International Journal of Languages and Culture

Publisher's Home Page: https://www.svedbergopen.com/

SvedbergOpen
DISSEMINATION OF KNOWLEDGE
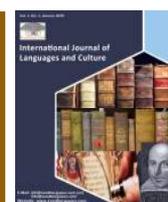
Research Paper

Open Access

Dialectal Variation and Mutual Intelligibility of Hindi in Delhi

Isha Bhola[1]* iD

[1]Master's Student at Department of Linguistics, University of Delhi, Delhi, India. E-mail: Ishabhola2310@gmail.com

Abstract

This paper aims to explore the extent of mutual intelligibility among different Hindi dialects in Delhi, the socio-cultural factors influencing variations, and language use. Language is an essential aspect of human communication, and its variations within a linguistic community often shape interactions and cultural identities. In the Indian context, Hindi is a primary language spoken across the country, with numerous regional dialects reflecting the diverse linguistic landscape. Delhi, the capital and a cosmopolitan city, presents a unique environment for studying mutual intelligibility in inter-dialectical variations of Hindi. The study employed a mixed-methods approach, combining qualitative and quantitative data collection techniques. Questionnaires and surveys were conducted across Delhi, where speakers were selected by random sampling method. Questions were administered to a sample of Hindi speakers that tested their comprehension level and analyzed factors influencing the inter-dialectical variations. This study establishes that socioeconomic factors and geographical factors are the primary reason behind the current state of the dominant language in Delhi and the rate of mutual intelligibility among the community. In this paper, 4 Dialects have been studied— Khari Boli, Bhojpuri, Haryanvi, Bagheli and Punjabi-influenced Hindi. Findings reveal a rich tapestry of inter-dialectical variations in Delhi's Hindi-speaking community. While Standard Hindi (also known as "Khari Boli") remains the most understood and widely used dialect due to its prominence in media and education, variations such as Bhojpuri, Haryanvi, Punjabi-Influenced Hindi, and Bagheli are actively present and exhibit distinct linguistic features. Although these regional dialects demonstrate differences in phonetics, vocabulary, and grammar, there exists a significant degree of mutual intelligibility, enabling effective communication among speakers.

Keywords: Dialectical Variation, Hindi Dialects, Intelligibility, Socioeconomic factors, Mutual intelligibility testing

1. Introduction

Mutual intelligibility is the extent to which speakers of one language can understand speakers of another language, without any prior knowledge or formal training in the other language according to Hockett (1958). Therefore, the number of words and phrases that a listener can infer correctly from a speaker, is the rate of mutual intelligibility among the interlocutors. The opinion of Fishman (1965), mutual intelligibility is "*the degree to which speakers of one language or dialect can understand speakers of another language or dialect without the need for significant adjustments in their linguistic behavior*".

* Corresponding author: Isha Bhola, Master's Student at Department of Linguistics, University of Delhi, Delhi, India. E-mail: Ishabhola2310@gmail.com

How is mutual intelligibility and language variation connected? Mutual intelligibility and variations in language are intricately linked. While variations within a language or closely related dialects promote higher mutual intelligibility, significant differences in pronunciation, vocabulary, and grammar can decrease comprehension. Understanding the connection between these concepts is essential for linguists, language policymakers, and educators to facilitate effective communication and promote linguistic harmony in diverse linguistic communities.

Weinrich (1953) explains mutual intelligibility as "the ability of speakers from different language varieties to comprehend each other with some ease, even though they may use different linguistic forms." Therefore, the notion of mutual intelligibility to exist includes the consideration of different linguistic forms that are prevalent.

The interconnection between linguistic varieties and mutual intelligibility is a continuum where mutual intelligibility can be asymmetrical in nature. As Trudgill (2003) claims that mutual intelligibility might not be equal between different linguistic varieties, the calculation of factors behind such asymmetrical causes can be performed. The primary factors that impact the mutual intelligibility rate are whether the particular linguistic dialect variety is adjacent in geographical nature or the socioeconomic factor such as social mobility, urbanization, and class distinctions affecting the communication discipline.

2. Impact of Language Variations on Mutual Intelligibility

- **Linguistic Similarity:** The more closely related two language varieties are, the higher the mutual intelligibility. Variations within a language tend to maintain a higher degree of mutual intelligibility, making communication easier.

- **Phonological Variations:** Differences in pronunciation, phonetic features, and accents can affect mutual intelligibility. For example, speakers from different regions may have difficulty understanding each other due to distinct phonetic nuances.

- **Lexical Variations:** Variations in vocabulary, regional expressions, and loanwords can either enhance or hinder mutual intelligibility. Similar words aid comprehension, while unfamiliar terms may lead to confusion.

- **Grammatical Variations:** Differences in grammar, word order, and sentence structures can influence mutual intelligibility. Shared grammatical rules facilitate understanding, while significant discrepancies may cause misunderstandings.

The Indo-Aryan language, Hindi, has multiple dialects. The classification of all the dialects is performed on a geographical basis by CIIL Mysore:

Western Hindi Dialects	Dialects are spoken in states like Rajasthan, Haryana, and parts of Madhya Pradesh, such as Marwari, Haryanvi, and Malvi.
Eastern Hindi Dialects	Dialects are spoken in Bihar, Jharkhand, and parts of Uttar Pradesh, such as Bhojpuri, Magahi, and Awadhi.
Central Hindi Dialects	Dialects are spoken in the central region of India, particularly in the Bundelkhand and parts of Madhya Pradesh, such as Bundeli.

