



LANGUAGE SNAPSHOT

Kanauji (Uttar Pradesh, India) – Language Snapshot

ANU PANDEY, Indian Institute of Technology Kanpur, IN, anu.pandey45@gmail.com

ABSTRACT

Kanauji is an endangered and low-resourced variety of Western Hindi that is spoken by approximately 7 million people across many regions in Uttar Pradesh, India, including Kanpur, Kanauj, Auraiya, Etawah, Farrukhabad, Hardoi, Kanpur, Kanauj, Pilibhit, Mainpuri, and Shahjahanpur. Native speakers of Kanauji use Kanauji, Khariboli (vernacular Hindi), and Standard Hindi in different social contexts, simultaneously switching between Standard Hindi, Khariboli, and Kanauji. Kanauji is spoken more in rural regions than in urban centers. Although Kanauji is still prevalent in homes and neighborhoods, it lacks representation in domains such as education, literature, and media, where Khariboli and Standard Hindi are used instead. The migration of youth towards cities for socio-economic mobility and higher education leads them to adopt the lingua franca of Khariboli and Standard Hindi. Consequently, Kanauji is now an endangered variety that requires urgent documentation efforts. In recent years, there have been ongoing efforts to document the language. The author's PhD project, which began in 2017, represents another attempt to describe and document the language, social situation, and culture of the Kanauji community. This paper provides a succinct overview of Kanauji, encompassing its geographical location, demographics, sociocultural aspects, as well as prior and current research.

Keywords: Hindi variety; low-resourced language; Kanauji, Khariboli, Uttar Pradesh

Language Documentation and Description is a peer-reviewed open access journal published by the Aperio. © 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons AttributionNonCommercial 4.0 International License (CC-BY-NC 4.0), which permits unrestricted distribution, reproduction and adaptation in any medium, provided the original author and source are credited, and that the material is not used for commercial purposes. See <https://creativecommons.org/licenses/by-nc/4.0/>.



 OPEN ACCESS

Language Name: Kanauji (Dehati), Kannauji

Language Family: Indo European > Indo Aryan > Western Hindi > Unclassified

ISO 639-3 Code: bjj

Glottolog Code: kana1281

Population: 7,093,740 (John & Varghese 2021)

Location: Kanpur, Kanauj, Auraiya, Etawah, Farrukhabad, Hardoi, Pilibhit, Mainpuri, and Shahjahanpur districts in Uttar Pradesh, India.

Vitality rating: Endangered (Eberhard et al. 2024)

1. GEOGRAPHY

Kanauji is spoken in the state of Uttar Pradesh in northern India in the districts of Kanauj, Kanpur, Farrukhabad, Hardoi, Auraiya, Etawah, Mainpuri, Shahjahanpur, and Pilibhit (see [Figure 1](#)). Kanauji's territory extends approximately between latitudes 26.5 to 28.5 degrees north and longitudes 70 to 80.5 degrees east. The Kanauji-speaking population is estimated to be around 7,093,740 according to the 2001 census (John & Varghese 2021: 1).

