

4 RESEARCH METHODOLOGY

This chapter discusses the overall methodology of the present study. The following sections describe the theoretical framework that has influenced the research design for the survey. The chapter is arranged in the following manner: Section 1 reviews the rationale behind using mixed methodology. Then, the detailed account of fieldwork conducted in several phases and its various components such as survey, sampling, details of the respondents etc., is provided in section 2. Finally, section 3 describes the techniques used for the analysis of the field data.

Currently, there is no universally accepted model for the study of language shift and maintenance, as the spectrum of variables and its relative values in various cultural and social contexts are too wide and distinct to accommodate in a single framework. There are two approaches to undertaking such sociolinguistic studies: the quantitative and the qualitative approaches. The quantitative method follows a deductive approach where the already existing theories or proposed hypotheses are scrutinised to obtain the set of postulates or assumptions that could be tested based on the collected data. The qualitative approach follows an inductive approach where the observations are used to propose new theories. The qualitative approach is appropriate when the goal is to comprehend the level of supporting a language enjoys from its speakers and the behavioural patterns and the language belief of its speakers.

The qualitative approach has often been employed in language shift and endangerment studies for languages with a small population size. A few examples of such studies are Marlow & Giles (2008), Sicoli (2011), Orcutt-Gachiri (2011), and Lüpke (2015), Odango (2015). These studies have explored the factors related to the attitude and ideology of the speakers towards their respective languages and ideological and attitudinal factors in language shift. On the other hand, the quantitative approach is useful for conducting studies on languages with a large speaker size

population. For example, Abtahian et al. (2016) favour the quantitative approach over qualitative to study language shift among millions of speakers in Indonesia. In sociolinguistic and language vitality investigation, some studies have used the survey method, especially to study the beliefs and attitudes of the speakers (see chapter 3). For example, the national Indigenous Language Survey (NILS) of the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) have used a survey method of analysis (AIATIS, 2005; 2014). The quantitative data helps analyse large-scale data from which the result can be deduced to prove or disprove the hypothesis (Rasinger, 2013). The advantage of using the survey method for data collection is that it is easily manageable, measurable, easy to compare across respondents (Brown, 2001; Paulhus

4.1 RATIONALE BEHIND CHOOSING MIXED METHODOLOGY

Initially, a pilot survey was conducted on a sizeable number of participants to get some idea about the speakers regarding their beliefs and opinion about the Maithili language. The focus was mainly on the questions like: what changes have occurred in the status of Maithili after its inclusion in the ES? Also, what is the current situation of the Maithili language in Bihar? The questionnaire survey method was used in the pilot study, which helped get attitudinal data from the large population size. However, it was not effective in understanding the language ideology, which is a major aspect of the linguistic situation of Bihar. During the survey, it was found that the respondents were randomly answering (choosing options) without paying much attention to the questions. In the course of interaction, it was perceived that only a handful of people could read and write in Maithili. In the questionnaire survey, most respondents rated themselves high on their literacy skills, i.e., reading and writing. In survey research, there is a high chance of results getting affected due to social desirability or social prestige, as "results represent what the respondents

report or believe, rather than what they actually feel or believe" (Dornyei and Taguchi, 2010, p.8). Another problem pointed out in survey research is that the closed-ended questions do not give an in-depth insight into the problem, as it "fails to penetrate the meaning and understandings of people who are sampled" (Baker, 2006, p.224). These shortcomings of the quantitative method were overcome using the mixed-method approach that involved focused group discussions and unstructured interviews along with a survey questionnaire.

The mixed-method approach is an amalgamation of both qualitative and quantitative methods (Angouri, 2010). The mixed-method approach is a common practice in the field of Sociolinguistics to study the maintenance or shift of a language (Sridhar, 1988; Hamid, 2005). The triangulation process (Denzin, 2017) helps coalesce the results from research (Angouri, 2010). Combining qualitative and quantitative approaches helps to "create a synergistic evaluation project, whereby one method enables the other to be more effective and together both methods would provide a fuller understanding of the evaluation problem" (Mertens & Hesse-Biber, 2013, p.7). Often the mixed-method approach is favoured for attitudinal research, which is a subjective factor and varies from one community to another or from one culture to another. Previously, much research has been done using a quantitative approach that has failed to consider the subjective factor in language attitude study. The mixed-method approach helps in getting a response from a large sample size of the population through the quantitative method and gives a more in-depth insight through qualitative analysis. Also, the techniques of focused group discussion and interviews provide flexibility to improvise questions during the discussion to get maximum insights.

