

CORPORUM

Journal of Corpus Linguistics

journal homepage: <https://journals.au.edu.pk/ojsrcr/index.php/crc>

A Corpus-Based Investigation of Linguistic Commodification in Pakistani Online Travel Discourse

Sidra 

Lecturer, Azad Jammu & Kashmir University of Bimber
fatma310396@gmail.com

ARTICLE INFO

Keywords:

Critical Discourse
Analysis, Digital
Branding,
Language
Hierarchies,
Linguistic
Commodification,
Tourism
Discourse,
Pakistan

ABSTRACT

This study investigates the linguistic commodification of tourism discourse in Pakistan's digital landscape. It draws on Critical Discourse Analysis (CDA), Linguistic Commodification Theory (Heller, 2010), and Word Frequency Analysis (Baker, 2006) to examine how English dominates tourism branding and shapes national identity. The analysis utilises a 10,000-word corpus from government and private tourism websites to examine how language constructs narratives of place, culture, and identity. The findings reveal that English functions as a language of prestige, reinforcing linguistic hierarchies that marginalize Urdu and regional languages. Additionally, tourism branding employs romanticized and commodified portrayals of landscapes, history, and culture, often prioritizing Western validation and market-driven narratives. These strategies contribute to a neoliberal tourism economy that packages experiences for elite consumption, while local voices and authenticity are diminished. The study calls for more inclusive linguistic practices in tourism branding to ensure equitable and diverse cultural representation in Pakistan's digital tourism discourse.

1. Introduction

The term linguistic landscape (LL) refers to the visible display of language in public spaces—such as signage, advertisements, and brochures—which reflect underlying societal hierarchies and power relations (Gorter & Cenoz, 2023). In tourism contexts, the linguistic choices made in promotional material shape how destinations are represented and how visitors and locals experience them. These choices can either reinforce dominant linguistic ideologies or promote greater inclusivity. Scholars have increasingly examined how language functions within tourism as a communicative and symbolic resource. Concepts such as linguistic commodification (Heller, 2010) and tourism discourse (Jaworski & Pritchard, 2005) have provided critical frameworks for analyzing how economic rationalities, global market demands, and postcolonial legacies shape language. In these discussions, English often emerges as a privileged language, a gateway to international visibility and consumption, while local and indigenous languages are frequently marginalized (Pennycook, 2017; Heller, Pujolar, & Duchêne, 2014).

Within this body of work, the digital sphere presents new dimensions for exploration, particularly as tourism promotion increasingly relies on online platforms. Although studies on linguistic landscapes and commodification have been conducted in various contexts (e.g., Ahmed et al., 2024; Gorter & Cenoz, 2023; Heller et al., 2014), research on digital tourism discourse in South Asia remains limited. In the case of Pakistan, recent scholarship has begun to engage with language use on tourism websites, pointing to patterns of dominance and erasure that warrant deeper investigation (Choudhri & Shehzad, 2024; Hussain et al., 2025). These studies highlight the need to further examine how language use in tourism interacts with identity, access, and cultural narrative questions in a rapidly evolving digital landscape.

This study responds to that need by investigating the representation and role of English in Pakistan's online tourism discourse. By focusing on how language choices function within this setting, the study seeks to contribute to broader conversations on linguistic hierarchies, commodification, and the politics of representation in multilingual societies. Therefore, the current study is guided by the following research objectives:

- To analyze the positioning of English in the digital landscape of tourism in Pakistan and its implications for linguistic commodification.
- To examine how linguistic choices construct Pakistan's identity in tourism marketing.
- To explore the exclusion of local languages of the domestic travelers in tourism narratives.

2. Literature Review

Language's role in tourism branding has received extensive attention in sociolinguistics, discourse analysis, and linguistic landscape research. Scholars have investigated how language can build national identity, shape cultural narratives, and market destinations to global audiences (Jaworski & Pritchard, 2005; Thurlow & Jaworski, 2010). However, in multilingual societies, tourism branding frequently reflects linguistic inequalities, with global prestige languages like English dominating over national and regional languages (Dann, 1996; Heller, 2010; Pennycook, 2017). This phenomenon, known as linguistic commodification, demonstrates how language can be transformed into an economic resource strategically used to attract international tourists (Heller, 2010; Duchêne & Heller, 2012). Recent research has consistently found English to be the dominant language in linguistic landscapes (LL), particularly in tourism spaces (Hasni et al., 2022; Gorter & Cenoz, 2023). Hasni et al. (2022) have conducted a comprehensive review of emerging themes in LL and tourism research from 2014 to 2022, identifying English as the primary language in multilingual tourism settings. Previous studies reveal how English consistently emerges as the dominant linguistic force in public spaces, tourism sites, and digital landscapes, often dominating local and national languages (Santos Rovira, 2025; Urry, 1990). Similar patterns have been observed in various tourism contexts. English signage, advertisements, and digital branding occupy the most prominent spaces, while indigenous languages are relegated to secondary or less visible positions (Datang, 2023; Ruzaitė, 2017).

