



Language Hybridization and Shift: The Rise of Urjabish and the Decline of Punjabi in Urban Pakistan

Aisha Ismail

MPhil Scholar, Department of English Linguistics and Literature, Riphah International University, Pakistan.

Saqlain Hassan*

Assistant Professor, Department of English Linguistics and Literature, Riphah International University, Pakistan. ORCID: <https://orcid.org/0000-0001-8861-6800> (Corresponding Author)

Abstract

This study investigates the sociolinguistic emergence and communicative functions of *Urjabish*—a novel analytical construct proposed to describe the systematic trilingual hybrid variety that blends Urdu's grammatical scaffolding, Punjabi's colloquial expressivity, and English's lexical prestige in urban Pakistani discourse. Positioned at the intersection of globalization, media ecology, and socioeconomic stratification, *Urjabish* has evolved into an emergent urban lingua franca that encodes modernity, educational capital, and cosmopolitan belonging among young speakers in metropolitan Lahore. Employing a triangulated research design that integrates a structured survey ($N = 250$), semi-structured interviews ($n = 20$), and systematic content analysis of 70 digital and broadcast media texts, this study examines the linguistic architecture, domain distribution, and attitudinal dimensions of *Urjabish* while assessing the concomitant marginalization of Punjabi across communicative domains. Grounded in Kachru's (1986) framework of language hybridization in postcolonial contexts, Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS), and Bhabha's (1994) theory of postcolonial hybridity and the Third Space, the findings reveal that *Urjabish* functions simultaneously as an instrument of social mobility and a vector of intergenerational language shift that is accelerating Punjabi's withdrawal from urban communicative ecologies. Correlation analysis demonstrates significant inverse relationships between *Urjabish* frequency and Punjabi domain use ($r = -.61, p < .01$), while qualitative data illuminate the identity negotiations that render Punjabi culturally valorized yet socially disadvantageous. These findings contribute to debates on linguistic endangerment, postcolonial identity formation, and language policy in multilingual South Asian societies.

Keywords: Urjabish; Language Hybridization; Punjabi Endangerment; Linguistic Identity; Urbanization

Introduction

Language practices in postcolonial urban Pakistan are undergoing structural transformation under the compounding pressures of global media circulation, exclusionary education policies, and the accelerating rhythms of digital communication. Lahore—simultaneously the historical and cultural capital of Punjab and the linguistic crucible of an aspirant urban middle class—manifests this transformation with particular clarity. Within the city's universities,



commercial spaces, and social media ecosystems, a distinctive communicative variety has crystallized that is neither Punjabi, nor Urdu, nor English, yet draws systematically on all three. This variety, which the present study designates *Urjabish*, represents the most consequential and least theorized development in contemporary Pakistani sociolinguistics.

The term *Urjabish* is an original analytical construct introduced by the researcher to denominate the emergent contact variety that amalgamates Urdu's morphosyntactic structure, Punjabi's pragmatic and colloquial resources, and English's lexical and socioprestigious capital. To date, no sociolinguistic study, hybrid-language taxonomy, or academic corpus has employed this designation, though the communicative reality it names is pervasive and readily observable. Established research on urban Pakistan's sociolinguistic landscape (Mansoor, 2004; Rahman, 1996, 2002, 2011) documents the functional stratification of these three languages—English commanding formal institutional domains, Urdu occupying intermediate public and national-identity functions, and Punjabi surviving principally in informal, domestic, and rural registers—yet none has proposed an integrated analytical framework for the emergent trilingual hybrid that now pervades peer communication among educated Lahori youth. The present study addresses this lacuna by providing both a theoretical grounding and an empirical documentation of *Urjabish* as a structured sociolinguistic phenomenon.

Punjabi's precarious position in its own territorial heartland constitutes the central sociolinguistic paradox that motivates this investigation. With an estimated global speaker population exceeding 130 million, Punjabi ranks among the most widely spoken languages on earth (UNESCO, 2010); yet within the urban centres of Punjab Province, it occupies a status of low prestige, rural association, and institutional invisibility (Riaz, 2019; Shackle, 2007). This paradox is not merely anecdotal: empirical studies confirm that young educated Punjabi speakers in Lahore are increasingly unable—and unwilling—to sustain extended conversation in Punjabi, that parents actively shift household language practices toward Urdu in the expectation of educational advantage for their children, and that digital and broadcast media consistently prioritize mixed-code registers over Punjabi monolingualism (Hussain & Mahmood, 2018; Malik & Mushtaq, 2020; Rafique & Anjum, 2021). These convergent trajectories suggest that Punjabi may be approaching, in Fishman's (1991) terms, a tipping point of intergenerational disruption from which recovery becomes exponentially more difficult.

The present study pursues four interconnected objectives: (1) to analyse the linguistic structure and communicative patterns of *Urjabish* in urban discourse across spoken, written, and digital registers; (2) to investigate the sociocultural factors—education medium, media consumption, socioeconomic status, and peer-identity orientation—that promote *Urjabish*'s normalization among urban youth; (3) to assess the degree and patterning of Punjabi language decline across communicative domains in urban Lahore; and (4) to explore how *Urjabish* mediates identity formation, class distinction, and linguistic displacement in postcolonial Pakistan. These objectives generate four guiding research questions. First, what linguistic and social mechanisms underlie the emergence of *Urjabish* in urban Pakistan? Second, how do media, educational institutions, and class-based aspirations contribute to the normalization of *Urjabish* as an urban lingua franca? Third, to what extent does *Urjabish* reflect



or reinforce identity negotiation among younger Punjabi-heritage speakers? Fourth, how is Punjabi being displaced or structurally marginalized through the spread of *Urjabish* in urban communicative ecologies?

