The Dual Faces of Pashto: Analyzing Diglossia and Vernacular Diversity Using Ferguson's Framework and the Strategies to Explore the Regional Varieties

Qasim Ullah[1*], Shams Ur Rehman[2], Saqib Raza[3]

Abstract

This study provides a comprehensive analysis of Pashto language varieties through the framework of Ferguson's diglossia model, focusing on Educated Standard Pashto, Regional Standard Pashto, and Prestige Standard Pashto. The research examines how these variants reflect and diverge from classical diglossic criteria. Educated Standard Pashto, which developed in the early 20th century, is characterized by its formal usage in education and media. While it exhibits features of a High form, such as a rich lexicon of paired terms and limited literacy, it does not fully meet Ferguson's criteria due to the absence of a religious text and minimal grammatical differences from Low Pashto. Despite its prestige in formal contexts, its impact is somewhat mitigated by the low literacy rate in Afghanistan, with many illiterate speakers acquiring aspects of this variant informally through media exposure. Regional Standard Pashto, used as a common dialect among Pashtuns from various tribes, emerges in cities like Kandahar, Jalalabad, and Kabul. Each of these regional standards reflects local tribal influences and adaptations, with Kandahar and Jalalabad Pashto being shaped by their respective tribal dialects. In contrast, Kabul's Pashto incorporates elements from multiple regional varieties and is notably influenced by Dari, the predominant language in the capital. This regional variation facilitates communication among Pashto speakers from diverse backgrounds. Prestige Standard Pashto, primarily represented by Southern Pashto, holds significant cultural and historical status due to the Durrani Confederacy's role in Afghan history. Although its prestige is largely symbolic, reflecting Pashtun ethnic unity rather than a formalized linguistic tradition, it plays a crucial role in the standardization efforts of Educated Standard Pashto. The study shows Pashto partly aligns with Ferguson's diglossia model, mainly through lexical differences and limited literacy, but lacks in areas like significant grammatical divergence and a central religious text. It highlights the complexities of Pashto's regional variations and the media's role in linking High and Low Pashto.

Keywords: Standard Variety, Diglossia, Ferguson's Framework, Vernaculars

1. Introduction

Pashto is spoken by approximately 40 million individuals residing on both sides of the border between Pakistan and Afghanistan, known as the Durand Line—a boundary established in 1893 by an agreement between Afghanistan and British India. This line delineated the southern limits of Afghanistan and partitioned Pashtun lands between Afghanistan and British India, now Pakistan since the 1947 Partition. Pashto is the language of the Pashtuns, who were instrumental in founding the Afghan state in 1747; in British India, they were referred to as Pathans, an anglicized plural form of Pashtuns. In Afghanistan, Pashto is a primary language, alongside Dari or Afghan Persian, and is spoken by about 60% of the population, particularly in the areas south of the Hindu Kush mountain range. There are also significant Pashto-speaking communities in northern and northwestern Afghanistan, where Pashtuns were relocated and settled in the late 19th century. In Pakistan, Pashto is the regional language of around 20–25 million people, with the majority residing in the North-West Frontier Province (NWFP), Baluchistan, or the Federally Administered Tribal Areas (FATA).

The Pashto language, spoken by the Pashtuns, is often referred to by its speakers as Pushto. This contrasts with the names used by others: Persians call them Afghani, while in Urdu and Hindi literature, they are known as Paṭhani and Pathans, respectively. In Afghanistan, Pashto is one of the two official languages, with about 60% of the Pashtun majority speaking it, while Dari is the other official language. In Pakistan, an estimated 20-25 million people speak Pashto. Additionally, there are Pashto-speaking communities in Tajikistan, and among diaspora populations in countries such as the UK, Canada, Germany, and the Gulf states (Hussain & Khan, 2023). According to Rahman's 2006 study, Pashto speakers make up 15.42% of Pakistan's total population, with 73% of the population in Khyber Pakhtunkhwa speaking it as their first language (Rahman, 1995). The city of Karachi, Pakistan's southern economic hub, also has a significant number of Pashto speakers (Habibullah & Barbara, 1996). Some individuals in Iran also speak Pashto as their primary language (Rahman, 1996).

Regarding dialectal differences within Pashto, research has shown that while phonological variations are present, other differences are not substantial enough to classify Pashto into distinct dialects (MacKenzie, 1959). Most Pashto dialects are mutually intelligible among the Pashtuns, but the Yousafzai dialect stands out for its widespread use among poets, novelists, and media professionals. This dialect is not only prevalent in Khyber Pakhtunkhwa but also holds significance for Pashto speakers and readers internationally due to its literary and rhetorical prominence.

[1*] MS Scholar, Department of English, Kohat University of Science & Technology. qasimkhab1@gmail.com
[2] M.Phil. Scholar, Department of English, Kohat University of Science & Technology
[3] Lecturer in English, Govt Degree College, Darra Adam Khel, TSD Kohat

From a strictly genetic perspective, Pashto, an Indo-European language, is classified within the northeastern group of Iranian languages (Coyle, 2014). The morphosyntactic structure of Pashto remains relatively consistent across its various dialects. The primary basis for classification lies in phonological features, specifically the pronunciation of the /x ˘/ and /gˇ/ sounds. These consonants exhibit regional variations, forming a prominent isoglossic line that is easily noticeable in the written script.