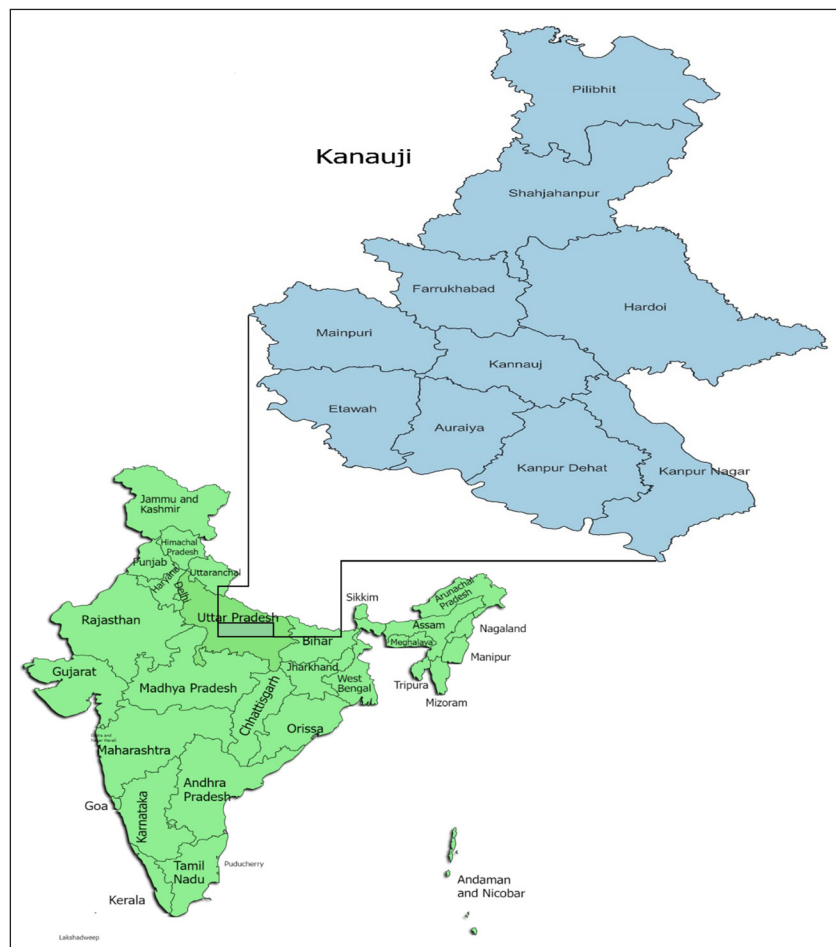


Figure 1: Kanauji-speaking regions of Uttar Pradesh, India. Map by the author.

2. LANGUAGE STATUS AND SOCIO-CULTURAL LIFE

The label *Kanauji* was coined by Grierson (1968[1916]) in the Linguistic Survey of India. Grierson categorizes the language as belonging to the Western Hindi group along with Bundeli, Braj Bhasha, and Hariyanvi, all varieties of Standard Hindi. According to Tripathi (1977: 22), Kanauji has multiple names referencing the Kanauj district in Uttar Pradesh where it is spoken: *Kannauji*, *Kanauji*, and *Kanujiya*. In the Farrukhabad district, it is called *Kanujujiya*. Kanauji has three sub-variants: Kanauji Proper, Tirhari, and Transitional Kanauji. The form that is considered “pure” and “proper” is spoken in Kanauj and Farukhabad.

Most Kanauji speakers identify themselves as speaking *Dehati* ‘village variety’ (John & Varghese 2021: 11). According to John & Varghese, Kanauji speakers are bilingual in Hindi. Although Kanauji is common in many domains of daily life, a shift toward Standard Hindi is being observed, especially among young and educated speakers. Education has played a primary role in enhancing the Hindi proficiency of Kanauji speakers, improving their ability to speak and read complex materials such as books, newspapers, and magazines in Hindi.

Kumar (2009: 5) explains that in this area spoken Hindi is known as *Khariboli*, sometimes also called *Hindustani*. This language emerged as the primary language for communication and lingua franca in North India due to the interaction and blending of different Indian and foreign language speakers. Khariboli was not regarded as a suitable medium for official instructions in feudal courts or for standardized use in creative endeavors, but it began to operate as the oral vernacular for the common people and has become a language of broader communication in the north (Verma 1933).

Standard Hindi and Khariboli have shared a classic diglossic relationship since Indian independence. While Khariboli is used as a lingua franca across northern India, it is Standard Hindi that holds the status of the official and literary language.¹ Kanauji speakers prefer the more prestigious Khariboli and Modern Standard Hindi in casual and formal situations, respectively. Since Khariboli is the lingua franca, using it helps Kanauji speakers blend in with urban Hindi speakers and communicate effectively with speakers of other languages. Standard Hindi is perceived as a pathway to attaining economic gains and higher prestige.

Kanauji is restricted to close-knit settings like households and local communities and is dispreferred in public and formal contexts. The use of Kanauji in public is a marker of lower status, implying rural background, lower economic class, and lack of formal education. Educated youths feel ashamed to use Kanauji in the presence of non-Kanauji speakers, and educated youth often use Kanauji to look down upon or make fun of someone. (Dwivedi & Kar 2016: 110). This suggests they harbor a negative attitude toward their mother tongue. Despite this, when asked if they would like to see Kanauji introduced into formal education settings, rural youth replied that they would (Pandey 2023a). There is evidence that university students use Khariboli in inter and intra-group communication, while they speak Kanauji with their older family members and grandparents.

1 North-Central India is referred to as the “Hindi belt” or the “Khariboli belt”, reflecting differing degrees of diglossic complementation across multiple speech varieties (Khubchandani 1972). Bilingual speakers in the north-central region retain their regional or caste dialects of Western Hindi or use another language for informal communication within their speech group (Khubchandani 1972, 1978). For formal communication, they prefer to use Standard Hindi.