English is more than just a functional language; it represents modernity, global appeal, and economic viability (Crystal, 2003). Scholars argue that using English in tourism branding is a deliberate strategy to market destinations within the global tourism economy, rather than a response to international communication needs (Datang, 2023). Local languages are marginalized in tourism linguistic landscapes, where English is dominant and visually and spatially prioritized over indigenous languages (Hasni et al., 2022). The research found that English signs, advertisements, and promotional materials were frequently larger, more prominent, and in primary positions, whereas local languages appeared in secondary or less visible areas. This symbolic inclusion reinforces local languages' subordinate status in digital tourism discourse, echoing concerns raised in previous LL studies about the erasure of indigenous languages in favor of global prestige languages (Heller, 2010; Pennycook, 2017). Research on Pakistan's linguistic landscape (LL) has highlighted English's dominance in public and commercial spaces, but its role in digital tourism branding is largely unknown. Abbass and Shibliyev (2024) investigated English prevalence in Islamabad's commercial signage, finding that while English dominates urban areas, local languages remain visible in smaller towns. This pattern may extend to Pakistan's tourism discourse. Similarly, studies on Pakistani ecotourism websites have examined linguistic strategies in environmental branding. However, their emphasis on ecolinguistics rather than language accessibility makes their connection to linguistic commodification in tourism branding more indirect. Choudhri & Shehzad (2024) significantly contribute by analyzing persuasive discourse on Pakistan's official tourism websites, shedding light on how language constructs ideological narratives in tourism promotion. While these studies help contextualize language use in Pakistani public and digital spaces, they leave a gap in understanding how English, Urdu, and regional languages function within online tourism branding, as well as the extent to which language is commodified in shaping Pakistan's tourism identity—a gap that this study seeks to address.

While LL research has traditionally focused on physical signage, billboards, and public information displays, studying digital linguistic landscapes—such as tourism websites, online advertisements, and social media marketing—has received far less attention. This gap is especially relevant in multilingual contexts like Pakistan, where English coexists with Urdu and other regional languages. However, little is known about how linguistic choices in digital tourism branding relate to broader socioeconomic and ideological trends, or how they may affect accessibility and inclusivity for various linguistic communities. This study aims to address this gap by investigating the linguistic landscape of Pakistani digital tourism branding, particularly emphasizing the role of English, Urdu, and regional languages in shaping tourism discourse. It offers useful insights to policymakers, tourism professionals, and language planners working to develop more inclusive and representative tourism strategies.

3. Theoretical Framework

This study adopts Critical Discourse Analysis (CDA), Linguistic Commodification Theory, and Word Frequency Analysis (WFA) to investigate how language is positioned, marketed, and negotiated in online tourism discourse. Critical Discourse Analysis (CDA), specifically Fairclough's (1995) three-dimensional model and Wodak's (2008) discourse-historical approach, is used to investigate the structural patterns of language in tourism branding. Fairclough's model investigates language use's textual, discursive, and social dimensions, revealing how linguistic choices influence perceptions of Pakistani tourism and national identity. Wodak's approach contextualizes it to Pakistan's colonial history and neoliberal tourism strategies. These aspects of CDA enable us to critically assess how language choices reinforce power dynamics and accessibility barriers.

Linguistic Commodification Theory (Heller, 2010) is used to evaluate how language is used as an economic asset in tourism marketing, particularly the notion of "market-oriented multilingualism" (Heller, 2010, p.108) in which languages are displayed for commercial purposes rather than genuine representation. By examining how English is positioned as the default tourism language, we highlight the exploitation of linguistic resources for financial gain. Word Frequency Analysis (WFA) (Baker, 2006) is employed to quantify linguistic patterns in tourism discourse. Using Python for data extraction and AntConc for statistical analysis, we identify the frequency and distribution of languages. This method enables us to empirically validate our CDA findings. Combining discourse analysis with corpus linguistics provides a systematic and data-driven examination of linguistic hierarchies in Pakistani tourism branding.

4. Research Methodology

This study is conducted within the critical discourse analysis (CDA) paradigm, which aligns with social constructivism (Berger & Luckmann, 1966). It assumes that language is not neutral but shaped by social, economic, and political power structures (Fairclough, 1995). The research follows an interpretivist approach, recognizing that discourse shapes and is shaped by historical and ideological contexts (Wodak, 2008). A corpus of 10,000 words was compiled from ten major Pakistani tourism websites to ensure a representative sample of tourism marketing discourse in the government and private sectors. These websites were chosen based on their credibility, market influence, geographical representation, and visibility. Government-run websites, such as Pakistan Tourism Development Corporation (PTDC) and Khyber Pakhtunkhwa Tourism Board, provide information about official tourism narratives, whereas private-sector platforms, such as Serena Hotels and Guide to Pakistan, reveal corporate branding strategies. Websites focusing on regional tourism, such as Visit Gilgit Baltistan and Sindh Cultural Tourism, were included to ensure a comprehensive understanding of linguistic patterns. Only actively maintained and widely accessed websites were chosen to ensure that the corpus accurately represented dominant linguistic trends in contemporary Pakistani tourism marketing. Table 1 lists the selected websites categorised by type, including government, private, luxury, and regional tourism initiatives.

Table 1

Tourism Websites Selected for Corpus Construction

Category	Websites Analyzed
Government Tourism Websites	Pakistan Tourism Development Corporation (PTDC), Khyber Pakhtunkhwa Tourism Board
Private Travel Operators	Chlo Pakistan, getout.pk, The Wanderer Pakistan
Luxury Travel and Hospitality	Serena Hotels, Tour my Pakistan, Guide to Pakistan
Regional Tourism Initiatives	Sindh Cultural Tourism, Visit Gilgit Baltistan

A corpus size of 10,000 words was chosen based on Paul Baker's (2006) recommendations for discourse analysis studies, which suggest that small-to-medium corpora (5,000–100,000 words) are effective for CDA-driven linguistic analysis. The dataset was extracted using Python tools (BeautifulSoup, Scrapy) and manually cleaned to remove non-textual elements, duplicate content, and formatting errors. The final cleaned corpus contained precisely 10,000 words, resulting in a manageable yet comprehensive dataset for linguistic investigation. Frequency patterns were identified using AntConc's word list and keyword list functions, which provided raw and relative frequencies across the corpus. Words were considered significant if they appeared at least 10 times, following Baker's (2006) threshold for identifying salient lexical items in small corpora. This allowed the study to focus on recurring terms that play a prominent role in constructing tourism discourse.

