

Khortha in Jharkhand: A Socio-linguistic Analysis of Language Use, Variation, and Identity

Siddharth Sanket

Centre for Linguistics, Jawaharlal Nehru University, New Delhi – 110067
siddharthraj1302@gmail.com

Abstract: This paper presents a sociolinguistic analysis of Khortha, an Indo-Aryan language spoken primarily in Jharkhand and adjoining regions of Bihar and West-Bengal. Situated within one of India's linguistically diverse regions, Khortha occupies an important position in the communicative practices of communities with different linguistic backgrounds. Despite its widespread use and substantial speaker population, the language has received comparatively limited attention in the mainstream linguistic scholarship and continues to be discussed within competing classifications as a dialect or variety of other Indo-Aryan languages.

Drawing on census data and existing linguistic and sociolinguistic literature, this study examines the status and function of Khortha in Jharkhand, with particular to its role as a Lingua Franca, the emergence of regional varieties through language contact, and the relationship between language use, identity, bilingualism, and language attitudes. The discussion also considers the influence of neighbouring languages and social factors that shape the use and prestige of Khortha across different sections of its speech community. The analysis indicates a significant tension between language's functional importance in everyday intercommunity communication and its relatively low social prestige, particularly among sections of the economically mobile and educated population. The paper argues that sustained documentation, mother-tongue-based educational initiatives, and wider institutional and cultural support are essential for strengthening the recognition, transmission, and long-term maintenance of Khortha in Jharkhand.

Keywords: Jharkhand, tribal language, tribal group, lesser-known language, Khortha language.

I. INTRODUCTION

1.1. Khortha Language:

Khortha is a language spoken primarily in the state of Jharkhand. It is also spoken in some parts of the neighboring states of Bihar and West Bengal. According to the census report (Census, 2011), Khortha is considered to be a dialect

of Hindi. This is also true for many such lesser-known languages spoken in northern part of India such as Braj, Awadhi, Kannaui, etc. This area is also referred sometimes as the "Hindi Belt".

Khortha is the second largest language spoken in the state second to Hindi. The total population of Khortha speakers is 8.04 Million (Census, 2011) and growing as the number of speakers have increased from 4.7 Million (Census, 2001).

When we talk about the geographical extent of Khortha, we see that it is spread throughout the state. Ohdar (2012), identifies that out of 24 districts in Jharkhand, 15 of the districts are heavily inhabited by the Khortha population. The districts such as Garhwa, Palamu, Latehar, Chatra, Hazaribagh, Ramgarh, Bokaro, Dhanbad, Koderma, Deoghar, Dumka, Godda, Pakud, and Sahebganj are densely populated with native Khortha speakers.



Map 1: Political Map of Jharkhand State.

1.2. Languages in Jharkhand:

The state of Jharkhand is very unique. In the linguistic landscape, we may call Jharkhand to be a "Sprachbund", i.e., "Linguistic Area". The reason to say this is that the state of Jharkhand is home to languages from three distinct language families.

In Linguistics we explore "India as a Linguistic Area" due to

the diversity we have in terms of the number of languages this land boasts (Approximately 1700). Jharkhand in itself can be called a “Mini Linguistic Area”. There are languages from Indo-Aryan language family, Austro-Asiatic language family and Dravidian Language family. The languages that are present in the region are indicated in the following figure (Figure 1).

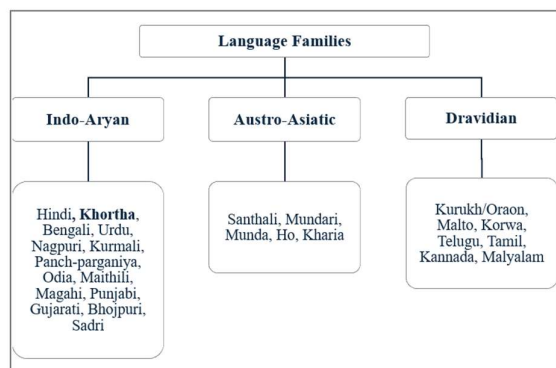


Figure 1: Languages from different language families in Jharkhand. Adopted from Aman et.al (2020).

1.2.1. Austro-Asiatic Languages:

The languages from Austro-Asiatic language families include Santhali, Mundari, Ho, and Kharia. There are 32 scheduled tribes present in Jharkhand, which contribute to more than one-fourth of the total population of Jharkhand. Out of these 32 tribes, 8 are classified under primitive tribal groups. Majority of the tribal groups are from Austro-Asiatic language family, but there are some that are from Dravidian language family as well.

