

# Linguistic Variation and Change in Bhojpuri

**Ranjan Kumar**

Assistant Professor of English ANS College, Barh

## ABSTRACT

Bhojpuri is a Bihari language, spoken primarily in the western part of the state of Bihar and the eastern part of the state of Uttar Pradesh. As per Grierson (1903), two copular forms have been attested in Bhojpuri in the Present Tense: *-h* and *-ba*. *H* was the minority form that was restricted in use, and *ba* was the dominant form. Over the years, due to mobility and contact with the neighbouring areas of Hajipur and Patna, it is proposed that *-h* forms have seen a rise in the daily use and frequency of speech. We have taken a diachronic sociohistorical perspective to deal with the issue of dialect contact and change in this paper. Multivariate analysis is done as a standard statistical tool to analyse different socio-regional factors.

**Keywords:** Dialect contact, Regional Variation, Dialect Boundary and Dialect Continuum

## INTRODUCTION

There are only a few studies that have been specifically oriented towards historical sociolinguistics in India. These include Satyanath (2003, 2006, 2010, 2011). Other more recent and important sociolinguistic studies focusing on Indian speech communities are Satyanath and Laskar (2008), Dey (2010), Sunny (2013), and (Suokhrie, 2015 & 2020). These all make use of some dimension of time (apparent or real) and space. For instance, some of these studies have focused on diffusion of sound shifts across Assam and Bengal (Satyanath, 2010, 2011), contact and diffusion of morphosyntactic features from Assam to the neighbouring regions (Satyanath, 2018, 2021), mobility and contact (Dey, 2010; Bhattacharya, 2017; Imchen, 2017; Satyanath and Bhattacharya, 2018; Sharma, 2017; and Suokhrie, 2016 & 2020).

Satyanath (2015b; 2021; in press) has indicated that Variationist Sociolinguistics has so far confined itself to Western contexts and English. A majority of the Western settings, according to Satyanath, represent monolingual and monocultural settings, which are very different from the Indian settings and other settings in Asia-Pacific, which are more multilingual and multicultural and present their dynamics. Various studies on Indian speech communities, as mentioned above, present factors that are locally relevant, and many of these are unique to India. Moving on to dialect geography, which the present study addresses, such studies have emerged essentially from the West.

The paper is a variationist analysis of copula in Bhojpuri. Bhojpuri is a Bihari language spoken by approx. 37.8 million population in India, primarily in the western part of the state of Bihar and eastern part of the state of Uttar Pradesh and some adjacent regions of Madhya Pradesh (Ethnologue). Currently, it is not an official language, it has been kept under the umbrella term 'Hindi' and is one of the dialects of Hindi under the 8th schedule of the Constitution of India. However, it is the most popular mass media language and language of the cinemas in Bihar.

Bhojpuri is the most westerly dialect of the Bihari language (Grierson 1903).

The Bhojpuri-speaking region covers an area of approx seventy-three thousand square kilometres in India and Nepal (Dinesh 2007). The Bhojpuri-speaking regions have the Awadhi-speaking regions to the West,

the Nepali-speaking regions to the North, the Magahi, Maithili and Bajjika-speaking regions to the East and the Magahi and Bagheli-speaking regions to the South (based on Bhojpuri Ethnologue 2009).

According to Grierson (1903), Bhojpuri has the following sub-dialects: **Standard Bhojpuri** is spoken mainly in the districts of Shahbad, Ballia and Ghazipur and the Doab of Gandak and Gogra in the state of Bihar.

**Northern Bhojpuri** is common in the areas of Gorakhpur, Deoria, Kushinagar, and Maharajganj districts and Basti regions in the state of Uttar Pradesh.

