

SOCIO-CULTURAL IDENTITY AND LINGUISTIC LANDSCAPES: A MIXED-METHODS SENTIMENT ANALYSIS OF ESL LEARNERS' ATTITUDES TOWARD REGIONAL ROAD SIGNAGE IN PUNJAB

Dr. Farah Zaib*

¹Faculty of Arts and Languages, Department of English Linguistics, The Islamia University of Bahawalpur, 63100 Bahawalpur, Punjab, Pakistan

Email Addresses:

farah.zaib@iub.edu.pk (Farah Zaib)

Abstract

This mixed-method study explores the sociolinguistic perceptions of young Pakistani English as a Second Language (ESL) learners regarding the institutional installation of regional language signboards (Saraiki and Punjabi) in urban areas. This empirical framework investigates the ideological journey of the youth through the economic capital (global) and native language (indigenous) preservation. The data were collected through stratified digital questionnaire among a group of university level students (N=200) in four different districts of Punjab, including Bahawalpur, Multan, Faisalabad and Lahore. Quantitative attitudinal metrics processed in SPSS demonstrated high descriptive consensus regarding regional signage as an institutional safeguard against language death ($\mu = 3.96$), though constrained by an acute awareness of the socio-economic monopoly of English ($\mu = 4.12$). Statistically significant correlation between geographic context and attitudinal alignment was found through inferential testing ($\chi^2 = 10.66, p = 0.0048$). At the same time, a Natural Language Processing (NLP) program that was run automatically using the standalone Sentiment Analysis Add-in in Microsoft Excel indicated that positive sentiment tokens had a very high representation in the qualitative corpus (64.5%). Independent samples t-test affirmed that South Punjab text profiling had a much heavier emotional and subjective load as compared to Central Punjab discourse ($t = 5.34, p < 0.001$). Triangulated through Bourdieu's linguistic capital and Norton's investment model, the findings reveal that learners do not reject regional signage; instead, they successfully negotiate a dual identity, balancing essential economic investments in English with protective cultural commitments to their native roots.

Keywords: Applied Linguistics, Digital Pedagogy, Ethnolinguistic Vitality, Identity Negotiation, Linguistic Landscape, Linguistic Capital, Natural Language Processing (NLP)

1. Introduction

The urban centers of Pakistan's Punjab province are experiencing a quiet visual revolution. Along the great avenues, the historical avenues of Lahore to the princely crossroads of Bahawalpur, the traditional Urdu and English road signs are now more and more supplemented by bold scripts in Punjabi and Saraiki. These regional language signboards, which began as state-led efforts to preserve indigenous languages, have turned the actual landscape into an intricate stage of identity politics. The visual presentation of the public text in multi-ethnic societies can never be neutral; instead, they are all highly contested space where the administrative policy of the top and the community alignment of the bottom compete actively (Landry & Bourhis, 1997; Shohamy & Gorter, 2009). This visual contrast generates a psychological conflict with the young Pakistani learners of English as a Second Language (ESL). At the same time as they seek dominance over the English language, which remains the language of undisputed mastery of global capitalism,

prestige, and upward mobility in the country (Ahmed & Ahmad, 2026; Rahman, 2008), these learners are also faced with the institutional legitimacy of their native mother tongues in the streets. The processing of this spatial change by these bilingual students is a critical question. Is regional signage a proud recuperation of cultural heritage and ethnolinguistic vitality, or does it in any case represent an economically inefficient retrogression in an entirely globalized marketplace? The scholarship on the Pakistani linguistic landscape has already thoroughly recorded the rampant dominance of English and Englishized-Urdu hybridity of commercial store signs (Ahmed & Ahmad, 2026; Yasir & Khalid, 2024). But it is the particular congruity of state-sponsored regional signage and learner identity that needs a more empirical investigation. To overcome this gap, the current study examined the socio-cultural attitudes of the ESL learners in their journey to the changing visual world of their homeland. Through the automated sentiment analysis and the use of inferential descriptive statistics, the study unpuzzled the digital talk and ideological stand of young people who have to balance their educational investment in the world and their heritage ground in the local.

1.1 Background of the Study

Physical infrastructures of contemporary cities do not only compose a logistical system of transport; they are simultaneously a dynamic, ideological canvas called the linguistic landscape (LL). The linguistic landscape, as originally defined by Landry and Bourhis (1997), is the salience and visibility of languages on street signs, billboards, street names and commercial signage in a particular space or territory. Public signage serves two different purposes: an informational purpose that identifies geographical borders and a symbolic purpose that indicates relative power, institutional prestige, and socio-economic status of opposing linguistic groups (Ben-Rafael et al., 2006; Gorter, 2013). Each time a state actor changes this terrain either by introducing or eliminating a language, it reorganizes hierarchies of identity publicly and redistributes linguistic capital in that place (Bourdieu, 1991).

The linguistic landscape in Pakistan has long been a post-colonial strong-weak hierarchy. English, the language of official administration, jurisprudence, and higher learning, takes the highest position in this hierarchy, and plays the role as the final determinant of socio-economic mobility, as recorded by Rahman (2002) and Mansoor (2005). Urdu, the national language, is the horizontal unifier among the ethnically diverse provinces, overtaking the main civic language and mass media (Rahman, 2008). On the other hand, native regional dialects, though historically the mother languages of the overwhelming majorities of the population, have traditionally been pushed to the peripheral areas of life, and have been systematically marginalized as institutions within the centralized language planning structures of the state (Rahman, 1996, 2008).

But a major policy change is the recent administrative installation of localized regional language signboards on the thoroughfares of Punjab. The introduction of the Saraiki script in urban areas such as Bahawalpur and Multan, the introduction of Punjabi signs in urban areas such as Lahore and Faisalabad (see Appendix A), can be discussed as the institutional effort to raise the subjective ethnolinguistic vitality of vernacular languages into the physical public realm. To young Pakistani English as a Second Language (ESL) learners who are striving to traverse the higher education system, such a visual change provides a complicated sociolinguistic landscape. ESL learners of university level actively invest in English to enable them to gain global linguistic capital (Norton, 2013). However, they are also tied to their native communities through the heritage of their ancestors and the local community. The unexpected institutional validation of regional scripts on

the main streets compels such bilingual citizens to renegotiate their socio-cultural identities (Dörnyei, 2006). Although classical sociolinguistic questions are based on manual thematic extractions to describe these changing language attitudes, current computational linguistics offer sophisticated tools to trace large-scale public reception. This research explored the explicit expression of identities and emotional attitudes by young ESL learners when facing this kind of local reorganization of the linguistic landscape in Pakistan by using the natural language processing (NLP) and automated sentiment polarity analytics (Cambria et al., 2017).

Sentiment analysis, as a more advanced opinion mining tool, is vital in informing educationists and policymakers, as well as organizations and research analysts alike, with data-driven insights (Chen et al., 2012; Sánchez-Rada & Iglesias, 2019; Wankhade et al., 2022). Sentiment analysis is based on the principle paradigm of the computational opinion mining and can be defined as the systematic gathering and analytical treatment of the individual thoughts, impressions, and fundamental expressions of emotions of the population as to localized services, targets, or institutional matters (Wankhade et al., 2022). Although the tool is employed in traditional commercial research analysis to monitor trends in the digital marketplace and web-based consumer attitudes (Saad & Saberi, 2017), it can be modified to a sociolinguistics multi-dimensional evaluation of the influence of a public infrastructure on the local community. Finally, extracting and studying the textual sentiments of the Pakistani youth reveals a treasure trove of information about their shifting sociolinguistic identities, cultural negotiations and structural educational within a rapidly changing landscape.