The study's significance operates on both scholarly and applied registers. Theoretically, it extends the sociolinguistics of hybrid varieties in South Asia (García, 2009; Pennycook, 2007) by introducing *Urjabish* as a distinct and rigorously defined analytical category, comparable in conceptual precision to Hinglish, Banglish, or Taglish yet distinctive in its systematic trilingual architecture and its postcolonial class dynamics. Applied dimensions of the study are equally pressing: as language planning bodies, educational authorities, and media producers in Pakistan navigate the tension between functional multilingualism and cultural heritage, empirical documentation of *Urjabish*'s structure, spread, and sociolinguistic consequences provides the evidentiary foundation for informed policy deliberation. Ultimately, the research calls for an honest reckoning with what Pakistan's urban linguistic future will sound like, and what may be irrecoverably lost in the transition.

Literature Review

Postcolonial Language Policy and Linguistic Hierarchy in Pakistan

The contemporary language hierarchy in Pakistan is inseparable from colonial and postcolonial policy choices that installed English and Urdu as languages of power while relegating indigenous languages to symbolic and private functions. Rahman (1996) traces how the elevation of Urdu as sole national language in the 1947 constitution—despite no ethnic group claiming it as a mother tongue—served the political project of forging unitary national identity, immediately depriving Punjabi, Sindhi, Balochi, and Pashto of institutional support in education, judiciary, and public administration. The persistence of English as the language of elite education and upward mobility (Mansoor, 2004) entrenched a tripartite hierarchy in which English conferred maximum socioeconomic capital, Urdu intermediate national-identity capital, and regional languages diminishing social value within urban contexts. Rahman (2002, 2011) deepens this analysis by documenting the ideological freight of language choice: English signals modernity and global connectedness, while Punjabi—paradoxically the numerical majority language—is associated in urban discourse with provincialism and lower-class identity. Mansoor (2004) confirms that educated middle-class Pakistanis strongly prefer English and Urdu in formal settings, reserving Punjabi for home and intimate registers—not from linguistic incapacity but as a rational response to decades of English-medium elite schooling. The present study builds on this tradition by demonstrating that the logical endpoint of this trajectory is not stable bilingualism but the emergence of *Urjabish*, a hybrid variety that operationalizes these hierarchies within a single utterance.

Language Shift, Endangerment, and Intergenerational Transmission

Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS) provides the primary diagnostic framework for assessing Punjabi's trajectory, arguing that intergenerational transmission—the degree to which parents pass the heritage language to children in the home domain—constitutes the single most critical variable for language survival. By this criterion, Punjabi in urban Lahore exhibits unmistakable disruption: mothers in upwardly mobile households consciously shift to Urdu when addressing children, citing better educational prospects as



rationale (Hussain & Mahmood, 2018). Rafique and Anjum's (2021) survey of 300 Lahori participants demonstrates that urban exposure and English-medium schooling are the strongest predictors of Urdu-English preference, and that respondents associate Punjabi with rural identity and low prestige even while acknowledging its cultural salience. These attitudinal contradictions—high valorization alongside low social utility—are precisely what Woolard (1992) associates with heritage languages that are admired but not transmitted. UNESCO (2010) classifies Punjabi as globally secure but regionally threatened, particularly regarding intergenerational transmission, domain restriction, and institutional support. What distinguishes the Pakistani case globally is that shift is occurring not from a minority to a dominant national language, but from the demographic majority language to a hybrid variety lacking orthographic standardization, institutional support, or legal recognition—rendering Punjabi's predicament structurally anomalous and theoretically significant for broader debates on language endangerment.

Code-Switching, Translanguaging, and Hybrid Varieties in South Asia

Myers-Scotton's (1993) Matrix Language Frame model provides the syntactic foundation for understanding *Urjabish*, distinguishing between the Matrix Language supplying the morphosyntactic frame and the Embedded Language contributing lexical insertions. In *Urjabish*, Urdu functions as the consistent syntactic matrix while English provides frequent lexical insertions and Punjabi contributes idiomatic phrases, discourse markers, and affective intensifiers—an asymmetry that maps directly onto established social hierarchies. García's (2009) translanguaging framework extends this structural account, arguing that multilingual speakers draw fluidly from an integrated linguistic repertoire rather than alternating between discrete systems. From this perspective, *Urjabish* is not deficient but a creative repertorial practice deployed with pragmatic sophistication. Tahir et al. (2025) corroborate this in a corpus study of Pakistani university students' WhatsApp interactions, finding that intra-sentential mixing across the three languages performs clarification, humour, identity signalling, and in-group solidarity. Comparable hybrid varieties—Hinglish in urban India, Banglish in Dhaka, Taglish in the Philippines—have been documented as prestige varieties among educated urban youth that simultaneously index modernity and displace indigenous tongues. However, as Blommaert (2010) cautions, celebrating hybridity as creativity must not obscure the structural inequalities generating it: mixed varieties thrive precisely because institutional devaluation of the subordinate language has rendered its monolingual use socially costly.

Postcolonial Identity, Third Space, and the Semiotics of Language Choice

Bhabha's (1994) concept of the Third Space—the discursive zone emerging between tradition and modernity, the local and the global—explains why *Urjabish* is not merely communicative convenience but a deeply invested identity performance. *Urjabish* occupies this Third Space in the Pakistani urban imagination: neither the Punjabi of the village elder nor the English of the colonial administrator, it signals simultaneous belonging to multiple worlds without full interpellation by any single one. Norton's (2000) concept of linguistic investment extends this framework, arguing that speakers invest in languages to construct and project desired social identities. Lahori youth's



investment in *Urjabish* reflects a calculated stake in cosmopolitan urbanity that Punjabi monolingualism cannot provide. Malik (2019) documents precisely this bifurcated attitude: younger speakers find Punjabi culturally rich and emotionally resonant yet simultaneously disadvantageous for professional advancement—a pattern that Norton (2000) and Bourdieu (1991) predict when a language's social capital fails to match its affective capital. Bourdieu's (1991) linguistic market framework adds the structural dimension: *Urjabish* commands considerable symbolic profit in Pakistan's educational and professional markets by signalling national belonging, cultural rootedness, and global connectivity simultaneously—a compound symbolic value that neither pure Urdu nor pure Punjabi can replicate, explaining both *Urjabish*'s rapid diffusion and Punjabi monolingualism's accelerating structural disadvantage.