In this paper, the inter-dialectical variations present in the Delhi region are studied, for that purpose, a database of four dialects and one mixed language has been created—Khari Boli, Bhojpuri, Bagheli, Haryanvi, and Punjabi-influenced Hindi. According to the 2015 Census, the population of Delhi is comprised of Punjabis at a major level, which can be associated as a reason for Punjabi and Hindi to be the lingua franca of Delhi. Hindi dialects have been studied previously which explains the diversity of the Hindi language including—Rajendra Singh (1991); Jayaseelan (2001). One study on the mutual intelligibility of Hindi dialects has also been studied by Arpita Ray (2014). Delhi's linguistic landscape is characterized by a diverse population that represents various regions of India. People from different states and linguistic backgrounds reside in Delhi, making it a stage of Hindi dialects and languages from other parts of the country. As the capital city of India, Delhi has historically been a major centre of cultural, political, and economic activities. Its central location attracts people from all over the country, resulting in a convergence of diverse dialects. Therefore, Delhi serves as a relevant and accurate subject for analysing Hindi dialect variation and intelligibility.

1. What is the rate of mutual intelligibility of the speakers of different dialects of Hindi among the residents of Delhi?

2. What role does geographical and socioeconomic factors influence the variations in language use of Hindi in the Delhi region?

For this research paper, the mixed approach of quantitative and qualitative methodology is used, and the participants for this research has been taken from the North-West Delhi, North Delhi, North-East Delhi and South Delhi. The methodology was inspired by Haugen's questionnaire method. For providing a theoretical framework, compilation and comparison of socioeconomic statistics and geographical data has been performed in this study.

3. Methodology

For collection of data, a method inspired from prominent researchers has been used along with the usage of linguistic

Table 1: Collection of Data

S. No.	Method	Prominent Researchers	Example
I	Sociolinguistic Interviews	William Labov and Surveys	The classic study of the New York City Department Store (Labov, 1966)
ii	Linguistic Atlas and Dialectology	Hans Kurath	Linguistic Atlas of New England (Kurath, 1939), which mapped the dialectal variation in the New England region of the United States
iii	Quantitative Analysis and Corpus Linguistics	Douglas Biber	Longman Grammar of Spoken and Written English (Biber *et al.*, 1999), which analysed language variation across different registers
iv	Code-Switching and Language Contact	Penelope Gardner-Chloros	Code-switching in multilingual communities (Gardner-Chloros, 2009)
v	Ethnographic and Participant Observation	Monica Heller	Ethnographic methods to study language variation, identity, and power in various multilingual communities (Heller, 2007)

corpora published, socioeconomic statistics of Delhi and ethnic group statistics of Delhi. Few methods used for studying language variation and some prominent researchers associated with these methods:

For this study, opinion testing- the questionnaire and surveys method are used as it is considered efficient means of gathering large amounts of information from sizeable sample volumes. This method was chosen for the purpose of it getting completed a at a self-paced and convenient time of the participants.

3.1. Methodology for this Research

For calculating the degree of mutual intelligibility among Hindi speakers in Delhi, a questionnaire was designed that contained questions which tested the comprehension level of different dialects and interviewed their knowledge about a few words and phrases from various dialects of Hindi. The questionnaire included free speech phrases, word lists, and queries about their awareness about different dialects. Each question was assigned a point for correct answer, the correct answer being the complete awareness of all dialects, words/phrases and comprehensibility of each dialect. The answers will be analyzed, percentage will be calculated and then overall degree will be interfaced by scoring system. For putting a theoretical framework for analysing the socioeconomic factors involved behind the dialectical variations of Hindi in Delhi, brief study of economic, political, and ethnic group statistics has been discussed.
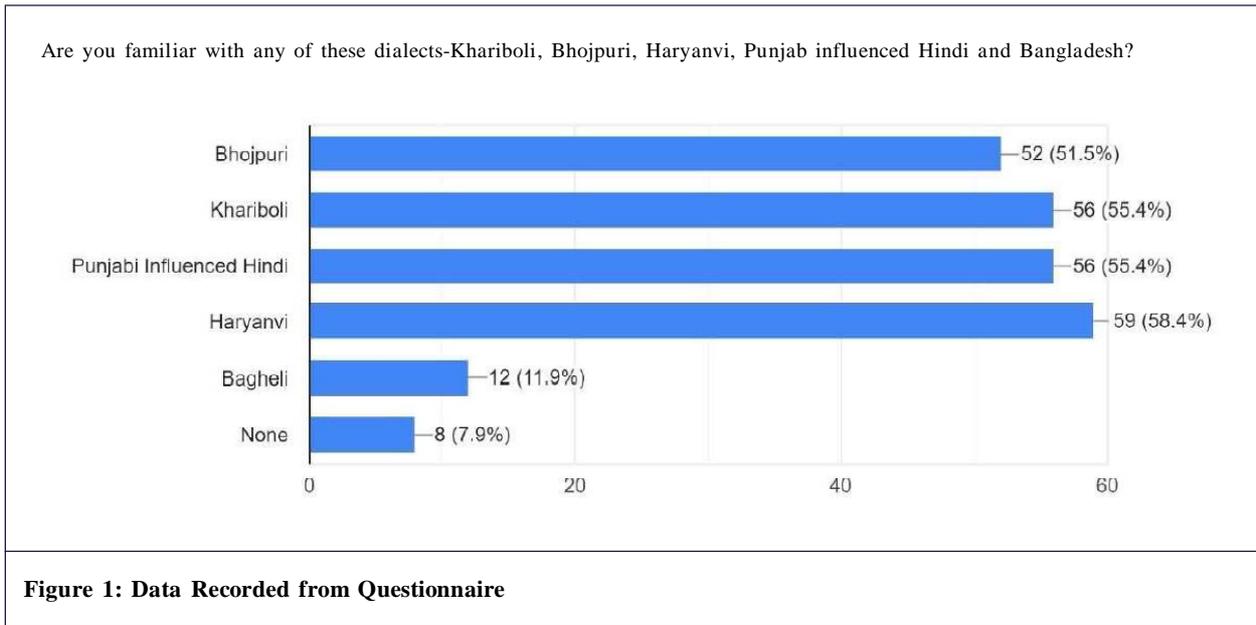
3.2. The Participants of the Study