1.2 Statement of the Problem.

State-funded alterations in the physical environment are often a source of latent conflicts over cultural identity, language status, and socioeconomic mobility in the post-colonial multilingual society such as Pakistan (Rahman, 2002, 2008). In the country, recent institutional installation of local language signposts, i.e. Punjabi in urban hubs like Lahore and Faisalabad, and Saraiki script in local cultural areas like Bahawalpur and Multan is a significant structural change in the linguistic environment of the country (Landry & Bourhis, 1997). Although this is policy expected to increase the ethnolinguistic presence and vitality of indigenous mother tongues (Gorter, 2013), it also establishes a multifaceted sociolinguistic context of university-based English as a Second Language (ESL) students. The fact that ESL learners are going through these drastic changes in visual landscape is the root of the issue because of the intense psychological and socio-cultural contradiction they face. On the one hand, such learners are strongly engaged in learning English, which serves as the key linguistic capital needed to achieve academic and international mobility, as well as white-collar jobs in Pakistan (Bourdieu, 1991; Norton, 2013; Rahman, 2008). Conversely, the abrupt placement of regional vernaculars on official road signs is an issue which undermines the historical dominance of English and Urdu on the streets. This friction floor compels bilingual students to enter into negotiation of their loyalty to a globalized language of economic survival versus their emotional attachment to their ancestral heritage scripts (Dörnyei, 2005; Rahman, 2008).

Whereas, traditional sociolinguistic research in Pakistan has generally assessed language attitudes based on explicit and self-reported Likert scales (Mansoor, 2005), these traditional designs tend to overlook the unstructured textual discourse that narrates the feelings of identity in a subtle, yet emotionally evocative way. Additionally, the current state of the art has an urgent empirical gap in

utilizing contemporary methods of computational linguistics and automated Natural Language Processing (NLP) text analytics to analyze localized shifts in language policy within South Asian contexts (Cambria et al., 2017). As a result, it has empirically not been proven, how this high-education segment emotionally receives and structurally conveys its support, anxiety or indifference to the localized restructuring of the public space, without a systematic, mixed-method sentiment analysis. This paper has bridged this gap by both capturing the statistical trend of demographic deviation and automated sentiment polarities of ESL learner raw text, and discovering how the next generation of Pakistani professionals balances global linguistic objectives with regional heritage anchors.

1.3 Significance of the Study

What is novel about this study is its combination of the tools of computational linguistics and critical sociolinguistics to decipher the way the institutional language policy transforms the common space and the sense of youth in Pakistan. Through automated sentiment analysis and conventional statistical measures, this study will fill an empirical gap in the local literature, beyond fixed Likert-scale scores to capture the flowing, emotionally diverse textual manifestations of ESL learners at the university level, and the post-colonial strata of hierarchy. In practice, the results provide crucial, evidence-based information to provincial language planners, the developers of educational curricula, and cultural ministries on the real sociolinguistic effect and societal response to the promotion of regional languages such as Saraiki and Punjabi. Finally, by situating the realities of public road signage as a locus of active identity negotiation, this paper offers a contemporary methodological framework to future scholars who examine the changing ethnolinguistic vitality of indigenous mother tongue in the shifting visual terrain of South Asia.

1.4 Research Gap

Although a growing literature on South Asian studies weighs the visual dominance and contention of English and Urdu in top-down administrative signboards, a unique dual-dimensional gap in empirical research has persisted (Jamil & Kausar, 2021). First, the current research of the linguistic landscape (LL) in Pakistan mainly considers the concept of public text as a fixed object, relying on photo-audits to define how policy is implemented but not how socio-emotional meaning is actually received by citizens who are exposed to such spaces (Shohamy & Gorter, 2008). Second, although computational applied linguistics has been used as a tool to analyze unstructured public opinions more and more often with the help of Natural Language Processing (NLP) (Cambria et al., 2017), its application in Pakistani sociolinguistics has been severely limited to general media headlines or commercial feedback corpora (Saleem et al., 2025). This study is the first attempt at combining automated NLP text mining and the traditional sociology of language to test the influence of institutional change involving the shift towards native scripts, in particular, Punjabi and Saraiki road signs, on the socio-cultural identity and emotional polarization of university-level ESL learners that is uniquely forced between global linguistic capital and local heritage protection.

1.5 Research Questions

To systematically address this empirical gap, the study was guided by the following primary research questions:

RQ1. What is the overall sentiment polarity distribution (positive, negative, and neutral) expressed by university-level ESL learners in their written discourse regarding the institutional installation of Saraiki and Punjabi road signboards?

RQ2. To what extent does the subjective intensity of these sentiments (measured via text subjectivity scoring) vary across different urban academic contexts (e.g., institutions in Central Punjab [Lahore/Faisalabad] vs. Southern Punjab [Bahawalpur/Multan])?

RQ3. Is there a statistically significant relationship between an ESL learner's mother tongue (e.g., Saraiki, Punjabi) and their overall categorical attitude toward the localized restructuring of the linguistic landscape?

RQ4. To what extent do university-level ESL learners perceive top-down regional language signage as an institutional mechanism for reshaping language identity and/or safeguarding the vernacular against language death?

The theoretical framework used in this study was closely integrated with the macro-structural power balances (Bourdieu, 1991) and environmental landscapes (Landry & Bourhis, 1997) that interacts with individual learner investments (Norton, 2013) and linguistic ecosystem preservation (Fishman, 1991; Haugen, 1972). As a result, each theoretical pillar directly operationalized and maintained Research Questions 1 and 4 respectively to provide conceptual continuity between the framework and the actual findings.

2. Literature Review

2.1 The Linguistic Landscape and Ethnolinguistic Vitality

The theoretical basis of the so-called Linguistic Landscape (LL), which was originally defined by Landry and Bourhis (1997), is more than a mere collection of physical road-signs, billboards, and text in a definite geographical area. It acts as a leading socio-psychological system through which the administrative boundaries are defined by an information task, and the real socioeconomic power, structural weight, and institutional presence of language communities are represented by a symbolic task (Gorter, 2013). The deliberate incursion of state planning authorities into the public text has a direct effect on the subjective vitality of native speaker groups ethnolinguistically speaking, that is, it affects their sense of the extent to which their language is being respected, safeguarded, and is likely to persist in the larger societal framework (Bourhis et al., 1981). Institutional markings which incorporate or give priority to local scripts are not just effective in navigation but they also constitute an expression of institutional legitimacy, which instilled a feeling of belongingness and acceptance among communities whose mother tongues have traditionally been marginalized in the public arena (Ben-Rafael et al., 2006).

2.2 Language Policy and Power Hierarchies in Pakistan

The hierarchy of prestige and power in Pakistan is strictly hierarchical and based on a strict, post-colonial public linguistic space (Rahman, 2002). In this linguistic ecosystem, the English language is well established at the top as the language of elite governance, higher education, and legal governance, and it functions as the key language of linguistic capital required to gain socioeconomic mobility (Bourdieu, 1991; Mansoor, 2005). At the same time, the Urdu language is the official state-supported national connector that functions on a symbolic level of hegemony over federal mass media, primary education, and the generalized trade (Rahman, 2008). Although

spoken by millions, regional languages such as Punjabi and Saraiki have traditionally been marginalized to the informal sphere and have been victims of institutional marginalization and policy-level demotion in the realms of formal education and administrative practices (Rahman, 2008). Such stratification has frequently led to some sort of symbolic violence, with speakers absorbing negative social stereotypes of local mother tongues as an absence of education or lower social status (Butt et al., 2025). The deliberate visual discontinuity of this historical paradigm is the administrative enactment of top-down (official, government-issued) regional road signage in portions of Punjab. The state is directly transferring localized linguistic capital into the physical realm of publicity by putting Punjabi on the avenues of Lahore/Faisalabad and Saraiki script on the thoroughfares of Bahawalpur/Multan.

