International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

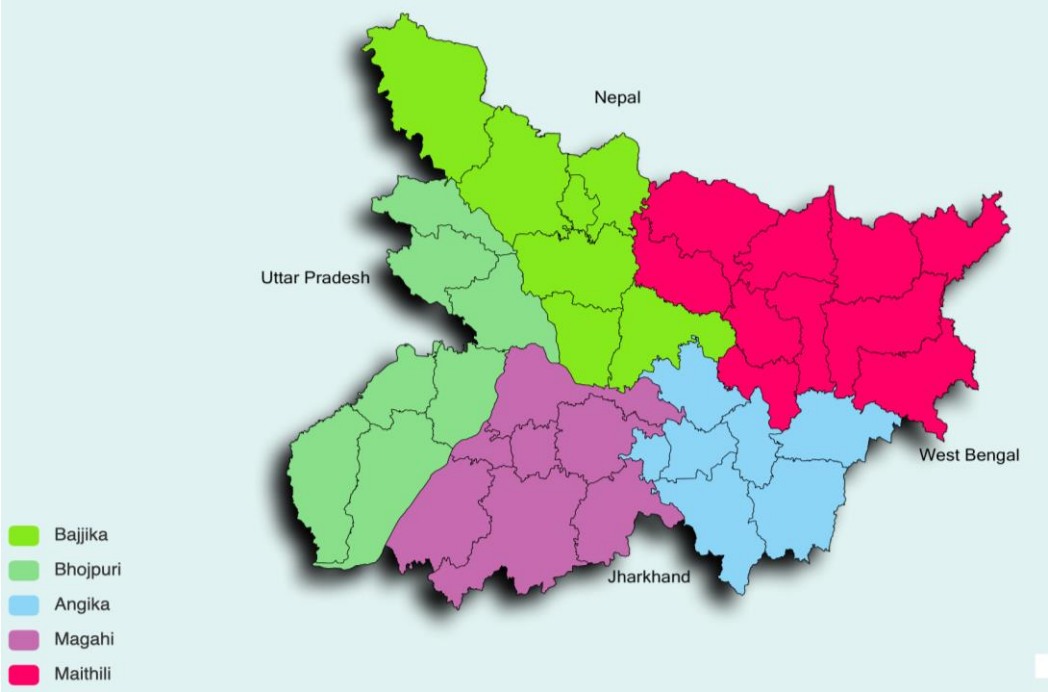
Dialect Contact and Change in Copular Forms in Bihar

Ranjan Kumar

Assistant Professor of English, Department of English, A.N.S College Barh

ABSTRACT

Unlike most Western contexts, urbanization in Asian settings including India is not a unidirectional process (McGee, 1977). People continue to maintain close ties with the homelands through kinship and marriage practices thus impacting the outcomes of contact in ways that might run contrary to the larger expectations. Moreover, multilingual settings may not necessarily be centred on a single dominant centre of power or a prestige variety (Author, in press). In this paper, we report findings from our study of dialect contact in Bihar. Bihar is one of the eastern states of India which has been traditionally home to three closely related Indo-Aryan varieties, namely Magahi, Maithili and Bhojpuri occupying distinct regions as shown on the Map:



1. Linguistic Map of Bihar

Two additional varieties have emerged over time known as Angika and Bajjika. Increased internal mobility across the state due to urbanization has resulted in increased contact across the dialect regions. The focus of the study is on four present copular forms: h-, ch-, chik- and ba- and their distribution across five varieties.

The findings suggest that contact over time has resulted in the redistribution of the copular forms across dialect regions resulting in newer variation and change in the use of *be* forms. There is evidence of an

increase in the use of one of the copular forms in 4/5 varieties among the younger age group; this, however, has not eliminated the use of other forms thus maintaining the distinctness of each of the varieties.
The findings are based on sociolinguistic interviews from 42 speakers representing male (19) and female (23); two age groups (25-30 and 55-60 years old) and 10 dialect regions representing the five varieties as well as secondary data from the historical texts. The data was coded for several internal and external factors such as region, generation, mobility etc. and subjected to multivariate analysis.

Keywords: Geographic Variation, Dialect Boundary, Linguistic Continuum

INTRODUCTION

The present work is a socio-historical comparative study of the five 'Bihari' languages, namely Maithili, Magahi/Magadhi, Bhojpuri, Angika and Bajjika. It is a synchronic as well as diachronic investigation of the nature of relationships among the major languages of Bihar, and the new changes that have taken place since the first linguistic survey of India by Grierson (1903)1.

Grierson (1903) divides Bihar into three main geographic dialect areas namely Maithili, Magahi and Bhojpuri. Maithili is the easternmost dialect of the Bihari language. It is mainly spoken in the Darbhanga district and also in the (British) districts of Champaran, Muzaffarpur, Purnia, Monghyr, north and west of Santhal Parganas and Nepal Tarai (Grierson, 1903: 13). The Maithili speech community is described as essentially a rural community, agrarian and not very mobile (Grierson, 1903: 04).