All these tribal groups have their own language. One of the key distinguishing factors for these tribal groups is their language. Language is the element that has kept the groups closely tied. Despite the impact and enforcement of various languages, these tribal languages have survived, and they have a thriving population as well. One of the languages, Santhali has also been included in the eighth schedule of the Indian constitution in 2003. It also has its own script known as Ol Chiki. It did not have a separate script up until the 1900s. It was written in Devanagari, Bengali and Roman. It was in 1925 that Pandit Raghunath Murmu created the Ol Chiki script for the language (Murmu, 2002).

1.2.3. Dravidian Languages:

The Dravidian languages are Kurukh, which is also known as Oraon, and others such as Malto, Korwa, Telugu. This is the second tribal group that exists in the state of Jharkhand. The Kurukh language which is the largest in terms of population in this language family and is primarily spoken in the districts of Gumla, Simdega, Lohardaga, and Ranchi. The

tribe that speaks this language as a native tongue are identified as Oraons and hence this language is also sometimes referred as Oraon.

Malto language is spoken by the tribes residing in the areas of Raj-Mahal hills of the Santhal Pargana. The tribes that speaks this language are identified as Sauria Paharias and Mal Paharias (Singh, 2014).

1.2.2. Indo-Aryan Languages:

The languages that are present from the Indo-Aryan language family are Hindi, Khortha, Bengali, Urdu, Nagpuri, Kurmali, Panch-parganiya, Odia, Maithili, Magahi, Punjabi, Gujarati, Bhojpuri, and Sadri.

Because of the neighboring states of Bihar, West Bengal, Odisha, and Uttar Pradesh, the languages such as Bengali, Odia, Urdu, etc. are spoken in the area as well. The areas closer to the borders of these states show the most influence of these languages.

In the year 2011, the government of the Jharkhand gave official status to 9 tribal and regional languages namely Santhali, Oraon, Mundari, Ho, Kurukh, Sadri, Khortha, Panch-parganiya and Kurmali. The speakers of other tribal languages which do not have a script are also trying to develop scripts for their languages that would help in the preservation efforts for the languages.

1.3. Khortha as a Lingua Franca:

According to the Oxford Dictionary of English, a Lingua Franca is “a language that is adopted as a common language between speakers whose native languages are different.” This is precisely true for Khortha in the state of Jharkhand. This is due to the fact that the region (Jharkhand) is home to groups that have distinct languages which are not mutually intelligible to each other.

The language use of the tribal groups have to be studied in a much-detailed way. Majority of the population of Jharkhand is bilingual if not multilingual. This is especially true for the tribal groups present in the state. When it comes to language use, the tribal groups use their own native languages intra-community, meaning they use their own language when they are within their community. But since there are different groups with distinct languages, they cannot use their own native language to communicate.

So, there is a tendency of using Khortha as their first language by many indigenous tribal communities of the state (Dangi 2012; Ohdar 2012). Also, when asked, the members of these groups report Khortha as their first language. One of the reasons behind this could be the infrequent use of their own native language which is restricted to communications inside the community.

II. VARIETIES OF KHORTHAS

We saw that Khorthas has become a Lingua Franca for the state, and it plays a crucial role in overall communication in this diverse region.

Since there are many languages that are present in the state of Jharkhand, the most commonly used language second to Hindi is Khorthas, and it has grown significantly from just a medium of communication to facilitate communication among different indigenous groups to having different variations of the same language. Apart from this, the languages of the neighboring states of Bihar, West Bengal, Odisha, Chhattisgarh, And Uttar Pradesh namely Magahi, Bhojpuri, Maithili, Bengali, etc. have also influenced the Khorthas language. These variations have led to the development of multiple varieties of Khorthas that are very distinct from each other.

The first works in Khorthas language has been done by the renowned linguist and scholar George Abraham Grierson, who founded the Linguistic Survey of India and did an extensive mapping of the linguistic diversity of India in his huge volumes of books that give us a detailed account of almost all the languages that are present in the land of India. It is still one of the most extensive and elaborate account of the linguistic heritage that is still used and cited today. It is considered as the 'holy grail' in linguistics not just in India but throughout the globe. It is in his fifth volume in 1903, that we see the first description of Khorthas language.

In his fifth volume titled "Indo-Aryan Family", he has described Khorthas language to be a dialect of Magahi Language. The name that he gave to this variety was called "Eastern Magahi". According to Grierson (1903), Magahi was the language spoken in the old country of Magadha. Apart from this, Magahi was also spoken all over in the Gaya District and the district of Hazaribagh². Both were also Kingdoms before the British rule in India. Magahi was also spoken in the district of Palamu as well. While he was mainly concerned with the scope of Magahi in this region, he also stated that while there did not exist many dialects of Magahi in the region, (present state of Jharkhand), there was a dialect of Magahi which was mixed with other languages in the area. And this variety was named "Eastern Magahi". He finds that,

"...the Magahi so affected possess distinct peculiarities entitling to be classed apart as a sub-dialect, which I have named Eastern Magahi."