4.2 MULTIPLE TRIPS TO FIELD

The fieldwork was carried out in different phases between March 2018 - Dec 2018. However, the very first round was carried out in July 2016 for the orientation. The professors and writers of Maithili were interviewed, who were directly involved in the movement for Maithili's inclusion in the ES. The interaction helped in contextualising the language use pattern in Darbhanga and Madhubani. The first phase of the field trip mainly helped in rapport building and understanding the language use pattern of the speakers by obtaining their linguistic history. Fishman's domain analysis was the basis for the formation of the questionnaire. Apart from domain analysis, the questionnaire also aimed to understand whether the attitude guides the linguistic choice of the speakers.

In the second phase (March 2018 - April 2018), a pilot survey was carried out on 115 respondents from both districts. A final survey followed it after incorporating some modifications based on the suggestions and feedback received from the experts (Maithili writers, professors and linguists). In the final questionnaire, the variable of caste was added, and a few questions were eliminated. The questionnaire was sent to three linguists across India through email to check the validity of the content and based on their feedback and comments, the questionnaire was administered for the pilot survey. Apart from linguists, some Maithili scholars were also consulted, and their feedbacks and suggestions were also incorporated in the final questionnaire. The questionnaire was prepared in English, which was later translated into Maithili with the help of two resource persons (a writer and a professor) of Maithili.

A questionnaire was administered, preceded by an informal interaction with every participant to maintain data accuracy. Then, an interview administered questionnaire survey was conducted to get the response from the respondents. Each respondent was targeted individually in

the survey process that started with a small interview followed by an interviewer-administered questionnaire. Two local persons (Rakesh Kumar and Chandan Kumar, PhD Scholars from Lalit Narayan Mithila University) also accompanied and extended their help throughout the fieldwork.

In the third phase of fieldwork (July 2018), I participated in the ‘Maithili script Workshop’ organised by CIIL (Central Institute of Indian Languages) Mysuru, at Darbhanga, where I mainly interviewed teachers to gain an insight into the condition of Maithili in education. Although I chose my respondents mainly working in schools of Madhubani and Darbhanga districts, I also interacted with a few teachers from other districts such as Bhagalpur, Samastipur, and Muzaffarpur to gather information about the implementation of Maithili education in schools and colleges. It was found that there is a scarcity of Maithili teachers and the unavailability of textbooks and better teaching aid. These interviews were mainly focused on understanding the government’s attitude towards Maithili, i.e., factor 7 (discussed in chapter 5). In the last and final phase of fieldwork (October - December 2018), I mainly conducted interviews (unstructured, focused group discussion and in-depth interviews with local people, which was also recorded. During my fieldwork, I also attended Vidyapati Mahotsav organised in Patna and Janaki Mahotsav in Madhubani.

Independent of the fieldwork, I also visited- the Janakpur district of Nepal, where I got the first-hand experience of Maithili’s visibility in signages such as road signs, shop names etc., across the district. This was a stark contrast from the scenario in Darbhanga and Madhubani districts, where such signages were negligible. Though the signages and their significance are out of the scope for the present study, they pose an interesting study that could be tak

4.3 THE SURVEY

The survey questionnaire comprised of 38 questions divided into the following sections:

1. background (13 questions), 2. Language proficiency for Maithili & Hindi (4 questions each), 3. Attitudes and Beliefs of the speakers (10 questions), 4. Use of Maithili language in the community (11 questions). The survey was designed to elicit responses related to language proficiency, attitude and usage of speakers. There were 38 questions overall, out of which items measuring language proficiency and attitude were on a Likert scale (Likert, 1932). The use of a Likert scale is common in studies that measure attitudes (Rasinger 2013). The remaining questions used dichotomous responses and multiple options. A survey with four sections was designed that covered questions related to the topics provided in the table below:

Table 4.1 Vitality Factors and the Scope of Description

	Vitality Factors	No of questions
1	Demographic details	13
2	Domains and functions	11
3	Knowledge of language	8
4	Attitudes and Beliefs	10

4.4 SAMPLING

The data was collected from two districts, which records the highest number of Maithili speakers. As per the 2011 census, Maithili was returned by 84.1 and 72.7 percent of the population in Madhubani and Darbhanga districts. The selection was made on the basis to ensure the representation of both urban and rural populations. The urban population was covered in the questionnaire survey in both the districts, while the interviews targeted both rural and urban population. According to the 2011 census, the total percentage of the urban population in Darbhanga and Madhubani is 9.7% and 3.6% of their total population, respectively. Of this total

urban population in the Darbhanga, the highest percentage (52.6%) is concentrated in Darbhanga sub-district. Similarly, the Madhubani sub-district has the highest urban population of 25.6 percent in Madhubani. Therefore, data was collected from these two sub-districts ($N = 179 + 201 = 380$). The respondents were selected using a simple random sampling technique. In the next step, wards and panchayats were selected from Darbhanga and Madhubani, and the household survey was carried out in the selected panchayats and wards.

