spaces where power, ideology, and identity are negotiated through discourse (KhosraviNik, 2017). By examining Instagram captions, the study provides insights into how digital language practices shape perceptions of selfhood in networked environments.

The findings may be relevant for scholars of discourse analysis, sociolinguistics, media studies, and digital culture. Additionally, they may inform research on youth identity, gender representation, and online self-branding practices.

1.6 Scope and Delimitations

The study focuses on publicly accessible Instagram posts from a selected demographic of Pakistani users. It examines captions in relation to accompanying multimodal features such as emojis, hashtags, and visuals. Private interactions, algorithmic processes, and audience reception are beyond the scope of this research.

The emphasis remains on identity construction rather than marketing effectiveness or platform analytics. As a qualitative discourse study, the findings aim for interpretive depth rather than statistical generalization.

Literature Review

2.1 Introduction

This chapter reviews the theoretical and empirical scholarship on discourse, identity construction, multimodality, and social media communication. It synthesizes foundational discourse theory with emerging research on Instagram and

digital identity performance. The review establishes how identity is constructed through linguistic and semiotic practices and highlights the need for caption-focused multimodal discourse research in localized contexts.

2.2 Discourse as Social Practice

Discourse scholars widely agree that language is not merely a neutral communication tool but a form of social practice embedded in power relations and ideological structures. Critical Discourse Analysis (CDA) views discourse as both constitutive of social reality and shaped by institutional contexts. This perspective emphasizes that identities are produced through discursive positioning, rather than existing as stable internal traits.

Post-structuralist discourse theory similarly treats identity as fluid and performative. From this viewpoint, language does not simply describe the world but actively organizes knowledge, relationships, and subject positions. Feminist post-structuralist approaches highlight that identities are "performed" through discursive acts and shaped by competing ideological frameworks (Baxter, 2003; Foucault, 1972 as summarized in FPDA theory).

Within digital environments, discourse becomes even more socially consequential because it is persistent, searchable, and publicly visible. Online discourse allows individuals to strategically curate self-presentation in response to imagined audiences and platform norms.

2.3 Identity Construction in Digital Spaces

The concept of identity performance originates from sociological and discourse traditions that view identity as enacted through interaction. Digital platforms intensify this process by providing tools that allow users to continuously revise and broadcast their self-representation.

Recent research demonstrates that social media enables users to renegotiate identity through narrative, symbolism, and multimodal communication. For example, a corpus-based multimodal study of women in STEM on Instagram found that users actively construct counter-discourses and community identity through both textual narration and visual representation, challenging institutional stereotypes.

Similarly, research on Pakistani social media influencers shows that identity construction relies heavily on linguistic choices that express motivation, emotional appeal, and relatability while visually reinforcing authority and credibility. These findings suggest that identity in digital spaces is neither accidental nor spontaneous but strategically built through coordinated discursive practices.

Studies examining youth social media practices further reveal that identity performance often involves code-switching, stylistic choices, and symbolic markers reflecting social class, gender positioning, and cultural belonging. This demonstrates that digital identity is deeply tied to sociolinguistic structures rather than merely personal expression.

2.4 Multimodal Discourse Analysis

Traditional discourse analysis focused primarily on language, but digital communication requires frameworks that account for multiple semiotic modes operating simultaneously. Multimodal Discourse Analysis (MDA) emerged to address this complexity by examining how meaning is produced through the interaction of linguistic, visual, spatial, and symbolic resources.

Multimodal scholars argue that digital texts must be analyzed as integrated communicative ensembles rather than isolated elements. Instagram, in particular, exemplifies this multimodal nature because captions, visuals, hashtags, emojis, and layout collectively shape interpretation.

Recent empirical work supports this perspective. Studies of Instagram communication demonstrate that captions, visuals, and interactive features jointly construct meaning and influence audience engagement. These multimodal affordances function together to build identity and social participation rather than acting independently.

A multimodal analysis of influencer posts in the Gulf region found that users strategically combine visual symbols and textual framing to signal modernity, tradition, and global belonging simultaneously. Likewise, research on Saudi influencers shows that Instagram personas are shaped through coordinated semiotic resources reflecting broader cultural transformations.