The participants of this study were 102 residents of Delhi. Male and female from 4 districts of Delhi. They were asked to give answers to the questionnaire provided to them. The age range of the participants were from 15 to 50, and it was ensured that all the participants are aware of the Hindi language and were literate since they were asked information about dialects and their comprehension level.

3.3. The Questionnaire and Procedure

After enquiring about personal data of the participants, the followed question asked the participants if they are familiar with the dialects- Khari Boli, Bhojpuri, Haryanvi, Bagheli and Punjabi influenced Hindi.

The calculated percentage of the participants' answers:

Are you familiar with any of these dialects-Khariboli, Bhojpuri, Haryanvi, Punjab influenced Hindi and Bangladesh?

Dialect	Value
Bhojpuri	52 (51.5%)
Khariboli	56 (55.4%)
Punjabi Influenced Hindi	56 (55.4%)
Haryanvi	59 (58.4%)
Bagheli	12 (11.9%)
None	8 (7.9%)

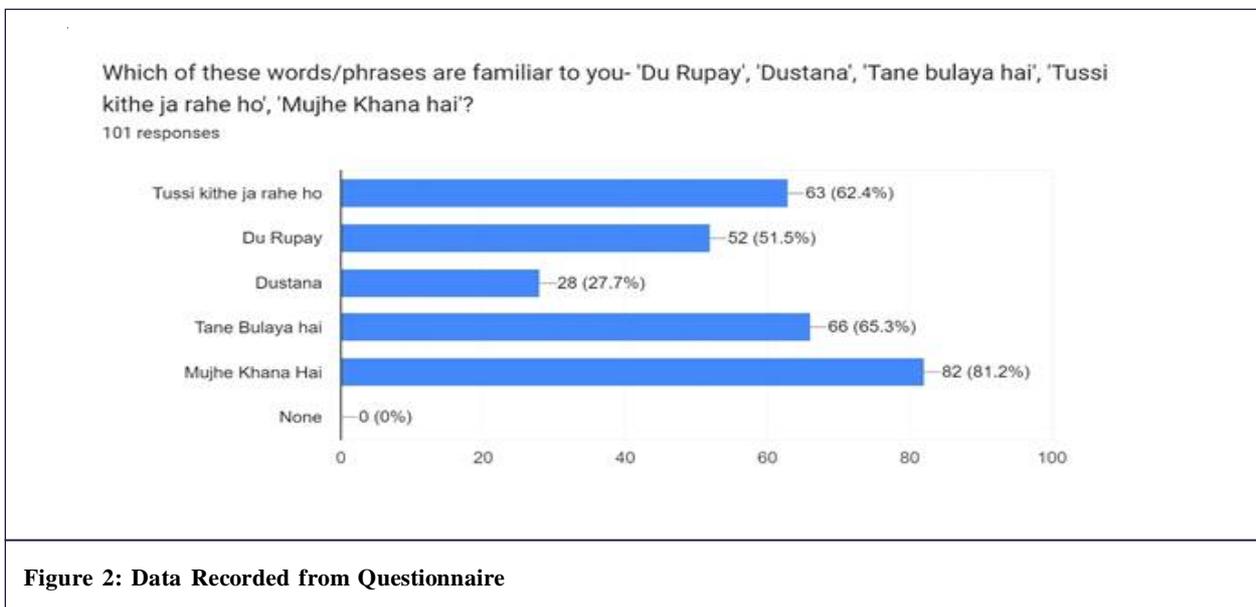**Figure 1: Data Recorded from Questionnaire**

The next question mentioned a few phrases and words from the aforementioned dialects, and participants were asked to recognize the phrases/words from the list.