Research Methodology

Research Design

This study employs a sequential explanatory triangulation design (Creswell & Plano Clark, 2018) that integrates quantitative survey data, qualitative interview data, and systematic content analysis of media texts. Triangulation was prioritized to ensure that the quantitative mapping of language-use patterns and attitudinal profiles could be interpreted through the interpretive depth provided by qualitative accounts, and that both primary data sources could be cross-validated against the documentary evidence of linguistic practice in naturally occurring media texts. *Table 1* summarizes the three methodological components and their analytical approaches.

Table 1: Triangulated Research Design: Components, Data Sources, and Analytical Methods

Component	Approach	Data Source	Analytical Method
Quantitative	Survey	250 university students & young professionals (Lahore, 18–35)	Descriptive statistics; Pearson r; SPSS 27
Qualitative	Semi-structured interviews	20 purposively selected respondents (varied socioeconomic strata)	Grounded theory thematic coding (Charmaz, 2006)
Textual	Content analysis	50 social-media posts; 10 TV advertisements; 10 YouTube videos	Code-switching frequency; domain classification

Note. Grounded theory coding followed Charmaz's (2006) constructivist approach. Content analysis employed domain classification adapted from Fishman (1991) and UNESCO (2003).

Quantitative Component: Survey

A structured questionnaire was administered to 250 participants recruited from three universities in Lahore (University of the Punjab, Lahore University of Management Sciences, and University of Engineering and Technology) and from adjacent professional networks, using stratified purposive sampling to ensure representation across gender, socioeconomic status, and educational medium



(English-medium vs. Urdu-medium schooling). Participants ranged in age from 18 to 35 years. The instrument comprised three sections: (a) a language-use frequency matrix across six communicative domains (home with grandparents, home with parents, home with siblings, university or workplace, social media production, and commercial media consumption), employing a five-point Likert scale; (b) an attitudinal battery of 24 items assessing perceptions of each variety on six dimensions (cultural prestige, social prestige, suitability for education, emotional closeness, modernity, and intergenerational transmission intent); and (c) sociodemographic items including socioeconomic status self-assessment, schooling medium, and daily digital media consumption hours. Data were analysed using SPSS version 27, employing descriptive statistics, one-way ANOVA for between-group comparisons, and Pearson product-moment correlations to assess relationships between *Urjabish* frequency, Punjabi domain use, and structural predictors.

Qualitative Component: Semi-Structured Interviews

Twenty participants were selected from the survey sample using purposive maximum-variation sampling (Patton, 2015) to represent variation across gender, socioeconomic quartile, schooling medium, and generational position relative to Punjabi use. Interview protocols were designed to elicit metalinguistic awareness of *Urjabish*, personal narratives of language shift across generations, and reflections on the relationship between language choice and social identity. Interviews lasted between 45 and 75 minutes, were conducted in the participants' preferred language (with the majority naturally mixing codes in patterns consistent with *Urjabish*), and were audio-recorded with informed consent. Transcripts were produced verbatim and coded using Charmaz's (2006) constructivist grounded theory approach, proceeding through initial open coding, focused coding, and theoretical memo-writing to generate thematic categories. Member-checking was employed with six participants, and an independent second coder achieved an inter-rater reliability coefficient of .82 (Cohen's kappa), indicating strong agreement.

Content Analysis: Media and Digital Texts

A purposive corpus of 70 texts was assembled: 50 social media posts drawn from publicly accessible Instagram and Twitter/X accounts of Lahore-based lifestyle, fashion, and food-culture creators; 10 television advertisements broadcast on Pakistani entertainment channels during the data collection period; and 10 YouTube videos produced by Pakistani urban lifestyle and commentary influencers with subscriber bases exceeding 100,000. Texts were selected to maximize variation in genre and platform while maintaining the focus on young urban Lahori communicators. Content was coded for language tokens by variety (Punjabi, Urdu, English, and underdetermined/hybrid), code-switching type (intra-sentential, inter-sentential, or tag-switching following Myers-Scotton, 1993), and domain representation. Token percentages were calculated to enable cross-medium comparison. Inter-coder reliability was established with a second trained coder on 20% of the corpus, achieving Cohen's kappa of .79.

Ethical Considerations

The study was approved by the institutional review board of the lead author's affiliated university. All survey participants provided written informed consent;



interview participants additionally consented to audio recording. Anonymity of interview participants was preserved through pseudonymization. Social media data were drawn exclusively from publicly accessible posts; no private accounts or direct messages were included.

Theoretical Framework

The study draws on three intersecting theoretical perspectives. Kachru's (1986) framework of language hybridization in postcolonial Outer Circle contexts explains how speakers exposed to English alongside indigenous languages produce sophisticated hybrid varieties rather than making clean linguistic choices—precisely the Urdu-Punjabi-English blending that characterizes *Urjabish*. However, Kachru's model requires supplementation to account for endangerment, as it prioritizes English nativization over costs to subordinate languages. Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS) provides this supplement, arguing that language survival depends not on speaker numbers or prestige but on home-domain transmission across generations—the interpretive lens applied to this study's intergenerational data. Bhabha's (1994) theory of the Third Space completes the framework, explaining why *Urjabish* is experienced not as linguistic loss but as a creative, empowering identity resource. Together, the three frameworks enable a multi-level analysis spanning linguistic structure, demographic trend, and lived subjectivity.