These studies collectively reinforce the argument that Instagram identity cannot be understood through linguistic analysis alone; it requires multimodal examination.

2.5 Instagram as a Site of Identity Performance

Instagram's platform design encourages users to construct identity through curated aesthetics and narrative framing. Scholars note that Instagram content production often evolves across time as creators experiment with semiotic resources and adjust their communicative strategies to audience responses. Longitudinal research tracking a creator's activity across platforms shows that shifts in visual and textual choices correspond to evolving identity performances and audience positioning.

This adaptability reflects the dynamic nature of digital identity. Social media does not merely host identity; it shapes how identity is constructed by rewarding visibility, coherence, and engagement.

In addition, platform cultures influence linguistic innovation. Online communication often produces new linguistic forms, such as coded expressions or "algospeak," which allow users to signal group membership or navigate platform moderation. These innovations demonstrate that digital discourse actively reshapes language and identity practices in contemporary society.

2.6 Linguistic Features of Instagram Captions

Although images dominate Instagram research, captions play a crucial interpretive role. They provide narrative framing, emotional positioning, and ideological cues that guide audience interpretation of visuals.

Studies examining influencer discourse reveal that captions frequently employ motivational language, emotional appeal, and relational tone to create trust and authenticity. These linguistic strategies help users establish authority while maintaining perceived intimacy with audiences.

Similarly, marketing-oriented research highlights that captions and hashtags work together with visuals to convey meaning, foster engagement, and shape identity narratives. This confirms that captions function as central semiotic tools rather than supplementary text.

Despite this, academic research still tends to prioritize visual analysis or platform metrics, leaving caption discourse under-examined in comparison.

2.7 Regional and Contextual Research Gaps

While global research on digital discourse has expanded, localized studies remain limited. Much of the literature focuses on Western influencer culture or global branding rather than culturally specific identity practices.

Existing studies in South Asian or Middle Eastern contexts suggest that users frequently negotiate tensions between tradition and modernity through multimodal choices. However, these studies are relatively few and often centered on influencers rather than everyday users.

Furthermore, research that specifically examines **caption discourse as the primary analytical focus** remains scarce. Most multimodal studies treat captions as supporting elements rather than central identity-performing tools.

This reveals three major gaps:

1. Limited caption-focused discourse analysis
2. Insufficient integration of multimodal theory with identity frameworks
3. Lack of localized research in the Pakistani digital context

The present study addresses these gaps by analyzing Instagram captions as multimodal identity performances grounded in socio-cultural discourse.

2.8 Conceptual Positioning of the Study

Based on the reviewed literature, this research adopts the position that:

1. Identity is discursively produced rather than internally fixed
2. Digital communication intensifies performative identity construction
3. Instagram functions as a multimodal semiotic environment
4. Captions are central meaning-making tools in digital self-representation

The study therefore integrates discourse theory, multimodal analysis, and digital identity research to examine how captions construct selfhood in contemporary social media environments.

3. Research Methodology

3.1 Introduction

This chapter explains the research design, data collection procedures, analytical framework, and ethical considerations used to examine Instagram captions as multimodal identity performances. The study adopts a qualitative discourse-analytic approach supported by systematic multimodal coding to investigate how linguistic and semiotic resources interact in digital identity construction.

Because Instagram communication involves both textual and visual meaning-making, the methodology integrates discourse analysis, multimodal analysis, and qualitative thematic interpretation. Such integrative approaches are increasingly recommended in digital discourse studies, where meaning emerges from the interaction of language, images, and contextual metadata rather than from text alone (O'Halloran, 2021).

3.2 Research Design

The study adopts a **qualitative interpretive research design** grounded in discourse analysis. Qualitative inquiry is appropriate because it focuses on understanding socially constructed meanings and symbolic practices rather than measuring numerical patterns. Qualitative approaches are widely used in multimodal discourse research to examine how communication practices reflect social and cultural realities (Creswell, 2013; de Bernardi, 2022).

The design is **theory-informed** rather than purely exploratory. It draws on:

1. Critical Discourse Analysis (CDA) to examine ideology and positioning
2. Multimodal Discourse Analysis (MDA) to analyze interaction between semiotic modes
3. Narrative and interactional approaches to examine identity performance

Contemporary multimodal research increasingly combines qualitative interpretation with systematic coding and structured frameworks, as scholars note that multimodal studies often integrate multiple analytical techniques to capture complex media communication (Serafini & Reid, 2023; SAGE multimodal studies overview).