**Table 2: Texts Mentioned In the Questionnaire**

Texts	Dialect	English Gloss Translation
/ ɖu ɾupɛ /	Bhojpuri	Two rupees
/ t̪uśi kit̪ʰɛ ja ɾəhɛ ho /	Punjabi influenced Hindi	Where are you going?
/ dust̪ana /	Bagheli	Friendship
/ muʒɛ kʰana hæ /	Khari Boli	I have to eat
/ t̪əɳ́ɛ bu:laja hæ /	Haryanvi	You have been called

The texts:

Which of these words/phrases are familiar to you- 'Du Rupay', 'Dustana', 'Tane bulaya hai', 'Tussi kithe ja rahe ho', 'Mujhe Khana hai'?
101 responses

Phrase	Value
Tussi kithe ja rahe ho	63 (62.4%)
Du Rupay	52 (51.5%)
Dustana	28 (27.7%)
Tane Bulaya hai	66 (65.3%)
Mujhe Khana Hai	82 (81.2%)
None	0 (0%)

**Figure 2: Data Recorded from Questionnaire**

The calculated percentage of the participant' answers:

The next question asked the participants about their fluency or familiarity with any of the mentioned dialects.
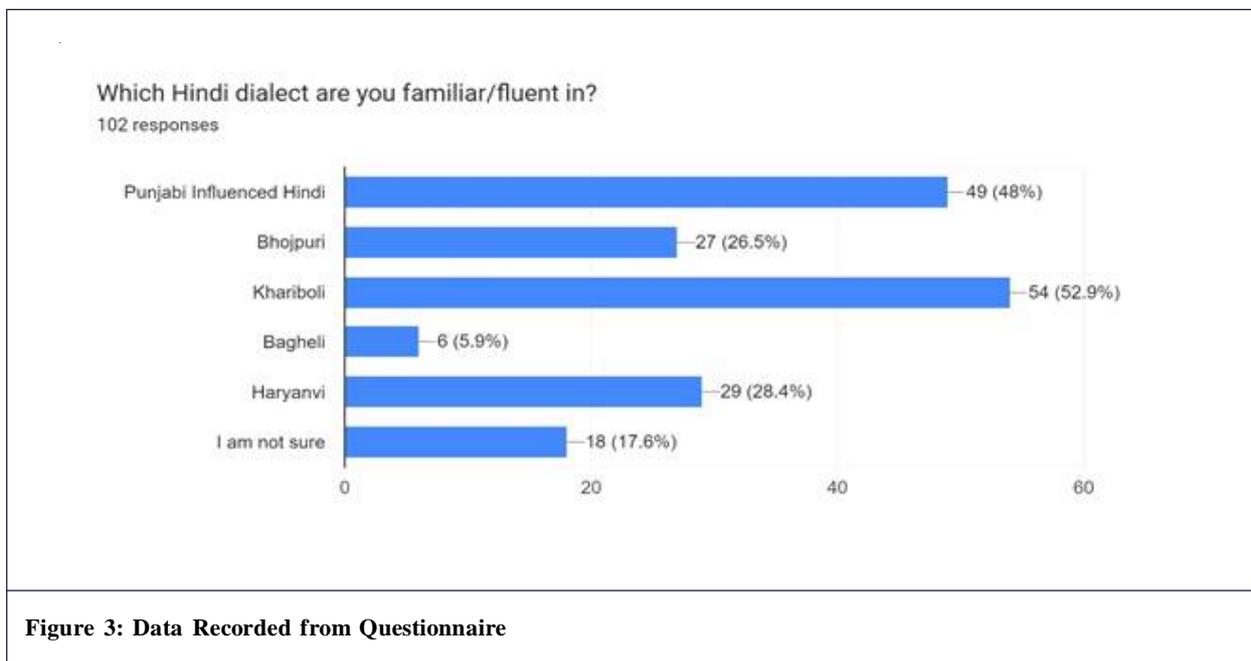
Figure 3: Data Recorded from Questionnaire

The calculated percentage of the participants' answers:

The next question asked the participants that which of the aforementioned dialects when spoken, they have a problem comprehending what speaker intends to say?
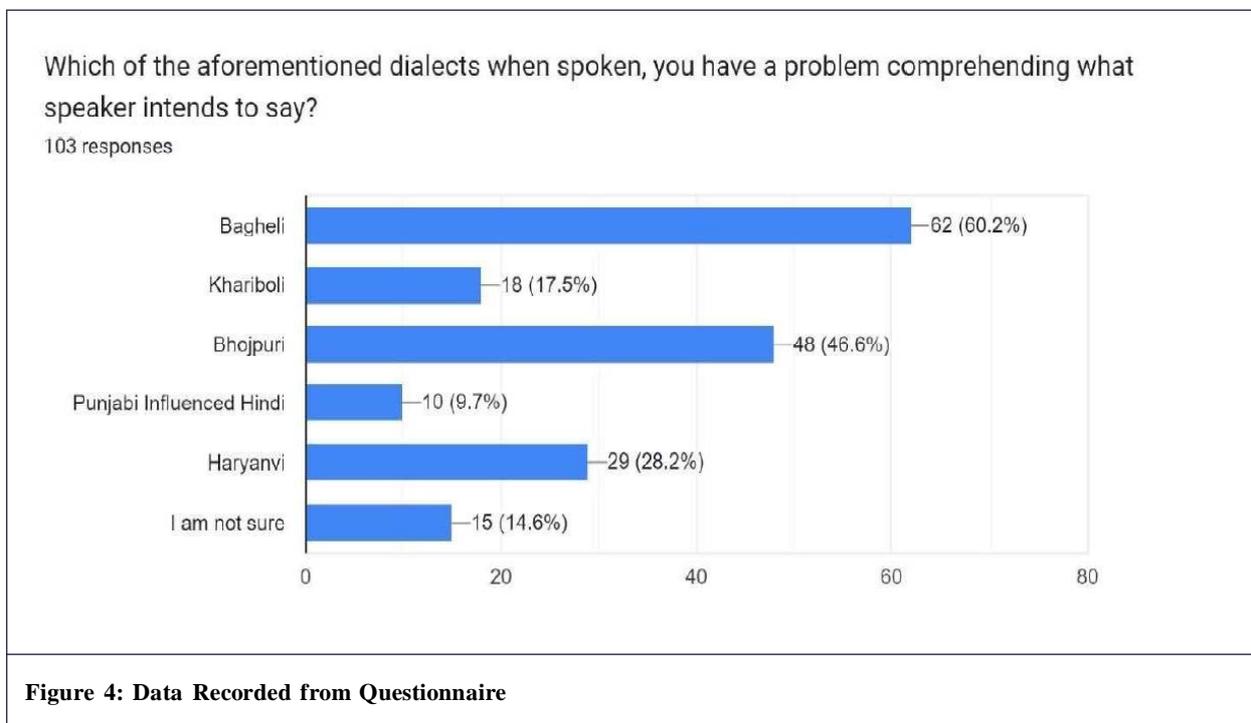
Figure 4: Data Recorded from Questionnaire

The calculated percentage of the participants' answers:

The next question asked the participants that which of the aforementioned dialects when spoken, you can easily comprehend what speaker intends to say?

The calculated percentage of the participants' answers:

Figure 5: Data Recorded from Questionnaire
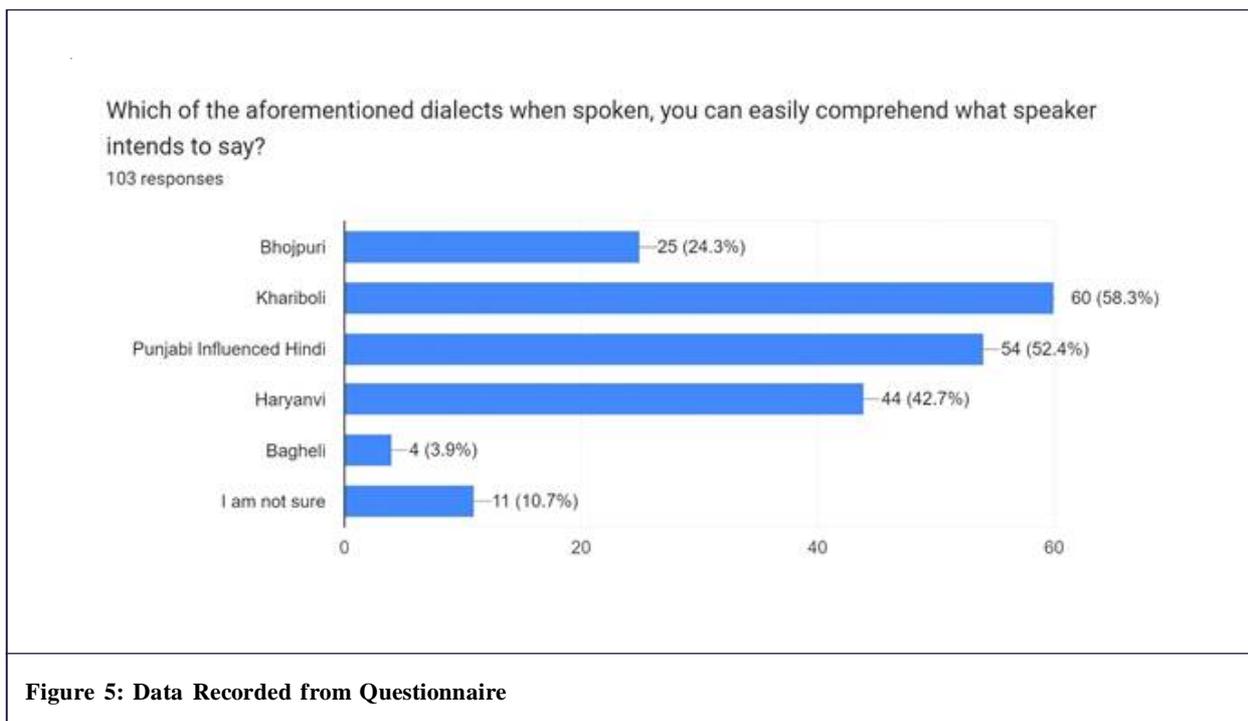
4. Procedure

As discussed in the preceding sections, the 102 participants were asked certain questions. These questions were designed to test the degree of mutual intelligibility the Hindi speakers in Delhi have. They involved commonly used texts from each dialect so that individuals are able to recognize the speech utterances. Their familiarity was tested with the direct question 'How many of these dialects are you familiar with?' and with the recognition test of words/phrases from the provided list.