3. LINGUISTIC FEATURES OF KANAUJI AND RELATED VARIETIES

Although the phonology and morphology of Kanauji and Hindi are quite similar, according to Dwivedi & Kar (2018: 206) Kanauji does have some phonological features that are distinct from Hindi. For instance, in Kanauji the palatal and velar nasal phones [ɲ] and [ŋ] are allophones of /n/ and thus do not have a phonemic status like they do in Hindi. Kanauji has processes of fortition, nasalization, vowel insertion, and reduplication when compared to Hindi; for example, Hindi *sku:l* ‘school’ corresponds to *ɪsku:l* ‘school’ in Kanauji; Hindi *luʈ* ‘theft’ is *luʈ-laʈ* in Kanauji. There are also differences in words ending with *r*, *l*, and *ɽ*. An example is *pipəl* ‘peepal tree’ in Hindi, which is pronounced *pipəɽ* ‘peepal tree’ in Kanauji.

Morphological similarities are observed between the Hindi and Kanauji lexicon. Question words are simultaneously inflected for number and gender, e.g., Hindi *kɛsa* and Kanauji *kəɪsa* ‘how.M.SG; Hindi *kɪt̪na* and Kanauji *kɪt̪ta* ‘how much/many.M.SG’. Full and discontinuous reduplication are observed in both languages’ question words; e.g., Hindi *kəbhi-na-kəbhi* and Kanauji *kəbhəũ-na-kəbhəũ* ‘some time or the other’. Synchronic morphophonological variation in question words was observed between different regional varieties of Kanauji, suggesting a need for more extensive documentation (Pandey 2023b).

4. EXISTING LITERATURE

Many works briefly discuss the areal distribution and position of Kanauji with respect to Khariboli. These include Shapiro (1989), Masica (1993), and Kachru (2006), among others. There is work on Kanauji folk songs (Santram 1975) and on the sociology of Kanauji literature (Tripathi 1977). Pandey (1978) provides a basic description of the linguistic characteristics of Kanauji, including vowels and consonants, nouns, pronouns, adjectives, and question words; Tripathi (1977) provides a comparative overview of the grammar of the varieties spoken in the lower Ganga plain region. Trivedi (1990) accounts for echo words in Kanauji. Recent work on the linguistic repertoire of Kanpur sheds light on Kanpur as the confluence of several varieties of Hindi (Chaturvedi 2015). Mishra & Bali (2010, 2011) have developed a small-scale NLP application in Kanauji.

Dwivedi & Kar (2016, 2017, 2018) describe the phonology and sociolinguistics of Kanauji of Kanpur. Singh (2016) provides a brief background, wordlist, and grammar of the Kanauji language variety. John & Varghese (2021) give a broad sociolinguistic profile of Kanauji speakers of Uttar Pradesh. Their study outlines the present geographical distribution of the Kanauji language and its relationship with other Western Hindi varieties such as Bundeli, Braj Bhasha, and Hindi. It also documents the level of bilingualism in Kanauji speakers, their attitudes toward the language, and the extent of the language shift toward Hindi.

5. ONGOING WORK

The author’s work on Kanauji was begun during PhD studies (2017–2023) at the Indian Institute of Technology, Kanpur. The project endeavored to describe the structure and function of questions and question words in Kanauji and Standard Hindi. Kanauji data was collected primarily from Kanpur district, however, a brief overview of the morpho-phonological variation among different regional varieties of Kanauji was also provided. As the data was mostly collected during COVID-19, remote fieldwork using mobile interviews was adopted. Since most Kanauji speakers exhibit bilingual proficiency in Hindi, elicitation was done via questionnaires prepared in Hindi. Around twenty audio recordings of dialogues

were also collected; these were used to create a podcast featuring rituals of the Kanauji language (Pandey 2024). A quantitative sociolinguistic survey with rural school children and college youth in Kanpur studied bilingualism, bidialectalism, and extended diglossia among Kanauji speakers who use Khariboli and Standard Hindi in the overlapping domains of education, media, religion, the household, work, and more. The project represents a step towards bringing recognition to this variety.

