FWU Journal of Social Sciences, Winter 2025, Vol.19, No.4, 139-152
DOI: http://doi.org/10.51709/19951272/Winter2025/12

A Cross-Linguistic Study of Grammatical Nominalization in Pakistani Languages

Muhammad Nawaz
COMSATS University Islamabad, Islamabad Campus

Miki Nishioka
Graduate School of Humanities, University of Osaka, Japan

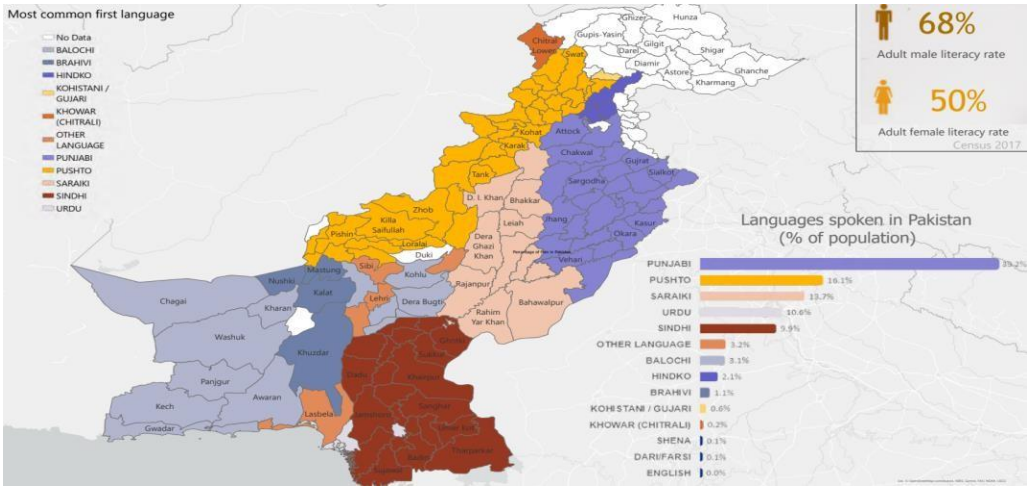
Saad Zahid
COMSATS University Islamabad, Islamabad Campus

This study compares the nominalization systems of Urdu, Pahari, Punjabi, and Hindko. It focuses on their morphological features and shared structures. The data were collected from five native speakers of each language, resulting in a total of twenty speakers. First, a list of Urdu nominalization constructions was compiled with the help of native speakers, based on Shibatani's (2019) framework, including relational, participle, genitive, and appositive constructions. Then, the five native Urdu speakers were recorded to verify the markers of compiled sentences. Subsequently, five speakers of each selected language translated the same Urdu sentences into their native languages. This process was documented through written and audio records. Urdu employs the marker *-wa:la:*, while the other languages use *-a:la:*. In relative constructions, Urdu uses *jo:*, whereas Pahari, Punjabi, and Hindko use *jirrha:*, which varies by gender and number. Additionally, Pahari and Hindko exhibit relative-correlative structures. All four languages utilize participles for nominalization and share genitive constructions. Urdu's genitive markers are *ka:, ki:, kay,* while Pahari has *nã:, nĩ:, nay,* and Punjabi and Hindko include *ɖa:, ɖi:, ɖay* — all of which agree in gender and number. Appositive constructions are generally similar, though Pahari uniquely features two forms: *kay* and *khay*. While each language maintains distinct morphological traits, the shared nominalization markers and genitive structures highlight their deep linguistic connections .

Keywords: nominalization, Urdu, Pahari, Punjabi, Hindko, morphological markers

Pakistan is a multilingual country with diverse languages spoken across its various regions. Urdu, Pahari, Punjabi, and Hindko are the Indo-Aryan languages spoken in Pakistan (The Indo-Aryan Languages, 2014). Urdu, spoken in India and Pakistan, belongs to the Indo-Aryan branch of the Indo-European language family. Pahari is spoken in the Pir Panjal Range of the Himalayas in Azad Jammu and Kashmir, Muree, and Hazara districts of Pakistan (Khalique, 2014). Punjabi is spoken in the Punjab region of Pakistan and India. Hindko is primarily spoken in Pakistan's Khyber Pakhtunkhwa (KP) province and in adjacent areas of Punjab by a small number of people (Sohail et al., 2011). Pashto is the primary language of KP, one of Pakistan's major regional languages (Sardaraz et al., 2021). Overall, around 2.5 percent of Pakistan's population speaks Hindko. Its dialects vary in their lexicon and syntax; however, the phonemic inventories are broadly similar (Nawaz & Afsar, 2019). The linguistic map of Pakistan illustrates the boundaries of various languages, as shown in Figure 1.

Correspondence concerning this article should be addressed to Muhammad Nawaz, Associate Professor, Department of Humanities, COMSATS University Islamabad, Islamabad Campus (mnawazfazal@gmail.com)



Around 70 Languages are spoken in Pakistan
Punjabi 39.23%
Pashto 16.11%
Saraiki 13.66%
Urdu 10.64%
Sindhi 9.86%
Balochi 03.08%
Hindko 02.11%
Brahui 01.10%
Others 3.22% (Pahari etc)

Figure 1: Linguistic map of Pakistan. (Source: Google Images)

The linguistic map indicates that Punjabi accounts for 39.23%, Pashto for 16.11%, Saraiki for 13.66%, Urdu for 10.64%, Sindhi for 9.86%, Balochi for 3.08%, Hindko for 2.11%, Brahui for 1.10%, and others for 3.22%. Pahari is included in the broader category of other languages. The four specified languages, Urdu, Pahari, Punjabi, and Hindko, share various commonalities and differences in their structures. Previous studies on the selected languages have explored their morphological, syntactic, and semantic aspects; however, they have neglected to examine how nominalization systems function across various linguistic levels and within Pakistan's typologically diverse linguistic region.

Generally speaking, nominalization systems, an important morphological aspect, have already been studied in major world languages such as English (Halliday & Matthiessen, 2004), Arabic (Ryding, 2005), and Spanish (King, 1992), as well as in crosslinguistic and typological works (Comrie & Thompson, 2007). Specifically, in South Asian languages, Nishioka and Kumar (2021, 2025) analyzed Bhojpuri, Magahi, and Maithili in Hindi, as well as the languages of Bihar, using grammatical nominalization according to Shibatani's theory. Such syntactic and morphological investigations enhance the pedagogical implications, preservation, and documentation of these languages.