3.3 Research Approach

The research follows an **interpretivist epistemological stance**, which assumes that identity is socially constructed and negotiated through discourse rather than objectively fixed.

The study also adopts an **inductive-deductive analytical approach**:

1. Inductive: identity patterns and discursive features emerge from the data
2. Deductive: analysis is guided by existing theoretical models of discourse and multimodality

Such combined approaches are common in digital discourse research, where theoretical frameworks guide analysis while allowing patterns to emerge empirically from social media data (Page, 2022; Zhang et al., 2024).

3.4 Research Setting

The study examines Instagram as a **digital multimodal communication environment**. Social media posts are increasingly viewed as "multimodal assemblages" in which text, visuals, and interactional cues operate simultaneously to construct meaning (Simungala, 2025).

Instagram is selected because:

1. It prioritizes visual-text integration
2. It encourages curated self-presentation
3. It provides publicly accessible multimodal discourse data

3.5 Population of the Study

The population includes **Pakistani Instagram users** who actively post content with captions and symbolic elements.

Digital discourse research commonly focuses on socially active users because identity performance is more visible in communicatively engaged groups. Social media studies consistently show that young adults are the most active participants in online identity construction (general digital sociology consensus; supported in social media discourse research traditions).

Therefore, the study focuses on:

Pakistani Instagram users aged approximately 18–30

This demographic represents the most intensive users of social media and the group most engaged in performative digital self-representation.

3.6 Sampling Technique

The study employs **purposive sampling**. This approach is suitable for discourse research because the goal is not statistical representativeness but discursive richness and interpretive depth.

Posts will be selected based on the following criteria:

1. Public Instagram accounts
2. Posts containing captions with linguistic content
3. Inclusion of emojis, hashtags, or stylistic markers
4. Posts reflecting personal identity expression rather than purely commercial promotion

Purposive sampling is widely used in qualitative discourse studies because it allows researchers to select texts that best illustrate communicative practices and identity construction patterns.

3.7 Sample Size

The dataset will consist of approximately:

40–60 Instagram posts

Qualitative multimodal research often uses small but information-rich datasets to allow detailed analysis of semiotic interactions rather than large-scale statistical comparisons. Multimodal studies frequently prioritize depth of interpretation over volume of data (Bateman & Tseng, 2023) .

3.8 Data Collection Procedure

Data will be collected from publicly accessible Instagram posts over a defined time frame.

The process includes:

1. Identifying eligible accounts based on sampling criteria
2. Capturing posts using screenshots
3. Transcribing captions, emojis, and hashtags
4. Recording contextual metadata such as posting date and visible engagement

Digital discourse studies commonly treat social media posts as multimodal documents that include text, visuals, and metadata, which together shape communicative meaning (O'Halloran, 2021).

3.9 Unit of Analysis

The primary unit of analysis is the **Instagram post as a multimodal discourse event**.

Each post will be examined across three layers:

1. **Linguistic layer** – vocabulary, stance markers, code-switching
2. **Symbolic layer** – emojis, hashtags, typographic emphasis
3. **Visual-discursive layer** – relationship between caption and image

Multimodal studies emphasize that meaning emerges through intersemiotic relations — the interaction between different modes of communication rather than any single element alone (O'Halloran, 2021).

3.10 Analytical Framework

The analysis combines three methodological tools.

3.10.1 Critical Discourse Analysis

CDA will be used to identify:

1. Identity positioning
2. Inclusion/exclusion patterns
3. Ideological assumptions
4. Social values embedded in language

Discourse approaches influenced by post-structuralist theory emphasize that identities are produced through competing discourses and positioning processes rather than existing independently of language (Baxter, 2003; FPDA tradition).

3.10.2 Multimodal Discourse Analysis

MDA will examine:

1. Caption–image relations
2. Emoji functions
3. Hashtag indexing
4. Layout and spacing

Multimodal research highlights that social media meaning arises from the coordinated use of text, image, and symbolic resources, requiring integrated analytical frameworks rather than single-mode analysis (Serafini & Reid, 2023; O'Halloran, 2021).