As far as Magahi is concerned, Grierson considers it a dialect of Maithili rather than a separate dialect based on the similarities it has with Maithili. Magahi is spoken in the districts of Patna, Gaya, Bhagalpur and Monghyr (present Bihar) and Hazaribagh, Ranchi, Manbhum, Singhbhum (now Jharkhand). It is also spoken in the west of Malda (West Bengal), bordering Bihar (Grierson 1903). Magahi is surrounded on the North and northeast by the various forms of Maithili, on the West by Bhojpuri, and on the southeast by the Bengali (Grierson, 1903: 30).

Grierson describes Bhojpuri as the westernmost dialect of the Bihari language. It is mainly spoken in the western parts of the state of Bihar and eastern Uttar Pradesh. It has the highest number of speakers among the three dialects. The Bhojpuri-speaking region is surrounded by Magahi, Maithili and other language regions.

However, Bihar has undergone a series of political and geographical changes over the past 100 years. For instance, there is a new State of Jharkhand carved out of earlier Bihar and new linguistic identities have emerged over time. Similarly, Nagpuria is no longer regarded as a dialect of Bhojpuri as it is in Grierson (1903). Most important, there has been much internal mobility over time.

Further, Maithili is now included in the Eighth Schedule of the Indian Constitution. Maithili is therefore a separate language now and not a dialect (of Bihari as described in Grierson, 1903). Both Magahi and Bhojpuri are now considered dialects of Hindi (Census 2001, 2011), in contrast with Grierson's classification of them as the dialects of Bihari. However, the speakers of these dialects strongly assert their separate identities. Bihar has further seen the rise of two new linguistic identities from within Maithili based on caste and region in Bihar, namely Angika and Bajjika.

1 Grierson's (1903) 66 specimen texts of the three Bihari languages of Maithili, Magahi and Bhojpuri have been coded for phonological, lexical and morpho-syntactic analysis and the results are compared with the modern data to give the present study a diachronic historical perspective.

The purpose of this study is to revisit Grierson's Geographical survey of the Bihari group of languages and understand the nature of the geographic variation as it existed in the region and explore the changes that have taken place since then as a result of various developments as outlined above. The study is inspired by The Atlas of North American English (ANAE)2 which is an analysis of socio-phonetic patterns across the major urban regional dialects of English across North America (Labov, Ash, and Boberg, 2003). However, the focus of the present study is on morpho-syntactic variation and change.

DIALECT BOUNDARIES AND DIALECT CONTINUUM

A dialect continuum is a continuum on the regional dialect axis in which the dialects which are geographically closer to each other are mutually intelligible to each other with the rate of intelligibility going down as the geographic distance gets larger. According to Chambers and Trudgill (1998): If someone travels from one village to another in a certain direction he/she notices linguistic differences that distinguish one region from another. Sometimes these differences will be greater, sometimes smaller but they will be cumulative. The more distant he/she goes from his/her starting point the greater the dialectal differences will become…This type of situation is known as a geographical dialect continuum (Chambers and Trudgill 2004: 5).

German dialect continuum is one of the popular dialect continuums on two recognized literary standards—standard Dutch and standard German having restricted mutual intelligibility. There are some dialects of the languages that form affinity. However today many gaps in intelligibility can be seen given the several centuries of influence that the standard languages have on the regional ones whereas earlier no such gap was found.

The Romance languages such as Portuguese, Spanish, Sicilian, Catalan, Occitan/Provencal, French, Sardinian, Romanian, Romansh, Friulan, other Italian, French, and Ibero-Romance dialects, and others—are another instance of a dialect continuum. The dialects show varying degrees of mutual intelligibility.

The Bihari group of languages based on my earlier study (Kumar, 2019) was found to form a dialect continuum at various levels in terms of various phonological and morpho-syntactic features but some of them also show discrete boundaries.

Real-time studies in sociolinguistics are generally socio-historical in nature. Labov's 1963 monumental work Martha's Vineyard has been re-studied by other researchers after him.

Blake and Josey (2003) performed an analysis of the variable (ay) by utilizing more recent acoustic and social techniques. After comparing the recent data and findings with that of Labov (1963), the study finds no more presence of /ay/centralisation, an indication that a possible reversal of the change is possible. (Blake, Renée; Josey, Meredith 2003).

Pope et al. (2007) conducted a more faithfully reproduced study following Labov's original survey methods and sampling procedure. On the contrary with what Blake and Josey have found in their study, Pope finds that the change was still in progress, with signs of decline in centralisation showing only among the youngest speakers. Pope and her colleagues concluded that Labov's construct of apparent time provides a valid representation of linguistic change (Pope, Meyerhoff and Robert 2007).

2 The **Atlas of North American English** provides the first overall view of the pronunciation and vowel systems of the dialects of the US and Canada. The **Atlas** re-defines the regional dialects of American English on the basis of sound changes active in the 1990s and draws new boundaries reflecting those changes. It is based on a telephone survey of 762 local speakers, representing all the urbanised areas of North America. For more information, see De Gruyter Mouton, 2005.