1.1. Phonemes in Yousafzai Pashto

The Pashto language comprises forty-one phonemes: twenty-nine consonants, eight vowels, and two diphthongs (Olson, 1996). To fully grasp the Pashto vowel system, it is important to consider the influence of Arabic vowels. Pashto has incorporated Arabic vowels and adapted pure Arabic vowels into its system. Despite this integration, some features present in Arabic are not utilized in Pashto. Unlike English, which has distinct vowel phonemes with clear alphabetic and phonetic representations, Pashto encounters challenges with certain sounds that do not have unique written symbols. In English, phonetic characters and written forms align closely, but Pashto lacks this direct correspondence, complicating the representation of its phonemes.

1.2. Pashto Speakers and the Importance of English Pronunciation

In Pakistan, English is the primary medium for official communication and is widely used in various sectors including education, administration, defense, diplomacy, banking, civil aviation, media, and IT. According to Rehman, understanding English phonology is crucial due to its stress-timed nature and the significance of intonation patterns in speech (Ghani, 2005). While Pakistani education focuses heavily on grammar and vocabulary from a basic level, phonology is often overlooked, even at the university level. This gap is also evident in Khyber Pakhtunkhwa's educational system, leading to difficulties for Pakistani students, especially Pashtun students, in comprehending native English speakers and pursuing higher education abroad. English holds a significant role in both official and academic contexts in this region, and interactions with English-speaking tourists are common in areas like Swat and Kashmir.

Rehman (2009) suggests that while achieving native-like pronunciation is not essential for foreign learners, intelligible pronunciation is necessary. Effective English language programs should integrate pronunciation instruction alongside vocabulary and grammatical structures, ensuring that learners can be understood by native speakers (Fraser, 2000).

In the eastern A zone (/maSreqi/), these sounds are pronounced as /g/ and /x/, respectively. For example, 'woman' is pronounced /khaza/ and 'beard' as /gira´/. Dialects in this region are termed 'hard,' reflected in the English transcription 'Pukhtu'. In contrast, the western C zone (/maWrebi/) features pronunciations of /Z/ and /§/, sometimes reduced to /S/ and /Z/ (e.g., Ghazni). Here, 'woman' is pronounced /§e Bja/ and 'beard' as /Zira´/. These are referred to as 'soft' dialects or 'Pushtu'. Despite these variations, both 'hard' and 'soft' dialects use the same script, allowing speakers to read with their pronunciation nuances. This script unity facilitates the formation of a standard Pashto that encompasses both A- and C-type dialects, whether 'hard' or 'soft,' from regions such as Kandahar or Peshawar. Additionally, another isogloss delineates a B zone, known as intermediary or central (/mandZaney/), which bridges the 'soft' and 'hard' Pashto varieties.

The pronunciation of certain consonants, like /xˇ/ or /S/ and /gˇ/ or /Z/, is notably distinct due to the unique vowel sounds in standard Pashto. This characteristic, called Waziri metaphony, is named after the Wazir tribes where it's commonly observed. It includes vowel changes such as /o/ replacing /a:/, /e/ replacing /o/, and /i/ replacing /u/. Transcribing this pronunciation would require a different script from traditional orthographic systems, which is why this variant of Pashto is not written. Consequently, speakers from this area use a different Pashto dialect when interacting with speakers from other regions (like zones A or C). This situation is a clear example of diglossia: they rely on a written version of their language that is more broadly recognized, though not necessarily more prestigious, since these dialects are important for cultural identity. Additionally, the Wanetsi dialect, found in Pakistani Baluchistan, represents another Pashto variant. This old form, which has been sparsely documented, is nearly unintelligible to other Pashtuns.

Diglossia is a relatively stable language situation in which, in addition to the primary dialects of the language which may include a standard (or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation (Fergusson, 1959). Ferguson identified four criterions found in all diglossic situations. These four criterions are:

- Perceived Superiority of High Form
- Sizeable Ancient Body of Literature
- Great Grammatical Differences between High and Low forms
- Lexicon has Paired Items between High and Low form
- Literacy Limited to a Small Elite

1.3. Morphology of Pashto

Pashto features intricate morphology compared to English. The language's morphemes are categorised into indeclinable morphemes and conjugatable morphemes. Indeclinable morphemes are fixed elements like particles, while conjugatable morphemes include nouns, pronouns, adjectives, and verbs.

Nouns Pashto nouns can change form to indicate number, case, and sometimes gender. For instance, the singular noun /hələk/ ('boy') becomes plural as /hələkan/ ('boys'). Nouns also have different cases, such as the direct case /hələkan/ ('boys') and the oblique case /hələkano/ ('boys'). Gender is sometimes indicated, with /ɻund/ ('a male blind person') being masculine and /ɻənda/ ('a female blind person') being feminine. The plural of masculine nouns often involves adding suffixes like /-ɑn/, while feminine plurals are formed by removing the suffix /-a/ and adding /-ee/. Compound nouns can be created by combining morphemes, such as /islam/ and /ɑbɑd/ to form /islamɑbɑd/ ('Islamabad').