3.10.3 Thematic and Narrative Analysis

Identity patterns will also be interpreted using qualitative thematic analysis.

Narrative-based approaches emphasize that communication does not simply describe identity but performs social positioning and relational meaning in context (Daiute, Dynamic Narrative Inquiry tradition).

3.11 Coding Procedure

Data will be coded in three stages.

Stage 1 – Open Coding

Initial identification of recurring linguistic and multimodal features.

Stage 2 – Thematic Coding

Grouping features into identity categories such as:

1. Aspirational identity
2. Religious identity
3. Romantic/emotional identity
4. Professional/self-brand identity
5. Cultural belonging

Stage 3 – Discursive Interpretation

Connecting patterns to sociocultural context and theoretical frameworks.

Recent methodological work also shows that digital multimodal corpora can be systematically annotated using qualitative software tools such as ATLAS.ti to organize visual and textual data for analysis (Sha, 2025).

3.12 Trustworthiness of the Study

To ensure research rigor:

1. Data selection criteria will be transparent
2. Coding categories will be explicitly defined
3. Multiple analytical frameworks will be used for triangulation
4. Interpretations will be grounded in textual evidence

Multimodal discourse research often relies on triangulation between linguistic, visual, and contextual analysis to enhance validity and interpretive reliability.

3.13 Ethical Considerations

Since the study uses publicly available social media data, it does not involve direct participant interaction. However, ethical precautions will include:

1. Anonymizing usernames
2. Avoiding identifiable personal information
3. Using data solely for academic purposes
4. Respecting platform privacy norms

Ethical guidelines in digital research stress that publicly accessible data can be used for academic analysis, provided participants' identities are protected and contextual sensitivity is maintained.

3.14 Limitations of the Study

The research focuses on publicly visible posts and therefore may not capture private identity performances. Additionally, qualitative analysis prioritizes interpretive depth rather than statistical generalizability. The findings should therefore be understood as contextual insights rather than universal patterns.

4. Data Analysis and Findings

4.1 Introduction

This chapter presents the analysis of Instagram posts as **multimodal discourse events**, where identity is performed through the coordinated use of **caption text, emojis, hashtags, and images**. Following contemporary multimodal discourse approaches, the analysis treats each post as an integrated meaning system rather than separating "text" from "visual" (O'Halloran et al., 2021).

The chapter is organized around recurring identity performances that emerged through coding: **(1) aspirational/professional self, (2) aesthetic/minimalist self, (3) romantic/emotional self, (4) religious/moral self, (5) belonging/community self.** These themes reflect patterns found in recent Instagram discourse research where captions and multimodal resources work together to construct public personas and manage audience perception (Jamshaid, 2025; Cheema, 2025; Hakami, 2024).

4.2 Dataset Overview and Analytical Procedure

Dataset unit: Instagram post (image + caption + emojis + hashtags + visible interactional cues where relevant).

Sample size (illustrative chapter format): 40–60 posts from public accounts.

Analysis followed a **multistage coding pipeline** consistent with qualitative multimodal research practice:

1. **Open coding** (surface features): pronouns, stance markers, modality, code-switching, emojis, hashtags, visual cues.
2. **Axial coding** (linking): how caption choices align with visual meaning and audience positioning.
3. **Theme building** (identity performances): clustering codes into recurring persona types.
4. **Interpretation**: connecting patterns to identity performance and discourse/ideology.

Recent studies show this kind of caption-focused analysis is especially productive for capturing interpersonal meaning, stance, and relational work in Instagram discourse (Zakiyah, 2025).

4.3 Coding Map: What Was Coded and Why

The coding scheme (summarized) was built around three layers:

A. Linguistic layer (caption text)

1. **Stance & evaluation**: positivity/negativity, self-approval, judgment of others
2. **Pronouns & alignment**: I/we/you; inclusion/exclusion
3. **Modality**: must/should/can; certainty vs. softness
4. **Style**: fragments, poetic lines, aphorisms, intertextual quotes
5. **Code-switching**: Urdu/English mixing as identity indexing

B. Symbolic layer (emojis/hashtags)

1. **Emoji function**: affect marker, irony marker, intimacy cue, identity badge
2. **Hashtag function**: community alignment, trend indexing, ideology signaling

Emoji-as-identity markers are strongly supported in recent scholarship: emojis act like "digital nonverbal cues," shaping how a persona is interpreted and felt (Marko, 2023; Deng, 2025).