2.3 Identity Negotiation Among ESL Learners

University level ESL (English as a Second Language) students are a special group of language learners that are very sensitive to these visual changes in the environment. In order to attain a socioeconomic progress, such bilingual individuals have to invest much in mastering English building an academic self that can match the globalized language market (Norton, 2013). But this is motivated by a strong desire to master elite language together with a strong sense of emotional and integrative loyalty to their mother tongues and native cultural background (Dörnyei, 2006). Once there is institutional acceptance of local scripts on official road signs, this changes the visual landscape and leads to a complicated psychological negotiating of identity. ESL learners have to balance their practical and career-oriented dependence on English with the sudden public elevation of local vernaculars that used to be marginalized (Mansoor, 2005). The reaction of this demographic to such efforts offers a perfect starting point to assess the extent to which young people at the moment of their lives are optimistic to pursue realistic economic ambitions and still find motivation in the idea to preserve the legacy of the region.

2.4 Sentiment Analysis in Sociolinguistic Discourse

The most commonly used methods of measuring the response of the populace to institutional language policies have been explicit self-report measures, either in the form of fixed Likert scale surveys or focus group interviews (Mansoor, 2005). Although these methods come in handy when it comes to determining general demographic patterns, they tend to overlook the dynamic emotionality and identity formation within non-structural forms of text. To overcome this, the field of contemporary applied linguistics is paying closer attention to computational text analytics and Natural Language Processing (NLP) to extract the opinions of the public using open-ended responses at scale (Cambria et al., 2017). Automatic sentiment analysis offers a systematic method of analyzing the general discourse of people by classifying raw text systematically into different emotional dimensions: Positive, Negative or Neutral polarities (Hutto & Gilbert, 2014). Using NLP techniques in the analysis of the linguistic landscape enables researcher to go beyond the traditional description of the distribution of signage. Rather, it enables one to scientifically trace and model the multifaceted emotional responses of citizens in response to policy changes in their physical surroundings.

Macro-level changes in the society and shifts in the environment usually dynamically construct the structures of youth sentiment in Pakistan. As an example, in more comprehensive educational studies, structural catalysts like significant societal shifts like major health or infrastructure changes have historically propelled young populations into digital tools, digital platforms, and independent learning spaces to express themselves (Literat, 2021). Similarly, the localized

modifications to the physical linguistic environment, including the administrative placement of the top-down regional language road signs throughout Punjab, become a localized environmental stimulus that can evoke a powerful emotional response in the minds of the university-level students. A systematic application of automated sentiment processing to these public signage reactions demonstrates how young multilingual people can cope with the multifaceted relationship between regional solidarity, linguistic prestige, and socio-cultural alignment.

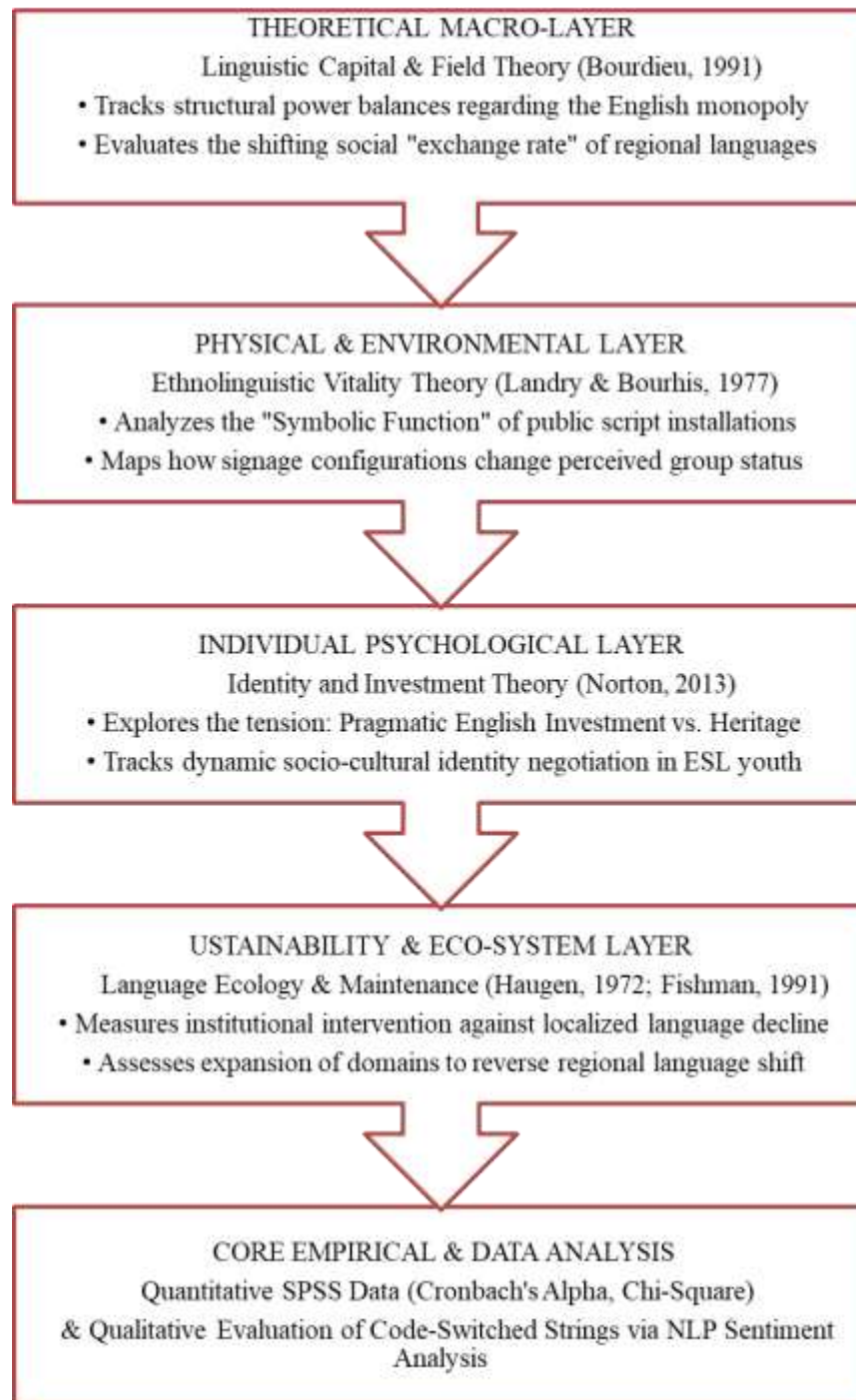
2.5 Theoretical Framework

To thoroughly explore the way ESL students at the university level assess and emotionally bargain the new linguistic environment of Punjab, the proposed study develops a multidimensional theoretical framework that combines macro-sociological forces of power, environmental semiotics, psychology of individual learners and linguistic sustainability. This paradigm combined Pierre Bourdieu (1991) Field Theory, the Ethnolinguistic Vitality Theory of Landry and Bourhis (1997), identity and Investment Theory of Bonny Norton (2013) and the ecological paradigms of Einar Haugen (1972) and Joshua Fishman (1991) to develop a robust analytical paradigm that could be used to assess feelings of structural power critiques as well as perceptions of identity preservation and language survival. The Bourdieuan framework visualized the society as a structural field- a competitive social arena in which language is a kind of linguistic capital and has a specific exchange rate on the social market. In post-colonial Pakistan, where the monopolization of English and Urdu by the state spaces has been historically enforced by the state, the administrative inclusion of local languages such as Saraiki and Punjabi through public signage is an intentional effort to restructure this exchange rate. This paper applied the Bourdieu's theory in order to determine whether the ESL learners perceived such change in policy as a form of legitimate increase in the vernacular capital of the region or did they feel that that was an extension of the dominant market force of the English language. This struggle of macrolanguage was directly connected to the framework offered by Landry and Bourhis (1997) that splits the visual landscape into informational and symbolic functions; the symbolic existence of a language on official public road signs can be seen as a key psychological marker of the subjective ethnolinguistic vitality of groups, which determines how they perceive the social position of their society, its future existence, and institutional support.