The present study examines the nominalization patterns of the four Pakistani languages. It explores how these patterns reflect sociolinguistic variation in South Asian languages. It aims to provide a deeper understanding of nominalization across languages. It also evaluates the applicability of Shibatani's (2019) framework in less-studied linguistic contexts. A data-driven approach supports this positioning and contributes to theoretical discussions on typology, language change, and South Asian linguistics in multilingual contexts.

Literature Review

Nominalization, in its basic meaning, refers to the phenomenon by which nominal expressions are derived. Nominalization turns something into a noun (Thompson, 1985). Nominal expressions can be derived from verbs and adjectives (Foong et al., 2011). They state that clauses can also be nominalized. Nominalization constructions can be categorized into participant vs. event, lexical vs. grammatical, and embedded vs. non-embedded.

Participant nominalization refers to first-order ontological entities, such as persons, places, or objects. These participants assume the role of agents, patients, locations, or instruments. Event nominalization refers to the nominalization of events. Lexical nominalization resembles non-derived nouns in terms of morpho-syntactic characteristics. Lexical nominalization receives genitive and plural markings, and the nominalizer may have scope over the entire clause (Comrie & Thompson, 1985). They further argue that clausal nominalization may not possess nominal features but verbal ones, such as person-number agreement. They cite the example of Korean, where nominalization constructions have both tense and case. Clausal nominalizations, as Foong et al. (2011) describe, are embedded as arguments of a matrix clause in most cases, though not always. Non-embedded nominalizations express speakers' attitudes. Key linguistic patterns in the regional grammatical distinctions of South Asian languages have been largely overlooked in previous research.

In this regard, Shibatani's (2017, 2018, 2019) typological works which has rich morphological marking specifically for South Asian languages may provide a more comprehensive account of nominalization according to the

nature of the specified languages' data, classifying it into lexical and grammatical categories, shown in the following flow diagram:

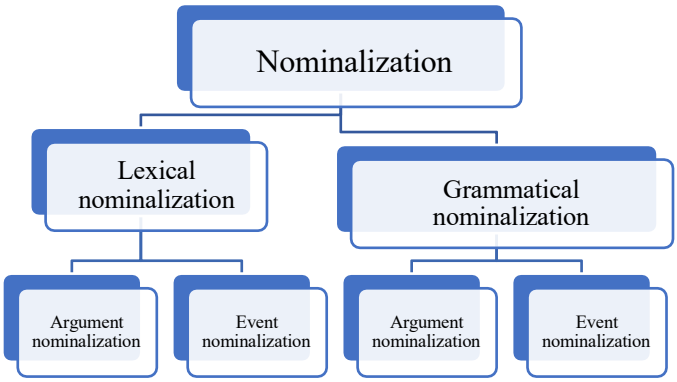


Figure 2: Categorization of nominalization based on Shibatani's typological works
Nishioka and Kumar (2021: 65)

Noun-based nominalizations denote entities related to the concept. They can have both activity and agent/instrument readings, while verbal nominalizations denote activities. Grammatically nominalized expressions denote concepts that are closely associated with their originals and are understood in context. The Gricean maxim of relevance also contributes to the interpretation of grammatical nominalizations (1975). They can function as modifiers in a noun phrase or an adverbial. Grammatical nominalizations include stem nominalizations and participles, which require a copula as an evidence marker.

Generally speaking, Shibatani's theory of nominalization is a development of functionalist linguistics, including cognitive linguistics. It solves conventional problems by reanalyzing various languages. Specifically, nominalization elucidates the cognitive motivation underlying syntactic structure and discourse function, including basic concepts such as "clause" and "sentence". This groundbreaking study examines word derivation and its relationship to relative clauses and embedded constructions across diverse languages worldwide. It demonstrates how neo-nominalization theory provides a unified view of phenomena traditionally treated separately in grammar, generative grammar, and linguistic typology.

According to Yoshio (as cited in Shibatani, 2019), more categories of grammatical nominalizations can include true grammatical, abbreviated, and clausal nominalizations. Event nominalizations denote abstract concepts and their related entities. Argument nominalizations denote event participants, using relatives, including subject and object nominalizations. Genitive case markers have varied functions and denotations that are understood in context. It can make possessive constructions in nouns and occur with non-finite verbal forms. The appositives take non-appositive forms, such as relatives and genitives. They also modify the head nouns. The actual denotation of argument nominalization is determined by context.

Studies have focused primarily on lexical nominalizations, despite the broader scope of grammatical nominalizations. This is due to continuous variation in the forms of grammatical nominalizations, as they exhibit structural properties similar to those of clauses. However, these typological works lack application to South Asian languages. These languages have numerous complex examples that can be integrated into the regional linguistic literature. This paper examines nominalization of the four Pakistani languages based on the following devices:

The concept of Relatives belongs to the category of central determiners. Words like which, who, whose, whatever, whichever, etc., are relatives (Greenbaum & Nelson, 2002). Relative clauses function as modifiers of noun phrases and other categories. The relatives may be restrictive or non-restrictive. These clauses have the same structure as Wh-question sentences, although they do not ask a question (Roberts, 2016). The relative clauses may be embedded in noun phrases (Chalker & Weiner, 1994). This study examines the structural and functional aspects of relative clauses in the selected languages.

The general concept of the Participle is defined as a non-finite form of the verb that ends in either -ing or -ed in the case of regular verbs. In the case of irregular verbs, the -ed participle may have a -en form or may have a change of vowel from its base form (Greenbaum & Nelson, 2002). Two general forms of participles are named as present and past,

but neither of these names is true. This is because they occur in various tenses, allowing them to refer to present, past, or future time. So, it is preferable to use the -ing and -ed forms of participles.

The concept of Genitives shows possession or belonging. They may be dependent and independent genitives (Greenbaum & Nelson, 2002). The possessive pronouns are the genitives of the personal pronouns. A generally applied strategy for making genitives in English is adding an apostrophe or the addition of a prepositional phrase. In contrast, genitive markers in South Asian languages nominalize complete clauses, as in possessive constructions.