- Grierson (1903)

Also, Grierson (ibid) states that there exists three 'enclaves'

of this eastern Magahi. So, he describes that this eastern Magahi is the language of the region, but it has also been influenced by the neighbouring languages like Bengali, Odia and Chhattisgarhi. They make these enclaves where we see that this eastern Magahi has developed into a distinct variety on its own. One in the native state of Mayurbhanja (Present day Odisha), and Bamra (Chhattisgarh), where this language was surrounded by Odia. This variety was called "Kurumali" or "Kurmali" by the locals. The second enclave was near the western Malda district of present-day West-Bengal. Grierson described this language to be called as "Khontai". The third and the last 'enclave' according to Grierson was the one which was surrounded on the North and West by Maithili and by Bengali in the East and South.

It has been more than a century to Grierson's work. While there has been a lot of change that has happened in the last century, but the work hasn't lost its relevance. The present scenario is different from what it was 100 years ago, but It has provided a solid base for research in languages and linguistics in India.

Other scholars such as Prasad and Shastri (1958), aligns with Grierson and identifies Khorthas to be a dialect of Magahi. But according to the ethnologue, Khorthas or Khotta is considered a dialect of Magahi (Language Code: mag) as well as Maithili (Language Code: mai). Other scholars such as Das (2013) considers it to be a dialect of Angika, which in turn is a dialect of Maithili itself.

So, we have a dilemma here, what should we consider Khorthas to be? From Grierson (1903) to Aman et al (2020), we see that both these claims are true. Prasad and Shastri (2012); Dangi (2012); Aman et al. (2020) argues that there are atleast 6 distinct varieties of modern day Khorthas.

Sr. No.	Varieties	Region	Neighboring Languages
1.	Deshwali [des ^h vali]	Places within the Patratu border to the western side of Palamu as in Daltonganj, Lohardaga, etc.	Magahi and Bhojpuri
2.	Ramgarhiya [ramgə ^h riya]	Ramgarh, Hazaribagh, Bhurkunda, Eastern Patratu, etc.	Kurmali and Nagpuri
3.	Parnadiya [pə ^h nədiya]	The northern side of Damodar river in and around Nai Sarai, situated opposite to Ramgarh.	
4.	Sikhariya [sik ^h əriya]	Jainamore, Parasnath Pahad, Topchanchi, Giridih upto Eastern Bengal borders.	Bengali
5.	Santhal Parganiya [sə ^h nt ^h əl pə ^h rgəniya]	Deoghar, Madhupur, Dumka, and other North-Eastern areas of Jharkhand Bordering to Bihar.	Maithili or Angika
6.	Golwari [golvari]	South of Ramgarh area from Kasmar, Kherachatta, Gola, Rajrappa, etc.	Kurmali

Table 1: Different speech varieties of Khortha along with the languages which are neighbouring those varieties, hence affecting the varieties. Adopted from Aman et al. (2020).

The table represents the speech varieties given by Aman et al. (2020). We can see that all these varieties are surrounded by different languages. And accordingly, we can see the distinct features of these varieties of Khortha. They are different in terms of the number of phonemes present in the language, the vocabulary used in the languages, the verbal endings, and the structure of the languages.

III. LANGUAGE ATTITUDE

In Jharkhand, Odisha and West Bengal, about 42 languages are considered endangered (Pankaj, 2010). When we talk about Khortha, we find that the speakers of Khortha can be divided into two groups based on their economic background. The first with relatively good economical background are limiting their use of Khortha language gradually. But the group with relatively low economical background are the one that use the language more.

The reason for this can be the low prestige of Khortha language and the imposition of Hindi and other major languages in the region. The first group with relatively high income see Khortha language as inferior to Hindi and because they are relatively more accessible to go for higher education, they chose not to use their own mother tongue and go for Hindi for its high prestige. The frequent use of Hindi and other languages along with English, provides a psychological boost for adopting these languages and since, there is no mother tongue education available in higher education institutions, this adoption becomes an inevitable compromise. On the other hand, the people who have just done their matriculation and are now involved in agriculture or other related jobs have still the usability of Khortha and they use it as a link language in the marketplace and their daily lives.