Collocational patterns were then analyzed using AntConc's collocate and concordance tools with a window span of ± 5 words. Collocates were included for interpretation only if they occurred at least five times within this span, ensuring consistent co-occurrence and reducing noise. Although mutual information (MI) scores were generated, frequency-based filtering was prioritised given the modest corpus size. This combination of quantitative thresholds and qualitative interpretation ensured that the analysis

remained both systematic and contextually grounded. To interpret the patterns revealed by the corpus analysis, this study draws on Fairclough's (1995) three-dimensional model and Wodak's (2008) discourse-historical approach. Additionally, Heller's (2010) concept of linguistic commodification provides a framework for understanding how language operates as a marketable resource in the context of tourism.

The study follows stringent methodological guidelines to ensure rigor and ethical transparency. Only publicly available data were used, and texts were analyzed in their original form without manipulation. While the study is limited to textual analysis and does not address visual or multimodal elements, it opens avenues for future research into how images, videos, and design contribute to the linguistic branding of tourism in Pakistan.

5. Results and Analysis

This section presents a corpus-based analysis of linguistic patterns found in Pakistani tourism websites, aiming to explore how language operates within the broader dynamics of identity, commodification, and representation. To organize the findings meaningfully, the section is structured around three thematic categories derived inductively from the data:

1. Identity construction in tourism marketing,
2. The commodification of language and culture, and
3. The exclusion of local linguistic communities.

Each theme is supported by illustrative examples and interpreted within the context of Pakistan's digital tourism discourse.

5.1 Identity Construction in Tourism Marketing

This section explores how national identity is presented in the discourse of Pakistani tourism websites. The goal is to examine how linguistic choices contribute to shaping particular images of Pakistan for tourism audiences.

5.1.1 *Romanticizing Landscapes and Cultural Heritage*

A dominant trend in Pakistani tourism marketing is the romanticization of landscapes and cultural heritage, which portrays Pakistan as an untouched, idyllic, and mystical destination. Tourism websites frequently use poetic, hyperbolic language to portray natural and historical sites as places of unparalleled beauty and eternal wonder. This is evident in The Wanderer Pakistan's promotional descriptions, which state: "Pristine coastal waters, marine islands, challenging sea cliffs, and beaches spread out over a 650km stunning coastline of the Arabian Sea." Similarly, Visit Gilgit Baltistan advertises its tourism offerings as follows: "A traveler's paradise nestled in the heart of untouched wilderness, where sky-high mountains meet crystal-clear lakes and lush alpine meadows." Such descriptions are not neutral representations of geography; rather, they are consistent with colonial and postcolonial tourism fantasies that portray non-Western spaces as exotic, unknown, and ripe for exploration. The emphasis on "pristine" and "untouched" landscapes creates a Western gaze that portrays Pakistan as a remote and adventurous destination, reinforcing Orientalist stereotypes. These linguistic choices promote Pakistan as a destination for adventure, escapism, and natural wonder, catering to tourists' desire for an authentic and visually appealing experience.

The romanticization of cultural heritage follows a similar pattern. Websites like Chlo Pakistan and Guide to Pakistan extol the beauty of Mughal architecture, Sufi shrines, and tribal traditions without digging deep into their historical complexities or contemporary sociopolitical significance. Cultural sites are frequently presented as static relics from the past, intended for visual appreciation rather than critical engagement. This phenomenon echoes with Britton's (1991) concept of the "Tourism-Industrial Complex," which claims that: "Tourism industries impose external frameworks upon local spaces, dictating what is valuable, visible, and economically viable, often at the expense of indigenous agency" (Britton, 1991, at p. 57). Both The Wanderer Pakistan and Visit Gilgit Baltistan emphasize the grandeur and mystique of Pakistan's mountain ranges, portraying them as natural wonders that exist outside of time and human intervention. For example, The Wanderer Pakistan states: "Imagine 4,500 peaks above 6,000 meters in an area about the size of Ireland." Similarly, it describes the Karakoram Mountains as: "A wall, 500 kilometers long and 150 kilometers wide, with more than 100 peaks above 6,100 meters (20,000 feet). "This region has the highest concentration of high peaks in the world. "These descriptions are linguistically constructed to evoke awe and sublimity, with phrases like, "a wall" and "the world's highest concentration of high peaks" depicting the landscape as vast and impenetrable.