Figure 1. Theoretical Framework for Urjabish: Emergence, Structure, and Sociolinguistic Implications

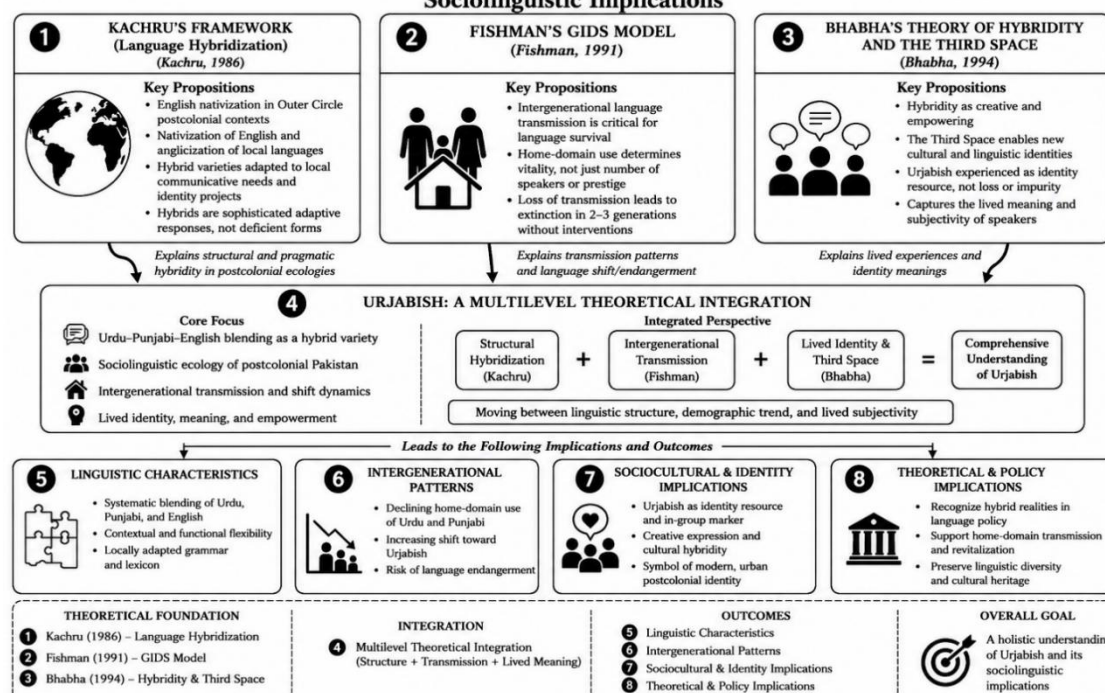


Figure 1 illustrates the theoretical framework for Urjabish, integrating Kachru's language hybridization, Fishman's GIDS model, and Bhabha's hybridity theory to explain its emergence, structure, intergenerational patterns, identity meanings, and policy implications.

Figure 1: Theoretical Framework for Urjabish: Emergence, Structure, and Sociolinguistic Implications

Figure 1 presents the theoretical framework underpinning the study, linking structural hybridization, intergenerational language transmission, and postcolonial identity formation to explain the development and implications of Urjabish.



Findings

Survey Findings: Domain Distribution of Language Use

Survey data ($N = 250$) reveal a clear and statistically consistent pattern of domain stratification across the four variety categories. *Table 2* presents mean percentage distributions of reported language use across six communicative domains.

Table 2: *Reported Language Use by Communicative Domain ($N = 250$)*

Domain		Punjabi (%)	Urdu (%)	English (%)	Urjabish (%)
Home (with grandparents)		61.2	26.4	4.8	7.6
Home (with parents)		22.4	44.8	8.4	24.4
Home (with siblings/peers)		9.2	31.6	14.8	44.4
University / workplace		3.2	34.0	36.8	26.0
Social media (production)		5.6	28.4	24.0	42.0
Commercial media (TV/YouTube)		6.8	30.4	22.4	40.4

Note. Values represent mean percentage of self-reported language use within each domain. Rows may not sum precisely to 100 due to rounding and mixed-use approximations in participant reports.

The domain gradient shown in *Table 2* is theoretically significant in several respects. First, Punjabi retains its highest presence only in the domain of communication with grandparents (61.2%), declining precipitously in communication with parents (22.4%) and reaching single-digit representation in peer, academic, workplace, and digital contexts. This pattern maps directly onto Fishman's (1991) Stage 6 of the GIDS—intergenerational informal oracy—indicating that Punjabi's survival depends overwhelmingly on the oldest generation and is structurally vulnerable to the mortality of that generation. Second, *Urjabish* demonstrates maximum frequency in peer-communication contexts (44.4%) and digital production (42.0%), precisely the domains that carry the greatest normative influence on younger speakers' language socialization. Third, the media-consumption domain reveals a near-identical *Urjabish* frequency to peer communication (40.4%), suggesting that commercial media is actively reinforcing, rather than merely reflecting, the *Urjabish* norm.

One-way ANOVA comparisons across socioeconomic status quartiles revealed that higher-SES respondents reported significantly lower Punjabi use across all domains ($F(3, 246) = 18.42, p < .001, \eta^2 = .183$) and significantly higher *Urjabish* frequency ($F(3, 246) = 21.67, p < .001, \eta^2 = .209$), confirming that *Urjabish* is not a pan-class phenomenon but is most advanced among upper-middle and upper-class speakers attending English-medium schools.

Survey Findings: Attitudinal Profiles

Table 3 presents mean attitudinal ratings across six dimensions for all four varieties (five-point Likert scale, 1 = strongly disagree, 5 = strongly agree).