3.1. Language Identity

Language identity refers to the way individuals or groups perceive and express their affiliation with a particular language, which is often linked to their culture, ethnicity, or community. So, whenever an individual is asked about their language, the association that they have with that language in terms of the usability, acceptance and culture, we can learn its identity. And when it comes to Khortha speakers, they know that this is the language that they speak. And often they can distinguish between the different varieties mentioned above in the paper but naming them is totally different. The varieties are often associated with distinct groups that are confined to a particular geographical area. One such instance can be seen with the Parnadiya variety. The name of this variety can be broken down into two (par meaning 'other side' and nadiya means 'river'). This language is spoken by the people who live on the other side of the Damodar river in Ramgarh district.

3.2. Bilingualism

Bilingualism in India is very ordinary phenomenon. This is especially true for Jharkhand. Being this diverse, the people are bilingual if not multilingual. In this paper, we have discussed how Khortha and other languages are present in the area and how these languages interact with each other. So, most of the people can speak more than one language. For example, a person from indigenous community will natively speak its tribal language. Along with this, they will also speak the link language (In case of Jharkhand, the link language is Khortha), also they will speak Hindi as well

because if they had any kind of formal education, it would mostly be in Hindi. And in schools English would also be taught.

3.3. Status of Khortha:

Recently, there has been many efforts to help the Khortha language. Along with Khortha, other tribal languages are also given much importance in order to preserve them. But the aim of this paper is to talk about Khortha, so we will restrict ourselves to this language only. In that way this paper is also an effort to document Khortha and make people aware about this language.

In 2011, Khortha was given official status by the government of India. This was achieved after a long resistance from the different linguistic groups. Because prior to this only Hindi and Urdu were the official languages of the state. We saw how the identity of Khortha is also degrading in the youth as they are moving ahead with their higher studies. The higher income group no longer wants to get associated with the language due to its low prestige.

This can be improved if mother-tongue based education is implemented in the state as suggested by the New Education Policy 2020. This will help in creating awareness about the language and improving the overall status of the language. It is not that efforts are not taken by the state government. Recently they have implemented Khortha as a subject in the State's Public Service exam (JPSC).

The language is rich in folklores and songs; they not only are helpful in preserving the language but also are helpful in preserving the cultural richness of the language. There has been many efforts to create and publish texts in the language, but these are very niche publications with only a handful of readers. So, it is imperative to penetrate Khortha texts into the major population.

CONCLUSION

The study has examined the sociolinguistic position of Khortha within the multilingual landscape of Jharkhand, focusing on its role as a Lingua Franca, regional variation, language contact, and the relationship between language use and identity. The discussion shows that Khortha serves an important communicative function among speakers of different linguistic backgrounds, while its regional varieties reflect sustained interaction with neighbouring languages. At the same time, the language's widespread use does not necessarily correspond to equivalent social prestige or equivalent institutional recognition. The shift towards Hindi among sections of the educated and economically mobile population rises important questions concerning language attitudes, identity, and the continuity of Khortha across generations.

The sociolinguistic position of Khortha therefore reflects a

complex relationship between functional vitality and social marginalization. Its continued use in everyday and intercommunity communication demonstrates its significance within Jharkhand's linguistic ecology, but long-term maintenance requires more than recognition as a regional language. Systematic documentation, mother-tongue-based educational initiatives, and the wider development and dissemination of linguistic and cultural materials are important for strengthening its transmission and visibility. Further empirical research involving diverse speech communities, generations, and regional varieties would contribute to a more comprehensive understanding of Khortha's changing sociolinguistic profile. Such efforts are essential not only for the documentation of the language but also for recognizing the linguistic diversity and cultural heritage of Jharkhand.

REFERENCES

- [1] किशोर वड़त्या. (2026). समाज भाषाविज्ञान के दृष्टिकोण से खोरठा भाषा की एक व्याख्या. गवेषणा, 143, 85-95.
- [2] कुमार, ल., & संकेत, स. (2026). कन्नौजी भाषा और पुनरुक्ति. गवेषणा, 143, 96-104.
- [3] Abbi, A. (1994). Semantic Universals in Indian Languages. Shimla: Indian Institute of Advance Studies.
- [4] Abbi, A. (2001). A Manual of Linguistic Fieldwork and Structures of Indian Languages. Muenchen: Lincom Europa.
- [5] Aman, A., Dash, N. S., & Chakraborty, J. (2020). Designing a Linguistic Profile of Khortha: A Less Resourced Language Spoken in the State of Jharkhand, India. *Dialectologia*, 25, 25-43.
- [6] Aman, A., Dash, N. S., & Gope, A. (2020). A Preliminary Account on the Khortha Speech Prosody. *PalArch's Journal of Archeology of Egypt/Egyptology*, 6440-6458.
- [7] Aman, A., Dash, N. S., Chakraborty, J., & Gope, A. (2020). Consonants in Khortha. *Alternative Horizons in Linguistics: A Festschrift in Honour of Prof. Panchanan Mohanty*, 336-352.
- [8] Baker, M. C. (2013). On Agreement and its Relationship to Case: Some Generative Ideas and Results. *Lingua*, 14, 32.
- [9] (2011). Census of India. Government of India. New Delhi: Office of the Registrar General & Commissioner, India.
- [10] Chaturvedi, S. (2015). A Sociolinguistic Study of Linguistic Variation and Code Matrix in Kanpur. *Procedia - Social and Behavioral Sciences*, 107-115.