The comprehension level of individuals regarding each dialect separately was determined by asking direct and easy to understand questions, and the level of their knowledge was determined about the variations in dialects within their currently residing location.

The last round of the questionnaire was two interrelated questions for individuals to retell their classification which they consider correct about their knowledge of dialectal variation, along with identifying the rate of mutual intelligibility.

4.1. Scoring Procedure of Analysis

- 1 point is assigned to each successful recognition of the word/phrase.
- 1 point is assigned to each choice option that mentioned familiarity with each dialect.
- 1 point is assigned to each choice option that mentioned the individual is fluent and familiar with each dialect.
- 1 point is assigned to each choice option which stated the individual being able to comprehend the speaker of each dialect.

The points are added collectively together of all participants.

$$Mutual\ intelligibility\ score = \frac{(Selected\ choice\ raw\ score)}{(Maximum\ raw\ score)} \times 100$$

The scoring of the results is based on initially the words/phrase recognition test, which gave the percentage of each individual's mutual intelligibility within Delhi and as every resident had chosen more than one familiar word/phrase from the list, the mutual intelligibility can be identified with the respect to particularity of dialects. In other words, how many dialects can the Hindi speakers of Delhi comprehend without any formal or effortful attainment of knowledge about the dialects.

The direct questions provided the percentage of degree of comprehension that Delhi residents hold in terms of inter-dialectal variations. All the questions were interfaced in such a way that demonstrated the more spoken/comprehensible mutually dialects than the others. The validation of this particular finding collected from this practice will be validated by

investigating the social factors that influence the variations and are responsible for dominant dialects in the following sections.

5. Hypothesized Framework for Socioeconomic and Geographical Factors for this Paper's Findings

5.1. Geographical Factors

When groups of people are physically separated from each other, they develop their own linguistic patterns over time. On the other hand, regions with high levels of contact and interaction between different communities tend to experience more language mixing and borrowing, leading to linguistic convergence.

Some factors that cause Dialect Variation: Colonial History: The history of colonization can lead to language contact and linguistic changes. Colonizers often bring their languages to new territories, leading to the adoption of loanwords, grammatical structures, and phonetic influences. Trade and Urbanization: Urban centers, which serve as hubs for trade and commerce, often experience a mix of languages and dialects due to interactions between people from diverse backgrounds. This can lead to the emergence of urban dialects that differ from rural ones. Cultural Practices: Certain cultural practices, rituals, or beliefs unique to a region can be reflected in the language and contribute to dialect variation. Language Policy and Education: Government language policies and the education system can also influence dialect variation. Standardized languages taught in schools can lead to homogenization, while local dialects may persist in informal settings. Social Identity: Geographical factors can be linked to social identity. People may identify with their regional dialect as a way to express their belonging to a particular community or distinguish themselves from others. Migration Patterns: Migration patterns within a region can influence dialect distribution. For example, if people from one region move to another, they may bring their dialect with them, leading to the spread of particular linguistic features.

The geographical distribution demonstrates that Delhi region is an amalgamation of boundaries of Uttar Pradesh, Haryana, Uttarakhand, and Punjab. Therefore, the dialect variations induced due to language contact. Regions with active trade and commerce tend to experience more language contact. Traders and merchants from different linguistic backgrounds interact, leading to the exchange of words, phrases, and grammatical structures, resulting in language variation. This distribution validates the data received from the findings of this paper that prevalent dominant dialects among Hindi speakers in Delhi is from Uttar Pradesh, Haryana and Punjab.

5.2. Socioeconomic Factors

Socioeconomic status, often associated with income and social class, can affect language variation. People from different social classes may use different vocabulary, pronunciation, and grammatical patterns as markers of their identity and social group. Dominant language that prevails in a city, is directly impacted by the dominant social class, whether in politics, urbanization, education or migration.

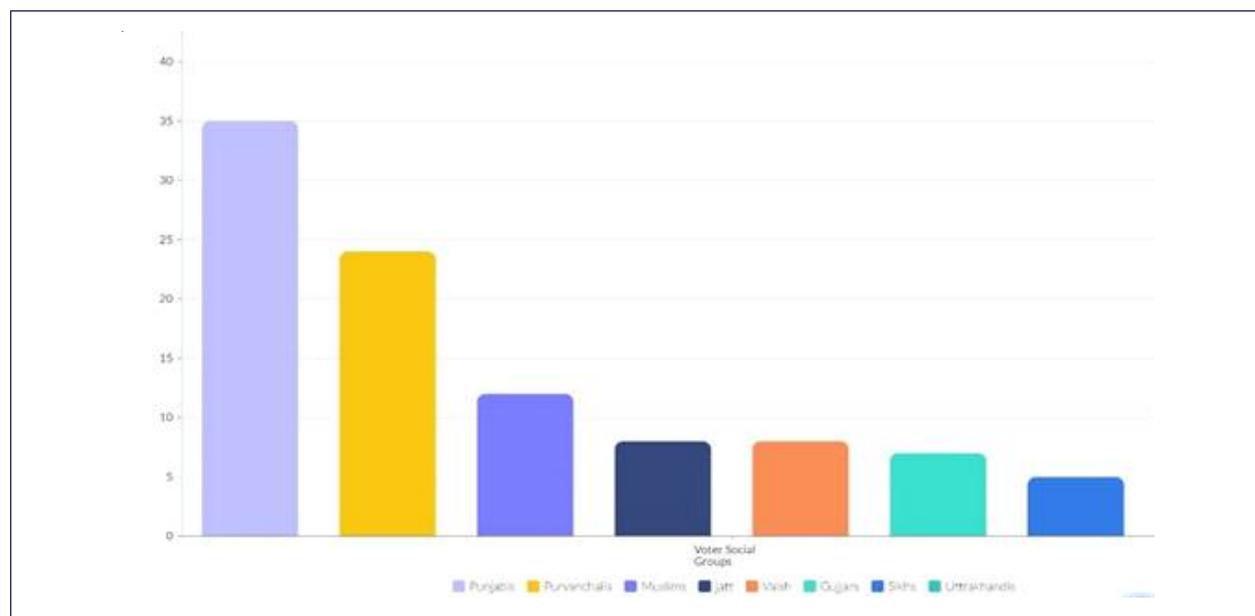Here is the data from 2015 Indian TV article that demonstrates the major social groups voter in Delhi.

