

## CHAPTER 3

### RESEARCH METHODOLOGY

This study is conducted on the ground of sociolinguistic analysis with regard to communicative efficiency among netizens in the context of online Hindu matrimony. In order to accomplish this kind of studies or researches, there is a need to have relevant research methods or approaches. As this study is very much related to the structure of language and its use in a particular domain, the methods of Structural Linguistics Analysis (Matthews, 2001) and Sociolinguistics Analysis have been selected to analyse the processed data meant for this study. The structural linguistic approach, a method that involves the classification of every element of a corpus at its different structural level beginning from phonemes, morphemes, lexical items, noun phrases, verb phrases and right up to sentence types (Searle, 2011) is used in this study. This approach identifies and analyses the lexical features and the occurrences of base forms and affixes in the morphological structures, syntactic structures, and also the formation of discourses with regard to different contexts or domains of use. The processed data are subjected to Discourse Analysis, and a detailed analysis of the discourses is presented in Chapter 4 (page 33). In this study, emphasis is given on the cohesion and coherence of the corresponding texts because these are the factors that determine the comprehensibility of the texts by the target readers. Cohesion and coherence of the discourses are analysed using a few selected cohesive devices as discussed in section 3.3.3 (page 28).

This is a data-oriented research, and the analysis is basically of qualitative nature. The processed data is also meant for the sociolinguistic analysis whereby structural

components that are defined through the structural linguistic method are then analysed for their correlation with that of the language use in socio-cultural contexts with regard to Hindu matrimony. In this study, the language used is described both by means of sociolinguistic analysis and discourse analysis to show why the netizens (providers of data for this study) chose to communicate in such a way, and that a particular language structure is related to its use in online Hindu matrimonial contexts.

### **3.1 Experimental Groups**

Random sampling is done on a population universe that comprised netizens who participated in the matrimonial websites. The samples' profile pages were clicked at random to obtain the required digital transcripts. Samples were chosen simultaneously from two active matrimonial websites, namely 'BharatMatrimony' and 'Matrimonials India'. These websites serve as the social network for a group of Hindu netizens seeking for marriage alliance, and also friendship to a certain extent. The webpage addresses of the corresponding websites are as the following:

(a) BharatMatrimony: [www.bharatmatrimony.com](http://www.bharatmatrimony.com)

(b) Matrimonials India: [www.matrimonialsindia.com](http://www.matrimonialsindia.com)

The samples were further distinguished into different experimental groups of concern, that is, by three different age groups and gender.

### **3.1.1 Age**

Digital transcripts that were used as the data in this study were extracted from the profile pages of samples who come under three age groups:

- (a) Group 1: 20 – 29 years
- (b) Group 2: 30 – 39 years
- (c) Group 3: 40 – 49 years

The age ranges were selected as the above owing to the fact that most candidates are seeking for marriage alliance during that age, hence are very much active in such social networks.

### **3.1.2 Gender**

Both male and female candidates' (samples) data is studied simultaneously in this study.

## **3.2 Data for the Study**

### **3.2.1 Data Collection**

The data, being written text in the form of digital transcripts, is extracted from the profile pages of 120 female and 120 male samples, and is chosen at random.

The manipulated variables in this study were age and gender, whereas variables such as religion (being Hindu) and education (being tertiary level) were made constant. Data on

other variables such as mother tongue, caste, profession, residential country and marital status (being single, divorcee, widow/widower, or even married) is not taken into account in this study.