5.1.2 *Historical Narratives and the Construction of National Prestige*

The historicization of identity is an important discursive strategy in Pakistan's tourism marketing, as it places the country in a prestigious lineage of ancient civilizations. This strategy can be seen in how The Wanderer Pakistan describes the country's historical past: "Home to the world's ancient civilizations. The history of this region in Pakistan dates back to 50,000 BC. The

earliest agricultural settlement was discovered in Mehrgarh, dating back to 7000 B.C., a thousand years before Mesopotamia." (The Wanderer, Pakistan). At the textual level, the use of chronological markers and comparative references ("a thousand years before Mesopotamia") establishes Pakistan's history as not only deep-rooted but also superior to other ancient civilizations. Positive lexical choices like "earliest," "discovered," "well-planned cities," and "sophisticated governance systems" reinforce the notion of advanced civilization and intellectual heritage. At the discursive practice level, this narrative reflects Pakistan's deliberate positioning within a global historical discourse, placing it alongside internationally recognized civilizations such as Mesopotamia and Ancient Egypt. This is significant because it challenges Eurocentric historical narratives, which frequently exclude or marginalize non-Western civilizations. However, the discourse fails to acknowledge historical discontinuities, colonial legacies, or sociopolitical conflicts that have shaped Pakistan's historical trajectory. At the socio-cultural level, this discourse serves as a national branding strategy, encouraging national pride and international recognition. It accords with Wodak's DHA framework, which contends that historical narratives in national discourse frequently serve strategic political purposes—in this case, reinforcing Pakistan's status as an ancient and resilient civilization amidst postcolonial challenges. However, this reflects a discursive omission of marginalized historical perspectives, such as ethnic communities' subaltern histories, which are rarely included in mainstream narratives.

5.1.3 *Western Validation of Pakistani Destinations*

One of the most common discursive strategies in Pakistan's digital tourism marketing is to rely on Western recognition to establish the country's legitimacy as a tourist destination. Both PTDC and The Wanderer Pakistan cite endorsements from Forbes, Condé Nast, and the British Backpacker Society, emphasizing Pakistan's status as one of the "coolest places to visit," a "top tourist destination," and a "under-the-radar travel gem." These references serve as external validations, establishing Pakistan as a worthwhile destination only after it has received recognition from Western media and travel authorities. It corresponds to the ideological shaping of discourse, in which power structures decide which voices are credible. Instead of focusing on Pakistan's inherent cultural, historical, and natural appeal, the discourse creates desirability through Western acceptance, reinforcing global tourism hierarchies in which destinations are validated through a Euro-American lens.

According to The Wanderer Pakistan, "the British Backpacker Society ranked Pakistan as the top adventure destination in 2018 and the third highest potential for 2020." This statement creates an authority-based appeal by using rankings and numerical placement to validate Pakistan's status. The use of precise statistics—such as the "92% decrease in terrorism-related incidents since 2013"—strengthens this quantification of safety, ensuring that Pakistan meets global (Western) security standards before being considered a viable destination. Wodak (2008) contends that past narratives of instability are actively rewritten to accommodate tourism promotion. However, this selective framing undermines local agency, portraying Pakistan as a passive recipient of Western approval rather than an active narrator of its own tourism identity.

5.1.4 *Culinary Discourse in Tourism: Building Identity through Food*

Culinary discourse is important in Pakistani tourism marketing because it helps to shape national and regional identities. Food is frequently portrayed as a symbol of cultural authenticity, with websites promoting local cuisines as essential travel experiences. For example, GetOut.pk describes Punjabi cuisine as: "Experience the spirit of Punjab through its lavish delicacies—Nihari, Lassi, freshly baked naan, and succulent BBQ treats prepared with age-old recipes." This description portrays food as a sensory and emotional experience, reinforcing a romanticized and idealized view of local culture. Similarly, Tour My Pakistan markets Balti cuisine as follows: "Enjoy local Balti food at homestays, taste hand-cooked organic meals in the shadows of the Karakoram, and sip traditional butter tea as the locals do." These representations establish an authenticity hierarchy, with rural and traditional foods positioned as more "genuine" than urban or globalized alternatives.

In addition, elite culinary tourism is becoming more prominent in high-end tourism marketing. Serena Hotels and other luxury tour operators promote five-star traditional dining experiences, emphasizing authenticity and exclusivity. This reinforces class-based exclusion, as only those with economic privilege have access to "true" cultural experiences. Fairclough (1995) observes that "discursive formations are always shaped by power structures," implying that what is considered authentic in tourism discourse is ultimately determined by capitalist hierarchies.

5.2 **Linguistic Commodification: Marketing Pakistani Landscapes as a Global Product**

This section outlines how lexical choices, phrase patterns, and globalized framings contribute to the commodification of Pakistani landscapes and culture. The following subsections discuss these elements in turn, drawing on examples from the corpus to examine the market-oriented functions of language in tourism marketing.

5.2.1 Lexical Choices and Discursive Strategies for Commercial Appeal

A close examination of lexical patterns in the corpus reveals a strong orientation toward commercialized and aestheticized language that frames Pakistan's landscapes and cultural heritage as marketable experiences. These choices are not merely stylistic; they reflect broader ideological frameworks that shape how the country is imagined and sold to a global audience. Table 2 presents high-frequency lexical items identified in the dataset, along with their discursive functions and implications for tourism discourse.