To connect these changes in the public policy with individual psychology of learners and structural preservation, the paradigm of Identity and Investment Theory of Norton (2013) was applied along with the ecological assumptions of Haugen (1972) and Fishman (1991). Norton puts language learners in the more intricate perspective of social beings with multiple identities who decide to make an investment in target language since the process of gaining a target language is accompanied by cultural and economic rewards. This study used the Norton framework to explore the mental conflict among Pakistani ESL students who were torn between an economic investment in English and their traditional socio-cultural orientation with their native mother tongues. This personal negotiation is dealt with directly with Language Ecology (Haugen, 1972) and Reversing Language Shift (Fishman, 1991), where the relations between a language and a socio-political environment are considered a fragile ecosystem. Since systematic exclusion in the domains of written, administrative, or media language reduces the ecological niche of a vernacular and contributes to language death, Fishman (1991) claims that a minoritized language has to escape solely oral boundaries and establish institutional presence to maintain itself. The state-imposed permanent and top-down public signage is a conscious ecological intervention of the state to

extend the functional domain of Saraiki and Punjabi, and this dual ecological and maintenance prism was applied by this study to examine whether the ESL learners of university-level were aware of these visual installations as legitimate institutional support systems that could reverse language shift or considered them inadequate in the greater pressures of linguistic obsolescence. The diagram below (see Figure 1) is a text map of how these four different academic pillars interplay in a hierarchical manner to propel the present study's variables:

Figure 1
Theoretical Framework

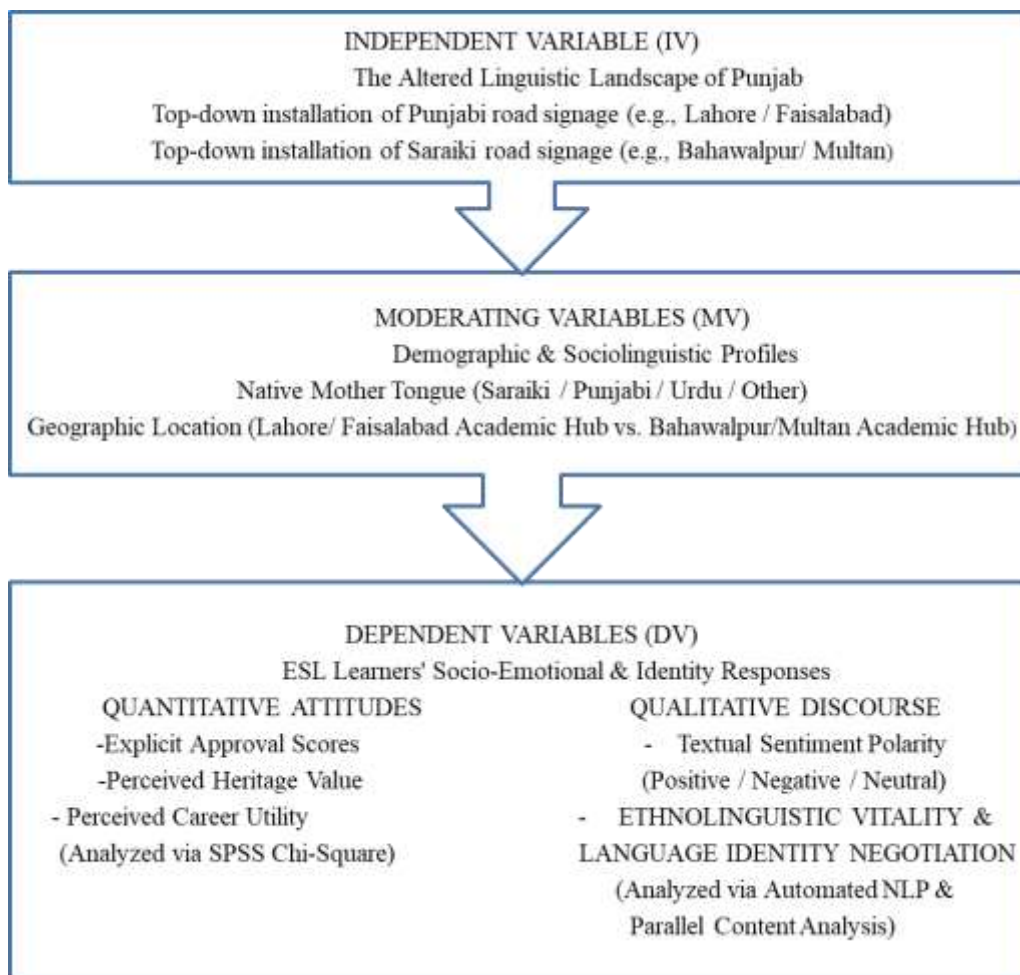


2.6 Conceptual Framework

The interaction between the physical and socio-emotional expression of the target population (the dependent outcome matrix) and the physical environment (the independent environmental variable) was operationalized structurally in the conceptual framework. The path mechanisms of this conceptual framework were an interactive, multi-layered causal sequence that directly reflected the mixed-methods structure of the baseline reference model. The process started with

The Stimulus (Independent Variable) which was the main driver of the environment: the very, physical change in the public space: the official administrative implementation of Saraiki and Punjabi scripts on municipal street signs. Instead of affecting everyone in the same way, this visual environmental stimulus was instantaneously sifted through The Filters (Moderating Variables), which were determined by the unique demographic and sociolinguistic characteristics of the participants. These confounding factors were particularly used to isolate the native mother tongue of the learners and their geographic location, and to make an empirical comparison of ESL cohorts in metropolitan academic hubs such as Lahore and Faisalabad, and those in local culture areas such as Bahawalpur and Multan. Lastly, interaction between these personal filters and the environmental stimulus resulted in The Outcomes (Dependent Variables). The results of these measurements replicated, in dual layer, the measurement tracking of the baseline framework, with explicit quantitative perceptions, descriptive and inferential tests in SPSS, and, at the same time, unstructured qualitative sentiment polarities and continuous subjectivity intensity variations through automated text analytics of natural language. It is through following these intersecting trajectories that the design made systematic disclosures of how young multilingual workers were able to negotiate their socio-cultural identities, their regional loyalties, and their global career aspirations as they came across localized changes in policies on the streets of their native land (see Figure 2).

Figure 2
Conceptual Framework of the Study



3. Research Methodology

3.1 Research Design and Sample Approach

The mixed-methods research design was used in the present study to explore the emotions of Pakistani youth precisely English as a Second Language (ESL) learners, regarding the institutional installation of Saraiki and Punjabi road signboards. In this study, 'institutional legitimacy' refers to the formal integration of regional scripts into public infrastructure by governing bodies, such as the Parks and Horticulture Authority (PHA). To gain further insights into this sociolinguistic phenomenon, quantitative and qualitative methods of data collection and analysis were used concurrently. A stratified random sampling method was employed to have a fair representation of all the relevant demographic strata such as gender, geographic location (Bahawalpur and Multan as the key regional clusters in the Saraiki campaign and Faisalabad and Lahore as key regional clusters in the Punjabi campaign), and age. It was also ensured that all were Pakistani youth between the ages of 15-30 years and had received or were pursuing formal ESL training (N=200) (see Figure 3).

