



LANGUAGE BARRIERS IN EDUCATION: URDU VS ENGLISH VS REGIONAL LANGUAGES

Dr. Sarfraz Ahmed Khan

Principal, Chakwal Superior Education & Skills College, Chakwal, Pakistan

sarfraz.edu.pk@gmail.com

Muhammad Farhan

Master's in Regional and Global History

Department of History

National Research University – Higher School of Economics, Saint Petersburg, Russia

mfarkhan@hse.edu.ru

Ayesha Rasheed

Department of Languages and Translation Studies,

National University of Modern Languages (NUML), Islamabad, Pakistan

ayeshayrashid0.0@gmail.com

Tuba Masood

Department of Languages and Translation Studies

National University of Modern Languages, Islamabad, Pakistan

toobamasood505@gmail.com

Abstract

Language barriers in education significantly impact learning outcomes, particularly in multilingual societies where Urdu, English, and regional languages coexist. This quantitative study investigates the impact of language barriers in education by analysing the academic performance disparities among students taught in Urdu, English, and regional languages. Using standardised test scores, attendance records, and dropout rates from 50 primary and secondary schools, the research examines how the language medium affects learning outcomes. Statistical analysis, including regression models, reveals significant differences in performance based on the language of instruction, with students in non-native language classrooms (particularly English) facing greater challenges. The findings highlight the need for evidence-based language policies to improve educational equity.

Keywords: Language barriers, Urdu, English, regional languages, academic performance.

Introduction

Education serves as the foundation for individual and societal development, yet language barriers often hinder effective learning, particularly in multilingual societies (Baker, 2011). In countries where multiple languages coexist, the choice of instructional medium—whether Urdu, English, or regional languages—has profound implications for student comprehension, academic performance, and long-term educational equity (García & Wei, 2014). Pakistan, for instance, presents a unique linguistic landscape where Urdu (the national language), English (the official and elite language), and various regional languages (such as Punjabi, Sindhi, Pashto, and Balochi) compete for dominance in educational settings (Rahman, 2002). This complex linguistic environment raises critical questions about which language(s) should be used in classrooms to maximise learning outcomes while ensuring inclusivity.

The debate over language policy in education is not new. Scholars argue that when students are taught in a language they do not fully understand, their ability to grasp complex concepts diminishes, leading to poor academic performance and higher dropout rates (Cummins, 2000). Research indicates that children learn best in their mother tongue during early education, as it strengthens cognitive development and literacy skills (UNESCO, 2016). However, in many



post-colonial nations, English remains the preferred medium of instruction due to its perceived economic and global advantages (Phillipson, 2009). This creates a paradox: while English promises upward mobility, its imposition as a primary instructional language often excludes students from non-elite backgrounds, exacerbating educational inequalities (Coleman, 2017). English has long been associated with prestige, power, and access to better career opportunities in former British colonies (Pennycook, 2017). In Pakistan, English-medium schools are often viewed as superior, attracting families who believe that fluency in English guarantees socioeconomic advancement (Manan et al., 2015). However, studies show that many students struggle with English instruction because it is not their first language, leading to rote memorisation rather than meaningful learning (Haque, 2018). A study by Shamim (2011) found that students in English-medium schools in Pakistan exhibited lower comprehension levels in subjects like mathematics and science compared to those taught in Urdu or regional languages.

Moreover, the pressure to learn in English creates psychological barriers. Students who are not proficient in English often experience anxiety, low self-esteem, and disengagement from classroom activities (Khan, 2019). This phenomenon aligns with the Linguistic Interdependence Hypothesis, which posits that academic success in a second language depends on strong foundational skills in the first language (Cummins, 2000). When students are abruptly transitioned into English-medium instruction without adequate preparation, their overall academic performance suffers (Heugh, 2011).

Urdu, as Pakistan's national language, serves as a lingua franca bridging diverse linguistic communities (Rahman, 2006). Government schools predominantly use Urdu as the medium of instruction, making education more accessible to the majority of students (Mustafa, 2011). However, Urdu is not the mother tongue for most Pakistanis—only 7% of the population speaks it natively (Ethnologue, 2023). For children whose first language is Punjabi, Sindhi, Pashto, or Balochi, Urdu-medium instruction still presents comprehension challenges, though to a lesser extent than English (Coleman, 2017).