The concept of Apposition refers to a relation between two noun phrases that have the same referent. These phrases are parallel to each other (Chalker & Weiner, 1994). In grammar, apposition occurs when two words or phrases are placed side by side in a sentence so that one defines or describes the other. An example is the phrase, "My dog Woofers". Here, "my dog" is in apposition to the name "Woofers".

In addition to the Shibatani's (2019) theory markers, the marker of *-wa:la:* is a particle used in Urdu and *-a:la:* in many other languages of Pakistan. These perform a variety of functions making nouns from verbs. It denotes agents of action, such as the -er inflection in verbs like clean, which performs a similar function. Thus, the syntactic roles of the like -wa:la: and -a:la: constructions are central to South Asian languages in general and Pakistani languages in particular, providing a regional linguistic analysis.

Method

The data were collected from 20 native speakers across the four selected Pakistani languages: Urdu, Pahari, Punjabi, and Hindko. The speakers were chosen at random to maintain linguistic bracketing and ensure a diverse representation of language use. First, a list of Urdu nominalization constructions was compiled with the assistance of Urdu native speakers. It consisted of four sentences for each of the four categories, including relative, participle, genitive, and appositive, based on Shibatani's (2019) framework. To verify the accuracy of these sentences, five native Urdu speakers were consulted and recorded, ensuring the correct identification and construction of nominalization markers in the language. Subsequently, the verified Urdu sentences were translated into Pahari, Punjabi, and Hindko by five native speakers of each language. Parallel patterns from Urdu were employed throughout the translation process to maintain cross-linguistic consistency (see Section 4). It was ensured that all selected speakers were also proficient in Urdu, thereby enabling reliable cross-linguistic comparisons. Their Urdu proficiency was informally confirmed through pre-session discussions on language background. The data collection process was documented in written and audio records for morphological and syntactic analysis. The data were analyzed using Shibatani's (2019) framework, which examines relative constructions, participles, genitives, and appositives across all four languages.

Results

The analysis of Urdu, Pahari, Punjabi, and Hindko is guided by Shibatani's (2019) theoretical framework, categorizing nominalizations into relative, participle, genitive, and appositive types. The relative markers of Urdu are analyzed using four examples, followed by those of Pahari, Punjabi, and Hindko. The descriptions of all four languages are then provided, followed by a tabulated summary of the findings.

Relatives

The data on relative markers are provided with the morphological description and an English gloss. The examples are labelled as a, b, c, and d. The data is provided with double inverted commas (""), and their English gloss with single inverted commas (''). The abbreviations used during the analysis are given after the reference list.

Relative in Urdu

a) "jo: chakki: ab bhi: chalt̪: hai."
which (REL) mill (F.SG.) now too work/function (IPFV.F. SG.) COP.PRS.SG.
'The mill which functions even now.'

b) "chakki: jo: ab bhi: chalt̪: hai."
mill(F.SG.) which (REL) now too work/function (IPFV.F. SG.) COP.PRS.SG.
'The mill which functions even now.'

c) “jo: ganã: chahʈa: hai.”
one (REL.) sing (INF.) want (IPFV.SG.) COP.PRS.SG.
‘one(person) who wants to sing.’
d) “jo: ho: chukka: hai.”
one (REL) be (STEM) complete (PFV.M.SG.) COP.PRS.
‘One (thing) which is done.’

Relatives in Pahari

a) “jirrhi: chakki: hun vi: chalnĩ: zi:.”
which (REL) mill (F.SG.) now too work/function (IPFV.F.SG.) COP.PRS.F.SG.
‘The mill which functions even now.’
b) “oh chakki: jirrhi: hun vi: chalnĩ: zi:.”
COR mill (F.SG.) which (REL.F.SG.) now too work/function (IPFV.F.SG.) COP.PRS.F.SG.
‘The mill which functions even now.’
c) “jirrha: ganey chhanã: za:.”
one (REL M.SG.) sing (INF) want (IPFV.M.SG.) COP.PRS.M. SG.
‘One(person) who wants to sing.’
d) “jirrha: hoi rhya za:.”
one (REL.M.SG.) be (STEM) complete (PFV.M.SG.) COP.PRS.M.SG.
‘One (thing) which is done.’

Relatives in Punjabi

a) “jirrhi: chakki: hunrr vi: chalʤi: ey.”
which (REL.F.SG.) mill(F.SG.) now too work (IPFV.F. SG.) COP.PRS.SG.
‘The mill which functions even now.’
b) “chakki: jirrhi: hunrr vi: chalʤi: ey.”
mill (F.SG.) which (REL.F.SG.) now too work (IPFV.F.SG.) COP.PRS.SG.
‘The mill which functions even now.'
c) “jirrha: ga:nra cha:nda: ey.”
who (REL.M.SG.) Sing (INF) want (IPFV.M. SG.) COP.PRS.SG.
‘One who wants to sing.’
d) “jirrha: ho chukia: ey.”
which (REL.M.SG.) be (STEM) finish (PFV.M.SG.) COP.PRS.SG.
‘One which is done.’

Relatives in Hindko

a) “jirrhi: chakki: hunrr bi: chalʤi:.”
which (REL.F.SG.) mill(F.SG.) now too work (IPFV.F.SG.)
‘The mill which functions even now.’
b) “oh chakki: jirrhi: hunrr bi: chalʤi:.”
that (COR) mill (F.SG.) which (REL.F.SG.) now too work (IPFV.F.SG.)
‘The mill which functions even now.’
c) “jirrha: ga:nra: chanda:.”
who (REL.M.SG.) sing (INF) want (IPFV.M.SG.)
‘One(person) who wants to sing.’
d) “jey kuj ho: gey/ga:.”
which (REL) be finished (PFV.)
‘One(thing) which is done.’