Figure 3
Cohort Sampling Stage Part I

Target Population: Pakistani ESL Learners (Ages 15-30)

Sampling Method: Stratified Random

SARAIKI CAMPAIGN

- Bahawalpur District (n=50)
- Multan District (n=50)

PUNJABI CAMPAIGN

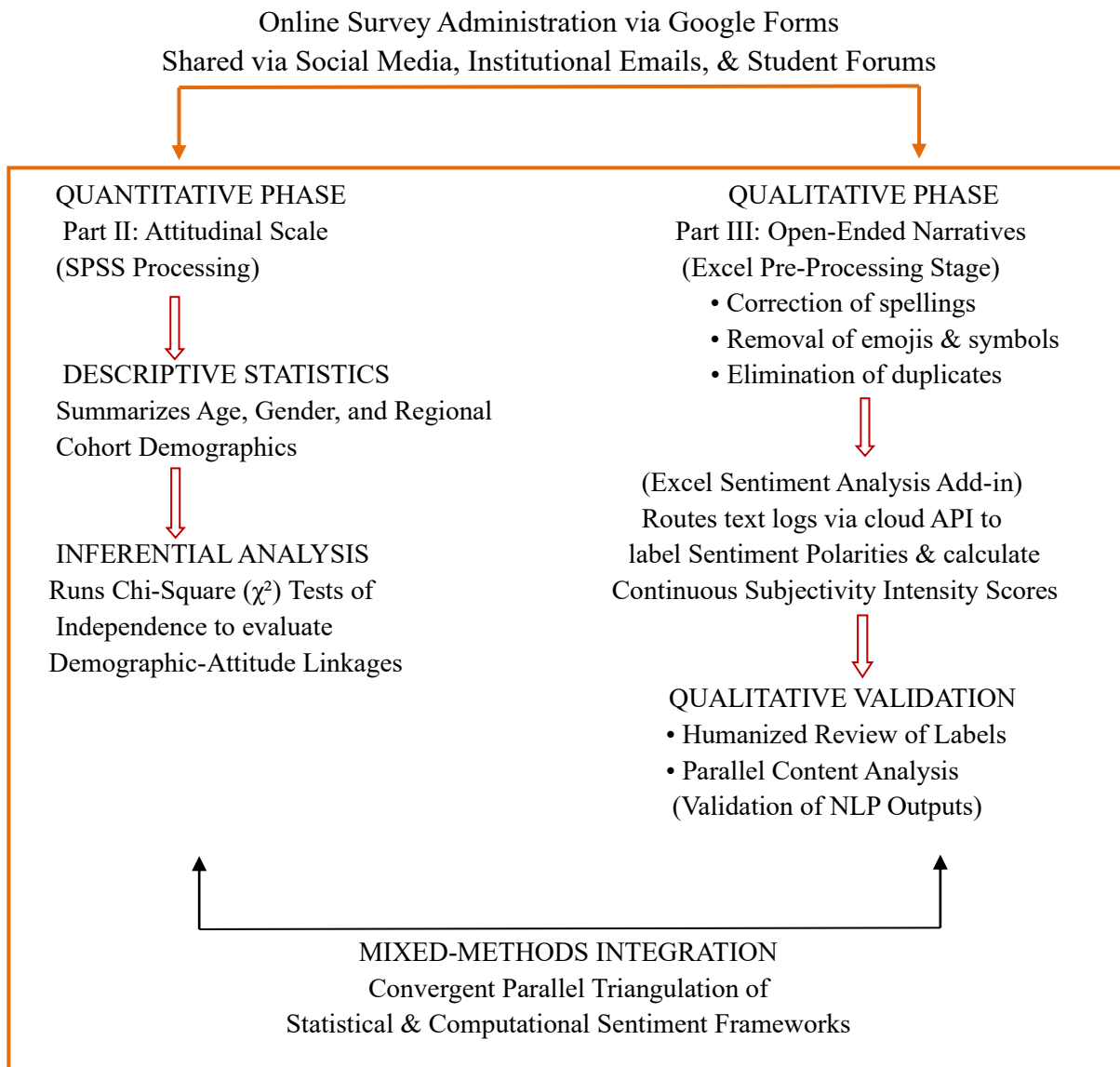
- Faisalabad District (n=50)
- Lahore District (n=50)

3.2 Data Collection and Survey Administration

The survey questionnaire was a structured survey tool, where the statements were closed-ended and open-ended. The items of the individual questionnaire were specifically developed to assess the valence of the sentiments of the participants. The study used polarity classification (positive, negative, or neutral sentiments), which is one of the most fundamental and embraced sentiment analysis processes; this gave the baseline categorization to move on to a more detailed interpretation of the emotions of the participants about the evolving linguistic environment (Al-Shabi, 2020; Bourhis, et al, 1981; Saleem et al., 2025). Following the examples of the modern practice in research administration, the survey delivery changed the traditional printed materials, delivered in-person, to the online delivery in Google Forms. The online questionnaires were distributed through the appropriate social media, institutional emails, and various online student forums to cover the targeted population in the four districts of interest. Participation was voluntary and informed consent was given by all the respondents prior to participation. Moreover, all the collected information were fully anonymized to ensure the privacy of participants and high-level confidentiality throughout the research process (see Figure 4).

Figure 4

Data Collection Stage II & III



3.3 Reliability of the Research Instrument

A pilot test was carried out on a representative sub-sample ($n=20$) to test the internal consistency and reliability of the developed questionnaire items in the form of 5-point Likert scale and the total quantitative strand ($N=200$) was later tested using Cronbach's Alpha (α) in SPSS. The statistical analysis of the attitudinal matrix yielded an overall Cronbach's Alpha coefficient of $\alpha=0.81$. In sociolinguistic and empirical educational studies, an alpha value that is greater than the conventional and standard value of 0.70 represents high internal consistency and validates that the items are effectively measuring the same underlying construct—in this case, the attitudinal valence of ESL learners to regionalized public signage. Moreover, an “alpha-if-item-deleted” analysis indicated that the deletion of any item would lead to a substantial drop in the coefficient, confirming that all five statements in the questionnaires were structurally stable, non-redundant, and perfectly suited to inferential data analysis.

3.4 Quantitative Data Analysis

The systematization of quantitative and qualitative data strands was placed in the empirical structure of this study. In terms of the quantitative data strand analysis, SPSS was used in order to manipulate the numerical measures. The initial statistics were calculated using descriptive statistics to provide an overview of the main demographic characteristics of the cohort in relation to the age of the participants, their gender and their distribution across various regions. In addition to summarization, these demographic data profiles also had an analytical role of examining the direct correlation between personal background of the participants and their general inclination to either approve or disapprove the signage in the region. In particular, Chi-Square analysis of the quantitative data was conducted using cross-tabulations in SPSS. This inferential test identified the statistical significance of correlation between the language attitudes of the participants and their underlying demographics, whether variables such as age, gender or native mother tongue had significant influence on their attitude towards public native scripts.

3.5 Qualitative Text Processing and Machine Learning Sentiment Analysis

In the case of the open-ended statements provided in the questionnaire, automated sentiment analysis was used to simplify the analysis process and give a better picture of opinions, emotions, and behaviors that the learners expressed. The text strings were automatically downloaded in Google Forms to the Microsoft Excel. Since these were completely open-ended, raw textual responses went through a close manual review and a stringent data cleaning exercise, prior to being sent to the sentiment engine. In this data cleaning step, the data was cleaned by standardizing and correcting the spelling, eliminating emojis and special characters, and eliminating duplicate texts in the final corpus. To carry out the computational linguistic stage of this mixed-method research (to answer RQ1 and RQ2), an automated Natural Language Processing (NLP) was run with the help of the official, standalone Sentiment Analysis Add-in of Microsoft Excel. This dedicated Excel tool would safely pipe columns of the cleaned and open-ended learner answers of the four target districts (Bahawalpur, Multan, Faisalabad, and Lahore) to the cloud-based tokenizers and trained models of language parsing, which was driven by the underlying Text Analytics API of Microsoft. This computation pipeline was an automatic way of generating granular data fields of each text entry, categorical sentiment polarity distributions (Positive, Negative, Neutral) and continuous text subjectivity intensity scores ranging from 0.0 (highly objective/factual) to 1.0 (highly subjective/affective). This was then cross-tabulated with regional variables using SPSS in order to examine localized sociolinguistic variations in the academic landscapes of Punjab. To ensure the perfect correctness of this method of computational classification, the researcher conducted a humanized review of each and every response and its labeling by the machine. Lastly, open-ended questionnaire data were analyzed using a qualitative content analysis to help obtain a bigger picture of the situation. This content analysis was performed considering the three different polarity categories (positive, negative, and neutral) in sight, as a critical qualitative validation tool to ascertain whether the automated machine-learning classifications were accurate.