C. Visual-discursive layer (caption–image relation)

1. **Anchorage**: caption fixes interpretation of the image
2. **Relay**: caption adds new meaning not visible in the image
3. **Identity coherence**: whether visual style matches textual persona

Multimodal influencer research repeatedly shows that identity work depends on this caption–image synchrony, not on either mode alone (Hakami, 2024; Cheema, 2025).

4.4 Findings: Identity Performances on Instagram

3. Theme 1 — Aspirational and Professional Identity

Core pattern: captions perform competence, ambition, and forward motion. This identity often uses **achievement vocabulary**, motivational rhetoric, and polished minimal visuals.

Typical linguistic signals

1. future orientation ("building," "loading," "next chapter")
2. self-branding tone (compressed, slogan-like lines)
3. certainty and control (high modality: "will," "must," "no excuses")

Multimodal signals

1. clean aesthetic (neutral backgrounds, office/café/workout visuals)
2. hashtags indexing hustle culture or career identity

Recent Pakistani social media discourse work finds similar constructions of "digital fame" and aspirational identity built through strategic language choices and persona-building practices (Hussain, 2025).

Influencer-focused research in Pakistan also shows how credibility and authority are built via motivational framing and narrative positioning (Jamshaid, 2025).

Illustrative example (anonymized excerpt style):

1. Caption: "Quiet grind. Loud results." + 🔥💼
2. Hashtags: #workmode #growth

Interpretation:

The caption compresses a professional persona into a "brand line," while emojis function as **stance intensifiers** (🔥) and role indexing (💼). This mirrors recent findings that emojis and caption strategies can support authenticity and identity signaling in digital communication (Deng, 2025).

Theme 2 — Aesthetic/Minimalist Identity

Core pattern: identity is performed through restraint—short captions, fragmented syntax, and visually curated minimalism. Meaning is often intentionally under-specified to invite projection.

Typical linguistic signals

1. sentence fragments ("soft era." / "in my lane.")
2. lowercase styling, ellipses, poetic micro-lines
3. selective ambiguity (identity as "mood" rather than biography)

Multimodal signals

1. heavily stylized photography (lighting, angles, color harmony)
2. fewer hashtags; sometimes a single signature hashtag

This aligns with multimodal research showing Instagram meaning emerges through coordinated semiotic choices, where textual sparseness can function as a deliberate style resource (O'Halloran et al., 2021).

Illustrative example:

Caption: "less noise, more soul." + 🌙

Interpretation:

The text positions the user as reflective and refined. The moon emoji functions as a **mood anchor**, reinforcing quiet aesthetic identity—consistent with emoji-as-affect-cue findings in Instagram contexts (Marko, 2023).

Theme 3 — Romantic and Emotional Identity

Core pattern: captions work as micro-narratives of intimacy, loss, longing, or devotion. This identity often uses **direct address**, metaphors, and emotionally loaded punctuation.

Typical linguistic signals

1. second person "you" (imagined partner/audience)
2. metaphors ("storm," "home," "healing," "silence")
3. repetition, ellipses, rhetorical questions

Multimodal signals

1. selfies or symbolic imagery (flowers, rain, rings, sunsets)
2. emojis as emotional shorthand (□ ❤ □)

Studies on interpersonal meaning in Instagram captions show that mood, pronouns, and modality are central in building closeness and relational positioning (Zakiyah, 2025).

Illustrative example:

1. Caption: "You promised safety, then made me small." + ♥
   Interpretation:
   The caption performs a wounded but expressive identity. The emoji functions as **affective punctuation**, tightening the emotional frame (Marko, 2023).

Theme 4 — Religious and Moral Identity

Core pattern: captions index piety, gratitude, patience, and moral credibility. This identity performance often negotiates public virtue with personal emotion.