Description of Relatives

The Pakistani languages use relative markers for nominalization, forming relative subordinate clauses. Urdu involves the relative marker *jo:* for making nominal expressions. This relative marker may be positioned before or after the noun phrase modified by the relative marker, as in 4.1.1a. However, the relative *jo:*’s position influences the sentence’s

emphasis. In 4.1.1a., the relative *jo:* at the initial position emphasizes *ab bhi (*even now*),* while placing the noun phrase at the initial position as in 4.1.1b. shifts the emphasis from *ab bhi* (even now) to *chakki:* (mill). The relative *jo:* is not expressive of gender and number, for instance, in examples 4.1.1a and 4.1.1d. The verb phrases like *chalt̪i:* determine the gender *hai* (works) and *chukka: hai* (is done). The *i* and *a* at the end of *chalt̪i:*, and *chukka:* determine the masculine and feminine genders, respectively. Moreover, the number (singular) is represented by these letters at the ends of verb phrases and the copula *hai,* which shows the singularity of the noun phrase. There is no apparent distinction between the relative standing of a person and a thing in Urdu, as seen in 4.1.1c and d, where the same relative marker is used for both person and thing.

Pahari, Punjabi, and Hindko share the same pattern except for having the embedded gender marker in the relative nominalizer, such as *jirrhi:* and *jirrha:* (in examples b. & d. of sections 4.1.2, 4.1.3, & 4.1.4) for feminine and masculine genders, respectively. The verb phrases independently represent the gender as well, as in *chalnɽ̃: zi:* in Pahari, where the *i:* of the verb phrase *chalnɽ̃:* and copula *zi:* determine the feminine gender. In 4.1.2d, in *hoi rhya za: (is done),* the *a*: at the end of the verb *rhya:* and copula *za:* determine the gender. Similarly, the number is determined by the helping verbs *zi:* and *za:,* representing singular feminine and masculine genders, respectively. Apart from this, Pahari experiences the same shift of emphasis by varying the positions of the relatives, as in 4.1.2a. and 4.1.2b. By shifting the position of the relative *jirrhi:* before and after the noun, the emphasis of the clause shifts from *chakki:* (mill) to *hunh vi:* (even now). Pahari shows the relative-correlative construction COR+NP+REL as in 4.1.2b., where the correlative *oh* (also functioning as a determiner) is incorporated to render more specificity and clarity to the noun *chakki:* (mill). This construction is absent from Urdu and Punjabi.

Punjabi uses the same relative markers*, jirrha:* and *jirrhi:,* to mark the singular masculine and feminine genders, respectively. In 4.1.3a, for instance, *jirrhi:* is the relative marker modifying *chakki:* (mill) and showing singular feminine gender. 4.1.3c. shows a nominal expression where the relative *jirrha:* represents the singular masculine gender. The copular verb *ey* does not show the gender in Punjabi. Moreover, if the relative *jirrha:* is shifted from the initial position to after the noun *chakki:*, the stress of the clause shifts from *chakki:* to *tho hunrh vi:* (even now). Punjabi shows no apparent distinction between the forms of relatives used to describe things and people. In 4.1.3c. and 4.1.3d., the same relative *jirrha:* is used to nominalize the expressions involving people (one person who wants to sing) and things (which is done), respectively.

In Hindko, the relative markers jirrha: and jirrhi: are used similarly to express gender and number, as in Pahari. However, the relative markers for people and things are distinguished forms, as shown in 4.1.4c. and 4.1.4d. In 4.1.4d, *jay/jay kuj* are the relative markers for specifying the thing(s). Additionally, Hindko exhibits a shift in emphasis similar to that of Urdu, Pahari, and Punjabi, as evidenced by the relative jirrhi position, as in 4.1.4a. and 4.1.4b. The relatives *jirrha:* and *jirrhi:* in Hindko demonstrate gender and number agreement, where *a:* and *i:* at the end of the words represent singular masculine and feminine genders. Hindko exhibits the relative-correlative construction COR+NP+REL, as seen in 4.1.4b. Like Pahari, the language uses the correlative marker *'oh'* before the noun.

Table 1
Summary Table of Relative Markers

Language	Relative Marker	Gender Marker	Number Marker	Emphasis Shift by Position
Urdu	jo:	None	None	Emphasis changes by position
Pahari	jirrhi: (fem.)	i: (fem.)	zi: (sing. fem.) za: (sing. masc.)	Emphasis changes by position
Punjabi	jirrhi: (fem.)	i: (fem.)	None	Emphasis changes by position
	jirrha: (masc.)	a: (masc.)	None	
Hindko	jirrhi: (fem.)	i: (fem.)	None	Emphasis changes by position
	jirrha: (masc.)	a: (masc.)	None	
	jay/jay kuj (thing)	None		

Participles

This section contains data on the participles of the languages and their analysis.

Participles in Urdu

a) "mera: pakka: hua: khanã:."
my (GEN.M.SG.) cook (PFV) COP.PFV. M.SG.food (M.SG.)
'My cooked food.'
b) "ab bhi chalʧ: hui chakki:."
nowtoo flow (IPFV.F.SG.) COP.PFV. F.SG/PL. mill (F.SG.)
'The mill which works even now.'
c) "chalʧ: chakki:."
work (IPFV.F.SG.) mill(F.SG.)
'Working mill.'

4.1.1 Participles in Pahari

a) "mari: pakki: roti:."
my (GEN.F.SG.) cook (PFV.F.SG.) food (F.SG.)
'my cooked food.'
b) "aja:n vi: chalnĩ: chakki:."
nowtoo work (IPFV.F.SG.) mill(F.SG.)
'The mill which works even now.'
c) "chalnĩ: chakki:."
work (IPFV.F.SG.) mill(F.SG.)
'Working mill.'

4.1.2 Participles in Punjabi

a) "maira: pakya: kha:nrã:."
my (GEN.M.SG.) cook (PVF.M.SG.) food (M.SG.)
'My cooked food.'
b) "ha:lay vi: chalʤ: chakki:."
nowtoo work (IPFV.F.SG.) mill (F.SG.)
'The mill which works even now.'
c) "chalʤ: chakki:."
work (IPFV.F.SG.) mill (F.SG.)
'Working mill.'

4.1.3 Participles in Hindko

a) "marra: pakey ɖa: khanrrã:."
my cook (PFV.M.SG.) COP. PFV. M.SG. food (M.SG.)
'My cooked food.'
"hunrr bi: chalʤ: chakki:."
nowtoo flow (IPFV.F.SG.) COP.PFV.f.sg Mill (F.SG.)
'The mill which functions even now.'
b) "chalʤ: chakki:."
work (IPFV.F.SG.) mill (F.SG.)
'Working mill.'