This present study reports on sociohistorical variation and change that have taken place in the group of the Bihari languages over a century, by revisiting Grierson (1903) during (2019-21).

TRANSMISSION AND DIFFUSION

Two different types of language change transmission and diffusion are distinguished in Labov (2007). Transmission according to Labov is the "unbroken sequence of native-language acquisition by children" and it is "reflected in the family tree model of change". And the transmission which is imperfect results in "internal language change" or "change from below" and which advances through generational incrementation. Diffusion, which is reflected in the wave model of language change, results from language contact among adults, who replicate detailed features of language structure poorly (Labov, 2007).

The present study deals with both the transmission of the linguistic forms across generations within a speech community over the years and the diffusion of certain morpho-syntactic features from one linguistic area and speech community to another linguistic area and speech community in dialect contact situations that are found between the Maithili and Bajjika, Bajjika and Bhojpuri, Maithili and Angika, Bhojpuri and Awadhi and Angika and Bengali and so on.

Bihari languages belong to the Eastern Indo-Aryan languages, mainly spoken in the states of Bihar, Jharkhand, West Bengal and Uttar Pradesh in India and also in Nepal. According to Grierson (1903: 8-9) there are three geographic dialects of the Bihari language, namely Maithili, Magahi and Bhojpuri. These dialects have many sub-dialects too. Currently, Maithili is the only Bihari language which is a scheduled language. The other two are considered as the dialects of Hindi. Recently due to Language movements against Maithili, two new linguistic identities have arisen, namely Angika and Bajjika. So unofficially there are five distinct mother-tongues present in the state of Bihar.

The classification of the Bihari languages can be done in the following way:

Indo-European
↓
Indo-Aryan
↓
Eastern
↓
Bihari
↓
Maithili
Magahi
Bhojpuri
Angika
Bajjika

The important characteristics of the Bihari languages are the fact that on several levels they form a dialect continuum and at a few levels they share discrete distinct boundaries. My earlier study (Kumar 2019) shows that in terms of phonology, all three dialects of the Bihari language have affinity. More than 70% of vocabulary is shared among the languages with Maithili vocabulary being more sanskritized than the

other two dialects. Overall Number and gender are not marked in all three dialects with Magahi and Bhojpuri showing some variation. Most of the pronouns are also shared across the dialects. The linear sequences of the morphemes in the verbal conjugation are the same. The case marking patterns are the same across the dialects where Nominative remains unmarked, Dative is marked by -ke, genitive by -k, locative by -me and ablative and instrumental by -se. Keeping these points in view there is an impression that three Bihari dialects are more or less linguistically similar and they form a dialect continuum (still to be tested further in the paper in the light of apparent time data).
However, verb patterns including agreement and honorificity3 separate the three languages in various ways. There are also both lexical similarities and differences.

The objectives of the current research are-

1. At first, the study provides a systematic corpus-based quantitative analysis of variation in the use and functions of various *be* forms in each of the five languages. This forms the basis for the next set of objectives listed in (ii) and (iii).
2. A second objective is to state the nature of relationships in terms of whether these languages form a continuum or discrete boundaries. My earlier study (Kumar, 2019a) suggested that the languages share many affinities at various levels, but some of them also show uniqueness.
3. A third objective is to account for changes over time and explanation in terms of contributory factors: contact-induced or internal: The Bihari group of languages have undergone several changes over the past 100 years. Grierson in his Linguistic Survey of India (1903: 8) classifies the "Bihari" language in three dialect areas of Maithili, Magahi and Bhojpuri. Since then due to socio-political and geographic changes that have occurred over the years, "Bihari" saw the emergence of two new identities namely Angika and Bajjika from within Maithili.

DATA AND METHODOLOGY

The present dissertation is based on two kinds of data:

4. The early 20th-century texts from the Linguistic Survey of India were compiled and edited by George Abraham Grierson in 1903. The data comprises 66 specimens representing the three dialects and sub-dialects of the Bihari language namely Maithili (and its sub-dialects Chhika-chhiki Boli which is known as Angika today, and western Maithili which is known as Bajjika), Bhojpuri and Magahi.

(ii) The apparent time conversational data on Maithili, Magahi, Bhojpuri, Angika and Bajjika, collected through interviews, questionnaires and narratives from Madhubani and Darbhanga-the Maithili regions; Bodhgaya and Patna-the Magahi regions; Vaishali and Muzaffarpur-the Bajjika regions; Chhapra and Sonpur-the Bhojpuri regions (Saran district); and Bhagalpur and Begusarai-the Angika regions in the state of Bihar. The data was collected from 42 speakers of 5 languages spoken in 10 language regions, representing two age groups generation and older generation. The goals of the selection of the language regions to be sampled in this dissertation are twofold: firstly, it intends to cover a wide range of dialect areas of Bihar; and secondly to investigate the dialect boundaries between the dialect regions.