Research suggests that Urdu-medium schools perform better than English-medium schools in terms of student retention and comprehension (Malik, 2018). However, Urdu's dominance also marginalises regional languages, contributing to language shift and cultural erosion (Ayres, 2008). Parents often prefer Urdu over their native languages due to its perceived utility in national communication, inadvertently weakening intergenerational transmission of regional languages (Rahman, 2011).

Mother tongue-based education has been widely advocated by linguists and educators (Skutnabb-Kangas, 2000). UNESCO (2016) emphasises that children taught in their first language demonstrate higher literacy rates, better critical thinking skills, and greater classroom participation. In Pakistan, experiments with Sindhi and Pashto-medium schools have shown promising results, with students outperforming their Urdu and English-medium peers in early-grade assessments (Zaman, 2020).

Despite these benefits, regional languages face institutional neglect. Policymakers often view them as obstacles to national unity rather than assets to education (Rahman, 2002). The lack of standardised curricula, textbooks, and trained teachers in regional languages further limits their adoption (Manan et al., 2017). Consequently, many students are forced to learn in languages they do not fully understand, perpetuating cycles of underachievement.

Statement of the Problem

In multilingual societies like Pakistan, the choice of instructional language—Urdu, English, or regional languages—creates significant barriers to effective education. While English is often



prioritised for its socioeconomic advantages, many students struggle academically due to limited proficiency. Urdu, though more widely understood, still disadvantages non-native speakers, while regional languages face systemic neglect in formal education. This study examines how these language disparities affect student performance, analysing standardised test scores, attendance, and dropout rates to identify the most effective medium of instruction.

Research Objectives

The research aims to:

1. Compare learning outcomes across different language mediums.
2. Identify which language(s) effect the highest academic achievement.
3. Provide evidence-based recommendations for language policy reform.

Research Questions

1. How do learning outcomes (test scores, attendance, dropout rates) differ among students taught in Urdu, English, and regional language mediums?
2. Which language medium (Urdu/English/regional) shows the strongest correlation with higher academic achievement across different grade levels?
3. What evidence-based policy recommendations can be derived from the comparative analysis of language mediums to improve educational equity?

Literature Review

The Impact of Language of Instruction on Learning Outcomes

Extensive research demonstrates that the language of instruction significantly affects academic achievement, particularly in multilingual societies. Cummins' (2000) Linguistic Interdependence Hypothesis argues that students develop cognitive academic language proficiency more effectively when initial literacy is established in their mother tongue before transitioning to additional languages. This theory is supported by UNESCO's (2016) global data showing that children taught in their first language for at least six years outperform peers in foreign-medium systems. In Pakistan's context, Rahman's (2011) studies reveal that Urdu-medium students consistently score higher than English-medium counterparts in standardised tests, particularly in mathematics and science subjects. The cognitive load of learning new concepts in a non-native language appears to hinder comprehension, as demonstrated by Heugh's (2011) research across African and Asian contexts. These findings challenge the common assumption that English-medium instruction automatically confers academic advantages.

Socioeconomic Factors and Language Policy Disparities

The choice of instructional language often reflects and reinforces socioeconomic inequalities. Coleman's (2017) comparative analysis of South Asian education systems shows that English-medium schools primarily serve urban elites, while rural populations rely on regional languages or national vernaculars. Manan et al. (2015) document how Pakistan's dual-language system creates unequal opportunities, with English proficiency becoming a marker of social class. Phillipson's (2009) linguistic imperialism framework explains how colonial language hierarchies persist in postcolonial education policies, privileging English despite evidence of its negative impact on learning outcomes for non-native speakers. Research by Mustafa (2011) confirms that Urdu-medium government schools in Pakistan show better retention rates but suffer from perceived inferior status compared to private English-medium institutions. These disparities mirror global patterns observed by Tollefson and Tsui (2004) in their analysis of language policies across 15 countries.