Description of Participles

The use of participles to derive nominal expressions is reflected in the data. Consider 4.2.1a from the Urdu data; the participle "cooked" is broken down into *pakka: hua:.* This nominal expression incorporates the contribution of *pakka:,* which generally means "cooked," and the additional use of *hua:*, which serves as a copula. PFV. Hence, Urdu has nominal expressions involving participles, such as the perfect participle and COP.PFV. tend to modify the head noun, that is, *kha:nã:* (meal). In 4.2.1b., the imperfect participle *chalʧ:* (working) and the perfect form of the copula *hui:* are used to modify the head noun. Both participles and the copula are gender markers. The letters *a:* and *i:* at the end of the words *pakka:* and *chalʧ:* represent the singular masculine and feminine genders, respectively. Participles can modify head nouns without the copula. For instance, in 4.2.1c, we observe a pattern of IPFV+N in nominalization.

The use of participles in Pahari follows the same pattern as in Urdu, except for the copula.PFV. Between the IPFV and the noun, it is not found in Pahari, as seen in 4.2.1a. and b. From 4.2.2a to 4.2.2c, the nominalized expressions like *pakki: roʈ̪i:* (cooked food) and *chalnɪ̃: chakki:* (working mill) exhibit a constant absence of the copula. PFV. Other than that, in 4.2.2a, the perfect participle *pakki:* (cooked) is modifying the noun *roti:* (food) by associating an attribute with the noun. In 4.2.2b., a participle *chalnɪ̃:* (working) modifies the noun *chakki:* (mill). This is an imperfect participle, demonstrating the progressive aspect. Other than that, these participles express gender and reflect number agreement. The participles in 4.2.2a, b, and c. represent singular feminine nouns. As noted above, the final-i marks the gender and number.

Punjabi employs participles with patterns similar to those in Urdu, but there is no regularity in the use of the copula.PFV. Between the head noun and the participle, both the patterns PFV/IPFV + COP. + N and PFV/IPFV + N are in application. In 4.2.3a, the perfect participle *pakya:* (cooked) is positioned before the head noun. In 4.2.3c, the pattern is the same as 4.2.3a, except for an imperfect participle before the noun.

Hindko also exhibits no regularity in using the copula between the noun and the participle. In 4.2.4a, the use of COP. PFV, additionally, marks the perfect aspect. The case is similar in 4.2.4b, where the COP is placed between the noun and the participle, marking the expression's progressive aspect. In 4.2.4c, the application of the copula is absent. Like Urdu and Punjabi, the number and gender are reflected by the participles and the copular verbs.

Table 2
Summary Table of Participles

Languages	Examples	Participles	COP.PFV.	Gender Marker	Number Marker	Position
Urdu	4.2.1a	*pakka: hua:* (cooked)	*hua:*	*a:* (masc.)	Sg.	Before/After noun
	4.2.1b	*chalʈ̪i: hui* (working)	*hui:*	*i:*(fem.)	Sg.	Before/After noun
	4.2.1c	*chalʈ̪i:* (working)	None	*i:* (fem.)	Sg.	Before noun
Pahari	4.2.2a	*pakki: roti* (cooked food)	None	*i:* (fem.)	Sg.	Before noun
	4.2.2b	*chalnɪ̃: chakki:* (working mill)	None	*i:* (fem.)	Sg.	Before noun
	4.2.2c	*chalnɪ̃:*(working)	None	*i:*(fem.)	Sg.	Before noun
Punjabi	4.2.3a	*pakya:* (cooked)	None	*a:*(masc.)	Sg.	Before noun
	4.2.3b	*chalɖ̪i: hoi* (working)	*hoi:*	*i:* (fem.)	Sg.	Before/After noun
	4.2.3c	*chalɖ̪i:* (working)	None	*i:* (fem.)	Sg.	Before noun
Hindko	4.2.4a	*pakey* (cooked)	None	*a:*(masc.)	Sg.	Before/After noun
	4.2.4b	*chalɖ̪i:* (working)	None	*i:* (fem.)	Sg.	Before/After noun
	4.2.4c	*chalɖ̪i:* (working)	None	*i:* (fem.)	Sg.	Before noun

Genitives

This section presents data on genitive markers and the analysis.

Genitives in Urdu

a) "Ahmed ki: kiʈ̪a:b."
Ahmed GEN.F.SG/PL. Book (F.SG.)

'Ahmed's book.'

b) "un logõ: ki: ʦ̪ehseen."
those people GEN.F.SG/PL. praise (F.SG.)
'The praise of those people.'

c) "ɖ̪abbay ka: (wa:la:) ɖ̪oodʰ."
pack GEN.M.SG. milk (M.SG.)
'canned milk.'

d) "soonay ka: waqʈ̪"
sleep (INF.OBL.) GEN.M.SG. time (M.SG.)
'The time to sleep.'

Genitives in Pahari

a) "Ahmed nɪ̃: kiʈ̪a:v."
Ahmed GEN.F.SG. book (F.SG.)

'Ahmed's book.'

b) "unã: lokã: nĩ: ʈareef."
those people GEN.F.SG. praise (F.SG.)
'The praise of those people.'

c) "ɖabbay nã: ɖuɖh."
pack GEN.M.SG. milk (M.SG.)
'canned milk.'

d) "suttay nã: waila:"
sleep (INF.OBL.) GEN.M.SG. time(M.SG.)
'The time to sleep.'

Genitives in Punjabi

a) "Ahmed ɖi: kiʈa:b"
Ahmed GEN.F.SG book (F.SG)
'Ahmed's book.'

b) "unhã: lokã: ɖi: ʈareef."
those people GEN.F.SG praise (F.SG.)
'The praise of those people.'

c) "ɖabbay ɖa:/a:la: ɖuɖh."
pack GEN.M.SG. milk (M.SG.)
'canned milk.'

d) "soonrr ɖa: waila."
sleep (INF.OBL) GEN/PTCL.M.SG. time (M.SG.)
'The time to sleep.'

Genitives in Hindko

a) "Ahmed ɖi: kiʈa:b."
Ahmed GEN.F.SG. book (F.SG.)
'Ahmed's book.'

b) "unã: logã: ɖi: ʈareef."
those people GEN.F.SG. praise (F.SG.)
'The praise of those people.'

c) "ɖabbay ɖa:/a:la: ɖuɖh."
pack PTCL.M.SG. milk (M.SG.)
'canned milk.'

d) "sainray ɖa: waila:."
sleep (INF) GEN.M.SG. time.
'The time to sleep.'