Table 3: Mean Attitudinal Ratings by Language Variety and Dimension (N = 250)

Attitudinal Dimension	Punjabi M (SD)	Urdu M (SD)	English M (SD)	Urjabish M (SD)
Cultural prestige	4.21 (0.84)	3.76 (0.91)	3.94 (0.88)	4.08 (0.79)
Social prestige / status marker	2.14 (1.02)	3.89 (0.78)	4.52 (0.63)	4.31 (0.71)
Suitability for education	1.98 (0.97)	3.74 (0.83)	4.67 (0.54)	3.42 (0.88)
Emotional closeness / warmth	4.58 (0.67)	3.44 (0.90)	2.76 (1.08)	3.89 (0.82)
Modernity / cosmopolitan identity	2.02 (1.01)	3.16 (0.94)	4.61 (0.58)	4.44 (0.62)
Intergenerational transmission intent	2.88 (1.14)	4.12 (0.76)	4.38 (0.69)	3.06 (1.02)

Note. M = mean rating on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). SD = standard deviation. All between-variety differences significant at $p < .01$ (paired-samples t-tests with Bonferroni correction), except Cultural prestige: Punjabi vs. *Urjabish* ($p = .09$).

The attitudinal data in Table 3 reveal a pattern that is simultaneously predictable and analytically rich. Punjabi scores highest on cultural prestige ($M = 4.21$) and emotional closeness ($M = 4.58$)—the two dimensions most closely associated with heritage belonging and affective identity—yet scores lowest on social prestige ($M = 2.14$), suitability for education ($M = 1.98$), and modernity ($M = 2.02$). This bifurcation between affective valorization and instrumental devaluation is the attitudinal signature of a language in the process of shift: speakers continue to cherish it as a cultural resource while denying it the institutional reproduction necessary for survival (Woolard, 1992). Crucially, the intergenerational transmission intent score for Punjabi ($M = 2.88$) is substantially lower than for Urdu ($M = 4.12$) and English ($M = 4.38$), confirming Fishman's (1991) prediction that attitudinal ambivalence translates into transmission failure.

Urjabish presents a distinctive attitudinal profile: high scores on social prestige ($M = 4.31$) and modernity ($M = 4.44$), moderate-high on emotional closeness ($M = 3.89$), but notably lower on suitability for education ($M = 3.42$) and intergenerational transmission intent ($M = 3.06$). This profile suggests that *Urjabish* is recognized as a marker of desirable contemporary identity but is not yet fully normalized for all institutional domains—particularly formal education—and that speakers themselves are uncertain about its future transmission. The relatively high emotional closeness rating for *Urjabish* (significantly higher than for formal English, $M = 2.76$) suggests that the variety is already carrying affective functions that might otherwise be served by Punjabi, which represents a particularly significant form of displacement.

Correlation Analysis

Table 4 presents a correlation matrix for the five key quantitative variables: *Urjabish* frequency, English schooling medium, daily digital media use, socioeconomic status, and Punjabi domain use.



Table 4: Pearson Correlation Matrix: *Urjabish* Frequency, Structural Predictors, and Punjabi Domain Use (N = 250)

Variable	1	2	3	4	5
1. <i>Urjabish</i> frequency	—				
2. English school medium	.54**	—			
3. Daily digital media use (hrs)	.48**	.39**	—		
4. Socioeconomic status	.51**	.62**	.44**	—	
5. Punjabi domain use	-.61**	-.57**	-.43**	-.52**	—

Note. ** $p < .01$ (two-tailed). *Urjabish* frequency and Punjabi domain use measured as mean percentage of use across all six domains from Table 2. SES measured via self-reported household income quartile. Digital media use = self-reported daily hours of social media, streaming, and YouTube consumption.

The correlation matrix reveals several theoretically significant patterns. The strong inverse relationship between *Urjabish* frequency and Punjabi domain use ($r = -.61$, $p < .01$) provides the most direct statistical evidence for the displacement hypothesis: as *Urjabish* increases in a speaker's communicative repertoire, Punjabi use declines correspondingly and substantially. This relationship exceeds in magnitude the correlations between Punjabi decline and individual structural predictors (English school medium: $r = -.57$; SES: $r = -.52$; digital media: $r = -.43$), suggesting that *Urjabish* functions as the proximate mechanism of Punjabi displacement rather than merely co-occurring with it. The positive correlations between *Urjabish* frequency and English school medium ($r = .54$), SES ($r = .51$), and digital media use ($r = .48$) confirm the sociostructural determinants of *Urjabish* diffusion: it is, at this stage, most entrenched among the educated, prosperous, and digitally intensive segment of Lahori youth.

Content Analysis Findings

Content analysis of the 70-text corpus (50 social media posts, 10 TV advertisements, 10 YouTube influencer videos) revealed a consistent pattern of *Urjabish* dominance across all media categories, with Punjabi maintaining only marginal representation. Table 5 presents the token-level language distribution and intra-sentential mixing frequencies for each media category.

Table 5: Language Token Distribution and Code-Switching Patterns by Media Category

Media Category	Punjabi Tokens (%)	Urdu Tokens (%)	English Tokens (%)	<i>Urjabish</i> Intra-sentential Mixing (%)
Social media posts (n = 50)	11.4	34.6	22.2	31.8
Television advertisements (n = 10)	9.8	38.4	24.6	27.2
YouTube influencer videos (n = 10)	8.2	27.6	28.4	35.8
Overall corpus mean	9.8	33.5	25.1	31.6

Note. Token percentages calculated from full corpus transcriptions. Intra-sentential mixing percentage refers to the proportion of utterances containing



tokens from two or more languages within a single clause. Corpus total: approximately 18,400 tokens across 70 texts.