Pronouns Pashto pronouns inflect for case, gender, number, and person. For example, /zə/ ('I') is the direct case, while /mɑ/ is the oblique case. Pronouns also vary by gender (/de/ for 'he' and /dɑ/ for 'she'), number (/de/ for singular 'he' and /dwəi/ for plural 'they'), and person (/zə/ for first person, /tə/ for second person, and /dɑ/ for third person).

Adjectives Adjectives in Pashto inflect for case, gender, and number. For instance, the adjective /ɻund/ ('male blind person') changes to /ɻandə/ for the plural direct case and /ɻəndo/ for the plural oblique case. Gender inflection differentiates between masculine and feminine forms. Adjectives can also express degrees through adverbs like /deer/ ('more'), /lɑ/ ('still'), and /ziɑt/ ('much'), which modify adjectives to reflect comparative and superlative degrees.

Verbs Pashto verbs typically end in /əl/ in their infinitive form, such as /wəhəl/ ('to beat'). Verbs inflect based on tense, mood, person, number, gender, voice, and aspect. Regular verbs are conjugated by removing /əl/ from the infinitive and adding appropriate suffixes. Irregular verbs, like /tləl/ ('to go'), have unique forms, such as /zəm/ ('I am going') for the present imperfective tense. Some infinitives are formed by combining nouns or adjectives with auxiliary verbs, like /ɣosa kəwəl/ ('to show anger') and /ɣosa keedəl/ ('to become angry').

1.4. Statement of the Problem

A standard form as referring to a prestige variety that cuts across regional differences, providing a unified means of communication (Crystal, 2011). Pashtu, a major language spoken in Afghanistan and Pakistan, showcases considerable linguistic diversity and complexity, notably through its diglossic nature and regional vernaculars. This diglossia involves a formal, standardized form used in official and written contexts alongside various regional vernaculars used in everyday conversation, with significant differences in grammar, vocabulary, and pronunciation between them. Although Pashtu's diglossic nature is widely recognized, a comprehensive analysis is needed to fully understand these dynamics and how regional varieties shape the language's overall landscape. Ferguson's diglossia framework offers a theoretical foundation but needs adaptation to suit Pashtu's unique sociolinguistic context. This research aims to explore how Ferguson's framework applies to Pashtu by examining the characteristics of its high (standardized) and low (vernacular) varieties, the sociolinguistic factors driving differences among regional vernaculars, and the methodologies for documenting and analyzing these varieties to enhance understanding of Pashtu's diglossic structure and regional diversity.

1.5. Significance of the Research

The significance of this study is to offer a comprehensive understanding of how Pashtu functions within its sociolinguistic context by examining diglossia, a phenomenon where distinct language varieties are used in different contexts. This study will utilize Ferguson's framework to elucidate how the high (formal) and low (informal) varieties of Pashtu are maintained, perceived, and utilized across various domains such as education, media, and everyday communication. Investigating Pashtu's vernacular diversity is also crucial for preserving the language's rich cultural and historical heritage, as it is spoken across various regions, each with its unique dialectal features. Documenting these regional varieties will aid in the development of educational resources, linguistic policies, and documentation efforts aimed at safeguarding Pashtu's diverse dialects.

The research findings will be instrumental in shaping language policies and educational strategies in Pashtu-speaking regions. By highlighting the dynamics of diglossia and vernacular diversity, the study will provide insights into the challenges and needs of Pashtu speakers, informing policymakers and educators about the importance of incorporating various Pashtu varieties into educational curricula and developing language policies that accommodate both high and low varieties. Furthermore, a detailed analysis of Pashtu's diglossia and vernacular diversity can enhance cross-cultural communication and mutual understanding among Pashtu speakers from different regions. Recognizing the linguistic and cultural nuances of different Pashtu varieties can foster a greater appreciation of the language's diversity and promote effective communication strategies in multicultural settings.

Applying Ferguson's framework to the study of Pashtu will also contribute to the theoretical understanding of diglossia and language variation. This application could refine Ferguson's framework and potentially lead to new theoretical insights and methodologies that apply to other languages and linguistic contexts. The research will lay the groundwork for future studies on Pashtu and similar languages experiencing diglossia and regional variation,

prompting further investigation into how diglossia affects language use and development, and how vernacular diversity shapes linguistic identity and communication practices. Additionally, this research can encourage comparative studies with other diglossic languages, enriching the broader field of sociolinguistics. Lastly, understanding the dual faces of Pashtu and its regional varieties can support community-driven initiatives aimed at promoting linguistic inclusivity and cultural pride, helping community organizations and activists advocate for the recognition and support of various Pashtu dialects and ensuring that all speakers have a voice in the preservation and promotion of their linguistic heritage.