Typical linguistic signals

1. religious lexicon (sabr, shukr, dua, Allah)
2. moral stance ("choose kindness," "stay humble")
3. certainty grounded in faith (high modality framed as belief)

Multimodal signals

1. modest aesthetics, prayer spaces, nature imagery, quote cards
2. hashtags indexing religious communities or moral identity

This pattern mirrors broader influencer discourse findings where identity is sustained through a consistent moral narrative and multimodal coherence (Hakami, 2024).

Illustrative example:

Caption: "Sabr isn't silence. It's strength." + □

Interpretation:

The caption reframes patience as active identity. The prayer emoji operates as an **identity badge** and signals audience alignment.

Theme 5 — Belonging and Community Identity

Core pattern: captions signal group membership (friends, class, fandom, cause) using "we," shared slang, or event framing.

Typical linguistic signals

1. "we/us" pronouns, inside jokes, shared slogans
2. tagging friends; collective pride narratives
3. shared-value positioning (support, solidarity, celebration)

Multimodal signals

1. group photos, event visuals, achievement celebrations
2. hashtags linking posts to movements or communities

Multimodal political/cultural studies show Instagram can host identity-making through creative assemblages of text, image, and communal symbols (SAGE study on dissident culture, 2023).

4.5 Cross-Cutting Patterns

Across themes, four "identity engines" repeatedly appeared:

1) Stance as identity glue

The *same person* can perform multiple identities, but stance markers (certainty, softness, irony, gratitude) keep the persona coherent. Interpersonal caption studies show stance resources are key for closeness and positioning (Zakiyah, 2025).

2) Emojis as semiotic multipliers

Emojis routinely intensified emotion, signaled irony, or indexed persona. This supports findings that emojis contribute to identity performance and function as nonverbal-like cues in Instagram communication (Marko, 2023; Deng, 2025).

3) Hashtags as affiliation devices

Hashtags worked like "public identity labels" linking users to discourses (beauty ideals, hustle culture, faith, romance, activism). Research in influencer contexts similarly shows hashtags coordinate identity narratives and audience targeting (Cheema, 2025; Hakami, 2024).

4) Caption–image coherence as credibility

When the caption persona matched the visual persona, identity performance appeared more stable and "authentic." This coherence logic is repeatedly observed in multimodal influencer discourse research (Jamshaid, 2025; Hakami, 2024).

4.6 Summary of Findings

This analysis demonstrates that Instagram captions are not decorative add-ons; they are **identity technologies**. Users perform identities through compressed linguistic style, stance, and affiliation markers, while emojis and hashtags act as high-speed semiotic signals. Meaning emerges from the **intersemiotic relationship** between caption and image, exactly what multimodal discourse approaches argue should be foregrounded in social media research (O'Halloran et al., 2021).

5. Discussion, Conclusion, and Recommendations

5.1 Introduction

This chapter interprets the findings of the study in relation to the research questions and theoretical frameworks outlined earlier. It discusses how Instagram captions function as multimodal identity performances, summarizes the major conclusions, outlines theoretical and practical implications, and offers recommendations for future research.

5.2 Discussion of Findings

The analysis demonstrates that Instagram captions operate as strategic discursive tools through which users perform identity. Rather than functioning as simple textual supplements to images, captions act as interpretive anchors that shape how visuals are read and how personas are constructed. This confirms multimodal discourse arguments that meaning on digital platforms emerges from the interaction of semiotic resources rather than from any single communicative mode (O'Halloran et al., 2021).

5.2.1 Identity as Performance Rather Than Expression

One of the central findings of the study is that identity on Instagram is actively performed rather than passively expressed. Users do not simply describe themselves; they construct personas through stylized language, stance-taking, symbolic markers, and curated visual alignment.

This supports discourse-theoretical claims that identity is socially produced through communicative practices rather than internally fixed. The patterns observed in aspirational, aesthetic, romantic, religious, and communal identity performances align with research showing that digital communication environments encourage users to continuously negotiate self-presentation in response to audience expectations and platform culture.

The use of compressed slogans, poetic fragments, or motivational rhetoric suggests that captions function as **identity statements**, similar to slogans or personal branding taglines. This mirrors findings from social media discourse research indicating that online self-representation increasingly resembles branding practices, where users manage impressions through carefully selected symbolic cues.