Figure 6: Major Social Groups Voter in Delhi in the Year 2015

According to the latest available Census data, approximately two-thirds of new migrants to Delhi originate from Uttar Pradesh (UP) and Bihar. Notably, the percentage of migrants from Bihar has seen an increase.

Currently, migrants constitute more than 40% of the population in three districts of Delhi— New Delhi, South, and South West. In the North West, North East, and South districts, each of them is home to over five lakh migrants from UP, and the North West, South West, and South districts house over two lakh migrants from Bihar.

Migrants from Haryana have a significant presence in the North West and South West areas of the city. Interestingly, in these two districts, migrants from Punjab and Haryana combined outnumber those from Bihar. Migrants from southern India, though fewer in number, are concentrated in the East regions like Mayur Vihar and Patparganj, as well as the central districts.

Migration and movement of people from one region to another can lead to language contact and the adoption of linguistic features from different dialects or languages. This process often results in language convergence and new variations. The statistical data of socioeconomic status of Delhi, validates the data of dialectical mutual intelligibility calculated. As it is clear that in findings of this paper, the percentage of Punjabi-influenced Hindi language, Haryanvi and Khari Boli dialect was relatively higher than the rest. This outcome can be associate its reasoning with degree of migrant population and dominant social groups voters in Delhi.

The media and technology accessible to individuals can influence language use. Exposure to various forms of media, such as television, internet, and social media, can introduce new words, phrases, and language trends into a person's speech. As demonstrated in the findings of this paper, the Haryanvi dialect is the most intelligible dialect among Hindi speakers in Delhi. Haryanvi has successfully made its presence count into Indian cinema (*Hindustan Times*, 2018), TV, Popular music albums and academia (Tribune, 2018).

Government language policies, such as official language status or language education initiatives, can impact dialect variation. Language planning efforts may promote the standardization of a particular dialect or suppress the use of regional variants. The rise of Punjabi influenced Hindi and the language Punjabi is a result of certain language policies that have been introduced for Delhi government schools. Right to Information (RTI) Act, has reported that of the 1,100 government schools in Delhi, Urdu and Punjabi languages are being taught in 284 and 282 schools, respectively (*Hindustan Times*, 2018). Under the National Education Policy (NPE), students in Hindi-speaking states have to learn a modern Indian language as the third language—apart from Hindi and English. In Delhi government schools, Sanskrit, Urdu and Punjabi are being offered as the third language.

Language policies can influence language use among migrants and newcomers to the city. If the city's language policy promotes a specific dialect, migrants might adapt their speech to conform to that norm, leading to linguistic convergence. In some cases, language policies may be designed to preserve the cultural heritage of specific linguistic communities. Such policies can promote the use and maintenance of local dialects, contributing to dialect preservation. Therefore, the rising use of Punjabi in Delhi can be associated as an outcome of this cause.

6. Calculation of Scores

Overall mutual intelligibility percentage of each testing question and of each inter-dialectal variation.

Table 3: Overall Intelligibility Rate per Each Test		
S. No.	Test	Overall Intelligibility (%)
1.	Familiarity Test	47
2.	Word/Phrase Recognition	58.2
3.	Fluency Test	33
4.	Comprehensibility Test	37.4

S. No	Dialect	Intelligibility Score (%)
Table 4: Overall Intelligibility Rate per Each Dialect		
1.	Khari Boli	64
2.	Bhojpuri	43.5
3.	Haryanvi	49.5
4.	Bagheli	12.5
5.	Punjabi Influenced Hindi	55.5

7. Findings and Discussion

This section will throw light on the degree of mutual intelligibility that Hindi speakers have who reside in Delhi and will discuss about the dialects that have been influenced maximum by the social factors.

1. The rate of overall dialectal intelligibility was the highest for the dialect Khari Boli 64%. The Punjabi-influenced Hindi dialect followed with second highest rate of 55.5%. The Haryanvi dialect holds the third highest rate of 49.5%. The dialects with the lowest rates are Bhojpuri 43.5% and Bagheli 12.5%.