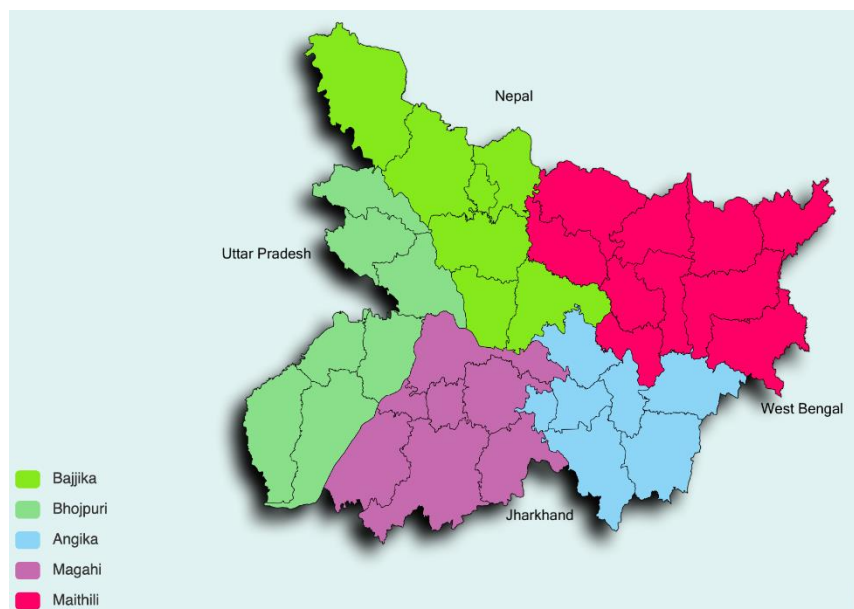
**Western Bhojpuri** is spoken in districts of Faizabad, Azamgarh, Jaunpur, Varanasi, the western half of Ghazipur and South-Gangetic Mirzapur in the state of Uttar Pradesh.

**Nagpuria Bhojpuri** is the southernmost dialect, spoken in the Chota Nagpur Plateau of Jharkhand, particularly in the parts of Palamau and Ranchi. It has more Magahi influence. It is sometimes referred to as ‘Sadari’ as well.

**Madhesi** is spoken in the district of Champaran in the state of Bihar.

Bhojpuri has been a widely used language in terms of speakers and geographical distribution, and that is the reason why even Grierson (1903) states that “the area covered by Bhojpuri is about fifty-thousand square miles. Natively, it has been spoken by more than 20 million populations as compared with 6 million Magahi and 10 million Maithili speakers. So, in terms of numbers, it is more important than the other Bihari dialects.”

The current study is based on a variety spoken in Saran district of the state of Bihar. Two dialect regions of Chhapra, a semi-urban area which is a typical Bhojpuri region and Sonpur borders the ‘Bajjikanchal’ and is a region very important from a dialect contact perspective:



**Map 1: Linguistic Map of Bihar (Bhojpuri)**

## RESEARCH METHODOLOGY

The present dissertation is based on two kinds of data:

1. 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. This paper covers only the Bhojpuri dialect of the Bihari Language.

2. The apparent time conversational data on Bhojpuri, collected through interviews, questionnaires and narratives from the Chhapra and Sonpur-the Bhojpuri regions (Saran district) in the state of Bihar. The data was collected from two age groups: the young generation and the older generation.

For collecting data, we have used the following data collection techniques:

- The questionnaires
- The sociolinguistic interviews

Questionnaires were designed, keeping verbal morphology in mind, and our purpose was to get as much information about the different verbal paradigms as possible. To fulfil this purpose, our questionnaire contains a list of lexical verbs which have been arranged according to their function in the language.

The interview method is considered an important technique for data collection in linguistics. In the present study, we also adopted the interview technique in data collection. We have used this technique for interviewing informants for several casual and conversational topics.

The dialect regions are discussed in brief below:

**Sonpur:** Sonpur is another important city situated on the bank of the River Gandak in the District of Saran. It is closer to Hajipur and is just 3 km away from Hajipur on the western edge. Sonpur is famous for Asia's largest cattle fair, which takes place in December every year. It is a historical and cultural place famous for Hariharnath Temple, with which some mythology of Lord Vishnu is associated.

Sonpur<sup>1</sup> is almost 25 km from Patna and, 58 km from Muzaffarpur in Bihar & 60 km from Chhapra, the headquarters of Saran district. The languages spoken in the area are Hindi and Bhojpuri.