Description of Genitives

The general pattern for nominalization in Urdu, using genitive markers, is NP + GEN + NP. In 4.3.1a, we observe this pattern, where the genitive function is performed by *ki:*. The genitive is placed between two noun phrases involving proper and common nouns. This genitive marker illustrates that the subsequent noun phrase has the noun placed before it. In 4.3.1b., the same pattern continues. The genitive marker *ki:,* placed between two nouns, modifies the noun *ʈehseen* (praise). The genitive *ki:,* used in these instances from the data, also marks the noun's gender following it. It marks the feminine gender that it modifies. In 4.3.1c, the genitive marker *ka:* between the noun phrases is not performing its usual function of demonstrating possession. The genitive *ka:* is associating an attribute to its subsequent noun *ɖoodh* (milk), just like *ed* in "packed milk" associates an attribute to "milk." In this case, the genitive and the noun phrase before it work like participle adjectives do in English. Additionally, the genitive marker ka also expresses the singular masculine gender. Likewise, in 4.3.1d, the genitive *ka:* does not show possession *per se.* The phrase conveys the meaning of "the time to sleep," involving infinitives.

Pahari has a general pattern of Urdu, except that the orthographic and phonetic realizations of genitive markers differ from those in Urdu. In 4.3.2a., and b., the genitive marker *nĩ:* occurs between the noun phrases Ahmed, *kiʈa:v*

(Ahmed, book), and *lokã:, ʈareef* (people, praise) respectively, and qualifies the noun phrases afterwards. However, in 4.3.2c and d, there is a variation in the function of the genitive marker. The genitive *nã:* in 4.3.2c, combined with the noun ɖabbay (pack), performs the attributive function like participle adjectives in English. Likewise, in 4.3.2d, the genitive *nã:* occurring between *suttay* (sleep) and *waila* (time) gives the meaning "the time to sleep". The genitives *nĩ:* and *nã:* used in Pahari also express gender and number agreement to the noun they modify. Genitive *nĩ:* is used in Pahari for singular feminine, while *nã:* is used for singular masculine.

Genitives *ɖa:, ɖi:,* etc., are used in Punjabi to derive nominal expressions. The genitives can also perform the attributive functions when combined with nouns, as in 4.3.3c and d. In 4.3.3c.; genitive *ɖa:* works in the same manner as discussed in the case of Urdu and Pahari. Similarly, 4.3.3d yields the same sense as in Urdu and Pahari. However, the Punjabi language also uses the *-a:la:* marker interchangeably with *ɖa:*, which serves the same function as *ɖa:* does here. The genitives *ɖi:* and *ɖa:* in Punjabi indicate both gender and number. The *ɖi:* is singular feminine, and *ɖa:* is singular masculine.

Genitives are also used in Hindko. *ɖa:, ɖi:* marks the gender and number, similar to Punjabi. These genitives show the possession of nouns before them. Besides this, genitive in 4.3.4c performs the function of English participle adjectives, as discussed in the case of Urdu. In 4.3.4d, the genitive *ɖa:* combined with *sainray* (sleep) functions like an English infinitive, and it would mean "to sleep." In 4.3.4d. *-a:la:* similar to Punjabi *-a:la:* performs attributive functions like *ɖa:* is performing here. In other words, it can be said that NP + GEN + NP results in participle + noun here.

Appositives

The data of the languages' appositives with the description is given in the following:

Appositives in Urdu

a) "mã: ka: apnay baiʈay se: kehnã: keh us ɖin school band ʈha:."
mother (F.SG) (GEN.) her son (M.SG.) to saying that (APP) that day school off was (COP.PST.M.SG.).
'Mother's telling her son that school was off that day.'

b) "yeh haqiqaʈ hai keh walidain apnay bachoon kayliey mehnaʈ kerʈay hain."
its fact (F.SG.) is (COP.SG.) that (APP) parents their children for hardworking do COP.PL.
'It is true that parents do hard working for their children.'

c) "un ka: helf keh voh paspa: nehi: hõ: gey."
theyGEN.M.SG. oath that (APP) they retreat not be COP.PL.
Their oath that they will not retreat.

d) "yeh khabar hai keh chor pakrra giya: ʈha:"
its news COP.SG. that(APP) thief catch be(PST.M.SG.) COP.PST.SG.
'It is news that the thief was caught.'

Appositives in Pahari

a) "ma:u nã: apny puʈray ki ɖasna: keh /khay school us ʈeharay band a:sa:."
mother (GEN.) her son to saying that(APP) school that day off COP.M.SG.
'Mother is saying to her son that school was off that day.'.

b) "yo sach za: keh ma: peyo apnay bachyã: nĩ: galla: mehnaʈ kernay zey."
its true COP.PRS.M.SG. that (APP) parents their children for hard working do (IPFV.PL.) COP.PRS.PL.
'It is true that parents do hard work for their children.'

c) "unã: na: helf keh oh pichey naey hatney."
theyGEN.M.SG. oath that (APP) they back not (PL.) move (M.PL.)
'Their oath that they will not retreat.'

d) "ey khabar zi: keh/khay chor pakrra: wa: asa:."

it news COP.PRS.F.SG. that (APP) thief catch COP.PST.M.SG.
'It is news that the thief was arrested.'

Appositives in Punjabi

a) "ma:u ɖa: apnray puṭray ki: ɖasnra keh us ḍiharay school band̰ si:."
mother her son to saying that (APP) on that day school off COP.PST.F.M.SG.
'Mother's telling her son that school was off that day.'

b) "ey haqiqat̰ ey keh ma: peyo apnray bachya:n lai mehnat̰ karḍay nain."
its true is (COP.PRS.SG.) that (APP) parents their children for hard working do (IPFV) are (COP.PRS.PL.).
'It is true that parents do hard work for their children.'

c) "unhã: ɖa: helf keh oh pichay nĩ: hattanrr gey."
their GEN.M.SG. oath (M.SG.) that (APP) they back not move will (COP.FUT.M.PL.)
'Their oath that they will not retreat.'

d) "eh khabar ey keh chor nappanr gya."
it news is (COP.PRS.SG.) that (APP) thief (M.SG/PL.) catch (INF) (COP.PST.SG.)
'It is news that the thief was arrested.'