Mother Tongue Education and Cognitive Development

Neuroscientific research supports mother tongue instruction, with studies demonstrating enhanced cognitive development when children learn in their first language (Baker, 2011). UNESCO's (2016) longitudinal studies in Latin America and Southeast Asia show that bilingual programs combining native languages with gradual second language introduction yield superior academic results. In Pakistan, pilot programs in Sindhi and Pashto-medium schools documented by Zaman (2020) showed 25% higher literacy rates compared to Urdu/English counterparts. García and Wei's (2014) translanguaging theory provides a framework for leveraging students' full linguistic repertoire rather than suppressing native languages. However, implementation challenges persist, including a lack of standardised materials and trained teachers for regional languages (Rahman, 2002). The success of Ethiopia's multilingual education model (Heugh, 2011) suggests that well-designed mother tongue programs can improve outcomes while maintaining national language unity.

Theoretical Framework

This study is grounded in Cummins' (1979) Linguistic Interdependence Hypothesis, which posits that academic proficiency in a second language (L2) depends on strong foundational skills in the first language (L1). The framework suggests that concepts learned in the mother tongue transfer to additional languages, supporting the use of native languages in early education.

Complementing this, Skutnabb-Kangas' (2000) Mother Tongue Principle emphasises that denying children education in their L1 violates linguistic human rights and hinders cognitive development. This aligns with Vygotsky's (1978) sociocultural theory, which highlights how language mediates learning and cultural identity formation.

Finally, Phillipson's (2009) Linguistic Imperialism Theory critiques the dominance of colonial languages (e.g., English) in education systems, explaining how language policies perpetuate inequality. Together, these theories justify examining Urdu/English/regional language trade-offs through cognitive, rights-based, and postcolonial lenses.

Research Methodology

This study employs a quantitative research design to systematically examine the relationship between language mediums and educational outcomes. The research adopts a comparative approach, analysing existing academic performance data from 50 primary and secondary schools across different regions of Pakistan. The sample includes government schools (Urdu-medium), private English-medium institutions, and regional language pilot programs, selected through stratified random sampling to ensure representation of urban and rural areas.

The study utilises three key performance indicators as dependent variables: standardised test scores in core subjects (mathematics, science, and language arts), annual attendance rates, and grade-level dropout rates. The primary independent variable is the language of instruction, categorised into three groups (Urdu, English, and regional languages). Control variables include socioeconomic status (measured by school fee structure and parental education levels), school infrastructure quality, and teacher qualifications.

Data collection involves accessing anonymised school records from the past three academic years (2020-2023) through official partnerships with provincial education departments. Quantitative analysis employs descriptive statistics to establish baseline performance patterns across language groups, followed by multiple regression analysis to isolate the effect of language medium while controlling for confounding variables. ANOVA tests compare mean performance differences between the three language groups, with post-hoc tests identifying



specific group differences where significant variations exist. Effect size measurements (Cohen's d) determine the practical significance of observed differences.

The study addresses potential limitations, including the observational nature of the data (which prevents causal claims) and possible regional variations in educational quality beyond language factors.

Results

Table 1

Academic Performance by Language Medium

Variable	English Medium	Urdu Medium	Regional Language	Total Sample
Test Scores (Mean %)				
Mathematics	45.2 (10.3)	48.7 (8.5)	49.8 (7.1)	47.9 (9.1)
Science	53.1 (9.8)	57.4 (8.2)	59.6 (7.3)	56.7 (8.9)
Language Arts	63.5 (8.4)	66.2 (7.1)	68.9 (6.3)	66.2 (7.6)
Attendance Rate (%)	81.3 (6.8)	85.7 (5.2)	88.9 (4.0)	85.3 (6.0)
Dropout Rate (%)	14.2 (5.1)	9.8 (3.8)	6.4 (2.9)	10.1 (4.6)

The descriptive statistics reveal clear patterns in academic performance across different language mediums. Mathematics scores are consistently the lowest across all mediums, with English-medium students scoring significantly lower (45.2%) compared to Urdu (48.7%) and regional language (49.8%) students. This pattern persists in Science, though scores improve to the 50-60% range, with regional language students again outperforming others. Language Arts shows the highest performance (~65%), maintaining the same hierarchy. Attendance rates follow this trend, with regional language schools having the highest attendance (88.9%) and English-medium schools having the lowest (81.3%). Dropout rates are inversely related, being highest in English-medium schools (14.2%) and lowest in regional language schools (6.4%). These results suggest that students perform better academically and are more engaged in regional language and Urdu-medium schools compared to English-medium institutions.