**Chhapra:** Chhapra is the main city and the headquarters of the Saran District. It falls near the junction of the Ghaghara River and the Ganges River.<sup>2</sup> The Road and rail connections to the neighbouring regions are well-connected, and the city is an agricultural trade centre. The languages spoken in the region are Hindi and Bhojpuri.

## QUANTITATIVE ANALYSIS OF BHOJPURI LANGUAGE

Bhojpuri has the following copular verbs: two copulas in the present, one in the past and a single copula in the future tenses:

**Present Tense:** *h* (restricted to 3rd person)

*ba* (inflects as **bani** for 1st and 2nd honorific person, and inflects for **bare** for 2nd person non-honorific and **bar-an** for 3rd person honorific)

**Past Tense:** *rah* (*rah* gets inflected for all persons and honorificity)

**Future Tense:** *ho* (*ho* gets inflected for person and honorificity)

### Copula Variation in Past and Future Tense

Past and future tenses each have a single copula, and thus, they are categorical and so will not be discussed in much detail as there is no variation in past and future tense copulas. *Rah* can get inflected for person and honorificity. The following table lists all the person and honorificity markings that the verb can take:

## COPULA VARIATION IN PRESENT TENSE

Now, we will move to the most interesting tense — that is, present tense, where we find variation between

<sup>1</sup> The information is based on <https://en.wikipedia.org/wiki/Sonepur>, accessed on April 25, 2021.

<sup>2</sup> The information is based on <https://en.wikipedia.org/wiki/Chhapra>, accessed on April 25, 2021.

two Copular forms. In the earlier study on Bhojpuri Grierson (1903), we find a *h~ba* variation, though we have very few tokens of copula from the text. The dialectal data from two important regions and the results from Grierson (1903) are the following:

Region	<i>h</i>	<i>ba</i>	Total
Shahbad (Bhojpur)	(25%)	(75%)	100
Saran (Chhapra)	(30%)	(70%)	100

**Table 1: Distribution of the Copulas across the Regions based on (Grierson 1903)**

So, the table and chart show that *h* and *ba* have been in variation with each other for a long time. We find that more or less both the dialect regions are behaving the same. The variation is persistent in both regions<sup>3</sup>. However, we don't have any information about the speaker, his/her background, gender, education and all except for the geographical distribution, so it's hard to relate the variation with any social factor. Based on Grierson (1903), we find the linguistic constraint on *h*. *h* is restricted to 3rd person only, while *ba* has a wider spread and can occur in all persons:

Person	<i>h</i>	<i>ba</i>	Total
1	0	1 (100%)	1
2	0	2 (100%)	2
3	4 (66.7%)	2 (33.3%)	6
Total	4	5	9

**Table 2: Distribution of the Copulas across persons as per (Grierson, 1903)**

One important linguistic constraint that *he* undergoes is a person. In Maithili, we saw that *h* was restricted to 3rd person only. In Bhojpuri, too the same pattern is observed. Here is the result:

Person	<i>h</i>	<i>ba</i>	Total
1	0	18 (100%)	18
2	0	11 (100%)	11
3	73 (25%)	219 (75%)	292

<sup>3</sup> The variation between *h~ba* is a stable variation, has been historically attested. (See Grierson 1903.)