Further, the respondents were categorised into different age cohorts as the study aims at assessing the intergenerational language patterns of Maithili speakers in terms of language proficiency and language preference. Note, in sociolinguistic survey the difference in usage in both the districts were almost the same. In contrast, it turned out in interviews the people of Darbhanga, especially in town are shifting to Hindi. Moreover, the language usage varied with regard to caste; and religion in this district compared to Madhubani. The interviews were conducted using purposive sampling. The village and panchayats were selected based on different features such as caste, religion and rural population to study the dynamics associated with these factors.

Age Cohorts

For the ease of analysis, the participants were classified into five age cohorts- 15 to 25 years ($n=170$), 26 to 35 years ($n=80$), 36 to 45 years ($n=45$), 46 to 55 years ($n=64$), 56 years and above ($n=21$), instead of generations.

Gender

Out of 380 participants, 52.4% were men ($n=199$) and 47.6% were women ($n=181$).

Caste

Out of 380 participants, 43.4% (n=165) belonged to general category, and 43.2% (n=164) and 13.4% (n=51) to OBC and SC/ST categories, respectively.

Religion

Among the participants, 86.6% (n=329) were Hindu, and 13.4 (n= 51) were Muslims.

Education and Employment

Among the participants, 28.4% (n=108) have been educated up to master's level, 30.8% (n=117) have done their bachelor and 40.8 % (n=155) graduation level. The majority of the participants have pursued their higher studies with Hindi as their medium of instruction, i.e., 81.1%, and only 18.9% (n=72) have completed studies with English as their medium of instruction. It is important to note that Hindi is the medium of instruction in government-run schools in Bihar and more prevalent than English or any other language. Employments include students 38.9% (n=148), government jobs 19.74% (n=75), farmers and labourers 9.21% (n=35), Homemakers 14.5% (n=55)

4.5 DETAILS OF THE PARTICIPANTS INTERVIEWED

In chapter 5, data collected from the interviews are analysed and discussed. The table below (4.2) provides details of all the participants interviewed. The identity of the participants is hidden to maintain anonymity and to maintain research ethics. Therefore, all the participants are assigned the letter 'P' followed by the serial number of responses to the survey in chronological order.

Table 4.2 Details of the Interviewee

ID	District	Age	Gender	Caste	Religion
P1	Darbhanga	68	Male	General	Hindu

P2	Madhubani	55	Male	General	Hindu
P3	Madhubani	63	Male	General	Hindu
P4	Madhubani	48	Male	General	Hindu
P5	Darbhanga	62	Male	General	Hindu
P6	Darbhanga	55	Male	General	Muslim
P7	Madhubani	45	Male	General	Hindu
P8	Madhubani	48	Female	General	Hindu
P9	Madhubani	49	Male	SC	Hindu
P10	Madhubani	19	Male	OBC	Hindu
P11	Darbhanga	21	Male	OBC	Hindu
P12	Madhubani	20	Male	OBC	Hindu
P13	Madhubani	35	Male	SC	Hindu
P14	Madhubani	41	Male	SC	Hindu
P15	Madhubani	43	Male	SC	Hindu
P16	Madhubani	36	Male	SC	Hindu
P17	Darbhanga	37	Male	OBC	Hindu
P18	Darbhanga	32	Male	OBC	Hindu
P19	Darbhanga	32	Male	OBC	Hindu
P20	Darbhanga	30	Female	Jolaha	Muslim

P21	Darbhangha	34	Female	Jolaha	Muslim
P23	Darbhangha	37	Female	Jolaha	Muslim
P24	Darbhangha	39	Female	Jolaha	Muslim
P25	Darbhangha	19 *	Female	OBC	Hindu
P26	Darbhangha	55	Female	OBC	Hindu
P27	Darbhangha	44	Female	OBC	Hindu
P28	Darbhangha	44 *	Female	OBC	Hindu
P29	Darbhangha	35	Female	SC	Hindu
P30	Darbhangha	40	Female	SC	Hindu
P31	Darbhangha	37	Female	SC	Hindu
P32	Darbhangha	37	Female	SC	Hindu
P33	Darbhangha	25	Male	General	Hindu
P34	Darbhangha	26	Male	General	Hindu
P35	Darbhangha	31	Male	General	Hindu
P36	Darbhangha	28	Male	General	Muslim
P37	Darbhangha	35-	Male	General	Hindu
P38	Darbhangha	45	Male	General	Muslim
P39	Madhubani	42	Male	General	Hindu
P40	Darbhangha	38	Male	OBC	Hindu