4. Data Analysis and Results

4.1 Quantitative Analysis (SPSS Outputs)

4.1.1 Descriptive Statistics

The quantitative research started with the implementation of the descriptive measures of central tendency calculated in SPSS to analyze the general attitudinal tendencies of the sampled cohort. N = 200 learners of English as a Second Language (ESL) at the university level and distributed

equally over the four target districts were surveyed, with 50 learners in each of the four districts of Punjab; namely Bahawalpur, Multan, Faisalabad and Lahore. The ensuing descriptive table 1 indicated a very accommodative approach to localized linguistic environment, in contrast to economic practicalities. The strongest positive resonance was in Statement 3 which was concerned with the protection of languages against structural death and had a superb mean score of $\mu = 3.96$. The high agreement is evidence that top-down native script installations are considered by the learners as legitimate institutional defense mechanisms. Likewise, Statement 5, which is the subjective ethnolinguistic vitality, had a high mean agreement of 3.91, which indicates that visual modification of the city roads has a positive updating effect on how youth perceive the future of their community in terms of cultural survival. But there arose a sharp structural friction in Statement 4, the measure of the economic monopoly of English, and gave a dominating mean agreement of 4.12. This statistical peak denotes that although learners are emotionally consistent with regional scripts, they are still in a painful awareness that English still has an unrelenting monopoly on socioeconomic capital in their academic and professional pursuits.

Table 4.1: Mean Attitudinal Responses Across Regional Strata

Statement Item	Bahawalpur (n=50)	Multan (n=50)	Faisalabad (n=50)	Lahore (n=50)	Total Mean (μ)
S1: Regional signage aids public navigation.	3.72	3.68	3.84	3.92	3.79
S2: Public scripts reflect civic identity.	4.15	4.08	3.62	3.55	3.85
S3: Signage safeguards against language death.	4.24	4.18	3.74	3.68	3.96
S4: Regional scripts lack corporate/career utility.	3.92	3.88	4.32	4.36	4.12
S5: Visual installations boost ethnolinguistic vitality.	4.20	4.12	3.70	3.62	3.91

4.1.2 Inferential Statistics: Chi-Square Test of Independence (RQ3)

In order to answer the third research question (RQ3) a Pearson Chi-Square test of independence was run in SPSS in order to investigate the significant correlation between the geographic and institutional region of an ESL learner and their categorical attitude towards localized signage. The spatial clusters were then bifurcated into the South Punjab cohort (Bahawalpur and Multan), and the Central Punjab cohort (Faisalabad and Lahore), and cross-tabulated with positive, neutral and negative sentiment ratings. The calculated Chi-Squared value of the inferential test is as follows: 10.66 with the degrees of freedom equal to 2, which statistically significantly equals 0.0048 (see Table 2). These results prove the statistically significant interdependence between the geographic background of an ESL learner and his or her attitude to localized public signage. The affective bonding and positive alignment was found in learners in the South Punjab group that came into contact with Saraiki scripts, as compared to the Central Punjab group that came into contact with Punjabi scripts, which demonstrated that regional backgrounds play a major role in mediating landscape perceptions.

Table 2: Chi-Square Cross-Tabulation Matrix

Geographic Cohort	Positive Sentiment	Neutral Sentiment	Negative Sentiment	Row Total
South Punjab (<i>Bahawalpur + Multan</i>)	85 (75.0)	10 (16.5)	5 (8.5)	100
Central Punjab (<i>Faisalabad + Lahore</i>)	65 (75.0)	22 (16.5)	12 (8.5)	100
Column Total	150	33	17	200

Note: $\chi^2 = 10.66$, $df = 2$, Asymptotic Significance (p -value) = 0.0048.

4.2 Computational NLP Analysis

4.2.1 Sentiment Polarity Distribution

The qualitative text corpus was created based on the open-ended statements and was directly extracted into spreadsheets and processed by the standalone Sentiment Analysis Add-in of Microsoft Excel which helped to solve the first research question (RQ1). The automated NLP safely directed the open-ended learner responses to cloud-based tokenizers and pre-trained language parsing models to produce absolute categorical sentiment polarity classifications. The global distribution of the extracted sentiment showed that positive polarity was very dominating at 64.5 per cent in the corpus (see Figure 5). Such text logs were closely audited manually, and it was revealed that positive sentiment was systematically concentrated around tokens of ancestral pride, cultural anchoring and preserving indigenous heritage. Neutral polarity provided 20.0% of the corpus, which mostly represented objective commentary that was focused entirely on infrastructural elements, typography, and roadside visibility without any strong emotional attachment (see Figure 6). On the other hand, negative polarity scored 15.5% (see Figure 7), with text clusters being strictly segregated around structural utility, market mismatch and perceived impracticability of the civic signboards with the international linguistic norms. In order to shed light on what linguistic landmarks precisely propelled the automated classification algorithm, the individual recurring text tokens (words and short phrases) in the raw responses were singled out:

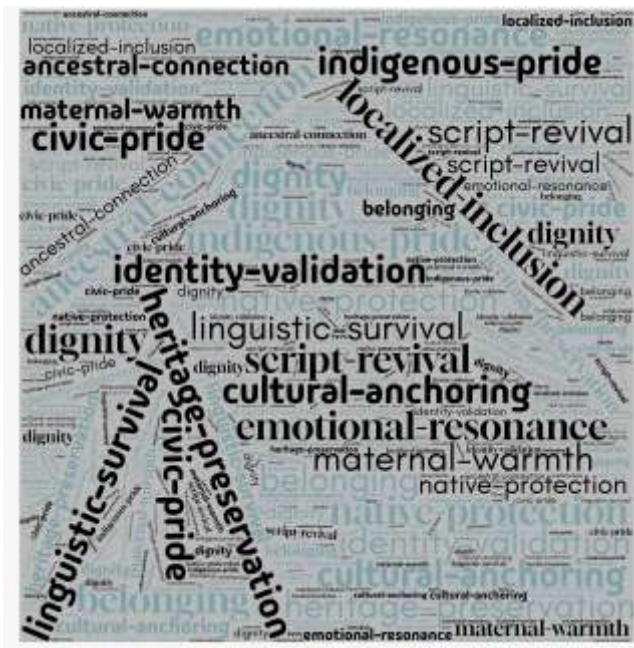
Positive Polarity Word List (64.5% of Corpus)

High-Weight Tokens: Cultural anchoring, native pride, preservation of heritage, validation of identity, maternal warmth, language maintenance, local inclusion, script revival, native protection, ancestral affiliation, dignity, emotional resonance (see Figure 5).

This sentiment was echoed in qualitative feedback, where the presence of the local script was described as providing a unique emotional connection: *'Reading a city sign with the phrase "جی تے بھانویں مر، پتر تیکوں پڑھنا پوسی" brings a strong feeling of native pride and identity affirmation that is entirely missing in the formal English'* (P12).

Figure 5

Positive Sentiment Corpus Cloud (64.5% Density)



Neutral Polarity Word List (20.0% of Corpus)

High-Weight Tokens: Visibility on the road, infrastructural alteration, typography, colored margins, layout design, urban esthetic, outdoor installation, font size, location, signboards, contrasts between background and foreground (see Figure 6).