1.6. Research Questions

• How does Ferguson's framework illuminate the functional distribution and interaction between Standard Pashtu and its regional vernaculars within the context of diglossia?

• What are the possible strategies to explore the regional varieties of the Pashtu language concerning with diglossic situation?

1.7. Research Objectives

• To analyze how Ferguson's framework defines the functional roles of Standard Pashtu and its regional vernaculars in a diglossic context.

• To develop strategies for enhancing the study and preservation of regional Pashtu varieties in a diglossic context.

2. Literature Review

Naheed (2015) also considers pronunciation crucial in distinguishing between the north-east and south-west dialects of Pashto, noting the significant influence of Urdu on the speech of Afghan migrants. This suggests possible interactions between Pashto and English phonemes.

Peter Roach (2009) asserts that English pronunciation should be learned in terms of phonemes rather than alphabetic letters. Ghani (2005) suggests that, generally, spelling systems mirror the pronunciation of a language, which is true for many languages. However, in English, spelling does not reliably guide learners to correct pronunciation due to the discrepancy between the 26 letters and the 44 distinct sounds—20 vowels and 24 consonants—found in Standard English (Received Pronunciation or RP). To accurately produce English speech sounds, one must understand English phonetics and phonology, focusing on the sounds themselves rather than their written representations, given the complex and often inconsistent nature of English spelling (Roach, 2009). Received Pronunciation is recommended as a model for non-native learners (Roach, 2009). Standard dictionaries in the UK use Received Pronunciation, which includes 24 consonant sounds categorized into plosives, fricatives, affricates, nasals, and semivowels, each varying by manner and place of articulation.

Tegey and Robson (1992, 1996) produced a series of comprehensive publications for foreign learners of Pashto, describing lexical diversity and phonetic variations in the standard dialect spoken by educated individuals in Kabul and eastern Afghanistan. While their main focus is on written skills, their materials also address pronunciation, using Roman script for transcription. They refer to Lorimer's 1914 work on Pashto to illustrate the fluidity in pronunciation and usage. Tegey and Robson advocate using a specific Pashto speaker's model to ensure intelligibility among Pashtuns, even if it may sound more urban and educated. Their "Pashto Conversation" manual emphasizes that learning Pashto's pronunciation is crucial.

Daniel Hallberg (1992) compiled data for Pashto research, with a primary aim to explore dialectal groupings and establish criteria for defining these groupings. Olson (1996) examined the vowels and consonants of East Afghan Pashto, noting considerable phonemic diversity across various dialects. Olson highlights that even within Pashto, individual pronunciation can vary significantly. He advises readers to rely on native pronunciation rather than written forms to avoid developing poor pronunciation habits, emphasizing the influence of Arabic on Pashto's phonetic features.

Herbert Penzl (1955) discusses Western loanwords in Pashto, noting that while many Western terms have been adopted, some pronunciations have been adapted to fit Pashto phonology. Morgenstern's extensive vocabulary list underscores the importance of pronunciation for semantic clarity and appropriate word usage, although it does not focus specifically on pronunciation (Georg M., 1927). The current study aims to address gaps in research on Pashto grammar, phonetics, and phonology, focusing particularly on phonetics and the overall phonology of Pashto.

3. Analysis and Discussion

3.1. Three Standard Pashto Varieties

3.1.1. Educated Standard Pashto

Pashto literacy can be traced back to the era of Ghaznawi Sultan Mahmood around 1000 AD (Hotak 2007, p. 24ff). Initially, Pashto was written in Arabic script, reflecting the widespread conversion of Pashtuns to Islam over the preceding 150 years. At that time, the Arabic script lacked distinct characters for Pashto's unique sounds, which were only added more than five centuries later. According to Hotak (2007, p. 25), Pir Roshan Bayazid is credited with introducing thirteen new letters to represent "difficult Afghan sounds," though his works have since been

lost. Khushal Khan Khattak, a prominent Pashto poet from the seventeenth century during the Mughal Empire, created his own version of the Arabic script. His poetry remains influential among Pashtuns today.

Efforts toward standardizing Pashto began in the twentieth century. In 1923, a literary society was established under Amir Amanullah Khan to develop Pashto orthography, resulting in an alphabet almost identical to the one used today, with only one exception. A more comprehensive literary society was formed in 1935, tasked with publishing educational materials, collecting vocabulary from various tribes, compiling dictionaries, creating a National Pashto Library, launching a literary magazine, and translating important works into Pashto. Various committees have since worked on standardizing grammar, and while some regional variations persist, Educated Standard or High Pashto has emerged over the past ninety years. According to a television interview with the Minister of Education (Najibullah, p.c.), there is a standardized curriculum and uniform textbooks are used across Pashtun-majority areas from grades one through twelve. This standardization has strengthened Pashtun unity and ethnic identity.

Educated Standard Pashto is not only taught in schools but is also broadcast on radio and television. International radio stations such as the BBC and Voice of America, as well as approximately five television stations based in Kabul, use Educated Standard Pashto. Many other stations broadcast partially in Educated Standard Pashto alongside Dari.