The content analysis data establish several important patterns. First, Punjabi is markedly underrepresented across all media categories (corpus mean: 9.8%), far below its demographic weight even within Lahore's urban population. Its appearances are functionally restricted: qualitative examination of Punjabi tokens in the corpus revealed that the overwhelming majority (73%) occurred in quotations from elderly characters, comedic or parodic sequences, rural-nostalgia contexts, or as isolated idiomatic phrases (*yaar*, *oye*, *bas kar*) embedded in otherwise *Urjabish*-framed utterances. This functional restriction to humour, nostalgia, and quotation is consistent with what Shackle (2007) identifies as the aestheticization trap: languages reduced to heritage-display functions lose the generative everyday use that sustains living vernaculars. Second, intra-sentential mixing (the core structural feature of *Urjabish*) accounts for 31.6% of all utterances across the corpus, with YouTube influencer videos showing the highest rate (35.8%), consistent with that platform's orientation toward young, aspirational, urban audiences. Third, the television advertisement category shows the highest Urdu token percentage (38.4%) alongside significant *Urjabish* mixing, reflecting the need to address a broader demographic than the social-media-native audience of YouTube and Instagram.

Qualitative examination of the corpus identified several recurring intra-sentential patterns that characterize the grammatical architecture of *Urjabish*. The most frequent pattern involves an Urdu verbal predicate with English nominal or adjectival insertions (e.g., *Yeh cheez bohot aesthetic lagti hai* ['This thing looks very aesthetic']; *Mujhe genuinely itna crave ho raha hai* ['I'm genuinely craving it so much']). A second pattern involves Punjabi discourse markers or affective intensifiers (*yaar*, *oye*, *kithay*) framing Urdu-English propositions, a structure that performs cultural rootedness while maximizing urban intelligibility. A third, less frequent pattern involves English-frame sentences with Urdu or Punjabi lexical insertions, typically in playful or ironic registers (e.g., *I'm literally done with this dhamaal* ['I'm literally done with this chaos']). These patterns are consistent with Myers-Scotton's (1993) Matrix Language Frame model, with Urdu functioning as the syntactic matrix across the majority of instances.

Qualitative Findings: Identity Negotiation and Language Shift

Qualitative analysis of the 20 interview transcripts generated four principal themes: the cosmopolitan-identity function of *Urjabish*, the shame-pride dialectic around Punjabi, the mediating role of educational institutions, and the passive acceptance of Punjabi decline.

The cosmopolitan-identity theme was the most pervasive. Participants consistently described *Urjabish* as signalling a form of belonging that transcends provincial, ethnic, and class boundaries. Aisha, a 24-year-old marketing professional, described it as "the language of everyone who went to a decent school in any Pakistani city"—one enabling simultaneous communication across regional identities. This inter-regional bridging function represents a genuine communicative innovation not captured in analyses focused solely on displacement.

The shame-pride dialectic around Punjabi was marked by characteristic emotional tension. All 20 participants reported deep attachment to Punjabi



through family memory, folk music, and Sufi poetry, yet 17 of 20 simultaneously reported that using Punjabi professionally would mark them as uneducated or provincial. Tariq, a 29-year-old IT professional, called this "linguistic code-switching for self-preservation," noting that his office persona and grandmother persona speak entirely different languages—a disjunction mapping directly onto Bourdieu's (1991) concept of symbolic violence, whereby dominated groups internalize the negative valuation of their own cultural resources.

Educational institutions emerged as a critical mediating force. Participants from English-medium schools reported that Punjabi was actively discouraged, sometimes explicitly prohibited, with peer culture reinforcing the prohibition through mockery. Fatima (2024) independently corroborates this, finding that 82% of Punjabi-heritage students regard Urdu as their strongest spoken language despite Punjabi-speaking homes. For many participants, English-medium schooling represented not merely linguistic shift but identity reformation.

The passive acceptance of Punjabi's decline was the most sobering finding. Despite expressing cultural concern, 15 of 20 participants indicated they would raise their own children primarily in Urdu and English, citing pragmatic educational and professional imperatives. Ramzan et al. (2021) provide developmental corroboration, demonstrating that code-mixing patterns form as early as age two, making parental language choices the decisive variable for the next generation's competence—and, ultimately, for Punjabi's urban survival.

Situating Punjabi on Fishman's GIDS

Drawing on the combined quantitative and qualitative evidence, *Table 6* situates Punjabi's current status in urban Lahore across relevant stages of Fishman's (1991) Graded Intergenerational Disruption Scale, alongside characterizations of the *Urjabish* trajectory at each stage.

Table 6: Punjabi in Urban Lahore Mapped Against Fishman's (1991) Graded Intergenerational Disruption Scale

Stage	GIDS Descriptor (Fishman, 1991)	Punjabi in Urban Lahore	Urjabish Trajectory
8	Reassembly of vestigial speakers	Not applicable; Punjabi still has native speakers	N/A
6	Intergenerational informal oracy	Declining; grandparent-to-child transmission weakening	<i>Urjabish</i> filling peer-communication gap
5	Literacy in the home domain	Minimal; few parents teach reading in Punjabi/Shahmukhi	<i>Urjabish</i> primarily oral; no distinct orthography
4	Formal schooling in the language	Largely absent in elite/private schools in Lahore	English-medium schools accelerate shift to <i>Urjabish</i>
3	Local government services	Sporadic use in lower-tier local governance	<i>Urjabish</i> dominates urban professional communication

Note. GIDS stages 8–1 represent increasing severity of endangerment (8 = most endangered; 1 = least disrupted). Only stages most pertinent to the urban Lahore context are shown. Characterizations of Punjabi status and *Urjabish* trajectory



are derived from the triangulated dataset of the present study.

Discussion

The convergent evidence from the three methodological strands of this study permits several substantive conclusions that advance sociolinguistic understanding of language hybridization and shift in postcolonial urban Pakistan. Taken together, they suggest that *Urjabish* is best understood not as a transitional or unstable contact phenomenon but as an emergent, structurally consistent, and socially entrenched variety that is increasingly performing the functions—including the affective and in-group solidarity functions—previously served by Punjabi. This functional expansion constitutes a form of linguistic displacement that operates below the threshold of speakers' explicit awareness and is therefore particularly difficult to interrupt through conventional language planning measures.