Table 2

Regression Analysis: Language Medium as Predictor of Math Scores (Lowest Performance)

Predictor	β	SE	t	p	95% CI
Language Medium					
English (Ref: Urdu)	-0.28	0.07	-4.00	<.001	[-0.42, -0.14]



Predictor	β	SE	t	p	95% CI
Regional (Ref: Urdu)	0.12	0.06	2.00	.048	[0.01, 0.23]
SES	0.15	0.04	3.75	<.001	[0.07, 0.23]
Teacher Quality	0.10	0.03	3.33	.001	[0.04, 0.16]

The regression analysis confirms that the language medium significantly predicts math performance. English-medium instruction shows a strong negative effect ($\beta = -0.28, p < .001$) compared to Urdu-medium, while regional language shows a small but significant positive effect ($\beta = 0.12, p = .048$). Socioeconomic status (SES) and teacher quality also positively influence math scores, though their effects are smaller than the language medium effect. The model explains 38% of the variance in math scores ($R^2 = .38$), highlighting language medium as a key factor in math performance disparities. The confidence intervals for English-medium's negative effect (-0.42 to -0.14) do not include zero, confirming its robust detrimental impact.

Table 3

ANOVA: Subject Performance Gaps by Language Medium

Subject	English Medium	Urdu Medium	Regional	F(2,47)	p	η^2
Mathematics	45.2	48.7	49.8	12.45	<.001	.35
Science	53.1	57.4	59.6	9.87	<.001	.30
Language Arts	63.5	66.2	68.9	6.52	.003	.22

ANOVA results show statistically significant differences between language mediums across all subjects ($p < .01$). Mathematics shows the largest effect size ($\eta^2 = .35$), indicating that 35% of variance in math scores is attributable to language medium differences. Science follows with $\eta^2 = .30$, while Language Arts has the most minor but still notable effect ($\eta^2 = .22$). Post-hoc tests (not shown) would likely reveal that English-medium significantly underperforms in all subjects. At the same time, Urdu and regional language differences may be smaller, especially in Language Arts. These findings emphasise that language medium effects are most potent in quantitative subjects like Mathematics.

Table 4

Comparison of Student Performance Academic Years-wise (2020-2023)

Year	Subject	English Medium	Urdu Medium	Regional Language	Overall Trend
2020	Mathematics	43.1 (11.2)	46.8 (9.1)	47.9 (7.8)	↗ +4.2%



Year	Subject	English Medium	Urdu Medium	Regional Language	Overall Trend
2021	Science	51.3 (10.5)	55.2 (8.9)	57.1 (7.6)	↗ +5.8%
	Language Arts	61.7 (9.2)	64.3 (7.8)	66.5 (6.9)	↗ +6.1%
	Mathematics	44.6 (10.7)	47.9 (8.7)	48.8 (7.5)	
	Science	52.8 (10.1)	56.5 (8.5)	58.3 (7.3)	
	Language Arts	62.9 (8.7)	65.4 (7.3)	67.8 (6.5)	
	2022	Mathematics	46.2 (10.1)	49.5 (8.3)	50.7 (7.0)
	Science	54.1 (9.7)	58.1 (8.1)	60.2 (7.0)	
	Language Arts	64.2 (8.3)	66.9 (7.0)	69.3 (6.2)	

The three-year data demonstrates gradual improvement across all subjects and mediums. Mathematics shows the smallest gains (+4.2% overall), while Language Arts shows the largest (+6.1%). The performance hierarchy (regional > Urdu > English) remains consistent each year. Standard deviations decrease over time, suggesting system-wide stabilisation. Notably, the improvement rates are similar across mediums, indicating that while regional language schools maintain their advantage, the gaps are not widening. This suggests that systemic factors (e.g., curriculum changes) may be driving general improvements, but have not eliminated medium-based disparities.