Table 2

High-frequency Lexical Choices and Their Ideological Functions in Tourism Discourse

Lexical Choice	Frequency (n)	Discursive Function	Function in Tourism Discourse
Dial us for adventure!	10	Creates an instant, action-oriented tourism call.	Presents tourism as an instant gratification commodity.
Premium campsites	12	Echoes elite glamping tourism.	Transforms nature into controlled, high-end luxury experiences.
colorful history of Pakistan	11	Reduces historical complexity to a simplified narrative.	Frames heritage as a digestible spectacle, erasing sociopolitical contexts.
Breathtaking landscapes	14	Frames geography as a visual spectacle.	Nature is reduced to passive aesthetic consumption, reinforcing the tourist gaze.
Switzerland of the East	13	Draws a Western comparison to legitimize appeal.	Reflects Eurocentric validation of non-Western landscapes.
Unrivaled beauty	11	Employs superlative language to enhance allure.	Frames nature as a commodified, capital-worthy asset.
Hill-station resorts (e.g., Nathiagali, Murree, Bhurban)	10	Lists popular high-altitude retreats.	Depicts landscapes as resorts, not inhabited places.
Exotic bazaars	10	Reinforces orientalist market tropes.	Otherizes traditional markets into tourist attractions.
Parathas	12	Appeals to sensory and culinary exoticism.	Portrays food as a consumable heritage object.
Tranquil escape	10	Frames Pakistan as a peaceful retreat.	Reflects Western escapism; ignores societal complexity.
cultural experience	13	Constructs heritage as a static spectacle.	Ignores sociopolitical change and local agency.
Hidden gems	11	Implies discovery and exclusivity.	Reproduces colonial exploration rhetoric; erases local knowledge.
Exclusive retreat	10	Promotes luxury tourism for elites.	Reinforces economic exclusivity over inclusive development.
Enchanting valleys	12	Romanticizes landscapes with nostalgia.	Suggests untouched paradise, ignoring real histories.
A paradise for adventurers	11	Highlights adrenaline tourism.	Isolates landscapes from sociopolitical realities.
Unparalleled beauty	14	Uses superlatives to intensify appeal.	Reinforces escapist fantasies about non-Western destinations.

Many of these phrases—such as breathtaking landscapes, exclusive retreat, or hidden gems—do more than describe; they help construct a highly stylized and marketable image of Pakistan. These expressions emphasize beauty, remoteness, and adventure, but they also risk simplifying the complex social and environmental realities that exist in these regions. Importantly, these patterns are not unique to this study. Choudhri and Shehzad (2024), in their recent work on Pakistan's official tourism platforms, observed similar tendencies—where promotional language tends to favor market appeal over inclusivity or grounded representation.

5.2.2 Collocations and Semantic Clusters: Reinforcing The Western Travel Imaginary

Beyond individual lexical choices, collocation analysis offers deeper insight into how recurrent word pairings construct ideologically loaded representations of Pakistan. Table 3 presents a selection of high-frequency collocations (appearing ≥ 5 times within a ± 5 -word span), along with examples from the corpus and their broader ideological implications.

Table 3

Notable Collocations in Tourism Discourse and Their Ideological Implications (Collocates appear ≥ 5 times within a ± 5 -word span)

Collocation	Raw Frequency	Example from Corpus	Ideological Implication
Unforgettable journey	9	“Treat your loved ones to an unforgettable journey.”	Supporting neoliberal tourism economies.
Mesmerizing views	11	“The whole family can enjoy the mesmerizing views.”	Nature is viewed as a visual spectacle, reducing ecological and cultural contexts to aesthetic appeal.
Adventure + luxury	7	“An exclusive retreat in the heart of the mountains.”	Reinforces tourism as a luxury product, distancing it from community-based experiences.
Secluded paradise	6	“Experience a secluded paradise in the mountains of Gilgit-Baltistan.”	Frames travel as exclusive, promoting Western consumerist ideals of escapism.
Premium hospitality	10	“Enjoy premium hospitality at our deluxe resorts.”	Commodifies traditional hospitality for elite consumption, aligning with capitalist tourism discourse.
Deluxe side of life	5	“Switch to the deluxe side of life.”	Encourages tourists to seek expensive lifestyles, reinforcing class-based travel identities.
Therapeutic pull	6	“Discover the therapeutic pull of the country.”	Sells national identity as a self-care experience, abstracting away historical and environmental complexities.
Regal hospitality	5	“Known for its regal hospitality and royal charm.”	Frames local traditions as elite service offerings, detaching them from everyday cultural contexts.
Quintessential hospitality	8	“Experience the quintessential Pakistani hospitality.”	Universalizes hospitality as a tourism marketing trope, erasing regional diversity and social realities.
Premium polo clubs	5	“Power-packed action at premium polo clubs.”	Reproduces colonial sport hierarchies, embedding leisure in elite colonial aesthetics.
Untouched beauty	14	“The untouched beauty of Naltar Valley is breathtaking.”	Constructs wilderness as devoid of human presence, supporting eco-tourism fantasies and erasing local life.
Exotic adventure	10	“Immerse yourself in the exotic adventure of northern Pakistan.”	Exoticizes local culture for escapist consumption, reinforcing orientalist binaries.

These collocational patterns are ideologically charged, reflecting the influence of capitalist and Eurocentric frameworks that shape what is rendered visible, desirable, and commodifiable in tourism discourse—while simultaneously obscuring or erasing alternative local perspectives. As Heller (2010) conceptualizes in her work on the commodification of authenticity, such language practices do not merely describe cultural and natural landscapes; they selectively curate them for international consumption, often at the expense of local agency, complexity, and empowerment.

5.2.3 *Experiential Commodification: The Marketization of Adventure and Culture*

A significant discursive trend in Pakistan’s digital tourism branding is the commodification of experience—where adventure, culture, and natural landscapes are not merely represented, but systematically packaged as pre-defined, consumable products. Rather than promoting organic interaction with local environments and customs, the discourse constructs tourism as a luxury-driven, pre-planned encounter emphasizing convenience, exclusivity, and spectacle. Table 4 below outlines recurring experiential themes drawn from the corpus and provides critical interpretations that reveal the ideological underpinnings of this commodified tourism discourse.