3.1.2. Evaluating Educated Standard Pashto Against Ferguson's Criteria

Table 5 assesses Educated Standard Pashto according to Ferguson's diglossia criteria. The criteria are listed in the first column, and the evaluation of Educated Standard Pashto against these criteria is provided in the second column. A checkmark indicates that the criterion is met, the Ø symbol denotes it is not met, and a ½ mark shows partial fulfillment.

Table 1. Ferguson's Diglossia Criteria and Pashto

Constraint	Pashto
Perceived Superiority of High Form	½
Sizeable Ancient Body of Literature	½
Great Grammatical Differences between High and Low Forms	Ø
Lexicon has Paired Items between High and Low Forms	✓
Literacy Limited to a Small Elite	✓

The table is taken from the Coyle (2014). Unlike Ferguson's High Arabic, Pashto lacks a religious text that is widely quoted and underpins its prestige. Although Khushal Khan Khattak's works are referenced, his grammar does not serve as the foundation for Standard Pashto. Despite this, Educated Standard Pashto retains a degree of prestige, especially in formal contexts, though it does not match the prestige of High or Literary Arabic. Consequently, Ferguson's criterion for perceived superiority is only partially met. The criterion related to historical literature is also partially met, given contributions from poets like Khattak and Rahman Baba from the seventeenth century. The significant grammatical differences criterion is not met, as Educated Pashto does not have the divergent standard grammar seen in Arabic. However, the criterion concerning paired lexical items is fully satisfied, with many terms having standard High and Low Pashto counterparts. A researcher's list of eighty-five standard words reveals numerous pairs where High and Low forms differ.

Despite the low literacy rate in Afghanistan, which stands at 31% (World Health Organization 2011), limiting full comprehension of Educated Standard High Pashto to a small segment of the population, some exposure to High Pashto occurs through radio broadcasts. With over 93% of households in Pashtun-majority provinces having radios, High Pashto vocabulary infiltrates Low Pashto varieties. This exposure, coupled with minimal grammatical differences between Educated Standard Pashto and Low Pashto, enables some illiterate individuals to acquire aspects of Educated Standard Pashto informally. Thus, while formal education remains inaccessible to many Pashtuns, increasing exposure through media may facilitate partial acquisition of High Pashto among the illiterate population.

In summary, although Educated Standard Pashto exhibits characteristics of a High form in diglossia, such as limited literacy and paired lexical items, the divide between High and Low Pashto is less pronounced due to media exposure bridging some gaps.

3.1.3. Regional Standard Pashto Varieties

In Afghanistan, Pashto exhibits numerous tribal and regional dialects due to its many tribes. Among these, Wardak Pashto is noted for being particularly challenging to understand. When Wardaks travel and interact with Pashtuns from other tribes, communication often requires a common dialect. This discussion explores "Regional Standard Pashto," which is the variant used by Pashtuns from different tribes when they meet. This includes a focus on the regional standards in Kandahar, Jalalabad, and Kabul.

3.1.4. Kandahar Regional Pashto

Kandahar, a central city in southern Afghanistan and the largest city in the Durrani Confederacy, serves as a meeting point for various tribes. The Pashto spoken here, known as Kandahar Regional Pashto, is a variant of Southern Pashto. The language of Kandahar reflects a minimal difference from the tribal varieties of the Durrani Confederacy, making it accessible and understandable across tribal lines. The dialects of tribes like the Populzai, Nurzai, and Barakzai are nearly identical to this regional standard.

3.1.5. Jalalabad Regional Pashto

Jalalabad, the capital of Nangarhar Province and a major city in eastern Afghanistan, is a hub for tribes from the Eastern Confederacy and beyond. The Pashto spoken here, referred to as Jalalabad Regional Pashto, incorporates elements from various tribes in the eastern region, including Mohmand, Shinwar, and Sapi. This variant shows some influence from neighboring Ghilji tribes, though this influence is relatively minor. The dialect spoken in Jalalabad is similar to Northern Pashto but remains distinct due to local variations and influences.

3.1.6. Kabul City Pashto Varieties

Kabul, the capital of Afghanistan, presents a unique linguistic landscape due to its diverse population. In areas like Kampani and Arzan Qemat, where various Pashtun tribes converge, regional varieties of Pashto develop. Kampani Regional Pashto blends elements from Wardak, Ghilji, and Southern Pashto, reflecting its diverse inhabitants. Arzan Qemat similarly features a mix of Pashto varieties. In Kabul city itself, Dari, spoken by the majority Tajik population, heavily influences the Pashto spoken by Pashtuns. This results in a Pashto dialect in Kabul that incorporates many Dari terms and expressions, affecting the tribal and regional varieties found there.

In summary, while each city has its own regional standard of Pashto, influenced by local tribal varieties and external factors such as Dari in Kabul, these standards facilitate communication among Pashtuns from different backgrounds.