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR AFFILIATIONS

Anu Pandey  orcid.org/0000-0003-0606-047X

Indian Institute of Technology Kanpur, IN

REFERENCES

- Chaturvedi, Sujata. 2015. A sociolinguistic study of linguistic variation and code matrix in Kanpur. *Proceedings of the Procedia: Social and Behavioral Sciences* 192: 107–115. DOI: <https://doi.org/10.1016/j.sbspro.2015.06.017>
- Dwivedi, Pankaj & Somdev Kar. 2016. Kanauji of Kanpur: A brief overview. *Acta Linguistica Asiatica* 6(1): 101–119. DOI: <https://doi.org/10.4312/ala.6.1.101-119>
- Dwivedi, Pankaj & Somdev Kar. 2017. On documenting low-resourced Indian languages insights from Kanauji speech corpus. *Dialectologia: Revista Electrónica* 19: 67–91.
- Dwivedi, Pankaj & Somdev Kar. 2018. Phonology of Kanauji. In Ghanshyam Sharma (ed.), *Advances in Hindi language teaching and applied linguistics*, 189–220. Munich, Germany: Lincom Europa.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2024. *Ethnologue: Languages of the World*. Twenty-seventh edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- Grierson, George A. (1968 [1916]). *Linguistic survey of India. Vol IX: Part I*. (First Published by Superintendent Government Printing, Calcutta). Reprint Delhi: Motilal Banarsidass.
- John, Mathews & Bezily P. Varghese. 2021. The Kanauji-speaking people of Uttar Pradesh: A sociolinguistic profile. *Journal of Language Survey Reports* 2021-012: 1–19. SIL International.
- Kachru, Yamuna. 2006. *Hindi*. Amsterdam: John Benjamins Publishing Company.
- Khubchandani, Lachman M. 1972. Mother tongue in multilingual societies. In A. Chandra Sekhar (ed.), *Economic and socio-cultural dimensions of regionalization*, 427–450. New Delhi: The Registrar General India.
- Khubchandani, Lachman M. 1978. Toward a selection grammar: Fluidity in modes of address and reference in Hindi-Urdu. *Indian Linguistics* 39: 1–24.
- Kumar, Uday. 2009. *Status of Hindi in India*. New Delhi: Readworthy Publications.
- Masica, Colin P. 1993. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Mishra, Diwakar & Kalika Bali. 2010. Hindi dialects phonological transfer rules for verb root Cə̀ə. Paper presented at the 13th Oriental Conference in coordination with the International Committee for the Coordination and Standardization of Speech Databases and Assessment Techniques (COCOSDA-2010). Kathmandu, November 2010.

- Mishra, Diwakar & Kalika Bali. 2011. A Comparative phonological study of the dialects of Hindi. Paper presented at the 17th International Congress of Phonetic Sciences (ICPhS XVII). Hong Kong, August 2011.
- Pandey, Anu. 2023a. Bidialectalism and extended diglossia in Kanauji speakers. *Studies in the Linguistic Sciences: Illinois Working Papers* 46: 53–81.
- Pandey, Anu. 2023b. Morpho-phonology of interrogative pronouns in Kanauji and Hindi. In Ghanshyam Sharma & John J. Lowe (eds.), *Advances in South Asian Linguistics*, 377–400. Germany: Lincom Europa.
- Pandey, Anu. 2024. Podcasts in Kanauji: Assisting language teaching and revitalization. *Language Documentation & Conservation* 18: 1–19.
- Pandey, Preeti. 1978. *Kanauji mein sangya pad rachna* [Noun word formation in Kanauji]. Lucknow, Uttar Pradesh: University of Lucknow master's thesis.
- Santram, Anil. 1975. *Kanauji lok-sahitya* [Kanauji folk literature]. New Delhi: Abhinav Prakashan.
- Shapiro, Michael C. 1989. *A primer of modern standard Hindi*. New Delhi: Motilal Banarsidass Publishing House.
- Singh, Premshankar. 2016. Kanauji. In Ganesh Devy & Badri Narayana (eds.), *Languages of Uttar Pradesh*, 258–263 (*People's Linguistic Survey of India* Volume 29, Part 1). Hyderabad: Orient Blackswan.
- Tripathi, Suresh Chandra. 1977. *Kannauji lok sahitya mein samaj ka pratibimb* [Reflection of society in Kanauji folk literature]. New Delhi: Rupayan Publication.
- Trivedi, Govind M. 1990. Echo Formation. In Shree Krishna (ed.), *Linguistic traits across language boundaries*, 51–81. Kolkata: Ministry of Human Resource Development.
- Verma, Dharendra. 1933. *Hindi bhasha ka itihas* [History of the Hindi language]. Allahabad: Khariboli Academy.

Submitted: 07 March 2023

Accepted: 13 May 2024

Published: 23 July 2024