Table 4

Experiential Tourism Narratives and Their Critical Interpretations

Experiential Narratives	Example from Corpus	Critical Interpretation
Adventure as a luxury	"Dial us for adventure! Mountain trekking, thrilling chair lift rides, premium campsites, meals by the lakeside..."	Adventure is packaged as a pre-curated commodity, preventing spontaneous exploration and local interactions.
Culture as spectacle	"Become a part of cultural folklore with us and see how the lives of these early settlers evolved."	Cultural identity is presented as a performative exhibit rather than a dynamic, lived experience.

Nature as curated beauty	"Trek through the mountain passes of these sky-high beasts, camp in the second-highest plateau of the world..."	Landscapes are aesthetically designed for passive consumption rather than being understood within their ecological or indigenous contexts.
Food as exotic indulgence	"Indulge in Nihari and Lassi, experience a burst of color in Punjab's festivals."	Culinary heritage is reduced to sensory pleasure rather than being investigated as part of a complex socio-historical framework.

In this kind of framing, authenticity becomes less about real cultural life and more about something to be packaged and sold. It's carefully designed for visual appeal, wrapped in polished narratives, and often detached from the everyday realities of the communities being represented. Phrases like "Dial us for adventure!" or "Become a part of cultural folklore with us!" aren't just catchy—they show how tourism is increasingly about offering ready-made, curated experiences rather than encouraging genuine exploration.

The problem with this, of course, is that it leaves little room for local voices. Communities are positioned more as service providers than storytellers, and the complex social, political, and environmental histories of these places are often left out of the picture. Everything becomes smooth and glossy—easy to sell, but harder to really connect with. In the end, it reinforces a model of tourism where experiences aren't discovered through meaningful interaction, but bought off the shelf—polished, predictable, and detached from the very people and places they claim to represent.

5.2.4 *The Tourism Industrial Complex and the Erasure of Local Agency*

The results of this dataset show that Pakistan's tourism discourse is deeply embedded in global capitalist tourism ideologies, in which landscapes and cultures are redefined as consumable experiences. Britton's (1991) Tourism-Industrial Complex is an effective framework for understanding how tourism industries impose market-driven narratives that prioritize external consumption over local representation. This study identifies three major forms of commodification in Pakistani tourism marketing: commercial commodification, Western imaginaries, and experiential commodification. These modes of linguistic and discursive framing highlight three major structural issues in Pakistan's tourism representation.

First, the selective visibility of culture and nature prioritizes aesthetic and consumable aspects while obscuring their deeper political, historical, and indigenous complexities. Tourism marketing narratives that emphasize scenic beauty and exotic charm present an idealized, decontextualized version of Pakistan's cultural and environmental diversity. Second, because tourism economies are extractive, the economic benefits of these representations are frequently directed toward external tourism industries and elite travelers rather than local communities. Third, the exclusion of local communities and languages from tourism marketing strengthens linguistic hierarchies that favor English and hyper-commercialized discourse. Heller's (2010) concept of linguistic commodification helps us understand how languages become economic capital rather than vehicles for cultural expression. The marketing discourse frames tourism experiences in Western-oriented linguistic registers, portraying tourists as consumers and local communities as passive hosts rather than active cultural agents with their own voices and narratives.

5.3 English as The Sole Medium: The Marginalization of Local Languages in Pakistani Tourism

The dominant use of English across Pakistani tourism websites reveals a striking pattern of linguistic exclusion. As shown in Table 5, nearly all major platforms rely exclusively on English, with little to no presence of Urdu or local languages such as Pashto, Sindhi, or Balti.

Table 5

Language Representation on Major Pakistani Tourism Websites

Website Name	English (%)	Urdu (%)	Local Languages (%)	Notes
Pakistan Tourism Development Corporation (PTDC)	100%	0%	0%	Government-run; English-only
Khyber Pakhtunkhwa Tourism Board	100%	0%	0%	No support for Urdu or local languages
Chlo Pakistan	100%	0%	0%	Private travel platform; English-centric
getout.pk	100%	0%	0%	English-only; no inclusion of Urdu or

The Wanderer Pakistan	100%	0%	0%	local languages
Serena Hotels	100%	0%	0%	Full reliance on English
Tour My Pakistan	100%	0%	0%	International branding; English-dominant
Guide to Pakistan	100%	0%	0%	No multilingual elements
Sindh Cultural Tourism	100%	0%	0%	Entirely in English
Visit Gilgit Baltistan	95%	0%	5%	No Sindhi or Urdu support Some Romanized Balti/Shina for tourist interaction

Note. Percentages reflect visible language use on the landing and primary sub-pages of each tourism website. "Romanized" refers to local languages written in the Latin script rather than native scripts.

Local languages are conspicuously absent from Pakistan's digital tourism marketing discourse. Across the dataset, most tourism websites use only English and, to a lesser extent, Urdu, with little to no representation of regional languages. This results in a linguistic hierarchy in which only globally and nationally dominant languages are considered relevant for tourism, leaving local languages invisible. The only notable exception is Visit Gilgit Baltistan, which includes a small section introducing Shina, one of the local languages of the region:

"Top Tips: Learn a few words in Shina, the local language.
Hello: Assalam-o-Alaikum
Goodbye: Khuda Yar
Thank you: Ju Nu
How much is it? Anu Kachak Kay Han"

This inclusion matters for two reasons. First, it recognizes the existence of a linguistic identity that is typically ignored in mainstream tourism narratives. Second, using Romanized Shina rather than the original script demonstrates a strategic adaptation for a larger, non-Shina-speaking audience. However, this minor inclusion does not significantly alter the dominant monolingual narrative of tourism discourse. Fairclough (1995) contends that power is frequently exercised through language choices—what is included and, more importantly, what is excluded. Tourism discourse reinforces the notion that indigenous languages such as Balti, Pashto, and Sindhi are peripheral or irrelevant to Pakistan's tourist and economic identity by excluding them.