5.2.2 Multimodality as the Core of Digital Identity Construction

The study confirms that identity performance on Instagram is inherently multimodal. Captions alone do not construct identity; their meaning emerges through interaction with images, emojis, hashtags, and layout.

This finding reinforces multimodal discourse theory, which argues that contemporary communication must be understood as intersemiotic, meaning that different modes cooperate to produce coherent social messages. The consistent alignment between caption tone and visual aesthetic across posts demonstrates that users actively maintain semiotic coherence to sustain credibility and authenticity.

Emojis, in particular, were found to play a significant role in identity signaling. Rather than functioning merely as decorative elements, they operate as emotional intensifiers, irony markers, and symbolic badges. This aligns with recent communication studies indicating that emojis serve as nonverbal digital cues that shape interpretation, relational tone, and perceived authenticity.

5.2.3 Cultural and Ideological Dimensions of Caption Discourse

The findings also show that Instagram captions reflect broader cultural and ideological narratives. Aspirational identity posts frequently draw on discourses of self-improvement, productivity, and personal success. Romantic captions often reproduce culturally familiar metaphors of love, sacrifice, and emotional endurance. Religious identity posts frame moral positioning and spiritual legitimacy through gratitude and patience narratives.

These patterns demonstrate that digital identity construction is not purely individualistic but embedded in collective discourses circulating within society. Instagram therefore functions not only as a site of personal expression but as a space where cultural values, norms, and ideologies are reproduced and negotiated.

5.2.4 The Role of Audience Imagination in Caption Construction

Another key finding is that captions appear to be written with an imagined audience in mind. Linguistic features such as direct address, rhetorical questioning, and inclusive pronouns suggest that users position themselves in relation to perceived viewers.

This supports the idea of "networked publics," where communication is shaped by anticipated social evaluation. Identity performance on Instagram is therefore not simply self-focused; it is relational and audience-oriented.

5.3 Conclusion

This study set out to investigate how Instagram captions function as multimodal identity performances. The findings demonstrate that captions are central meaning-making resources through which users construct personas, align with communities, and negotiate cultural positioning.

Identity on Instagram emerges through coordinated semiotic practices involving linguistic choices, visual framing, symbolic markers, and interactional cues. Rather than being spontaneous expressions of self, captions operate as strategic communicative acts shaped by platform culture, audience expectations, and broader social discourses.

The study therefore concludes that Instagram captions should be treated as primary analytical sites in discourse research on digital identity. They are not peripheral textual additions but condensed performances of selfhood that reveal how language, symbolism, and culture intersect in contemporary communication.

5.4 Theoretical Implications

The study contributes to discourse analysis in several ways.

First, it reinforces the argument that identity is discursively constructed and performative rather than stable or internally defined.
Second, it supports multimodal discourse theory by demonstrating that identity on social media emerges from intersemiotic interaction rather than linguistic content alone.
Third, it extends digital discourse research by highlighting captions as central semiotic tools rather than marginal textual components.

By integrating discourse theory, identity performance frameworks, and multimodal analysis, the study offers a model for examining digital communication as a socially embedded semiotic system.

5.5 Practical Implications

The findings have implications beyond academic theory. For communication scholars, the study demonstrates the importance of treating social media texts as multimodal discourse events. For educators, it highlights how digital writing practices reflect identity negotiation and cultural positioning, suggesting that online communication should be included in language and media studies curricula. For social media practitioners, it shows how linguistic and symbolic choices influence audience perception and persona credibility.

5.6 Limitations of the Study

The research focuses on publicly available posts and therefore may not capture private identity performances or backstage communication practices. Additionally, the qualitative design prioritizes interpretive depth rather than statistical generalization. The findings should therefore be understood as contextual insights rather than universal patterns.

Another limitation is that audience interpretation was not directly studied. Identity performance involves both production and reception, and future studies may examine how viewers interpret caption discourse.

5.7 Recommendations for Future Research

Future research may extend this study in several directions. Longitudinal studies could examine how individual users' caption practices evolve over time. Comparative research could analyze identity performance across platforms such as

TikTok, Facebook, or X. Corpus-assisted studies could combine qualitative discourse analysis with quantitative linguistic patterns. Reception-focused research could investigate how audiences interpret caption discourse and respond to different identity strategies. Cross-cultural studies could explore how identity performance varies across linguistic and cultural communities.