3 The verbs in all the three languages inflect for different honorificity grades such as NH, MH and H but the pronominal representing them are cross linguistically constrained. Maithili has three layers of pronouns in 2nd person while the other two dialects have two layers only: Nonhonorific and Honorific.

GOALS OF THE PRESENT PAPER

Covering the ten dialect regions makes it possible for us to broadly divide Bihar into five dialect regions, in much the way as in *The Atlas of North American English North America* as a whole is divided into multiple dialect regions. Our goal is to investigate the nature of relationships between the dialects and regions as a whole. The current work tries to define the overall morphosyntactic features of the concerned languages from a socio-historical dialectological point of view, however, a more detailed picture of the dialectological status of the Bihari speech communities is needed in future.

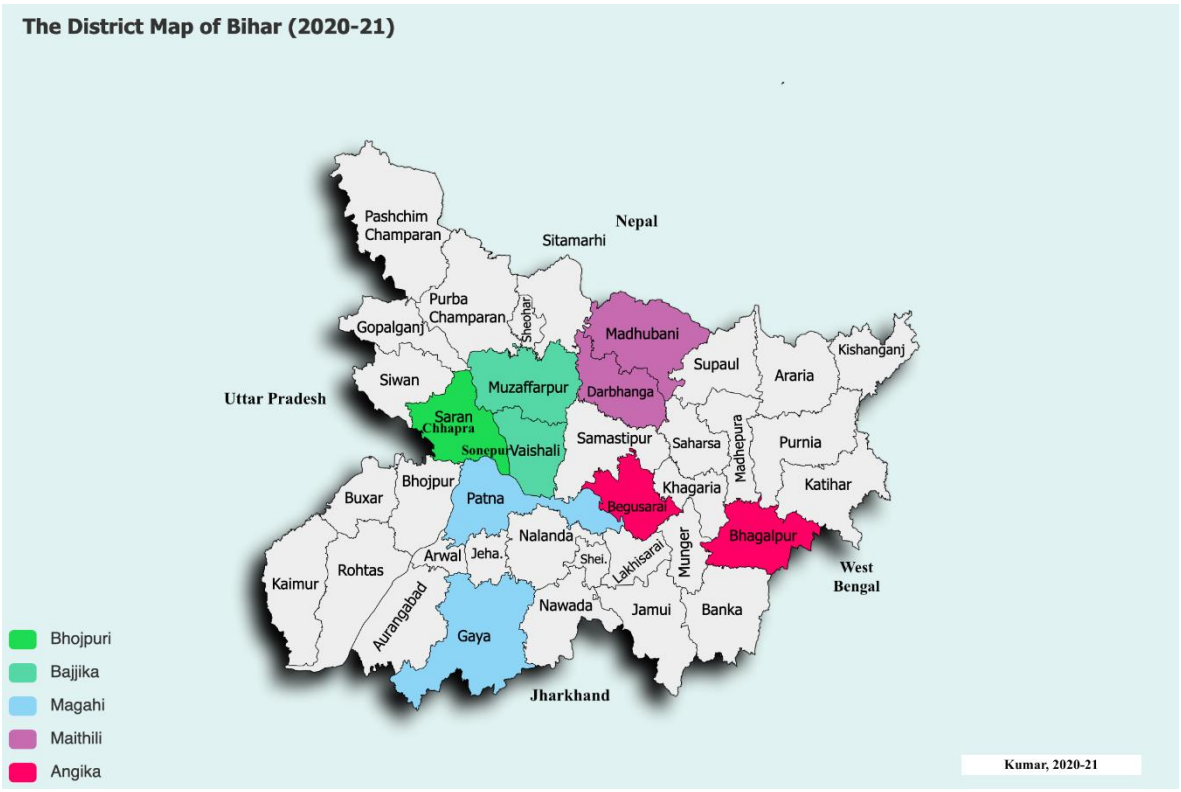
The present study has made use of the following types of data:

- We have conducted face-to-face sociolinguistic interviews first in selected medium-sized cities to narrow the gaps left by the earlier study (Grierson 1903)
- We have conducted a few telephone interviews in the areas where it was not possible to conduct sociolinguistic fieldwork for some reason.
- And finally, we also conducted some additional in-person interviews in certain communities which were near the dialect boundaries for the sake of getting more contact-induced features.

Thus the mixed methodology used in the current work allowed us to achieve both of our goals-

1. Sampling both a geographically broad set of communities.
2. And sampling communities near dialect boundaries in particular so that we can get more information.

The present study collected data from both urban and rural areas. The urban centres that we have taken under consideration are Patna, Muzaffarpur, Hajipur, Sonpur and Darbhanga cities whereas in rural settlements we have done fieldwork in several areas such as Bhagalpur, Madhubani, Chhapra and Begusarai as shown in the map below:



A total of 9 speakers were interviewed from the Maithili-speaking regions of Darbhanga and Madhubani; five speakers from the district of Madhubani and 4 speakers from Darbhanga. These two dialect regions

border each other. A total of 312 tokens of present-tense copulas in Maithili are statistically analyzed in the present study.