Appositives in Hindko

a) "ma:u ɖa: a:pnray puṭray ka: a:khnrã: keh us ṭairrhay school band̰ a:hey."
mother (GEN) GEN.M.SG/PL. her son to saying that (APP) that on that day school off COP.PST.M.SG.
'Mother's telling her son that school was off that day.'

b) "ey sach ṭhia: keh ma: pe: apnray bachyã: a:staṭay mehnat kerḍey."
its true COP.M.SG. that (APP) parents their children for hard work do+COP.PRS.M.PL.
'It is true that parents do hard work for their children.'

c) "unã: ɖa: heluf ṭhia: keh oh pichey nã: hetley."
their GEN.M.SG. oath COP.M.SG. that (APP) they back not move COP.FUT.PL.
'Their oath that they will not retreat.'

d) "khabar ey ḍi: keh chor pakri:/a: ga:"
news it is(COP.SG.) that(APP) thief catch goPFV.PST.M.SG.
'It is news that the thief was arrested.'

Description of Appositives

The data from Urdu exhibits the use of the appositives *keh*. Urdu, Punjabi, and Hindko share the same pattern for apposition. Appositive *keh* is used in these languages. Pahari uses the appositive *keh*, with the phonetic form *khay*, which has the same meaning. The use of *keh* and *khay* is not regular in Pahari.

Table 3
Summary of Appositives

Language	Appositive Form	Notes
Urdu	keh	Consistent use
Punjabi	keh	Consistent use
Hindko	keh	Consistent use
Pahari	keh, khay	No regularity in use

The *-wa:la:* in Urdu

This section discusses the *-wa:la:* marker using the data from the Pakistani languages:

***-wa:la:* in Urdu**

Consider the examples from Urdu incorporating the use of *-wa:la:*:

a) "ab bhi chalney wa:li: chakki:."
nowtoo work (INF.OBL.) PTCL.F.SG/PL. mill (F.SG.)
'The mill which works even now.'

b) "kabba:rr lainay -wa:la: a:ḍmi:."
scrap buy (INF.OBL) PTCL.M.SG. man (M.SG.)
'The man who buys scrap.'

c) "kaba:rr lainay -wa:la:."
scrap buy (INF.OBL.) PTCL.M.SG.
'The one who buys scrap.'

d) "kabba:rr -wa:la:."
scrap SUF.M.SG.
'The scrap man.'

e) "kabba:rr -iya (kabba:rriya)"
scrap man (SUF) scrapman
'The scrap man.'

***-wa:la:/-a:la:* in Pahari**

a) "aja:n vi: chalnay -a:li: chakki:."
nowtoo work (INF.OBL.) PTCL.F. mill
'The mill which works even now.'

b) "kachra: (k)hinḍay -a:la: banḍa:."
scrap buy (INF.OBL.) PTCL.M.SG. man.
'The man who buys scrap.'

c) "kachra: (k)hinḍay -a:la:."
scrap buy (INF.OBL.) PTCL.M.SG.
'The one who buys scrap.'

d) "kachray -a:la:."
scrap (M.SG.) PTCL.M.SG.
'The scrap man.'

***wa:la:/-a:la:* in Punjabi**

a) "ha:lay vi: chalanr r -a:li: `chakki:."
now too work (INF.OBL.) PTCL.F.SG. mill (F.SG.)
'The mill which works even no6325.w.'

b) "kabba:rr lainrr -a:la: banḍa:."
scrap buy (INF.OBL.) PTCL.M.SG. man (M.SG.)
'The man who buys scrap.'

c) "kab:rr lainrr -a:la:."
scrap buy (INF.Obl) PTCL.M.SG.
'The one who buys scrap.'

d) "kabba:rr -a:la:"
scrap SUF.M.SG.
'The scrap man.'

***-wa:la:/-a:la:* in Hindko**

a) "hunrr bi: chalnay -a:li: chakki:."

nowtoo work (INF.OBL.) PTCL.F.SG. mill (F.SG.)
'The mill which works even now.'

b) " kabba:rr kin-nay -a:la: banda:."
scrap buy (INF) PTCL.M.SG. man (M.SG.)
'The man who buys scrap.'

c) "kabba:rr kin-nay -a:la:."
scrap buy (INF.) PTCL.M.SG.
'The one who buys scrap.'

d) "kabba:rr i: (kabba:rri:)."
scrap man (SUF) (scrapman)
'The scrap man.'

Description of *-wa:la:/-a:la:*

In Pakistani languages, the suffix *-wa:la:/-a:la:* plays an important role in nominalization, transforming verbs into nouns that describe people or things associated with a particular action. This linguistic feature *-wa:la:* is consistently observed in Urdu, and *-a:la:* in Pahari, Punjabi, and Hindko, showing similarities and unique variations. The *-wa:la:/-a:la:* marker nominalizes verbs and creates nouns that mean 'the one who' or 'the person who'. This marker also follows gender agreement rules, with *-wa:li:/-a:li:* used for feminine nouns and *-wa:la:/-a:la:* for masculine nouns. For example, in Urdu, *kabba:rr lainay -wa:la:* translates to 'the man who buys scrap metal,' while *kabba:rr lainay -wa:li:* would refer to a woman performing the same action. Each language has distinct forms and usage patterns. Urdu, for example, employs variations such as -wa:la:, -wa:li:, and -iya, the latter of which is interchangeable with -wa:la: in specific contexts. In contrast, Pahari and Punjabi predominantly use *-a:la: and -a:li:, while Hindko has both -a:la: and -i:*. These differences highlight the linguistic diversity within the region. In addition, the frequency of use of *-wa:la:/-a:la:* and its specific vocabulary vary between languages. For example, *kabba:rr* is 'junk' in Urdu, while *kachra:* is used in Pahari. These lexical choices further distinguish the unique identity of each language. In summary, while the *-wa:la:* marker serves a similar grammatical function in Urdu, Pahari, Punjabi, and Hindko, the specific forms, frequencies, and vocabularies differ, reflecting the region's rich linguistic diversity. This interplay of commonality and variation highlights the dynamic nature of language and its capacity to adapt to diverse cultural contexts.