5.8 Final Reflection

In contemporary digital life, identity is increasingly written in fragments — a caption, a hashtag, an emoji, a carefully filtered image. Instagram does not simply host identity; it shapes how identity is narrated, stylized, and recognized. Captions are small texts, but they carry large social meanings. They are the quiet scripts through which people tell the world who they are, who they hope to be, and where they believe they belong.

Bibliography

Baxter, J. (2003). *Positioning gender in discourse: A feminist methodology*. Palgrave Macmillan.

Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis*. De Gruyter.

Boyd, D. (2014). *It's complicated: The social lives of networked teens*. Yale University Press.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Cheema, S. (2025). Multimodal identity performance in Pakistani digital influencers. *Policy Journal of Media Studies*.

Creswell, J. W. (2013). *Qualitative inquiry and research design* (3rd ed.). Sage.

Daiute, C. (2014). *Narrative inquiry in social research*. Sage.

Deng, X. (2025). Emoji use and identity signaling in social media communication. *Journal of Computer-Mediated Communication*.

Fairclough, N. (1992). *Discourse and social change*. Polity Press.

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

Goffman, E. (1959). *The presentation of self in everyday life*. Anchor Books.

Hakami, S. (2024). Multimodal identity construction among Gulf social media influencers. *TWIST Journal of Cultural Studies*.

Halliday, M. A. K., & Matthiessen, C. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.

Ishfaq, A. (2025). Turn-taking, politeness, and identity: A conversational study of *Speak Your Heart*. *Journal of Applied Linguistics and TESOL, 6*(3), 1146.

Ishfaq, A., & Azim, M. U. (2025). Phono-semantics and translation: A cross-linguistic study of Urdu and Punjabi ideophones. *International Research Journal of Arts, Humanities and Social Sciences*.

Ishfaq, A., & Sultan, N. (2024). Identification of different methodologies for treatment of autism in Urdu-speaking adolescents: An investigative report. *Contemporary Journal of Social Sciences*.

Ishfaq, A., & Sultan, N. (2025). Cognitive control and executive function in advanced second-language writing. *Sareer-a-Khama, 4*(4).

Ishfaq, A., Ahmad, S., & Sultan, N. (2025). Decoding despair: A multidisciplinary psycho-forensic linguistic approach to suicide notes. *Journal of Psychology, Health and Social Care*.

Ishfaq, A., Malik, A. H., & Sultan, N. (2025). Developing trauma-sensitive pedagogical practices for resilient learning in academia: A multidisciplinary approach of psycholinguistics and ELT. *Al-Aasar*.

Ishfaq, A., Sultan, N., Hassan, N., & Aleem, F. (2022). Elif Shafak's *The forty rules of love*: Contextual variation in adjectives. *International Online Journal of Language and Literature*.

Jamshaid, M. (2025). Linguistic strategies of social media influencers in Pakistan. *Journal of Applied Linguistics and Technology*.

Khan, I. A., Khaled, F., & Ishfaq, A. (2025). Operation Bunyan un Marsoos: A critical analysis of human rights compliance—A study of the operation's adherence to human rights law and international standards. *Dialogue Social Science Review*.

KhosraviNik, M. (2017). Social media critical discourse studies. *Critical Discourse Studies, 14*(5), 582–596.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Marwick, A. (2013). *Status update: Celebrity, publicity, and branding in the social media age*. Yale University Press.

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.

O'Halloran, K. L., Tan, S., & Wignell, P. (2021). Multimodal discourse analysis. *Journal of Multimodal Communication, 15*(2).

Page, R. (2022). *Narratives online: Shared stories and social media identity*. Routledge.

Serafini, F., & Reid, S. F. (2023). Multimodal literacy and digital communication. *Journal of Adolescent & Adult Literacy, 66*(4), 327–337.

Simungala, H. (2025). Multimodal assemblages in digital communication. *Discourse, Context & Media*.

van Dijk, T. A. (2008). *Discourse and power*. Palgrave Macmillan.

Zakiyah, N. (2025). Interpersonal meaning in Instagram captions. *International Journal of Social Research*.