This exclusion is not unintentional; it is consistent with historical patterns of linguistic marginalization in Pakistan, where English and Urdu have been privileged in official and elite discourses, while regional languages remain linguistically and politically marginalized (Rahman, 1996). The fact that tourism websites primarily use English as their primary medium reflects global linguistic hierarchies, with tourism tailored to an Anglophone and Westernized audience, reinforcing linguistic imperialism (Phillipson, 1992).

Language hegemony in Pakistani tourism discourse is reinforced by the corpus's heavy reliance on English vocabulary, grammar, and discourse structures. Because Urdu and native languages are completely missing, English is the only "legitimate" language to use when interacting with prospective tourists. Visit Gilgit Baltistan is the sole exception, utilizing Romanized Balti and Shina phrases for routine tourist interactions like greetings and transactional exchanges (e.g., purchasing food, requesting directions). This inclusion, however, is extremely selective and only serves a functional purpose; it is neither an attempt to maintain local linguistic identity nor a complete linguistic expression. Rather than being sincere attempts at multilingual inclusivity, these expressions are touristic conveniences. English is a strategic decision that corresponds to Pakistani tourism marketing with international neoliberal standards, rather than just being a pragmatic one. English is emphasized as the default language of legitimacy in tourism discourse by the exclusion of Urdu and indigenous languages, which implies a purposeful erasure of local linguistic identities. The use of Romanized Balti and Shina in Visit Gilgit Baltistan reflects a phenomenon that Pietikäinen et al. (2016) describe as "linguistic display without cultural depth" (p.23). Here, local languages appear only for economic utility rather than for authentic representation. Instead of using local scripts (e.g., Perso-Arabic or Tibetan-based orthographies), the phrases are presented in Romanized English, further reinforcing English as the dominant linguistic framework.

6. Discussion

This study aimed to critically examine how Pakistan's digital tourism discourse constructs national identity, commodifies cultural and natural assets, and marginalizes local linguistic practices. The results reveal not only recurring linguistic patterns but also the ideological frameworks that shape those patterns. The discussion below synthesizes these findings in light of the research

objectives and broader critical discourse studies literature.

One of the central findings of this study is the systematic romanticization and aestheticization of Pakistan's landscapes and heritage. As shown in Sections 4.1 and 4.2, lexical and collocational patterns (e.g., untouched beauty, exclusive retreat, mesmerizing views) reveal how tourism discourse promotes a curated image of Pakistan, rooted in Western travel imaginaries. These representations are not merely descriptive; they function ideologically. By framing Pakistan as a land of luxury, escapism, and visual pleasure, the discourse privileges marketable narratives while downplaying political, ecological, and historical realities. As Fairclough (1995) argues, discourse plays an active role in shaping social reality—and here, that reality is filtered through neoliberal tourism logics. The result is a commodified vision of place and identity, tailored for elite, largely Western audiences. The study also highlights the commodification of authenticity, a trend well-theorized by Heller (2010). Cultural identity is repackaged as a tourist product—something to be consumed rather than experienced meaningfully. Importantly, these discursive strategies have consequences. They reinforce class-based access to culture, where heritage is available primarily through curated experiences for those who can afford them. As seen in the experiential marketing language (e.g., Dial us for adventure!), spontaneity is replaced by packaged, performative narratives that strip communities of agency.

Perhaps most revealing is the near-total exclusion of Urdu and local languages from Pakistan's tourism websites. As Section 4.3 shows, English dominates—both lexically and structurally—across all major platforms. The few instances where local languages are included (e.g., Romanized Shina on Visit Gilgit Baltistan) serve functional, touristic purposes rather than affirming linguistic diversity. This linguistic exclusion corresponds with Phillipson's (1992) theory of linguistic imperialism, where global languages gain legitimacy at the expense of local ones. It also echoes Rahman's (1996, 2002) findings on the marginalization of indigenous languages in Pakistani education and governance. By failing to engage local linguistic identities, tourism discourse reinforces a monolingual, elite-centric narrative—one that overlooks the symbolic and cultural significance of language in constructing inclusive national identity.

Taken together, the findings point to a tourism discourse firmly embedded in global neoliberal logics, where landscapes, cultures, and even languages are positioned primarily as commodities to be consumed. While the strategic branding of Pakistan as a desirable tourist destination may serve economic objectives, this study highlights the risks of relying exclusively on externally validated, market-driven narratives. Such framings often marginalize local voices, flatten cultural complexity, and prioritize aesthetic appeal over authenticity.

In response, the study advocates for more inclusive, locally grounded forms of tourism storytelling—narratives that recognize linguistic and cultural diversity, resist reductive portrayals, and remain critically engaged with questions of history and power. As Wodak (2008) reminds us, discourse is not merely reflective—it is constitutive. It plays an active role in shaping national identity and determining whose stories are amplified, and whose are silenced.