2. The data revealed that lowest rate of mutual intelligibility for speakers in Delhi was the test of fluency. The average rate of successful chosen options in the opinion testing with fluency in dialects was 33%.

3. The overall mutual intelligibility rate among Hindi speakers of Delhi is 43.9 based on the findings of the survey conducted for this study.

The conclusion drawn after performing background research and interlinking the collected data:

1. The highest rate of dialectal intelligibility is of Khari Boli is the outcome of the highest number of migrants from Uttar Pradesh in Delhi, as the Khari Boli dialect is spoken in Uttar Pradesh vastly. UP is also the geographically closer state of New Delhi, multiple geographical factors impacting the dialectal variation.

2. The second highest rate is of Punjabi influenced Hindi, the responsible factors are Punjabis being the most dominant social voter group in Delhi, and numerous language policies introduced to have Punjabi being taught as third language in Delhi Government Schools.

3. Haryanvi being relatively higher in rate than other dialects is the result of the large number of migrants from Haryana, influence of technology and media on Delhi residents. Haryanvi being excessively popular on media has impacted the Hindi speakers to comprehend Haryanvi and has increased the dialectal intelligibility among Hindi Speakers.

4. Bhojpuri dialect is lesser in rate but holds a good degree of percentage in the distribution, indicating that this particular dialectal intelligibility is a result of progressive increase in number of migrants from Bihar over the years.

5. Bagheli dialect being the lowest in rate is associated with the reason that the region Bagheli is spoken in geographically further in distance than Delhi, the number of migrants from Baghelkhand region is less and the involvement of Bagheli dialect in media, technology. Education policies and urbanization.

8. Conclusion

In conclusion, this paper made an attempt to investigate the mutual intelligibility of Hindi among Delhi residents and the social factors involved in the dialectal variation in Delhi. It establishes that the dialectal intelligibility confirms to the trends seen in current socioeconomic status of Delhi. In this dialect continuum, the rates of intelligibility are highest in the geographically adjacent dialects and those regions from where the highest numbers of migrants are residing in Delhi. This paper integrates the importance of socioeconomic and geographical factors involved in the determination of dialectal variation and intelligibility. Apart from these factors, the motivations such as urbanization, media, education, etc have an impact on dialectal intelligibility. This paper establishes that the mutual intelligibly rate in the hot melting point of variations such as Delhi is 43.9%.

Acknowledgment

Funding

No funding was received for conducting this study.

Conflicts of Interest

The author has no conflicts of interest to declare that they are relevant to the content of this article.

References

Bell, A. (1984). *Language and Social Contexts*. Routledge.

Bernstein, B. (1971). *Language and Social Class*. Routledge & Kegan Paul.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan Bieber, E. (1973). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Bickerton, D. (1981). *Roots of Language*. Karoma Publishers.

Bloomfield, L. (1927). *Language Geography: A Plea for More*. University of Chicago Press

Eckert, P., Rickford, J. and Epstein, S.L. (Eds.). (2008). *Language in Geographic Context*. Stanford University Press.

Ferguson, C. A. (1964). Diglossia. *Word*, 15(2), 325-340.

Fishman, J. (1965). Who Speaks What Language to Whom and When? La Linguistique, 1(2), 67- 88. https://www.jstor.org/stable/30248773

Gumperz, J.J. (1982). *Language and Social Identity*. Cambridge University Press

Gardner-Chloros, P. and Weston, D. (2015). Code-Switching and Multilingualism in Literature. *Language and Literature, 24(3), 182-193.*

Heller, M. (Ed.). *Bilingualism: a Social Approach* (pp.1–21). New York: Palgrave Macmillan.

Hockett, C.F. (1958). *A Course in Modern Linguistics. New York: Macmillan.*

Jayaseelan, K. A. (2001a). IP-Internal Topic and Focus Phrases, Studia Linguistica, 55, 39-75.

Kachru, Y. (1965). *Hindi Dialects: A Sociolinguistic Approach*, 41(4). Linguistic Society of America.

Kurath, H. (1973). *Handbook of the Linguistic Geography of New England* (Second Ed.). New York: AMS Press. pp. xi–xii.

Labov, W. (1972). *Sociolinguistic Patterns*. University of Pennsylvania Press.

Labov, W. (1966) The Social Stratification of English in New York City (Washington, DC: Center for Applied Linguistics).

Milroy, L. (1980). *Social Factors in the Explanation of Language Change*. Blackwell Publishing.

Nichols, Patricia, C. (1976). *Linguistic Change in Gullah: Sex, Age and Mobility.* Ph.D. Diss., Stanford University.

Singh, R. (1991). Interference and Contemporary Phonological Theory. *Language Learning*, 41, 157-175.

Trudgill, P. (1986). *Dialects in Contact*. Blackwell, London.

Trudgill, P. (2003). *A Glossary of Sociolinguistics*. Edinburgh University Press, Edinburgh.

Walter, C. Sherzer. (1983). *Language Variation in North Central India*. Summer Institute of Linguistics

Weinreich, U. (1953). *Languages in Contact, Findings and Problems*, Linguistic Circle of New York, Michigan.

Williams, C.H. (1980). *The Geography of Language*. Indiana University Press.