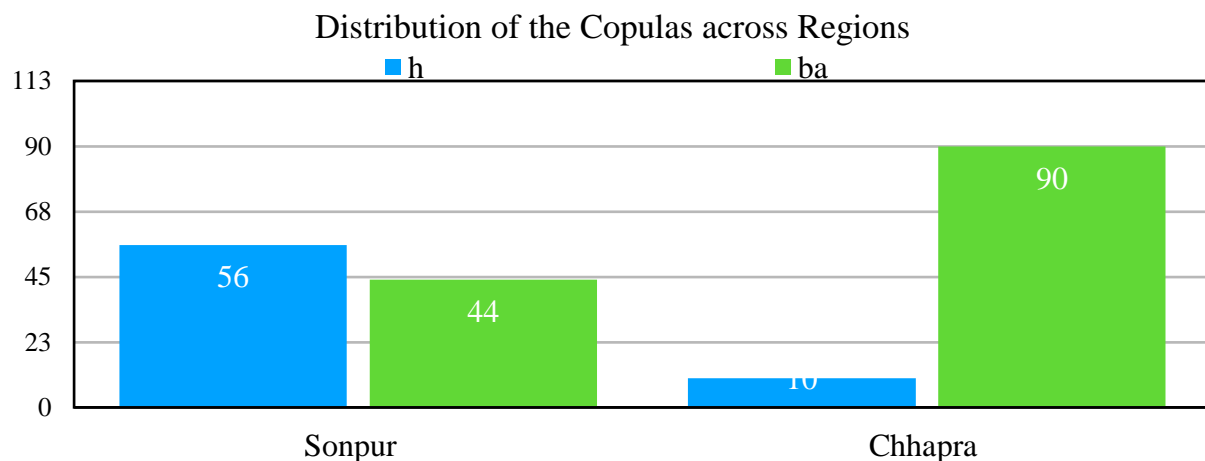
Total	73	248	321
-------	----	-----	-----

**Table 3: Distribution of the Copulas across persons in Bhojpuri (current study)**

The most important factor that conditions the variation is dialect geography. The region has played a very significant role in *h~ba* variation:

Region	<i>h</i>	<i>ba</i>	Total
Sonpur	51 (56%)	40 (44%)	91
Chhapra	22 (9.6%)	208 (90.4%)	230
Total	73	248	321

**Table 4: Distribution of the Copulas across Regions in Bhojpuri (Current Study)**



**Chart 1: Distribution of the Copulas across regions in Bhojpuri**

Though we have found that both *h* and *ba* have been in variation with each other over a long period of time, the variation is stable, but this stable variation is conditioned by many other social and extra-linguistic factors; most important of them is the geographical boundary or region as it is evident that Sonpur has more *h* (56%) than Chhapra (10%) and Chhapra has more *ba* (90%) than Sonpur (44%). Both regions behave differently.

The most important factor that influences the result is the linguistic boundary that Sonpur shares on the Eastern side with Bajjikanchal, an *h*-region.

One important factor is the cultural and social set-up of Sonpur and Chhapra. Sonpur is closer to the capital Patna and the so-called mini capital Hajipur, while Chhapra is far distant from both regions. Secondly, Chhapra is a relatively semi-urban area on the other hand, Sonpur is an urban area bordering Hajipur. Being the headquarters of the Eastern Central railway, Sonpur is more prone to inward and outward mobility.

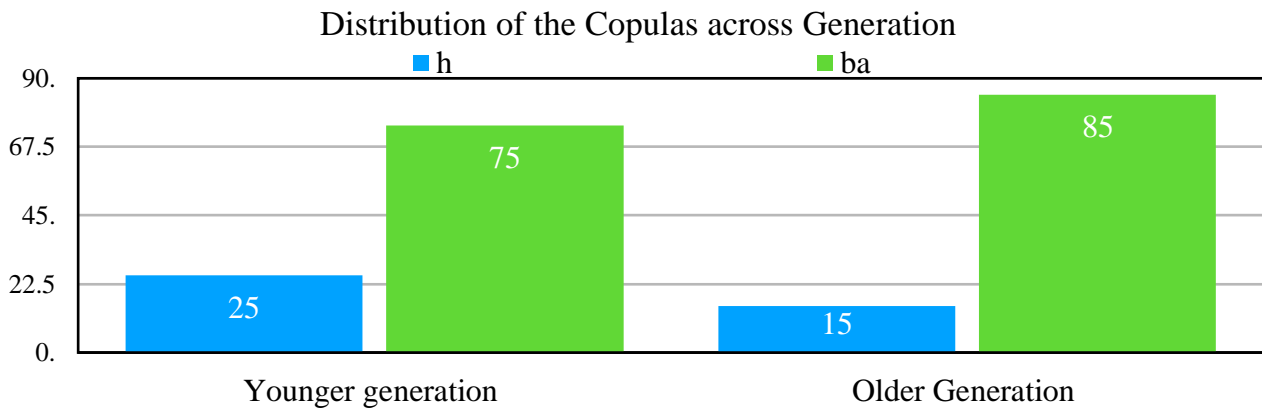
Chhapra is a traditional semi-urban centre while Sonpur is a modern urban centre, and because of being