### **3.2.2 Data Processing**

The data collected is subjected to an array of processes to enable effective analysis. Firstly, to secure confidentiality and to avoid direct indication of specificity, proper nouns including actual names of persons, including the candidates', places (places of living), castes, sub-castes and other such specific indicators that were totally not used in the discourse analysis were substituted with common nouns that denote the types. Example of the unprocessed data in comparison with the processed data is shown below:

The unprocessed data extracted directly from the profile page of the concerning netizen:

I am a technology project lead based out of New York. I grew up in Srirangam, Tamilnadu. I did my engineering in BITS Pilani, Rajasthan, and since then work has taken me to a few places around the World. Currently, I work with an investment bank in their front-office trading technology team. I consider myself liberal and easy going with a good appreciation of my roots and traditions. I like to search for new stimulating ideas - in gadgets, books, movies or shows. Besides, I have a serious hobby in photography. I do not have parameters to judge compatibility, I am looking for someone with the right fit and chemistry, though an independent person with a liberal outlook and a good sense of humor will be a great start.

Preferably someone that lives in the U.S

(Sample from a male in the age group 30 – 39)

The same data, after substituting the proper nouns with common nouns in brackets:

I am a technology project lead based out of [*country*]. I grew up in [*state*]. I did my engineering in [*college*], and since then work has taken me to a few places around the World. Currently, I work with an investment bank in their front-office trading technology team. I consider myself liberal and easy going with a good appreciation of my roots and traditions. I like to search for new stimulating ideas - in gadgets, books, movies or shows. Besides, I have a serious hobby in photography. I do not have parameters to judge compatibility, I am looking for someone with the right fit and chemistry, though an independent person with a liberal outlook and a good sense of humor will be a great start.

Preferably someone that lives in the [*country*]

Since the inclusion of proper nouns in the discourses is not significant in this study, by doing this substitution, the processed data now looks more biasfree and comparisons from one discourse to the other is done without much attention to the proper nouns. Samples of the processed data are presented in Appendix A.

The lexical items that calls for particular attention in this study, namely the adjectives and culture-bound words (or phrases), were observed for the frequency of their occurrence in the discourse in the corresponding experimental groups. The figures obtained here were used to determine the significant differences (if any) in the language use and style of the netizens between genders, and also among the various age groups tested. This included the identification of culture-bound words that significantly portrayed the degree of culture needs of the candidates and their (social) community. The data obtained from this observation is presented in Appendices B and C.

### 3.2.3 Sample Size

A subtotal of 40 samples of discourse representing the population universe of the selected age group were obtained for both genders, and this made up a grand total of 240 discourses altogether. Table 3.1 shows the segregation of the female and male samples according to their respective age groups and gender. Each sample provided one discourse for the analysis.

**Table 3.1: Segregation of female and male samples**

<b>Age groups</b>	<b>Group 1</b>	<b>Group 2</b>	<b>Group 3</b>
<b>Female</b>	40	40	40
<b>Male</b>	40	40	40
<b>Total</b>	80	80	80 = <b>240</b>

### 3.3 Application of the Research Methods Chosen (in Data Analysis)

The genre of the discourse or text is determined by mere reading of the text. A discourse analysis is done to determine the cohesion and coherence of the discourse. Whether there exists an agreement between the thought of the netizens (samples) with the language structure, and that with the language use is analysed. This sociolinguistic study is conducted with regard to the two factors mentioned, namely the language structure and language use.

### **3.3.1 Language Structure**

As for the language structure, four main aspects that were analysed are given below:

- (a) Lexical items
  - First person pronouns
  - Adjectives
  - Culture-bound words
- (b) Morphological structures
  - Noun structures
  - Verb structures
  - Adjectival forms
  - Adverbial forms
  - Particles
- (c) Phrases
  - Noun Phrases (NP)
  - Verbal Phrases (VP) (Adjectival and Adverbial phrases or forms)
- (d) Sentences
  - Complete sentences versus incomplete sentences
  - Simple, compound and complex sentence patterns

### **3.3.2 Language Use**

Aspects on how the language structure is related to (or correlated with) the language use is analysed in this part. Here, the correlation of the language structure and language use is seen through to identify the pattern of the discourse. This included analysis on how the structure is used to perform the language use, and how the text is interpreted by the

reader with reference to contextual, socio-cultural (alliance-seeking function) – the sociolinguistics division.