7. Conclusion

This study explored how Pakistani tourism websites construct national identity, commodify culture and landscapes, and marginalize local languages through their discourse. Drawing on a 10,000-word corpus and a critical discourse analysis framework, the findings show that tourism narratives are dominated by commercialized language and Eurocentric framing. High-frequency lexical choices such as "breathtaking landscapes" (14 occurrences), "authentic cultural experience" (13), and "unparalleled beauty" (14) reflect a consistent emphasis on aesthetic appeal, exclusivity, and spectacle. Similarly, collocations like "premium hospitality" (10) and "exotic adventure" (10) reveal how tourism discourse aligns with global neoliberal values, often at the expense of local representation. Quantitative data also revealed a striking absence of linguistic diversity: 100% of the websites analyzed used English as the primary language, with no inclusion of Urdu and minimal, functional use of Romanized regional languages. These patterns illustrate how linguistic and cultural authenticity are selectively presented for marketability.

Ultimately, the study argues that while tourism promotion is a valuable economic tool, it should not come at the cost of flattening identities or erasing local agency. More inclusive, multilingual, and locally rooted narratives can offer richer and more ethical representations of Pakistan's cultural and natural diversity.

References

- Abbass, G., & Shibliyev, J. (2024). English and local languages in linguistic structuration of the public space in Islamabad and the neighbouring town. *International Journal of Multilingualism*, 1-14.
- Ahmed, N., Zahid, M., & Anwar, M. N. (2024). Linguistic Landscape Analysis of Shop Signs: A Case Study. *Pakistan Languages and Humanities Review*, 8(1), 467-478.
- Baker, P. (2006). *Using corpora in discourse analysis*. Bloomsbury Publishing.
- Berger, P. L., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Anchor Books.
- Britton, S. G. (1991). Tourism, capital, and place: Towards a critical geography of tourism. *Environment and Planning D: Society and Space*, 9(4), 451–478. <https://doi.org/10.1068/d090451>
- Choudhri, S., & Shehzad, W. (2024). Linguistic projection of Pakistan's official tourism websites: Implications for tourism sustainability. *Kashmir Journal of Language Research*, 27(1), 91-115.
- Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.
- Dann, G. M. S. (1996). *The language of tourism: A sociolinguistic perspective*. CAB International.
- Datang, Frans Asisi (2023) "Signage in Public Spaces: Impact of Tourism on the Linguistic Landscape of Labuan Bajo," *International Review of Humanities Studies: Vol. 7: No. 1, Article 8*.
- Duchêne, A., & Heller, M. (2012). *Language in late capitalism: Pride and profit*. Routledge.
- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. Longman. London & New York.
- Gorter, D., & Cenoz, J. (2023). *A panorama of linguistic landscape studies*. Channel View Publications.
- Hasni, N. A., Shak, M. S. Y., Malik, N. A., & Anuarudin, A. A. S. (2022). Linguistic landscape of tourist spaces from 2014 to 2022: A Review. *International Journal of Academic Research in Business and Social Sciences*, 12(10), 1695-1708.
- Heller, M. (2010). The commodification of language. *Annual Review of Anthropology*, 39(1), 101–114. <https://doi.org/10.1146/annurev.anthro.012809.104951>
- Heller, M., Jaworski, A., & Thurlow, C. (2014). Introduction: Sociolinguistics and tourism–mobilities, markets, multilingualism. *Journal of Sociolinguistics*, 18(4), 425-458.
- Heller, M., Pujolar, J., & Duchêne, A. (2014). Linguistic commodification in tourism. *Journal of Sociolinguistics*, 18(4), 539-566.
- Hussain, M., Ali, S. S., Murtaza, M., & Shahid, S. (2025). *Ecolinguistic Analysis of the Evaluation of Nature on the Ecotourism Websites of Pakistan*. Research Consortium Archive, 3(1).
- Jaworski, A., & Pritchard, A. (2005). *Discourse, communication, and tourism*. Channel View Publications.
- Pennycook, A. (2017). *The cultural politics of English as an international language*. Routledge.
- Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.
- Pietikäinen, S., Kelly-Holmes, H., Jaffe, A., & Coupland, N. (2016). *Sociolinguistics from the periphery: Small languages in new circumstances*. Cambridge University Press.
- Rahman, T. (1996). *Language and politics in Pakistan*. Oxford University Press.
- Rahman, T. (2002). *Language, ideology, and power: Language-learning among the Muslims of Pakistan and North India*. Oxford University Press.
- Ruzaitė, J. (2017). The role of English in European tourism marketing: A linguistic analysis. *Journal of Language & Tourism*, 14(3), 201–219.
- Santos Rovira, J. M. (2025). Beyond Words: Navigating Lisbon's Linguistic Landscape in the Era of Mass Tourism and Migration. *Journal of Language and Social Psychology*, 0261927X251339207.
- Thurlow, C., & Jaworski, A. (2010). *Tourism discourse: Language and global mobility*. Palgrave Macmillan.
- Urry, J. (1990). *The tourist gaze: Leisure and travel in contemporary societies*. Sage Publications.
- Wodak, R. (2008). Introduction: Discourse studies—Important concepts and terms. In R. Wodak & M. Krzyzanowski (Eds.), *Qualitative discourse analysis in the social sciences* (pp. 1–29). Palgrave Macmillan.
- Wodak, R. (2008). The discourse-historical approach. In Pera, R., & Meyer, M. (Eds.), *Methods of critical discourse analysis* (pp. 63-94). SAGE.