3.1.7. Kabul Regional Standard Pashto

This section examines how different Pashtun dialects interact when speakers from various Pashtun confederacies converge in neutral environments like Kabul or university dormitories. Historically, mutual understanding was challenging, with significant communication barriers between speakers of Northern, Southern, and Wardak Pashto. However, advancements in the standardization of educated Pashto have improved comprehension among educated Pashtuns from different regions. Today, in Kabul—where Pashtuns from various areas frequently meet—the dominant dialect is referred to as Kabul Regional Standard Pashto.

To explore regional differences, small groups from each dialect area—Northern, Southern, and Wardak—compiled word lists of unique terms not found in Kabul Regional Standard Pashto. The assumption was that these lists would highlight distinct regional variations. The findings generally supported this assumption, revealing that Southern and Wardak Pashto speakers produced lists with terms that contrasted with Kabul Regional Pashto, while Northern Pashto speakers compared their terms with Southern Pashto instead of Kabul Pashto. This comparison suggested that Northern Pashto is relatively close to Kabul Regional Standard Pashto, but the Northern Pashto group's comparison with Southern Pashto indicated that they might view Southern Pashto as a regional standard. The analysis suggests two main conclusions. First, Northern Pashto is quite similar to Kabul Regional Standard Pashto, indicating a close relationship between the two. Second, Southern Pashto speakers seem to have a better understanding of Northern Pashto than vice versa. This could imply that Southern Pashto, due to its prestige, might influence Kabul Regional Standard Pashto more, though both dialects have significant mutual influence.

In summary, Kabul Regional Standard Pashto reflects a blend of regional influences, with Southern Pashto contributing notably due to its prestige and broader comprehension among its speakers. The influence of Dari, the predominant language in Kabul, further shapes this regional standard.

3.2. Prestige Standard Pashto

The Durrani Confederacy Pashtuns speak Southern Pashto, while the Eastern Confederacy Pashtuns speak Northern Pashto, making both Southern and Northern Pashto prestigious dialects. Southern Pashto holds greater prestige than Northern Pashto due to the Durrani Confederacy's pivotal role in the establishment and governance of Afghanistan. This added prestige is reflected in the simple designation of the Durrani region as "the South," a title no other region enjoys. Consequently, I propose that Southern Pashto be recognized as Prestige Standard Pashto.

There are two additional reasons for elevating Southern Pashto to a standard status. Firstly, over the past century, various literary committees have worked to define and standardize Educated Standard Pashto. As the most influential region, the South has contributed numerous members to these committees, leading to the adoption of Southern Pashto forms in educational materials and standardization efforts, thereby integrating Southern Pashto into Educated Standard Pashto.

The second reason, while largely symbolic, holds significant importance. The strong tribal identity among Pashtuns and their pride in their respective tribes. However, this tribal pride is matched by a strong ethnic identity among Pashtuns. Due to the unifying efforts of Ahmad Shah Durrani and other leaders, Pashtuns maintain a fierce loyalty to their common heritage, often disregarding the Durand Line established by the British in 1893, which

separates Pashtun lands between Pakistan and Afghanistan. Today, many Pashtuns advocate for an independent Pashtun nation. Within Afghanistan, Southern Pashto serves as a symbol of unity for Pashtun ethnic identity.

3.3. Ferguson's Model of Diglossia and Pashto Language: A Critique

Firstly, Ferguson identified the perceived superiority of the High Form as a key criterion in diglossic situations, where this form is viewed as more prestigious and refined compared to the Low Form. Secondly, the presence of a sizeable ancient body of literature in the High Form reinforces its elevated status and historical significance. Thirdly, there are often considerable grammatical differences between the High and Low forms, which further distinguishes their respective uses and structures. Fourthly, the lexicon in diglossic settings frequently includes paired items, with specific terms in the Low Form having their counterparts in the High Form, highlighting the divergent vocabularies. Lastly, literacy is typically limited to a small elite who are proficient in the High Form, while the general population tends to use the Low Form in everyday communication.

Arabic fulfills all the criteria of a diglossic language. Firstly, the High form is regarded as the prestigious variant, to the extent that the Low form is sometimes considered non-existent. In Arabic-speaking countries, those who haven't learned High Arabic may be deemed ignorant of the language, even if they speak a Low form natively. This perception equates High Arabic with the only legitimate form. The High form's superiority is often linked to religious significance, as the Quran is believed to have been written in High Arabic and to have existed before the world, implying that High Arabic predates all other languages. Secondly, there is a significant amount of ancient literature written in the High form. Thirdly, there are notable grammatical differences between the High and Low forms, such as the High form having three noun cases while the Low forms have none. Fourthly, the vocabulary includes many pairs of High and Low terms for similar concepts, marking an utterance as either High or Low based on the term used. Lastly, while Ferguson's definition notes that the High form is mainly acquired through formal education, he later clarifies that this education is accessible only to a small elite. This is also true for Arabic, where only a limited number of people have the opportunity to learn the High form.