closer to Hajipur, the society of Sonapur is more influenced by Bajjikanchal culture than Bhojpuri. So Sonpur, no doubt, keeps affirming with Bhojpuri, but it tends more towards the Bajjika features of speech: *h* is one of the clues along with the person markers like *-thin*, *-thun*, *-ai*, *ae* and *-ao*.

Apart from the geographical region, generation has stood very important. The variables seem to be age-graded:

Generation	<i>h</i>	<i>ba</i>	Total
Younger generation	60 (25.4%)	176 (74.6%)	236
Older Generation	13 (15.3%)	72 (84.7%)	85
Total	73	248	321

**Table 5: Distribution of the Copulas across age groups in Bhojpuri**



**Chart 2: Distribution of the Copulas across age groups in Bhojpuri**

Both the younger age groups (74.6%) and old age groups (84.7%) use *ba* forms more than *h* forms. But the youngsters (25.4%) use more *h* than the older age groups (15.3%). This looks like a case of age-graded variation.

We have found regional differences that *h* is mostly found in Sonpur than in Chhapra. As stated earlier, Sonpur is bound on the East by Hajipur, a Bajjika region, which is *h*-exclusive region. It is plausible that the greater number of tokens of *h* from Sonpur is a result of the dialect contact situation. Point to remember that Chhapra is a traditional representative Bhojpuri region, while Sonpur, in its set-up and geography, is closer to Hajipur (3-4km) and Patna (15-20 km). To understand the results better, the region is cross-tabulated with other social factors:

Region-Generation	Copula	Sonpur	Chhapra	Total
Young generation	<i>h</i>	44 (55%)	16 (10%)	60

	<i>ba</i>	36 (45%)	140 (90%)	75
Old generation	<i>h</i>	7 (64%)	6 (8%)	13
	<i>ba</i>	4 (36%)	68 (92%)	72
Total		91	230	321

**Table 6: Cross-tabulation of region and generation in Bhojpuri**

So it's the younger age groups who have more frequency of *h* than the old age groups overall. We find that in Chhapra, the difference between the younger (10%) and older age groups (8%) are almost blurred. On the other hand, in Sonpur, it's the young age groups (55%) who use more *h* than *ba*. Looking at the percentage, we find that it's the old age groups that use more *h*, but given that the only old speaker from Sonpur is a woman and women are almost ahead of men in the use of *h*, we find here an interaction between age and gender.

The younger generation in Sonpur uses significantly more *h* (55%) than the younger generation (10%) in Chhapra. Similarly, the old generation (64%) in Sonpur uses more *h* than that of Chhapra (8%).

Overall, it's the younger age groups who use more *h*; the variation seems to be age-graded.

### QUANTITATIVE/ MULTIVARIATE ANALYSIS

The factor groups selected as significant in the Varbrul statistical analysis of *h~ba* variation are the following:

1. **Region**
2. **Gender**

1.Region	<i>h</i>	Percentage	Total	Probability
Sonpur	51	56%	91	0.86
Chhapra	22	9.6%	230	0.32
2.Gender				
Male	20	15.3	131	0.34
Female	53	28%	190	0.60

**Table 7: Quantitative Analysis of Copula in Bhojpuri (present tense)**

Convergence at Iteration 6

Input 0.166

Log-likelihood = -129.671 Significance = 0.002

So undoubtedly region has greater effect and weightage as can be seen. Sonpur (0.867) has more probability of using *h* than Chhapra (0.32) (the possible reasons for that we have already discussed).

Similarly, gender has stood to be another significant factor group. Women (0.60) are more likely to use the *h* than men (0.347).