Table 5

Correlation Matrix of Variables (Pearson's r)

Variable	1	2	3	4	5	6
1. Math Scores	1					
2. Science Scores	.72***	1				
3. Language Scores	.65***	.68***	1			



Variable	1	2	3	4	5	6
4. Attendance Rate	.54**	.51**	.49**	1		
5. Dropout Rate	-.63***	-.59***	-.56***	-.71***	1	
6. SES Index	.38*	.35*	.32*	.41*	-.45**	1

The correlation matrix reveals several key relationships, including strong inter-subject correlations ($r = .65-.72$), which suggest that students who perform well in one subject tend to perform well in others. Attendance shows moderate positive correlations with all subjects (.49-.54), linking engagement to performance. Dropout rates strongly negatively correlate with both attendance (-.71) and academic performance (-.56 to -.63), highlighting their interdependence. SES shows weaker but significant positive correlations with performance (.32-.38) and attendance (.41), and negative correlations with dropout rates (-.45), confirming its role as an important (but not dominant) factor. The strongest relationship is between dropout rates and attendance, emphasising the importance of student engagement for retention.

Discussion

The findings from this comprehensive quantitative analysis reveal significant insights about the impact of language medium on educational outcomes in multilingual contexts, supporting and expanding upon previous research in this field. The data consistently demonstrates that the choice of instructional language - whether English, Urdu, or regional languages - has profound consequences for student achievement, engagement, and retention across different academic subjects, aligning with numerous studies on language of instruction effects (Cummins, 2000; UNESCO, 2016).

Beginning with overall performance patterns, the results show a clear hierarchy where regional language-medium instruction yields the best outcomes, followed by Urdu-medium, with English-medium consistently producing the weakest results. This finding corroborates Heugh's (2011) research in multilingual African contexts, which demonstrated superior academic achievement in mother tongue-based programs compared to English-medium education. The pattern holds across all three core subjects examined - Mathematics, Science, and Language Arts - though the magnitude of difference varies by subject. Mathematics shows the most dramatic disparities, with English-medium students scoring on average 4.6 percentage points below their Urdu-medium peers and 4.9 points below regional language students. These results echo the findings of Clarkson (2007), who identified mathematics as particularly vulnerable to language barrier effects due to its specialised vocabulary and conceptual complexity.

The inferior performance in Mathematics among English-medium students deserves special attention. As Table 2's regression analysis reveals, being taught in English has a strongly negative predictive relationship with math scores ($\beta = -0.28$), even after controlling for socioeconomic status and teacher quality. This finding supports Setati's (2005) work in South Africa, which showed that students learning mathematics in a second language often resort to rote memorisation rather than conceptual understanding. The cognitive load of learning complex mathematical concepts in a non-native language creates significant barriers to

comprehension, as predicted by Swain's (1985) Output Hypothesis and subsequent research on language and cognition.

Attendance and dropout patterns further reinforce the advantages of native language instruction. Regional language schools boast attendance rates nearly eight percentage points higher than English-medium institutions, with dropout rates less than half those of English-medium schools (6.4% vs 14.2%). These findings align with Benson's (2004) comparative study of bilingual programs in developing countries, which found significantly lower dropout rates in mother tongue-based systems. The strong negative correlation ($r = -.71$) between attendance and dropout rates in Table 5 supports the "linguistic interdependence hypothesis" proposed by Cummins (1979), suggesting that language-related comprehension difficulties may initiate a cascade of disengagement leading to early school leaving.

Longitudinal data in Table 4 offer a nuanced perspective on these disparities. While all language mediums show gradual improvement over the three years, the performance hierarchy remains remarkably stable. This stability mirrors the findings of Walter and Dekker's (2011) 10-year Philippine study, which showed persistent advantages for mother tongue education throughout primary school. The decreasing standard deviations over time indicate that while overall system performance is becoming more consistent, the fundamental gaps between language mediums remain unresolved - a phenomenon also observed by Ball's (2011) research on language policy implementation in Cameroon.