A total of 9 speakers from Magahi dialect regions were selected for the current study. Out of which 4 speakers were from the district of Gaya and 5 of Patna. Patna borders Vaishali, a Bajjika region to its north and so it is a dialect contact region. The total number of tokens of present tense copulas is 221.

A total of 9 speakers from the Bajjika regions were selected; five speakers were interviewed from the Hajipur region and four speakers from Muzaffarpur (which borders Darbhanga).

A total of seven speakers from the Bhojpuri-speaking regions were selected for the current study. The two cities under study are Chhapra which stands in the centre of the Saran district and Sonpur which stands on the eastern edge of the district of Saran and is bordered by Hajipur on the East, a Bajjika region. Four speakers from the Chhapra and 3 speakers from Sonpur have been selected for the current study.

The two Angika speech areas under study are Bhagalpur which borders the Bengali-speaking area on the eastern edge and Begusarai which is closer to the Bajjika speech regions. A total of 8 speakers were selected for the current study from the Angika regions: 4-4 from each dialect region. The total number of tokens of the present tense copulas is 385.

In total, data from 42 speakers have been used in the current study. Out of the 42 speakers, the number of male speakers is 19 and that of females is 23. The ten dialect areas under study represent five languages of Bihar, namely Maithili, Magahi, Bhojpuri, Angika and Bajjika. The ratio of language and region is two dialect areas per language.

The total duration of recording is 2,182 minutes; out of which 421 minutes have been utilised for the current study. 2530 tokens of be verbs have been taken for the current study. The study, overall uses 882 tokens of past tense copular verbs and 1648 tokens of present tense copular verbs.

VARIABLES UNDER STUDY

This is a quantitative socio-historical study of copulas (chh~h~chhik~ba in present) and (rah~chh-al~h-al in past) in the five languages of Bihar. The variables were coded for multiple factors such as types of copular constructions (existential, equative, adjectival, possessive and locational); person (1st, 2nd and 3rd), honorificity (honorific and non-honorific); humanness (human and non-human); age (old generation [55-60 years] and young generation [25-30 years]; gender (male and female); tense (present, past and future); languages (Maithili, Magahi, Bhojpuri, Angika and Bajjika); regions (Madhubani, Darbhanga, Chhapra, Bodhgaya, Hajipur, Bhagalpur, Sonpur, Muzaffarpur, Begusarai and Patna); mobility (no mobility, 1-2 years mobility, 3-4 years mobility and 5+ years mobility).

A Variationist Analysis of Copular Verb

The languages of the world have their respective copula typology. There can be just one verb like in English (to be) or French (être) or German (sein), or there can be multiple copular verbs like in Spanish/Portuguese (*ser* and *estar*), or some Bantu languages such as Navajo, Lakhota, Burmese, Thai, Khmer. Some languages have an affix that functions like a copular verb (mainly a suffix) for example in Turkic languages, Korean, Beja or Inuit languages, in which the suffix is attached to the noun. In some languages, it can be a linking demonstrative or pronoun like in Arabic (hiya) or Quechua (kay, which means ‘this’). Some languages of the world have a mixed strategy like Japanese, Polish, Czech and Slovak, in which, according to the context, different strategies may be used. A particle is used in Austronesian languages. Finally, there may not be any copular verb at all. Many languages have zero copula in some contexts, like Russian and Turkish.

Copula constructions have been the focus of many previous studies—for example, the early work of Meillet, (1906) and Benveniste, (1950). Other works include the many language-specific papers in the series of volumes edited by Verhaar, (1967-72). More recently semantically based studies such as Declerck, (1988) and Hengeveld, (1992), as well as Stassen's, (1997) are interesting typological studies of intransitive predication, including (some) copula constructions.

The copular verbs discussed in this paper include a variety of forms such as *h*, *chh*, *chhik* and *ba* in the present tense which function as copula in a variety of non-verbal predicates or as auxiliaries accompanied by another verb. However, this covers only the copular function of the *be* forms. As earlier mentioned, there are a total of four distinct be forms in present that function as copular verbs in the languages of Bihar as shown in the following table:

Table 1: Overall Distribution of the Copulas[4] across the five Languages (Present Tense)

Languages	*h*	*chh*	*chhik*	*ba*	Total
Bhojpuri	73 (22.7%)	0	0	248 (77.3%)	321
Maithili	29 (9.2%)	283 (89.8%)	0	0	315
Magahi	221 (100%)	0	0	0	221
Bajjika	374 (93.5%)	26 (6.5%)	0	0	400
Angika	6 (1.5%)	214 (54.7%)	171 (43.7%)	0	391
Total	703	523	171	251	1648