Discussions and Conclusion

Pakistani languages use relatives, participles, genitives, appositives, and the *-wa:la:* markers in their nominalization system. A key structural feature of Punjabi, Pahari, and Hindko is the embedded gender and number agreement within relative markers, which distinguishes them from Urdu and Hindi, where the relative marker "*jo*" remains uniform across animacy distinctions. Hindko and Pahari further exhibit relative-correlative constructions, thereby increasing their syntactic complexity relative to Urdu.

Nishioka and Kumar's (2021, 2025) discussion on Bhojpuri, Magahi, and Maithili provides insight into comparative patterns. Bhojpuri differentiates "who" vs "what" using distinct relative markers, akin to English. In contrast, Punjabi, Pahari, and Hindko employ gender-based distinctions, reflecting a morphosyntactic agreement system rather than an animacy-based opposition. This difference underscores that Western Indo-Aryan languages employ typologically distinct nominalization strategies compared with their Eastern Indo-Aryan counterparts.

Participle constructions also play a crucial role in nominalization. Urdu uniquely inserts a perfect participle of the copula (COP.PFV) between the modifying participle and the head noun, forming a structure like [IPFV/PFV + COP.PFV + N]. This syntactic behavior is absent in Bhojpuri, Magahi, and Maithili, as well as in Punjabi, Pahari, and Hindko, aligning these Pakistani languages with broader Western Indo-Aryan trends.

Furthermore, variations of *-wa:la:/-a:la:* in Punjabi, Pahari, and Hindko originally denote agency. In Urdu, the suffix *-iya:* can substitute *-wa:la:*, leading to lexical nominalization. According to Nishioka's study of lexical nominalization in Hindi (2020: 398-400), Hindi presents additional forms, such as *-ha:ra:*, as seen in *chuurri:ha:ra:* (bangle seller) or *lakarrha:ra:* (woodcutter). While these markers are productive to varying degrees, *-wa:la:* remains the more dominant form across Pakistani languages.

As seen in other nominalization studies of world languages, the analysis of Pakistani languages provides key insights into morphosyntactic and typological contexts, particularly for relative, participial, genitive, and appositive markers. This study identifies the *-wa:la:/-a:la:* nominalization markers and their functional roles, reinforcing language

preservation, pedagogical implications, and theoretical advancements in South Asian linguistics. The identification of these markers may help learners and teachers in learning these languages. Such markers are rarely found in the most-studied languages, such as English.

While nominalization patterns largely align with Urdu, subtle differences emerge upon closer examination. Future investigations could explore the distinction between nominal expressions like *ɖ̪uuɖ̪h -wa:la: maraɖ̪* "the milkman" and infinitive-based constructions such as *a:ne wa:la: pal* "the moment to come" versus *a:ne ka: pal* "the moment of coming". Similarly, structural ambiguities in participle constructions, such as *us larrke: ke: a:ʈ̪e: vaqt̪* "at the time when the boy comes" vs. *us ke: ka:m walli: company* "the company where he works," warrant deeper syntactic analysis.

Since Japanese lacks relative markers, relative clause equivalents often rely on participle non-finite constructions. Further examination of these cross-linguistic differences can refine our understanding of nominalization strategies in Indo-Aryan and other language families. Empirical corpus analysis and speaker testimony could significantly enhance future studies, offering broader insights into cross-linguistic variation in nominalization.

References

Cardona, G., & Jain, D. (2014). *The Indo-Aryan Languages*. (D. J. George Cardona, Ed.) London: Routledge Publishers.

Chalker, S., & Weiner, E. (1994). *Oxford DIctionary of English Grammar.* New York: Oxford University Press.

Comrie, B., & Thompson, S. A. (1985). Lexical Nominalization. *In Language typology and syntactic description* (pp. 349-398). Cambridge University Press.

Foong, H. Y., Karen, G. H., & Wrona, J. (2011). Introduction: *Nominalization strategies in Asian languages.* Retrieved from Research gate: research gate.net

Greenbaum, S., & Nelson, G. (2002). *An Itroduction to English Grammar.* London: Pearson Education.

Khalique, S. (2014). Pahari Language of Azad Kashmir: A Corpus Based Grammatical Description. Muzaffarabad, Azad Jammu and Kashmir, Pakistan.

Lothers, M., & Lothers, L. (2010). Pahari and Pothwari: Asociolinguistic survey. *SIL International.*

Malghani, F. A., Zhu, L., & Weijer, J. V. (2020). Saraiki. *Journel of the International Phonetic Association.*

Masayoshi, S. (n.d.). Nominalization. *In A Handbook of Japenese Syntax. Berlin*: Walter de Gruyter.

Muhammad Nawaz, A. A. (n.d.). *Kashmir Journal of Language Research* (Krl0.

Nawaz, M., & Afsar, A. (2019). A Descriptive Study of Hindko Segmental Features. *Kashmir Journal Of Langugae Research (Krl).*

Roberts, N. B. (2016). ANALYZING SENTENCES: *An Introduction to English Syntax.* New York: Routledge Publishers.

Sardaraz, K., Nusrat, A., & Ab Rashid, R. (2021). Conceptualization of time in Pashto language. *FWU Journal of Social Sciences, 15*(4), 92-113.

Shibatani, M., & Chung, S. Y. (2017). Nominal-based nominalization. *Japanese/Korean Linguistics,* 25, 63-88.

Shibatani, M. (2018). 12 Nominalization in crosslinguistic perspective. *Handbook of Japanese contrastive linguistics, 345-410.*

Shibatani, M. (2019). What is nominalization? Towards the theoretical foundations of nominalization.In Nominalization in languages of the Americas (pp. 15-167). John Benjamins Publishing Company.

Sohail, A., Rashid, H., & Raja, N. A. (2011). A Brief Introduction of Hindko Language. *Language in India,* 471-481.

Abbreviations used are as follows. F = feminine; M = masculine; COP = copula; OBL = oblique; PVF = perfective; IMPF = imperfective; ADV = adverb; PRS = present; PST=past ; FUT = future; SG = singular; PL = plural; 1SG = First person singular; 1FSG = First person female singular; 1MSG = First person male singular; CON=conjunction; AUX=auxiliary; GEN=genitive; APP= applicative; REL=relative; INF=infinitive; COR=correlate