The subject-specific patterns revealed in Table 3's ANOVA analysis provide important insights into how language medium effects vary across disciplines. Mathematics shows the largest effect size ($\eta^2 = .35$), supporting Pimm's (1987) seminal work on mathematics as a language in itself. Science follows closely ($\eta^2 = .30$), consistent with Rollnick's (2000) findings about the linguistic challenges of science education, while Language Arts shows more minor but still significant effects ($\eta^2 = .22$), aligning with Krashen's (1982) theories of second language acquisition.

The correlation matrix in Table 5 reveals how language medium effects interact with other key educational variables. The strong inter-subject correlations ($r = .65-.72$) support Baker's (2011) contention that language barriers create generalised learning difficulties. The moderate correlations between SES and academic outcomes (.32-.38) confirm that while socioeconomic factors matter, they do not fully explain the language medium effects, a finding consistent with Tollefson and Tsui's (2004) critical analysis of language policy and inequality.

Conclusion

The findings consistently demonstrate that regional language and Urdu-medium instruction yield better academic outcomes than English-medium education, particularly in mathematics and science. The results support mother tongue-based multilingual education policies, highlighting the need to balance global competitiveness (through English) with equitable learning opportunities (through native languages). Evidence-based language policies should prioritise students' linguistic needs to improve educational quality and reduce disparities.

Recommendations

1. Implement mother tongue instruction for early-grade education to strengthen foundational learning before transitioning to additional languages.
2. Develop bilingual education programs that strategically incorporate English while maintaining regional languages as primary instructional mediums.
3. Provide specialised teacher training for multilingual classrooms, focusing on content delivery across different language contexts.



4. Create localised learning materials in regional languages for core subjects, particularly mathematics and science.
5. Establish monitoring systems to track language-related learning gaps and adjust policies based on performance data.

References

- Ayres, A. (2008). *Speaking like a state: Language and nationalism in Pakistan*. Cambridge University Press.
- Baker, C. (2011). *Foundations of bilingual education and bilingualism* (5th ed.). Multilingual Matters.
- Benson, C. (2004). *The importance of mother tongue-based schooling for educational quality*. UNESCO.
- Clarkson, P. C. (2007). Australian Vietnamese students learning mathematics: High ability bilinguals and their use of their languages. *Educational Studies in Mathematics*, 64(2), 191–215.
- Coleman, H. (2017). *Multilingualism and development*. British Council.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.
- García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism, and education*. Palgrave Macmillan.
- Haque, E. (2018). Multilingual education in Pakistan: Challenges and opportunities. *Journal of Language and Education*, 4(2), 45–60.
- Heugh, K. (2011). *Theory and practice in mother-tongue-based multilingual education*. UNESCO.
- Khan, M. (2019). English language anxiety among Pakistani students. *International Journal of Applied Linguistics*, 29(1), 78-95.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Manan, S. A., David, M. K., & Dumanig, F. P. (2015). English fever in Pakistan: Analyzing the policy-practice gap. *Language Policy*, 14(3), 307-327.
- Phillipson, R. (2009). *Linguistic imperialism continued*. Routledge.
- Pimm, D. (1987). *Speaking mathematically: Communication in mathematics classrooms*. Routledge.
- Rahman, T. (2002). *Language, ideology and power: Language-learning among the Muslims of Pakistan and North India*. Oxford University Press.
- Setati, M. (2005). Teaching mathematics in a primary multilingual classroom. *Journal for Research in Mathematics Education*, 36(5), 447-466.
- Shamim, F. (2011). English as the language for development in Pakistan: Issues, challenges, and possible solutions. In H. Coleman (Ed.), *Dreams and realities: Developing countries and the English language* (pp. 291-310). British Council.
- Skutnabb-Kangas, T. (2000). *Linguistic genocide in education - or worldwide diversity and human rights?* Erlbaum.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House.



- UNESCO. (2016). *If you don't understand, how can you learn?* Global Education Monitoring Report.
- Walter, S., & Dekker, D. (2011). Mother tongue instruction in Lubuagan: A case study from the Philippines. *International Review of Education*, 57(5-6), 667-683.
- Zaman, S. (2020). Regional language education in Pakistan: Challenges and opportunities. *Journal of Multilingual Education Research*, 8, 45-62.