This is very much important. We can see the wide spread of *h* forms. *H* is the most widely used copular verb in Magahi and Bajjika followed by Bhojpuri. Maithili and Angika prefer *chh*. This indicates that the Bihari languages form a linguistic continuum so far as the *h* form is concerned; they are found in all the languages. What is more important here is that *h* is categorical in Magahi. It is the only present tense *be* verb in the language. Next *h* is almost nearly categorically found in the other neighbouring language, Bajjika. So Magahi and Bajjika more or less form affinities as we find that *h* is the dominant copula in both languages. Bhojpuri is another language in which we can see a good number of tokens of *h*. So far as Maithili is concerned it has some *h* newly introduced while Angika is the language that has only six tokens of *h*. So we find that the linguistic peak of *h* lies in Magahi followed by Bajjika. In the other three languages, the *h* form alternates with the other dominant copula of the concerned language with Bhojpuri retaining a good number of *h*; there is a decrease in Maithili and even more decrease, just marginally found

[4] All the Bihari languages have inflectional morphology. Tense-aspect-person-honorificity markers are attached to the main verbs, auxiliary verbs or copular verbs. The verbal paradigm follows the pattern: VP: VERB STEM+(ASPECT) + (AUX)+(TENSE)+(PERSON)+(HONORIFICITY).

in Angika. This is evidence that the *h* form is associated with and native to Magahi, Bajjika, and Bhojpuri to some extent; while they have been newly introduced in Maithili and Angika.

The other copula *chh* is non-existent in Bhojpuri and Magahi. So we can say that these two languages stand distinct as far as *chh* is concerned. Maithili has the most number of *chh* tokens, followed by Angika and Bajjika. Maithili stands as a linguistic peak so far as *chh* is concerned followed by Angika. Bajjika on the other hand has retained a few tokens of *chh*. There is an important point to note that Bajjika and Angika historically have been considered dialects of Maithili. And historical texts like Grierson (10903) attest more *chh* than the other copulas in these three languages. Currently what we find is that somehow Maithili and Angika retain the historical pattern while Bajjika has changed drastically over the years. It was historically a *chh*-dominant language but over the years it has become a *h*-dominant language. Some of the most significant factors that are important are —(1) The geographical reorganisation of the districts —earlier Muzaffarpur and Vaishali were part of the old Mithila but from the 1950s onwards they are no longer considered part of the Mithila. (2) The speaking region Vaishali is bounded on the South by Magahi, a *h* exclusive language region. Further Vaishali is connected to Patna by the historical Gandhi Setu bridge and so there is a dialect contact between both the speech communities.

Bajjika has however retained some *chh* as Muzaffarpur stands near Mithila (Darbhanga) an evidence that Bajjika is not completely out of the influence of Maithili. So we see that language boundaries condition the copula variation.

Angika follows the Maithili pattern partly with (54%) *chh* and partly it has its distinguishing copula *chhik* (44%). So unlike Bajjika Angika has not broken away its ties with Maithili. Broadly it uses the typical be verb *chh* which alternates with *chhik*. So the broad proposal that Angika arose out of Maithili as a non-Brahmin linguistic identity seems to be plausible.

Chhik is found only in Angika (Angika exclusive) and it might be further a marker of the new non-Brahmin Angika identity.[5]

Ba is restricted to Bhojpuri. *Ba* is indeed a Western Bihari feature to be found exclusively in the Bhojpuri language of Bihar and Awadhi language of Eastern Uttar Pradesh. Historically we have found an alternation between *ba* and *h* in Bhojpuri. *Ba* is the dominant pattern while *h* is the alternative pattern found in Bhojpuri. The *ba* form is what distinguishes the Bhojpuri from the rest of the languages of the group in the same way as *chhik* distinguishes the Angika language.

The conclusion is that all the Bihari languages form a linguistic continuum so far as some features are concerned (*h* is shared across all the languages). The three Bihari languages of Maithili, Angika and Bajjika further form a linguistic continuum as the *chh* form is shared across these languages. Considering *h* and *chh*, the main distinction among the three languages can be made about frequency of occurrence as discussed above.

Apart from these two shared be forms, we have found two other be forms— *chhik* and *ba* which form a discrete distinct linguistic boundary. *Chhik* is restricted to Angika while *ba* is restricted to Bhojpuri.

The final verdict is that Bihari languages form a linguistic continuum[6] so far as *chh* and *h* are concerned while they form discrete distinct linguistic boundaries so far as *chhik* and *ba* are concerned.

[5] More investigation in the area of caste and language identity is need in the area to consolidate the hypothesis.

[6] The findings of the paper match with some of the classical dialectological studies such as ANAE.