In contrast to Ferguson's High Arabic, Pashto lacks a religious text that is currently referenced and provides a foundation for the prestige of its Standard form. Although Khushal Khan Khattak's work is cited, his grammar does not form the basis for Standard Pashto. Despite the absence of a religious text, High Pashto enjoys considerable prestige and is used in formal settings, though this prestige is not as significant as that of High or Literary Arabic. It would not be accurate to claim that an illiterate, elderly Pashtun who has never attended school does not know Pashto. Consequently, the first criterion of Ferguson's diglossia is only partially satisfied in the case of Pashto. The second criterion is also partially met by poets like Khattak and Rahman Baba from the 17th century. The third criterion is not fully met, as Educated Pashto lacks a distinct standard grammar compared to Arabic. While there are subtle grammatical differences between High and Low Pashto, these are not as pronounced as in Arabic. However, the fourth criterion of Ferguson's diglossia is fully met by Pashto, with a rich lexicon of paired terms. A list of eighty-five standard words was created to test comprehension, though it is not exhaustive. For instance, the verb 'to sell' is 'plorəl' in High Pashto and 'xartsawəl' in Low Pashto. Some of these terms denote concepts absent in Low Pashto, while others are paired words or new forms of existing Low Pashto words with the same meaning. Educated words frequently appear in BBC news online. I personally experienced a sense of starting language learning afresh when reading the news in Pashto. In summary, while only one of Ferguson's first four criteria for a High form is fully met, two are partially satisfied, and one is not met. Diglossia is further evidenced by the fact that literacy is confined to an elite minority, with only 31% of Afghanistan's population being literate (World Health Organization 2011). Thus, a small fraction of the population fully grasps Educated Standard High Pashto, with over two-thirds of Afghans being uneducated and thereby excluded from High Pashto's spheres of use, such as government and media.

Nonetheless, this exclusion may not be complete. High Pashto vocabulary and grammar might be seeping into Low Pashto varieties. Radio news, which reaches all parts of the Pashtun region, plays a significant role. Over 93% of households in predominantly Pashtun provinces have radios, compared to about 34% with televisions likely due to unreliable electricity in rural areas (Hopkins, 2012). Battery-operated radios are a key source of information and entertainment, and they disseminate High Pashto vocabulary. Consequently, even illiterate villagers are exposed to High Pashto words. This exposure, combined with the minimal grammatical differences between Educated Standard Pashto and Low Pashto, may enable illiterate individuals to acquire some aspects of Educated Standard Pashto informally. Therefore, despite the majority of Pashtuns lacking formal education, an increasing number of illiterate Pashtuns are partially learning High Pashto.

3.4. Literacy and Regional Variability in Pashto Language

Low literacy levels contribute to the existence of both an Educated Standard Pashto and various regional standard Pashto forms. In countries with higher education levels, the differences between the educated and regional standards tend to diminish. An Educated Standard Pashto facilitates communication between speakers from diverse regions. However, in Afghanistan, due to widespread illiteracy, the Educated Standard Pashto is insufficient for effective communication among speakers from different tribes. As a result, a separate Regional Standard Pashto has emerged to address this communication need.

Educated men can relatively easily learn the Educated Standard Pashto, and those who travel often pick up the Regional Standard Pashto. However, women face greater challenges. Additionally, limited travel outside their local areas means they rarely acquire Regional Standard Pashto. Consequently, women encounter significantly higher barriers to mastering standard Pashto forms compared to men.

Educated Standard Pashto, which has been standardized over the past century, has been shaped by both Southern and Northern Pashto, with a notable impact from Southern Pashto due to the greater number of literary committee members from the South. Regional Standard Pashto, bridging regional differences in Kabul, is more influenced by Northern Pashto. Southern Pashto, symbolizing ethnic identity, has had the most significant effect on Prestige Standard Pashto.

3.5. Strategies to Explore the Regional Varieties

Exploring the regional varieties of Pashtu, a language with a rich tapestry of dialects and linguistic features, involves a multi-faceted approach. First, understanding the geographical distribution of Pashtu speakers is crucial. Pashtu is predominantly spoken in the Khyber Pakhtunkhwa province of Pakistan and the eastern regions of Afghanistan, but within these areas, distinct regional varieties emerge. For instance, the Pashtu spoken in the Peshawar region may differ significantly in pronunciation, vocabulary, and syntax from that spoken in the southern regions of Afghanistan, such as Kandahar.

To delve into these regional differences, one effective strategy is to engage with local communities and conduct fieldwork. Immersing oneself in the daily lives of Pashtu speakers allows for a firsthand experience of the language in its natural context. Interviews and conversations with native speakers can provide valuable insights into how regional variations manifest in everyday communication. Furthermore, analyzing recordings of spoken Pashtu from different regions can reveal subtle differences in accent, intonation, and rhythm, which are often lost in written texts.

Another important approach is to study linguistic research and existing literature on Pashtu dialectology. Academic papers, linguistic surveys, and ethnographic studies offer in-depth analyses of the structural and lexical variations across different Pashtu dialects. Engaging with these resources can help identify key characteristics and variations within the language. Additionally, collaborating with local language experts, such as linguists and educators who specialize in Pashtu, can provide a deeper understanding of the historical and sociopolitical factors influencing regional dialects. These experts often have access to comprehensive data and can offer insights into how historical migrations, cultural exchanges, and regional influences have shaped the language.