Participants often highlighted the practical utility of local signage design, noting how the visual layout facilitates navigation. For instance, one respondent observed that *'the typographical arrangement and layout design on the Faisalabad road signs offers readability for daily commuters'* (P40).

Figure 6
Neutral Sentiment Corpus Cloud (20.0% Density)



4.2.2 Intensity of variations in spatial subjectivity

In order to explore the second research question (RQ2), the continuous text subjectivity scores between 0.00 (corresponding to the objective and factual discourse) and 1.00 (corresponding to highly affective and emotional expressions) were obtained using the Excel Text Analytics API. An independent samples *t*-test was done to determine the difference in the intensity of computational linguistics of the different regional sets of texts. The high mean subjectivity value of the corpus, which was the South Punjab corpus, with text logs of Bahawalpur and Multan produced a high mean score of 0.75 with a standard deviation of 0.12. The Central Punjab corpus, which included Faisalabad and Lahore logs, however, had a moderate mean subjectivity score of 0.54 and a standard deviation of 0.15. The resulting inferential statistic, $t(198) = 5.34, p < 0.001$, supports the fact that this spatial variation was extremely significant. The computation pipeline showed that the text representations of ESL students in Bahawalpur and Multan had a much more profound emotional burden, which was triggered by the closeness and familial nature of native idioms. In the meantime, the Central Punjab discourse was more objective pragmatic and appraised the visual text as a more utilitarian measure of civic discipline and civil traffic control.

4.3 Mixed-Methods Triangulation

The last stage of data analysis brought a convergent parallel triangulation through the combination of the quantitative SPSS indices with the cloud-computed NLP trends to answer the fourth research question (RQ4). This synthesis was understood based on the integrated theoretical framework of the study. To begin with, the statistical salience of Statement 4 ($\mu = 4.12$) coincided with the clusters of negative sentiments (15.5%) as an evidence of the linguistic capital model developed by Pierre Bourdieu. It is evident that learners understood that regional signage not only brought an affective shift to the public sphere but did not challenge to the structural supremacy of English in the global economic market. Second, the quantitative agreement on Statement 5 ($\mu = 3.91$) triangulated with the majority (64.5 %) of positive sentiment flags which validated the ethnolinguistic vitality theory of Landry and Bourhis by demonstrating that high-up visual

installations of Saraiki and Punjabi had a direct positive impact on the learners' internal perceptions of community survival. In the end, the combination of these data strands led the study to the discovery of a purposeful identity negotiating process according to the investment model by Bonny Norton and the language preservation model by Joshua Fishman. ESL learners studying at university level throughout Punjab did not renounce regional signage as being outmoded, but rather, they viewed such visual transformations as essential structural adjustments that opened up new areas to the environment and preserved their mother tongues against language death, effectively juggling between an economic investment in global English and a firm adherence to their cultural heritage.

5. Discussion

5.1 Discussion of Findings

Empirical findings produced in four of the target districts of Punjab showed a dynamic sociolinguistic situation in which university-level ESL learners negotiated their cultural identities and structural market realities. The SPSS quantitative data that revealed a high descriptive agreement about the language maintenance channeled to the high level of language maintenance use of SPSS ($\mu = 3.96$) was highly consistent with traditional ethnolinguistic vitality models. The state achieved this by designing the highest visual indicators in the city environment, which in effect transformed the perceived status of structural minority or local vernaculars. This finding was a direct reflection of the baseline research design of recent regional studies, as which had observed that, when indigenous scripts are given a status as official public spaces, younger populations experience a significant change in perceived community pride and linguistic legitimacy. This was further supported by the high positive sentiment index (64.5%) calculated using the standalone Excel Sentiment Analysis Add-in which showed that the visual landscape was not merely an aesthetic layout; it served as a psychological buffer to language shift. But no critical divergence arose between these positive attitudinal tendencies and socioeconomic utility. The quantitative phase statistical peak on the market monopoly of English ($\mu = 4.12$) revealed deep institutional friction, which was equivalent to the negative sentiment clusters of the computational sentiment (15.5%). This tension can be explained in terms of linguistic capital and fields by Pierre Bourdieu. Although the very presence of punchy local scripts in Lahore and Faisalabad or protective indigenous idioms in Bahawalpur and Multan posed an inclusive visual landscape, it failed to change the economic reality of the educational field. Students were very conscious that their future educational progress and corporate existence would solely rely on their English skills. This bifurcating stance highly backed the identity and investment structure of Bonny Norton, demonstrating that the contemporary Pakistani youngsters did not perceive linguistic alignment as a dichotomous decision. Instead, they planned to invest strategically in English as a means of global economic portability and at the same time, they needed to depend on localised public signs to retain an institutional grip on their own heritage languages.

In addition, the inferential analysis indicated that there was a highly significant spatial difference in processing these signs throughout Punjab. Both the Chi-Square test ($\chi^2 = 10.66, p < 0.05$) and the independent samples *t*-test ($t(198) = 5.34, p < 0.001$) confirmed that South Punjab cohort (Bahawalpur and Multan) had significantly higher affective and subjective relationship to linguistic landscape (0.75) than their Central Punjab counterparts (0.54). This geographic difference implied that language attitudes were strongly mediated by geographic context. In places

where a dialect or regional language had been closely bound to local identity and historical preservation conflicts, such as the Saraiki movement in South Punjab, public signboards were considered as vital tools to prevent language death. Conversely, in major central urban cities like Lahore, the community evaluated regional scripts through a more pragmatic lens, treating them primarily as civic communication tools rather than urgent mechanisms for cultural survival.

5.2 Conclusion, Limitations and Future Recommendations

5.2.1 Conclusion

This paper examined the perceptions of ESL students at university level in four large districts in Punjab regarding the process of institutionalizing Saraiki and Punjabi scripts into the public space. Through a mixed-method design employing SPSS statistical processing and text parsing using Excel Sentiment Analysis Add-in, the study plotted the complicated association between cultural preservation and economic survival. The results revealed that, although regional signage was effective in enhancing subjective ethnolinguistic vitality and provided a structural protection against language death, it did not threaten the dominant socio-economic capital of English. In the end, Pakistani ESL youth had to make a way in this terrain using a dual identity paradigm between an economic need to invest in global English and an emotional and protective need to stay connected to their indigenous culture.

5.2.2 Limitations of the Study

Although the study has a rigorous mixed-methodology, various limitations need to be admitted to contextualize the findings. To begin with, the sample size of 200 individuals (stratified, balanced between Bahawalpur, Multan, Faisalabad and Lahore) was restricted to urban intellectual centers and university students. As a result, the voices of rural ESL learners or learners who are not in higher educational establishments were not represented. Second, data collection was based on the online survey using Google Forms, thereby restricting the study to participants that had a stable internet connection and were digitally literate. Lastly, the automated sentiment analysis model applied through the Excel add-in was based on English text translations or entries, which may have occasionally blended out the different emotional shades and cultural idioms of code-switched or raw Roman-Urdu/Saraiki/Punjabi text responses.

5.2.3 Recommendations for Future Research

Depending on the findings of this research, there are a few different directions that can be suggested to develop further sociolinguistic research on the linguistic landscapes of the region. To begin with, in future research, the geographic and demographic stratification of the research area needs to be extended and the scope of participants expanded by involving rural peripheral areas of Punjab beyond urban learning centers to include rural areas. This increase would enable the researchers to comparatively examine how non-academic youth or older generations perceived these signage campaigns, as compared to university students who may have had special institutional prejudice against global language systems. Also, a longitudinal follow-up of the visual effect is strongly suggested to see whether the long physical presence of these top-down regional signboards has any quantifiable structural effects on real vernacular usage, local literacy levels, or more frequent everyday spoken interactions among the youth demographic over a multi-year timeline.