Quantitative Analysis of Copula in Maithili, Angika and Bajjika

Since *chh* has been found in the three languages namely Maithili, Angika and Bajjika and it alternates with one another copular verb in these languages, the Varbrul program will select just these three languages. So below are the statistical details and analysis and comparison of the three languages namely Maithili, Angika and Bajjika:

The factor groups that have been eliminated while stepping out are the following:

1. Person
2. Honorificity
3. Humanness
4. Language
5. Gender

The factor groups that have been selected as significant are the following:

1. Predicate types
2. Region
3. Generation

Factors	Copula	Probability	Tokens	Percentage	Total
1. Predicate-types					
One place predicates	*chh*	0.78	164	75.6	217
Two place predicates		0.41	359	40.8	880
2. Region					
Madhubani	*chh*	0.98	200	99.0	202
Darbhanga		0.90	83	75.5	110
Bhagalpur		0.68	142	58.0	245
Begusarai		0.67	72	51.4	140
Muzaffarpur		0.20	24	14.7	163
Hajipur		0.01	2	0.8	237
3. Generation					
Older generation	*chh*	0.72	363	70.1	518
Younger generation		0.29	160	27.6	579

Table: 2 Quantitative Analysis of Copula in Maithili, Angika and Bajjika

Convergence at Iteration 5
Input 0.286
Log-likelihood = -881.957 Significance = 0.004
We can see that *chh* is more likely to occur in one-place predicates (0.78) than two-place predicates (0.41) in contrast with *h* whose pattern is just the reverse of this.
We find the region to be significant as *chh* is more likely to occur in Madhubani (0.98) and Darbhanga (0.90) than in Bhagalpur (0.68) and Begusarai (0.67). More or less these four areas form a linguistic continuum and the languages show affinity. *Chh* is partially found in Muzaffarpur (0.20) and alternates with the dominant form *h* while coming to Hajipur (0.01) *chh* almost loses its existence. So Geography stands very significant. The four areas of Madhubani, Darbhanga, Bhagalpur and Begusarai have a good number of *chh* while Muzaffarpur as it is closer to Darbhanga, due to its geographic proximity retains some *chh*, while Hajipur almost is *chh*-less as it is fairly distant from Muzaffarpur and Darbhanga.
So taking *chh* as a feature we can say that the region stands significant. Angika and Maithili regions form affinity and follow the Mithila pattern while Bajjika breaks away from Maithili and follows the Magadh pattern while retaining some of the Maithili patterns, especially in Muzaffarpur.

Conclusion and Discussions

The major findings of the study are that for some of the features such as the present tense copula *h* the Bihari languages form a continuum and for some of the features such as *ba* they form distinct discrete boundaries. In the present tense, we have found that the distribution of the copulas are conditioned by copula types; *h* is more likely to occur in adjectival copular construction (0.69) followed by possessive (0.52). Then it occurs more likely in existential (0.42) followed by locational construction (0.36) and equatives (0.36). So *h* is an all-purpose copula that occurs in all the copular classes with only differences in frequency and the likelihood that certain constructions like adjectivals and possessives are more likely to use *h* than other constructions.
The next factor group that constrains the distribution of the copulas is person, 3rd person is more likely to use *h* (0.55) than any other person. 2nd person and 1st person are less likely to have *h*: (0.27) in 1st person and (0.13) in 2nd person. There has been a relation between copular verbs and 3rd person and it has been historically attested in Grierson (1903).
Another factor that constrains the copulas is Humanness. Overall *H* is more likely to be used in human contexts (0.62) than non-human contexts (0.38). Further, it is subject to variation and true of Magahi, Bajjika and Bhojpuri from where most of our *h* data have come from, but in Angika and Maithili *h* is more restricted to 3rd person nonhuman contexts.
Language itself has stood as a significant factor. *H* is present everywhere in the language. This indicates that the Bihari languages form a linguistic continuum. What is more important here is that *h* is categorical in Magahi (0.99). Next, *h* is nearly categorically found in Bajjika (0.91). So, Magahi and Bajjika more or less form a closeness. Bhojpuri is another language in which we find a good probability of the occurrence of h (0.22). So far as Maithili is concerned it has some *h* —a newly introduced copula (0.10) while Angika is the language that has only six tokens of *h* (0.04). So, language has stood significant and the frequency and likelihood of the occurrence of *h* is subject to cross-linguistic variation.
So far as *chh* is concerned it is the representative form of Maithili (90.7%) followed by Angika (55.6%) and is partially found in Bajjika (6.5%). So more or less we see that considering *chh* in mind, it is plausible to think that Maithili, Angika and Bajjika form a linguistic continuum.

The next important finding is that Language interacts with the region. In other words, we can say that linguistic variation is conditioned strongly by region. The *h* form is found everywhere but not equally. It is largely concentrated in the two Magahi regions of Patna (0.97) and Bodhgaya (0.99) and two Bajjika regions of Hajipur (0.99) and Muzaffarpur (0.90) and one Bhojpuri region of Sonpur (0.67). In the rest of the regions, *h* is marginally found: Darbhanga (0.34), Madhubani (0.01), Begusarai (0.04) and Bhagalpur (0.01).

Chh on the other hand is more likely to occur in Madhubani (0.98) and Darbhanga (0.90), followed by Bhagalpur (0.68) and Begusarai (0.67). More or less these four areas (representing the languages: Maithili and Angika) form a continuum. *Chh* is partially found in Muzaffarpur (0.20) and alternates with the dominant form *h* while coming to Hajipur (0.01) *chh* almost loses its existence. The next factor group that has stood significant is the generation. Overall we find that *h* is more used by the younger generation than the older one. The youngsters (0.62) have more probability to use *h* than the old age-groups (0.32). This indicates that all the languages of Bihar are undergoing a generational change. However the "change in progress" is restricted to Maithili and Angika.