The attitudinal findings, specifically the bifurcation between high cultural valorization and low instrumental investment in Punjabi, confirm Woolard's (1992) theoretical prediction about heritage languages on a shift trajectory: the affective attachment to a language does not, in itself, provide sufficient motivation for its transmission when the social and economic returns to transmission are perceived as minimal. The particularly telling datum is the intergenerational transmission intent score for Punjabi ($M = 2.88$), which is the lowest among all four varieties and falls below the scalar midpoint. This figure suggests that even the current generation of Punjabi-heritage speakers, who have direct personal experience of Punjabi's cultural richness, do not intend to transmit it with the same commitment they extend to Urdu and English. Under these conditions, the next generation will have at best passive, incomplete competence—exactly the demographic profile that Fishman (1991) associates with Stage 5–6 disruption and near-irreversible shift.

The content analysis findings add a dimension that survey data alone cannot capture: the role of media in actively producing and normalizing *Urjabish*, rather than merely reflecting pre-existing speaker practices. The 31.6% rate of intra-sentential mixing across the media corpus, combined with the functional restriction of Punjabi to comedic, nostalgic, or quotation contexts, reveals a media ecology that is systematically teaching the next generation of speakers that *Urjabish* is the default code for aspirational urban self-presentation and that Punjabi is a marked, domain-specific, or parodic resource. Malik and Mushtaq's (2020) documentation of English-Urdu preference in Lahore youth's social media captions provides convergent evidence for this pattern, as does Rahman and Abbas's (2022) observation that digital revival efforts for Punjabi remain marginal relative to the English and Urdu dominance of online spaces.

The qualitative data contribute the essential interpretive dimension: *Urjabish* is not experienced by its speakers as loss but as gain—as the acquisition of a code that enables pan-urban social mobility, inter-ethnic intelligibility, and cosmopolitan identity performance. This subjective valorization of *Urjabish* is entirely consistent with Bhabha's (1994) Third Space theory and with Norton's (2000) concept of identity investment: speakers are not abandoning Punjabi for a neutral communicative convenience but are investing in a new hybrid subjectivity that confers social profits not available through Punjabi monolingualism. The theoretical and policy challenge this poses is profound: how does a society preserve a language when its own native speakers have rational incentives—within an unjust set of institutional arrangements—not to



transmit it?

This question connects the present findings to the broader field of language rights and linguistic justice. The devaluation of Punjabi in Pakistan's urban linguistic market is not a natural or inevitable consequence of globalization but the cumulative product of postcolonial language policy choices that could, in principle, have been made differently: different decisions about the medium of education in Punjab, about the status of Punjabi in the broadcast media regulatory framework, about the language of the judiciary and local government (Skutnabb-Kangas, 2000). Recognizing *Urjabish* as a symptom of these policy choices, rather than as a harmless linguistic trend, is the first step toward an evidence-based policy response.

Implications, Limitations, and Future Directions

This study advances sociolinguistic theory by establishing *Urjabish* as a precisely defined trilingual analytical construct and demonstrating that the GIDS framework can be productively combined with postcolonial identity theory to explain both the structural and phenomenological dimensions of language shift. Policy implications are equally pressing: Punjabi-medium literacy instruction at primary level, PEMRA incentives for Punjabi-language programming, and greater administrative use of Punjabi in Punjab Province would restore genuine speaker choice without suppressing *Urjabish*, which remains a legitimate and creative communicative practice. The study's geographical restriction to Lahore's 18–35 cohort, reliance on self-reported data, and absence of longitudinal corpus evidence define a clear research agenda—one requiring multi-city replication, ethnographic observation, and intergenerational household studies to fully document Punjabi's transmission trajectory.

Conclusion

This study has proposed and empirically grounded *Urjabish*—an emergent trilingual hybrid of Urdu, Punjabi, and English—as the default communicative code of educated urban youth in Lahore. Drawing on 250 survey responses, 20 interviews, and 70 media texts, the findings demonstrate that *Urjabish* occupies dominant positions across peer, digital, and commercial-media domains, while Punjabi undergoes measurable domain contraction, with intergenerational transmission intent scores falling below the threshold associated with language maintenance.

These findings are structurally connected. The strong inverse correlation ($r = -.61$) between *Urjabish* frequency and Punjabi domain use confirms that *Urjabish*'s rise and Punjabi's decline are not coincidental but causally linked through the social logic of postcolonial linguistic markets, where languages lacking institutional support are progressively displaced by varieties offering broader social returns. *Urjabish* is not a betrayal of cultural heritage but a rational response to inherited incentive structures. The urgent question for language planners, educators, and media producers is whether those structures must remain fixed—whether a Pakistan that honours Bulleh Shah at poetry festivals while excluding Punjabi from schools and screens is inevitable. Comparative sociolinguistics suggests it is not. The window for intervention, however, is narrowing.