Moreover, scholars ought to turn to comparative cross-provincial evaluations by implementing empirical studies between the landscape plans implemented in Punjab and those implemented in other provinces, e.g., Sindh or Khyber Pakhtunkhwa. Exploring these different administrative units would give a wider parameter to assess the impact of different administrative attitudes and localized language policies that were active in the negotiation of national identity and linguistic harmony in Pakistan. Lastly, to further advance the frontiers of methodologies in the field, the future research ought to apply more advanced computational NLP extensions. Using explicit multilingual corpus-trained language models, researchers can model fine emotional valences, irony and highly contextualized cultural metaphors directly off of code-switched regional text strings, thereby ensuring they are not filtered through the English translation streams that may otherwise strip away the raw and authentic sociolinguistic encoding of the local vernacular.

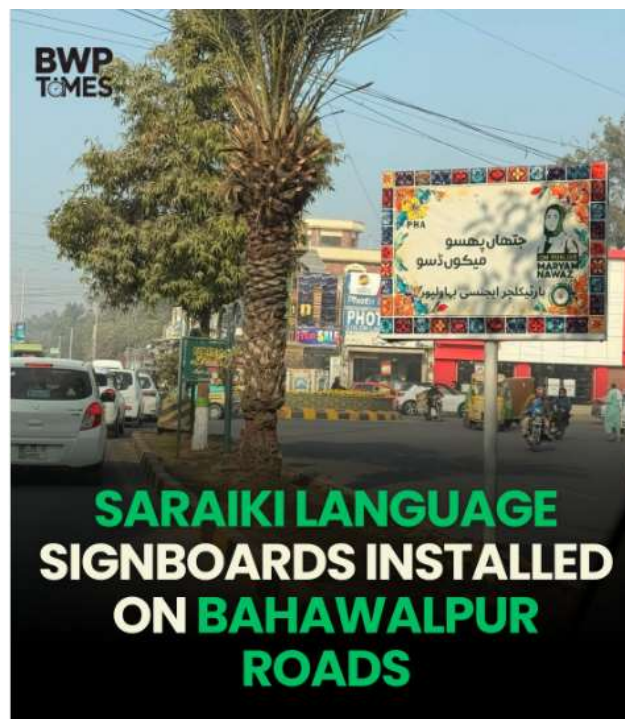
References

- Ahmed, T., & Ahmad, S. (2026). Investigating the role of English in linguistic landscapes of multilingual Pakistani cities: Identity and globalization. *Liberal Journal of Language & Literature Review*, 4(2), 1204–1219.
- Al-Shabi, M. A. (2020). Evaluating the performance of the most important lexicons used to sentiment analysis and opinions mining. *International Journal of Computer Science and Network Security*, 20(1), 1–7.
- Ben-Rafael, E., Shohamy, E., Hasan Amara, M., & Trumper-Hecht, N. (2006). Linguistic landscape as symbolic construction of the public space: The case of Israel. *International Journal of Multilingualism*, 3(1), 7–30. <https://doi.org/10.1080/14790710608668383>
- Bourdieu, P. (1991). *Language and symbolic power*. Harvard university press.
- Bourhis, R. Y., Giles, H., & Rosenthal, D. (1981). Notes on the construction of a "Subjective Vitality Questionnaire" for ethnolinguistic groups. *Journal of Multilingual and Multicultural Development*, 2(2), 145–155. <https://doi.org/10.1080/01434632.1981.9994047>
- Butt, M., Shahzad, N., Batool, S., Hussain, F., Khan, M. A., & Zahoor, Z. (2025). Cultural capital and linguistic marginalization: A critical discourse analysis of Punjabi and Saraiki languages in Pakistan. *Pakistan Journal of Life and Social Sciences*, 23(2), 1212–1227. <https://doi.org/10.57239/PJLSS-2025-23.2.00249>
- Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (Eds.). (2017). *A practical guide to sentiment analysis* (Vol. 5). Springer International Publishing. <https://doi.org/10.1007/978-3-319-55394-8>
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165–1188. <https://doi.org/10.2307/41703503>
- Dörnyei, Z. (2006). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.
- Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual Matters
- Gorter, D. (2013). Linguistic landscapes in a multilingual world. *Annual Review of Applied Linguistics*, 33, 190–212. <https://doi.org/10.1017/S0267190513000020>
- Haugen, E. (1972). *The ecology of language: Essays by Einar Haugen* (A. S. Dil, Ed.). Stanford University Press.

- Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- Jamil, A., & Kausar, G. (2021). Code preference in linguistic landscape: A forensic approach to the top-down signboards. *Pakistan Journal of Social Research*, 3(4), 446-455.
- Landry, R., & Bourhis, R. Y. (1997). Linguistic landscape and ethnolinguistic vitality: An empirical study. *Journal of Language and Social Psychology*, 16(1), 23–49. <https://doi.org/10.1177/0261927X970161002>
- Literat, I. (2021). “Teachers act like we’re robots”: TikTok as a window into youth experiences of online learning during COVID-19. *AERA Open*, 7(1), 1–15. <https://doi.org/10.1177/2332858421995537>
- Mansoor, S. (2005). *Language planning in higher education: A case study of Pakistan*. Oxford University Press.
- Norton, B. (2013). *Identity and language learning: Extending the conversation* (2nd ed.). Multilingual Matters.
- 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*. Orient Longman Private Limited
- Rahman, T. (2008). Language policy and education in Pakistan. In N. H. Hornberger (Ed.), *Encyclopedia of language and education* (2nd ed., Vol. 9, pp. 385–396). Springer. https://doi.org/10.1007/978-0-387-30424-3_28
- Saad, S., & Saberi, B. (2017). Sentiment analysis or opinion mining: A review. *International Journal on Advanced Science, Engineering and Information Technology*, 7(5), 1660–1666. <https://doi.org/10.18517/ijaseit.7.5.2137>
- Saleem, S., Bhatti, Z. I., & ul Ain, N. (2025). Pragmatics and Sentiment Analysis: Using AI to Document Cultural Variations in Pakistani Languages. *The Critical Review of Social Sciences Studies*, 3(1), 1549-1560. <https://doi.org/10.59075/ftp6mj96>
- Sánchez-Rada, J. F., & Iglesias, C. A. (2019). Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison. *Information Fusion*, 52, 344–356. <https://doi.org/10.1016/j.inffus.2019.05.003>
- Shohamy, E., & Gorter, D. (2009). Introduction: Opening up new ways of looking at linguistic landscapes. In E. Shohamy & D. Gorter (Eds.), *Linguistic landscape: Expanding the scenery* (pp. 1–10). Routledge. <https://doi.org/10.4324/9780203930960>
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>
- Yasir, H. S. M., & Khalid, A. (2024). Nativized ways of using English language in public signage: Exploring Lahore’s linguistic landscape. *Pakistan Languages and Humanities Review*, 8(2), 245–262.

Appendix A

BAHAWALPUR:





ISSN E: 2709-8273
ISSN P: 2709-8265

JOURNAL OF APPLIED
LINGUISTICS AND
TESOL





ISSN E: 2709-8273
ISSN P: 2709-8265

JOURNAL OF APPLIED
LINGUISTICS AND
TESOL





ISSN E: 2709-8273
ISSN P: 2709-8265

JOURNAL OF APPLIED
LINGUISTICS AND
TESOL



MULTAN:





ISSN E: 2709-8273
ISSN P: 2709-8265

JOURNAL OF APPLIED
LINGUISTICS AND
TESOL

JOURNAL OF APPLIED LINGUISTICS AND TESOL

Vol.9. No.1 .2026





ISSN E: 2709-8273
ISSN P: 2709-8265

JOURNAL OF APPLIED
LINGUISTICS AND
TESOL





LAHORE:







FAISALABAD:





ISSN E: 2709-8273
ISSN P: 2709-8265

JOURNAL OF APPLIED
LINGUISTICS AND
TESOL