Besides, the existence of differences (if any) between genders and age groups in terms of the text presentation style is of concern too in this study.

### **3.3.3 Aspect of Cohesion**

#### **3.3.3.1 Lexical Cohesion**

Analysis is also done to determine lexically cohesive group of words. Words that represent the same topic, and that are semantically related were extracted from the discourse using the following cohesive devices (Salkie, 1995):

(a) **Repeating key words (content words)**

- Important keywords or phrases were pointed out via summarisation of the discourse or checking if the discourse or texts still appears coherent without those words
- Key words were also determined via observing patterns of word delivery

(b) **Using synonyms**

- This denotes different word classes with a related meaning
- A particular word is chosen, then the discourse is checked for the existence of one or more synonyms
- Owing to the fact that words closest to their overall meaning are considered synonymous (in any text), the register differences were taken into account too. In this study, the observation is done with



regard to formal register in comparison with that of the colloquial register.

(c) **Superordinates and Generals**

- The occurrence of common nouns that follow the proper nouns in the discourse (hyponymy) at word level and phrase level is observed
- General words with expression were compared and contrasted against those without expression

(d) **Opposites and Related Words**

The existence of any one or more of the following four kinds of opposites is found as classified below.

- Binary opposites (absolute opposites) – opposites that are absolutely incompatible with each other, such as *male* and *female*
- Non-binary opposites (not absolute) – opposites that are different items altogether such as *shovel* and *axe* (this type is usually used in a text to contrast one subject or object from the other)
- Binary opposites (not absolute) – also known as antonyms in which there is a lack of a clear dividing line; for example *simple* and *complex*
- Converseness – applies to the words used in expressing the same subject differently, for example *parents* and *children*

### 3.3.3.2 Other Kinds of Cohesion

Besides the lexical cohesion, other kinds of cohesion were also tested via the following cohesive devices (Salkie, 1995):

(a) **Substitutes**

- Whether a substitute is used to create a strong link between one part of a text with an earlier part is observed
- Sentences or phrases occurring with words such as *one*, *do* and *so* contributed to cohesion
- The noun substitutes, verb substitutes and clause substitutes were compared and contrasted against each other

(b) **Ellipsis**

- The practise of ellipsis, that is, leaving out a word, or phrase and morphological form is also of concern in this study

(c) **Reference words**

- The words that play an important role in creating cohesion in a discourse, which are also the words that determine the full meaning of the discourse are identified
- Reference words are derived upon or extracted based on text reference and situation reference
- Common reference words to be identified include personal pronouns (including the object forms & possessive forms)
  - *I*, *you* and *we* – (1<sup>st</sup> and 2<sup>nd</sup> person) – situation reference
  - *He*, *she*, *they* and *it* – (3<sup>rd</sup> person) – text reference and situation reference
- Demonstratives such as *this*, *that*, *there* and *those*

- The comparative constructions are emphasised, namely general comparisons versus specific comparisons