-
- [11] Comrie, B. (1989). *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell Publishing.
- [12] Das, P. K. (1997). *Agreement in Hindi-Urdu*. New Delhi: Jawaharlal Nehru University.
- [13] Das, P. K. (2013). Ergativity in Khortha: An Enigmatic Phenomenon. *The Journal of South Asian Studies*, HUFs, South Korea, 18, 321-350.
- [14] Das, P. K. (2017). Case and Agreement in Khortha. *Aligarh Journal of Linguistics*, 7(1).
- [15] Das, P. K. (2018). Ergativity and Agreement in Kinnauri. *Interdisciplinary Journal of Linguistics*, 11, 36-53.
- [16] Dash, N. S., & Aman, A. (2015). Generation of a Dialect Corpus in Khortha used in Jharkhand India: Some Empirical Observations and Theoretical Postulations. *The Journal of Advanced Linguistic Studies*, 151-171.
- [17] Dwivedi, P., & Kar, S. (2016). Kannaui in Kanpur: A brief overview. *Acta Linguistica Asiatica*, 6, 101-119.
- [18] Dwivedi, P., & Kar, S. (2017). On Documenting Low Resourced Indian Languages: Insights from Kannaui Speech Corpus. *Dialectologia*, 19, 67-91.
- [19] Grierson, G. A. (1916). *Linguistic Survey of India (Vol. 9) Indo-Aryan Family, Central Group Part 1*. Calcutta: Superintendent Government Printing.
- [20] Grierson, G. A. (1927). *Linguistic Survey of India, 1927*. Calcutta: Central Publication Branch.
- [21] John, M., & Varghese, B. P. (2021). The Kannaui-speaking People of Uttar Pradesh: A Sociolinguistic Profile. *Journal of Language Survey Report*.
- [22] Katamba, F. (1993). *Morphology*. New York: St. Martin's Press.
- [23] Kaul, O. N. (2008). *Modern Hindi Grammar*. Springfield: Dunwoody Press.
- [24] Kellogg, S. H. (1875). *A Grammar of Hindi Language*. London: Trubner & Co.
- [25] Kishore, V., & Sanket, S. (2025). *Influence of Telugu on Lambadi Language (2nd Edition ed., Vols. VBUSS Series - 11)*. New Delhi: A.K. Publications.
- [26] Kumari, A. (2021). Folklore and Language: Ergativity and Agreement of Khortha in Jharkhand. *International Journal of Humanities, Engineering, Science and Management*, 2, 35-44.
- [27] Masica, C. P. (1991). *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- [28] Mohanan, T. (1994). *Agreement Structure in Hindi*. Stanford: CSLI Publications.
- [29] Murmu, G. (2002). Development of Santhali Language, Literature and Its Recognition (Language/Script Movement). In N. H. Itagi, & S. K. Singh (Ed.), *Linguistic Landscaping in India*. Central Institute of Indian Languages and Mahatma Gandhi International Hindi University.
- [30] Pandey, A. (2024). Kannaui (Uttar Pradesh, India) - Language Snapshot. *Language Documentation and Description*, 24(1).
- [31] Prasad, B. N., & Shastri. (1958). *Linguistic Survey of the Sadar Subdivision of Manbhum and Dalbhum (Singhbhum)*. Patna: Bihar Rashtrabhasha Parishad.
- [32] Priya, S., & Singh, R. (2018). Khortha, A Dying Language and the Urgency to retain its Pure Variety. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 5(1), 171-190.
- [33] Ray, B., & Kumar, A. (2022). Person Agreement in Khortha. *Interdisciplinary Journal of Linguistics*, 15, 85-98.
- [34] Saksena, S. (1981). Verb Agreement in Hindi. *Linguistics*, 14, 467-474.
- [35] Singh, P. K. (2014, December 20). Nurturing Diversity in Jharkhand: Role of Mother Tongue. *Economic and Political Weekly*, 49(51), 17-19.