P41	Darbhanga	40	Male	OBC	Hindu
P42	Darbhanga	53	Male	OBC	Hindu
P43	Darbhanga	31	Male	OBC	Hindu
P44	Darbhanga	20	Male	OBC	Hindu
P45	Darbhanga	25	Male	OBC	Hindu
P46	Darbhanga	19	Male	OBC	Hindu
P47	Darbhanga	60	Male	OBC	Hindu
P48	Madhubani	16	Male	OBC	Hindu
P49	Darbhanga	16	Male	OBC	Hindu
P50	Darbhanga	15	Male	OBC	Hindu
P51	Madhubani	16	Male	OBC	Hindu
P52	Madhubani	17	Male	General	Hindu
P53	Darbhanga	55	Female	General	Hindu
P54	Darbhanga	52	Female	General	Hindu
P55	Darbhanga	54	Female	General	Hindu
P56	Darbhanga	55	Male	OBC	Muslim
P57	Darbhanga	56	Male	OBC	Muslim
P58	Darbhanga	48	Male	General	Hindu
P59	Madhubani	57	Male	OBC	Muslim

P60	Madhubani	54	Male	OBC	Muslim
P61	Madhubani	17	Female	SC/ST	Hindu
P62	Madhubani	19	Female	SC/ST	Hindu

Audio Recording Equipment

For the recording of interviews, Sony IC recorder, model no. ICD-PX470 was used, and

4.6 METHODS OF RESULTS INTERPRETATION

The analysis of the results provides a holistic picture of the information elicited through surveys and interviews along with the UNESCO's framework factors. The statistical analysis has been provided mainly for factors 1, 4, and 8, whereas interview excerpts have been provided throughout Chapters 5 and 6 to illustrate all the factors. The nuanced analysis of the interviews in the form of thematic analysis has been used for factor 8, which assess the attitude of Maithili speakers toward their language. The following table (4.3) outline how each factor is analysed:

Table 4.3 Aspects and Methods of Study

Factor	Aspects of study	Methods of study
Intergenerationally	<ul style="list-style-type: none"> • Speakers number by age cohort • Intergenerational picture of Attitudes and Beliefs • Trends in Maithili language proficiency 	Synthesising various sections of the survey and interviews.
Domains and functions	<ul style="list-style-type: none"> • Responses indicating in what domains Maithili exists • Use of Maithili with interlocutors 	Quantitative using the survey Qualitative field observations
Access to Materials	<ul style="list-style-type: none"> • Availability of textbooks in Maithili 	Qualitative using interviews and field observations
Attitude	<ul style="list-style-type: none"> • Speakers attitude towards Maithili • Prestige 	Quantitative using survey Qualitative interviews and field observations

4.6.1 Thematic Analysis

For the analysis of interview verbatims, thematic analysis has been used. Coding was done by identifying keywords, terms, and phrases, which helped create themes. Themes were created

from the data, and the analysis process has been aided by using Nvivo software which helped simplify the process of data analysis.

4.6.2 Creation of Indices from the Dependent Variables

In this section, the calculation of the indices based on the two prime aspects, namely, the speakers' attitude towards the Maithili language and their proficiency in Maithili and Hindi languages, are discussed. The respective indices are,

1. Attitude Index (AI)
2. Proficiency Index (PI)

The statistical analysis is categorised into four main themes, which are as follows:

1. Language Proficiency Index
2. Use of Hindi and Maithili in different domains
3. The attitude index is created by coalescing different items measuring the general attitude of the speakers towards Maithili.
4. Language preferences of the speakers in the region

Maithili proficiency index (MPI) is created by transforming four ordinal variables: speaking, understanding, reading, and writing skills. The MPI is generated by summing the mean scores for all four skills and then dividing that total value by five, as a five-point Likert scale has been used in the questionnaire survey.