Though generation and mobility have been eliminated while stepping down, we have seen in the cross-tabulation tables that age groups have a good amount of effect on the *h~ba* variation; the youngsters use more *h* than the older age groups. Mobility, too, has worked like a catalyst for the variation. The more mobility implies the leaning towards the colloquial variant *h*.

### BROAD OBSERVATIONS

- Both *h* and *ba* are all-purpose copulas. *Ba* is the dominant form, while *h* is the somewhat equally competitive form, though it is less in frequency.
- One important linguistic constraint on *h* is that its domain is restricted to 3rd person only, while *ba* gets inflected for all persons and honorificity.
- Both *h* and *ba* are widespread across all genders and age groups among people who are mobile or immobile, giving a hint that this is unlike Maithili in which *h* is a newer form introduced by young mobile women; here is a case of stable variation which is conditioned by other social and regional factors like region, age, gender and mobility.
- It's the younger age group comparatively who use more *h* than the older age groups. Overall, both groups use more *ba* than *h*.
- It's the women rather than men who use more *h* than *ba*.
- Finally, we find variations within an individual that are the outcome of maintaining social ties with the people of the other linguistic areas and partly because of the mobility in the other area.
- Of course, mobility has worked like a catalyst to the variation, but it's not the cause of the variation as it is in the case of Maithili.
- Language boundaries have become very important; Sonpur has more *h* than Chhapra because Sonpur borders Hajipur, and Chhapra is quite far away from Hajipur.

### CONCLUSION AND DISCUSSIONS

We have seen that the variation between *h~ba* is conditioned by many linguistic and sociolinguistic factors and further by regional boundaries. The real locus of variation lies in the 3rd person, where *h* (25%) is restricted and varies with *ba* (75%). In another person, *ba* is categorical. Again, we have found that while *h* gets inflected for all person and honorificity grades, *h* is restricted to 3rd person non-honorific contexts in Chhapra, but *h* has a wider scope in Sonpur, which is closer to Hajipur and thus follows the Hajipurian pattern to most extent.

The most significant factor that has emerged is the Linguistic boundary or region. We find that the representative feature of Magadh and Vaishali *h* is found more in the area of Sonpur (56%), which is very close to Hajipur, than Chhapra (9.6%), which is not adjacent. For *ba*, on the other hand, the pattern gets reversed, and Chhapra emerges as a significant conditioning factor, as we can see that Chhapra (90.4%) has significantly more *ba* than Sonpur (44%).

We have found that gender interacts with geography: Women in Sonpur use more *h* than men; the probability is (0.60) for women, and it is just (0.34) for men. This is further conditioned by region as we can see that the linguistic difference for *h* between men (9%) and women (10%) in Chhapra is almost blurred. In Sonpur, on the other hand, it's significant. Women use more *h* (74%) in contrast with men (32%) only. Therefore, we find an interaction between gender and region. This even gives a clue that



unlike Chhapra, which is a stable and cultural landscape, Sonpur is the headquarters of the ECR (East-Central Railways) zone and thus is more urban and open to mobility than Chhapra.

Then we find that age groups interact with geography; in Chhapra again, there is not much difference between the old generation and the young generation as both the groups use more *ba* than *h*, while in Sonpur, we find that both the groups use more *h* than *ba*. So again, the region has stood to be a significant conditioning factor on the copulas.

We also find gender and age groups interacting with each other: Younger men use more *h* than older men, and younger women use more *h* than older women.

Not all individuals behave the same: it's the three individuals from one region, Sonpur, who use more *h* than the individuals from Chhapra, a case of intra-speaker variation conditioned by region.

Finally, we have found that mobility has worked only as a catalyst to the current variation *h~ba*. It's not like what we have observed for Maithili, in which *-h* tends to be introduced by younger women with mobility history.

Our closing statement is that there is no change in progress in the Bhojpuri region. Both the copulas *h* and *ba* have been through history, as the historical data from Grierson (1903) shows. What we find is that *h* is largely conditioned by region. *H* is mostly found in Sonpur than in Chhapra. Sonpur is bound on the East by Hajipur, a Bajjika region, which is *h*-region. It is plausible that the greater number of tokens of *h* from Sonpur is a result of the dialect contact situation and more because both Sonpur and Hajipur share the socio-economic norms.