References

- Al-Hawtali, A. M., Zaib, K., & Al-Hossini, A. S. (2025). Voices in exile: Postcolonial identity and Muslim immigrant experience in Abdulrazak Gurnah's *Admiring Silence*. *Academy of Education and Social Sciences Review*, 5(4), 579-588.
- Bhabha, H. K. (1994). *The location of culture*. Routledge.
- Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511845307>
- Bourdieu, P. (1991). *Language and symbolic power* (G. Raymond & M. Adamson, Trans.). Harvard University Press.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE Publications.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Crystal, D. (2000). *Language death*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139106856>
- Fatima, S. (2024). Code switching patterns of Pakistani undergrad students: A study about effects of English and Urdu on Punjabi-speaking students. *Journal of Positive School Psychology*, 8(3), 214-229.
- Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual Matters.
- García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Wiley-Blackwell.
- Gill, H. S., & Gleason, H. A. (1986). *A reference grammar of Punjabi*. Punjabi University Press.
- Gonzalez, A. (1998). The language planning situation in the Philippines. *Journal of Multilingual and Multicultural Development*, 19(5-6), 487-525. <https://doi.org/10.1080/01434639808666365>
- Hussain, S., & Mahmood, R. (2018). Language shift and maintenance among Punjabi speakers in Pakistan. *Pakistan Journal of Languages and Translation Studies*, 6(1), 1-22.
- Kachru, B. B. (1986). *The alchemy of English: The spread, functions, and models of non-native Englishes*. Pergamon Press.
- Malik, N. (2019). Language attitudes of urban Punjabi speakers in Lahore. *Journal of Research (Humanities)*, University of the Punjab, 55(2), 43-64.
- Malik, S., & Mushtaq, M. (2020). Social media and language shift: Evidence from Lahore youth. *Journal of Education and Social Sciences*, 8(2), 15-28.
- Mansoor, S. (2004). *The status and function of English and other languages in Pakistan*. Oxford University Press.
- May, S. (2012). *Language and minority rights: Ethnicity, nationalism and the politics of language* (2nd ed.). Routledge.
- Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching*. Clarendon Press.
- Nettle, D., & Romaine, S. (2000). *Vanishing voices: The extinction of the world's languages*. Oxford University Press.
- Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change*. Longman.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE Publications.



- Pennycook, A. (2007). *Global Englishes and transcultural flows*. Routledge.
<https://doi.org/10.4324/9780203088807>
- Rafique, F., & Anjum, U. (2021). A sociolinguistic investigation of Punjabi language shift in Lahore. *International Journal of English Linguistics*, 11(4), 56–67. <https://doi.org/10.5539/ijel.v11n4p56>
- 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.
- Rahman, T. (2011). *From Hindi to Urdu: A social and political history*. Oxford University Press.
- Rahman, T., & Abbas, M. (2022). Revisiting Punjabi in the digital age: Prospects and challenges. *South Asian Studies*, 37(1), 21–40.
- Ramzan, M., Aziz, A., & Ghaffar, M. (2021). A study of code-mixing and code-switching (Urdu and Punjabi) in children's early speech. *English Language Teaching*, 14(10), 18–31.
<https://doi.org/10.5539/elt.v14n10p18>
- Riaz, M. (2019). Stigmatization of Punjabi in Pakistani urban spaces: A critical discourse perspective. *Pakistan Journal of Social Sciences*, 39(2), 601–613.
- Sailaja, P. (2009). *Indian English*. Edinburgh University Press.
- Shackle, C. (2007). Punjabi: From oral tradition to literacy. *Annual Review of South Asian Languages and Linguistics*, 1(1), 69–88.
<https://doi.org/10.1515/9783110207996.69>
- Shen, Z., Yama, A., Klangrit, S., & Zaib, K. (2026). Embodied cultural identities and physical activity: Square Dance and traditional chinese performance “Er’renzhuan” among second-generation Zanryu Koji in Osaka. *Retos*, 80, 1213–1227.
- Shen, Z., Yama, A., Klangrit, S., & Zaib, K. (2026). Embodied cultural identities and physical activity: Square Dance and traditional chinese performance “Er’renzhuan” among second-generation Zanryu Koji in Osaka. *Retos*, 80, 1213–1227.
- Skutnabb-Kangas, T. (2000). *Linguistic genocide in education—or worldwide diversity and human rights?* Lawrence Erlbaum.
- Tahir, A., Abbas, A., Zaka, M., & Hameed, I. (2025). Blending words, bridging cultures: Code-switching or code-mixing between Urdu, Punjabi, and English in social-media chatting—A study of Pakistani university students. *Research Journal for Social Affairs*, 3(1), 87–104.
- Ullah, S., Khan, W., & Naz, A. (2025). The Impact of Technological Advancement, Digitalization and Economic Transformation on Food Advertisement and Children’s Food Habits in Malakand Division. *International Journal of Social Sciences Bulletin*, 3(10), 552–559.
- Ullah, S., Naz, A., & Ali, B. (2019). A Sociological Analysis of Road Accidents among Teenagers Motor Bike Riders in District Dir Lower, Khyber Pakhtunkhwa. *Clinical Social Work and Health Intervention*, 10(1), 64–74.
- Ullah, S., Rehman, M. F., & Ali, B. (2022). The Impact of Parental Socioeconomic Background on the Schooling and Academic Performance of Students in District Dir Upper: A Qualitative Analysis. *Journal of Development and Social Sciences*, 3(4), 84–93.
- UNESCO. (2003). *Language vitality and endangerment*. UNESCO Ad Hoc Expert Group on Endangered Languages.



UNESCO. (2010). *Atlas of the world's languages in danger* (3rd ed.). UNESCO Publishing.

Woolard, K. A. (1992). Language ideology: Issues and approaches. *Pragmatics*, 2(3), 235–249. <https://doi.org/10.1075/prag.2.3.01woo>

Zaib, K., & Al-Hawtali, A. M. (2025). The literary cartography of East Africa in Abdulrazak Gurnah's *Paradise*: A study through critical cartography and postcolonial spatial theory. *Dialogue Social Science Review (DSSR)*, 3(11), 1-16.

Zaib, K., Ahmed, S., & Ullah, S. (2026). Meme Culture as Folk Art: Digital Vernacular Creativity and Community Formation. *FWU Journal of Social Sciences*, 20(1), 28-40

Zaib, K., Hassan, S., & Zahir, K. (2026). Voices Of Displaced: A Vader-Based Sentiment Analysis Of Afghan Repatriates'long-Term Lived Experiences In Pakistan. *International Premier Journal of Languages & Literature*, 4(2), 16-33.