In summary, a comprehensive exploration of Pashtu's regional varieties involves fieldwork, analyzing linguistic data, and consulting with experts. By combining these strategies, one can gain a nuanced appreciation of the rich diversity within the Pashtu language and its regional dialects.

4. Findings

Partial Alignment with Diglossia Criteria: Pashto exhibits some characteristics of Ferguson's diglossia model but does not fully align with all criteria. Specifically, Pashto partially meets the criteria for perceived superiority and historical literature but falls short in terms of significant grammatical differences and the presence of a widely recognized religious text. Unlike High Arabic, which is underpinned by religious texts and a distinct grammatical structure, Pashto lacks a religious foundation for its prestige and has only minimal grammatical divergence between its High and Low forms.

Prestige and Standardization: Educated Standard Pashto, which has been developed and standardized over the past century, reflects a degree of prestige in formal contexts, similar to Ferguson's High Form. However, its prestige is not as pronounced as High Arabic due to the absence of a foundational religious text and less pronounced grammatical differences. Despite this, Educated Standard Pashto is used in education and media, which helps bridge some gaps between High and Low Pashto varieties.

Lexical Pairing: Pashto meets Ferguson's criterion of paired lexical items, with many terms having distinct High and Low forms. This lexical pairing is a key feature of diglossic languages and is evident in the rich vocabulary of Educated Standard Pashto, which includes numerous terms that contrast with their Low Pashto counterparts.

Literacy and Language Use: The limited literacy rate in Afghanistan contributes to the persistence of both Educated Standard and Regional Standard Pashto forms. While Educated Standard Pashto is primarily used in formal and educational settings, Regional Standard Pashto emerges to facilitate communication among different tribal and regional groups. This reflects a partial alignment with Ferguson's model, where literacy is confined to a small elite and the general population uses the Low Form for everyday communication.

Media Influence: Media exposure plays a significant role in bridging the gap between High and Low Pashto varieties. Radio broadcasts and other media channels disseminate Educated Standard Pashto vocabulary, allowing even those with limited formal education to acquire some aspects of the High Form informally. This exposure contributes to a blending of High and Low Pashto features, demonstrating an adaptive dynamic within the diglossic framework.

Regional Variability: The diverse regional varieties of Pashto highlight the language's adaptability and complexity. Regional standards in cities like Kandahar, Jalalabad, and Kabul show how local linguistic and

cultural factors influence the development of Pashto dialects. This regional variation complements the broader framework of diglossia by addressing specific communication needs among different Pashtun groups.

References

Coyle, D. W. (2014). Placing Wardak among Pashto varieties.

Crystal, D. (2011). *A dictionary of linguistics and phonetics*. John Wiley & Sons.

Fraser, H. (2000). *Coordinating improvements in pronunciation teaching for adult learners of English as a second language*. DETYA (ANTA Innovative project).

Ghani, U. (2005). Problems of English Pronunciation among Pashto Speaking

Hallberg, G. D. (1992). Sociolinguistic Survey of Northern Pakistan. Volume 4. Pashto , Waneci, Ormuri. National Institute of Pakistan Studies, Quaid – I - Azam University, Islamabad, Pakistan.

Hopkins, Nancy, ed. 2012. A Survey of the Afghan People - Afghanistan in 2012. Kabul, Afghanistan: The Asia Foundation.

Hussain, A., Ali, Z., & Khan, T. (2023). Articulation of English vowel sounds by Pashto speakers: A contrastive phonetic analysis. *Erevna: Journal of Linguistics and Literature*, *7*(1), 14-30.

MacKenzie, D. N. (1959). A standard Pashto. *Bulletin of the School of Oriental and African Studies, University of London*, *22*(1/3), 231-235.

Naheed. Q. (2015) Language Drop a Line to Change the Jargon: A Socio Linguistic Study of Afghani and Pakistani Dialect of Pashto Language, Pashto Volume 44 No 650s.

Olson, Randall. B. (1996). Speaking Afghan Pashto, Interlit foundation inc., Peshawar, Colorado Press, USA.

Penzle, H. (1955). A Grammar of Pashto: A Descriptive Study of the Dialect of the Kandahar, Afghanistan. American Council of Learned Society,Washington, D. C.

Rahman, G. (2009). A Comparative Study of Pashto and English Phonology for English Language Teaching and Learning , MS Thesis, UMT, Lahore, Pakistan.

Rahman, T. (1996). Language and politics in Pakistan. Karachi: Oxford University Press.

Rehman, T. (1995). The Pashtu language and identity formation in Pakistan. Contemporary South Asia, 4(2), 151-170.

Roach, Peter. (2009). Phonetics & Phonology (4th ed.). Cambridge: Cambridge University Press

Students at Graduate level and Remedial Measures.(Master's thesis). All amaIqbal Open University, Islamabad, Pakistan

Tegey, Habibullah, and Barbara Robson. 1996. A Reference Grammar of Pashto. Washington, D.C.: Distributed by ERIC Clearinghouse.