Proficiency Index (PI)

$$PI = \text{mean (all skills in the respective language)} \div 5$$

Actual example: Maithili Proficiency Index, MPI = 4.70 (understanding) + 4.37 (speaking) + 3.18 (reading) + 2.49 (writing) divided by 4 (the number of variables) gives the value of mean score as

3.69. The score is further normalised by dividing with the highest rating possible on the scale, i.e., 5, and the result is 0.74. The Hindi Proficiency Index (HPI) was created by combining all four skills in the Hindi language. Similarly, the Attitude Index for Maithili (AIM) was created by combining ten ordinal variables. However, some questions used in the survey have not been incorporated while creating indices; rather, descriptive analysis is provided for them. Those questions are discussed separately in chapter 6. Examples of these variables are language usage in different domains and the language preferences of the speakers (discussed in Chapters 6).

Cronbach's alpha reliability assessment was performed to check the internal consistency of the items, where a higher score signifies that the scale is internally consistent and the items are internally connected (DeVellis, 2016, p.34). Although the recommended value of alpha is 0.80 (Bryman, 2012), alpha values higher than 0.70 are considered to be good, and values higher than 0.60 are considered acceptable (DeVellis, 2016). However, any score of less than 0.60 is considered to be poor or unacceptable. The reliability score for all three indices was good (0.70). The following table below shows the indices, the number of items combined to create each index, and their alpha coefficients values.

Table 4.4 Reliability Score of the Main Indices

Index	No. of variables	α score
MPI	4	0.807
HPI	4	0.870
AIM	10	0.824

4.7 NORMALITY TEST

The normality test was conducted to establish how participants rated their language behaviours and attitudes and to understand whether the data is normally distributed. Depending on the nature of data, i.e., whether it is normally distributed or non-normally distributed, the parametric or non-parametric test is employed (Field, 2013). Non-parametric tests are applied for ordinal and nominal-type data, and parametric tests are employed for the interval or scale-type data. Along with these criteria, it is imperative to ascertain the distribution of data to decide the type of tests to be employed. For instance, when the data is distributed normally, parametric tests are used and vice-versa (ibid.). The normally distributed data returns the bell-shaped curve (Rasinger, 2013). However, in non-normally distributed data, the curve will either be too skewed (either right or left oriented in the graph). A considerable number of outliers are expected, which correspond to either too low or too high ratings as compared to the mean. Three levels of testing used to determine the nature of data distribution (i.e., normal or non-normal) are as follows:

1. When the mean and median are almost equal, the data is normally distributed (Rasinger, 2013). However, in the present study, none of the means and medians is very similar.
2. Visual inspection of the histograms for all variables and the respective indices also indicates the data to be non-normally distributed (Tabachnick and Fidell, 2013)
3. None of the variables has a skewness score close to zero. However, for a large sample, the deviation of skewness and kurtosis values from zero are insignificant, and only the shape of the distribution is essential to determine the distribution of data (Waternaux, 1976, and Tabachnick and Fidell, 2013).

4.8 CORRELATION TEST

When the data is non-normally distributed, non-parametric tests are applied, which involves Spearman rank correlation tests. Spearman rank correlation tests are appropriate to understand the relationship between variables (Field, 2013). Using the same principle, here correlations tests are used to see if the variables are interrelated. If the correlation coefficient value is greater than 0.5, it is considered a strong (positive/negative) correlation between the factors. If the value falls between 0.3 and 0.5, then it is a moderate correlation, and if it is between 0.1 and 0.29, the correlation between variables is weak (Cohen, 1992). The correlation tests in social sciences are applied to illustrate the association between variables rather than causality (Oppenheim, 1992). In the present research, the correlation test has been carried out to see whether variables are

4.9 SIGNIFICANCE TEST

Non-parametric tests have been used seeing nature of the data. Here, non-parametric tests have been applied to check the significance of different variables. Mann-Whitney U test is employed when there are two groups, whereas, when there are more than two groups, Kruskal-Wallis test is applied. These two tests are done to understand whether the differences between groups are significant or not (Rasinger, 2013; Field, 2013). In this case, the mean values for the transformed index score and mean ranks for the individual dependent variables are tested. In order to examine the effectiveness of Maithili's inclusion in the ES, the present study has tried to explore the three major aspects:

- i) Behavioural aspect (self-reported language proficiency)
- ii) Language use in different domains

- iii) Attitudinal aspect (functional and emotional) aspect of Maithili language speakers. The attitude here has been used in a general sense to understand the perception of Maithili speakers toward their language. The study operationalised the term 'effectiveness' in terms of rates of change in attitude and increase in the respondents' language proficiency and literacy proficiency.