Bibliography and References

5. Blake, R.; Josey, M. The /Ay/ Diphthong in a Martha's Vineyard Community: What Can We Say 40 Years After Labov? Lang. Soc. 2003, 32 (4), 451–485. DOI: 10.1017/S0047404503324017
6. Blaxter, T. T.; Trudgill, P. On Case Loss and Svarabhakti Vowels: The Sociolinguistic Typology and Geolinguistics of Simplification in North Germanic. J. Ling. Geogr. 2019, 7 (1), 1–13. DOI: 10.1017/jlg.2019.3
7. Britain, D. Dialect Contact and Phonological Reallocation: "Canadian Raising" in the English Fens. Lang. Soc. 1997, 26 (1), 15–46. DOI: 10.1017/S0047404500019394
8. Britain, D.; Trudgill, P. Migration, New-Dialect Formation and Sociolinguistic Refunctionalisation: Reallocation as an Outcome of Dialect Contact. Trans. Philol. Soc. 1999, 97 (2), 245–256. DOI: 10.1111/1467-968X.00050
9. Britain, D. Dialect Contact and past BE in the English Fens, 2001.
10. Chambers, J. K.; Trudgill, P. Dialectology (2nd ed., Cambridge Textbooks in Linguistics); Cambridge University Press, 1998. DOI: 10.1017/CBO9780511805103
11. Cheshire, J.; Kerswill, P.; Fox, S.; Torgersen, E. Contact, the Feature Pool and the Speech Community: The Emergence of Multicultural London English. J. Sociolinguistics 2011, 15 (2), 151–196. DOI: 10.1111/j.1467-9841.2011.00478.x
12. Grierson, G. A. 1883–87. Seven Grammars of the Dialects and Subdialects of the Bihari Language, Parts. I–VIII; Bharatiya Publishing House.
13. Grierson, G. A. 1903. Linguistic Survey of India, Vol. V. Indo-Aryan family, eastern group, Part II: Specimens of the Bihari and Oriya languages. Calcutta: Office of the Superintendent of Government Printing. Reprinted in 1968 by Delhi: Motilal Banarsidass.
14. Gumperz, J. J. Dialect Differences and Social Stratification in a North Indian Village. Am. Anthropol. 1958, 60 (4), 668–682. DOI: 10.1525/aa.1958.60.4.02a00050
15. Jones, M.; Tagliamonte, S. From Somerset to Samaná: Preverbal Did in the Voyage of English. Lang. Var. Change 2004, 16 (2), 93–126. DOI: 10.1017/S0954394504162029
16. Labov, W. Principles of Linguistic Change: Social Factors, Description, 1994. http://www.wiley.com (accessed 27-04-2020)

17. Labov, W. Transmission and Diffusion. Language 2007, 83 (2), 344–387. DOI: 10.1353/lan.2007.0082
18. Milroy, J. (1988a). Social Network and Prestige Arguments in Sociolinguistics. In Sociolinguistics Today: International Perspectives; Bolton, K., Kwok, H., Eds.; Routledge, 1991.
19. Pandit, P. B. 1972b. India as a Sociolinguistic Area. (Gune Memorial Lectures, 1968). Gareshkhind; University of Poona.
20. Payne, T. E. Describing Morphosyntax: A Guide for Field Linguists; Cambridge University Press, 1997. DOI: 10.1017/CBO9780511805066
21. Pope, J.; Meyerhoff, M.; Ladd, D. R. Forty Years of Language Change on Martha's Vineyard. Language 2007, 83 (3), 615–627. http://www.jstor.org/stable/40070904 (accessed April 29, 2021). DOI: 10.1353/lan.2007.0117
22. Sankoff, G.; Cedergren, H. Some Results of a Sociolinguistic Study of Montreal French. In Linguistic Diversity in Canadian Society; Darnell, R., Ed.; Linguistic Research, 1971.
23. Sankoff, G. The Grammaticalization of Tense and Aspect in Tok Pisin and Sranan. Lang. Var. Change 1990, 2 (3), 295–312. DOI: 10.1017/S0954394500000387
24. Satyanath, S. Genealogies of Sociolinguistics in India. J. Sociolinguistics 2021, 25 (5), 762–784. DOI: 10.1111/josl.12496
25. Satyanath, S. Language Variation and Change: The Indian Experience. In Globalising Sociolinguistics; Smarkman, D., Heinreich, P., Eds.; Routledge, 2015.
26. Tagliamonte, S.; Hudson, R. Be Like et al. Beyond America: The Quotative System in British and Canadian Youth. J. Sociolinguistics 1999, 3 (2), 147–172. DOI: 10.1111/1467-9481.00070