## BIBLIOGRAPHY AND REFERENCES

1. Bhattacharya, P. (2017). Variation and Change: A Case Study of Calcutta Bengali [PhD Dissertation]. University of Delhi.
2. Bhattacharya, T. (2016) inner/outer politeness in central māgadhan Prākṛit languages: Agree as labelling, *Linguistics Analysis*, 40, 1–40.
3. Dey, K. (2010). Silchar Bengali: A Sociolinguistic Study [PhD Dissertation]. University of Delhi.
4. 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
5. Kashyap, A.K. (2014) The Bajjika language and speech community. *International Journal of the Sociology of Language*, 227, 209–224.
6. Satyanath, S. (2003) On the maintenance of transplanted Indian languages overseas. In: *Indian Diaspora: In Search of Identity* (edited by R. Sharma & E. Annamalai). Central Institute of Indian Languages, pp. 85–104.
7. Satyanath, S. (2006) English in the new world: Continuity and change, the case of personal pronouns in Guyanese English. In: *The Structure of Creole Words: Segmental, Syllabic and Morphological Aspects* (edited by P. Bhatt & I. Plag). Max Verlag Niemeyer, pp. 179–200. DOI: 10.1515/9783110891683.179.
8. Satyanath, S. (2015) Language variation and change: The Indian experience. In: *Globalising Sociolinguistics* (edited by D. Smarkman & P. Heinreich). Routledge.
9. Satyanath, S. (2021) Genealogies of sociolinguistics in India. *Journal of Sociolinguistics*, 25, 762–784. DOI: 10.1111/josl.12496.
10. Satyanath, S. & Laskar, N. (2008) Lexicon in a contact language. In: *Linguistics* (edited by S. Morey

- & M. Post). North East Indian. New Delhi: Foundation/Cambridge University Press., 75–92.
11. Satyanath, S. & Laskar, N. (2009) Ethnicity, bilingualism and variable clitic marking in Bishnupriya Manipuri. In: Variation in Indigenous Minority Languages (edited by J. N. Stanford & D. R. Preston). John Benjamins Publishing Company, pp. 441–462.
  12. Sharma, D. (2005) Dialect stabilization and speaker awareness in non-native varieties of English1. *Journal of Sociolinguistics*, 9, 194–224. DOI: 10.1111/j.1360-6441.2005.00290.x.
  13. Sharma, R. (2008). Phonetic Realizations of Vowels in Indian English [MPhil Dissertation]. University of Delhi.
  14. Sharma, R. (2017). A Sociophonetic Study of Variation in Vowels Among Hindi / Punjabi-English Bilinguals in Delhi [PhD Dissertation]. University of Delhi.
  15. Shuy, R., Wolfram, W. & Riley, W.K. (1966). A Study of Social Dialects in Detroit [Final report], project 6-1347. Office of Education.
  16. Tagliamonte, S. (1998) Was/were variation across the generations: View from the city of York. *Language Variation and Change*, 10, 153–191. DOI: 10.1017/S0954394500001277.
  17. Tagliamonte, S. & Hudson, R. (1999) Be like et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics*, 3, 147–172. DOI: 10.1111/1467-9481.00070.
  18. Trudgill, P., editor (1978b). *Sociolinguistic Patterns in British English*. Edward Arnold.
  19. Trudgill, P., editor (1984a). *Applied Sociolinguistics*. Academic Press.
  20. Trudgill, P., editor (1984b). *Language in the British Isles*. Cambridge University Press.
  21. Trudgill, P. (1988) Dialect contact, dialectology and sociolinguistics. In: *Sociolinguistics Today: International Perspectives* (edited by K. Bolton & H. Kwok) (1991). Routledge.
  22. Trudgill, P. & Hannah, J. (1982). *International English*. Edward Arnold. Tse, J. K.