(d) **Connectives**

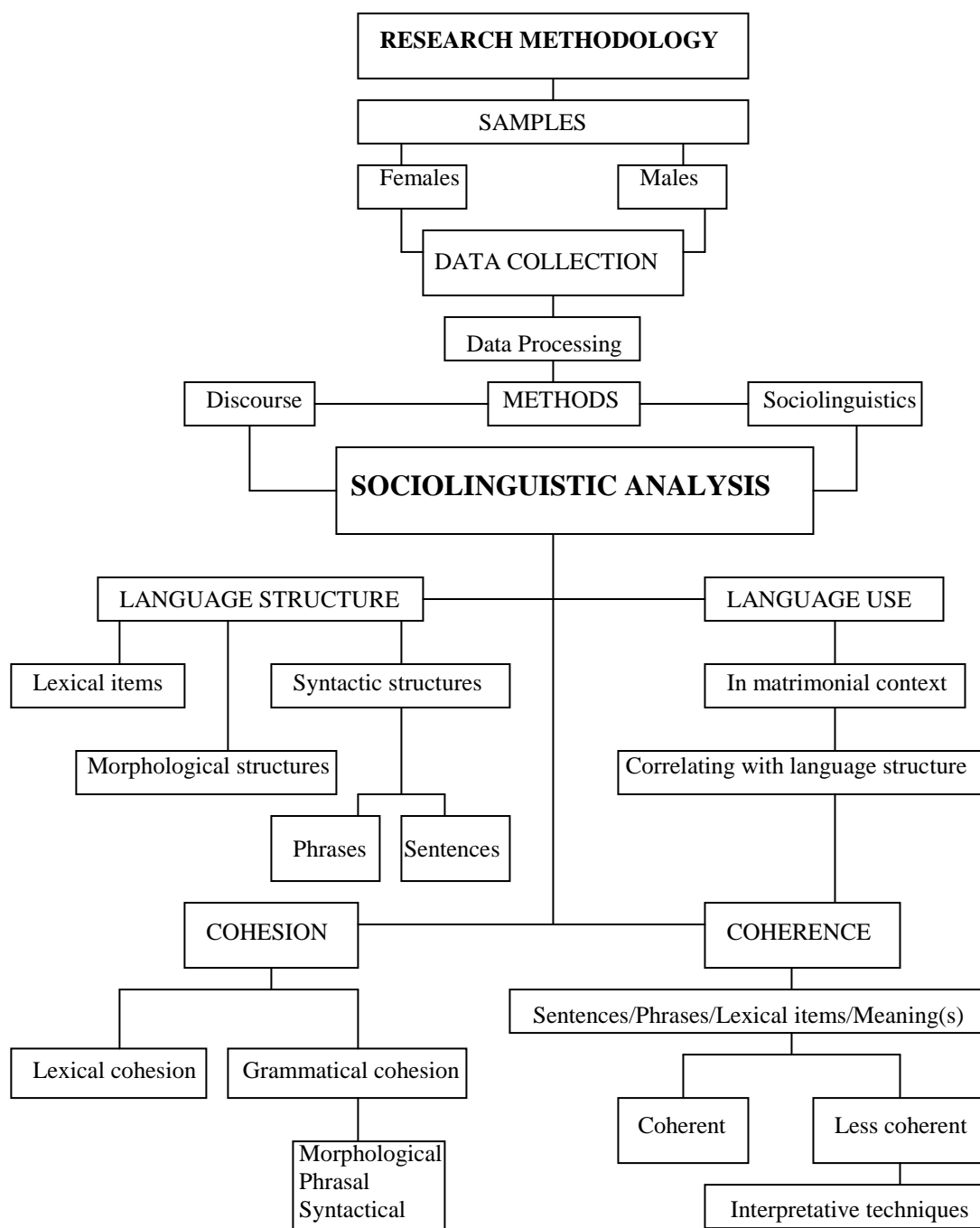
The following types of connectives are also looked for in the discourse as they play a vital role in the cohesiveness of a discourse:

- Addition such as: *and, also, in other words, in particular and for example*
- Opposition such: *as but, yet, in spite of, nevertheless and though*
- Cause such as: *therefore, because, as a result, hence and for the same reason*
- Time such as: *then, first, finally and just then*

### 3.3.4 Aspect of Coherence

Sentences (and phrases) that give rise to the coherency of a discourse are distinguished from that which did not. In other words, the presence of coherent sentences (and phrases) and also less coherent sentences (if any) will be identified and analysed with regard to the context of the discourse. Moreover, lexical items that made the discourse coherent will be discussed too. It is to be mentioned here, that starting from the key vocabulary almost all the grammatical structures and features have well-defined role to play in making the discourse (a narration, description, lecture, etc.) more cohesive and coherent from the points of view of ‘linguistic competence’ and ‘linguistic performance’.

Figure 3.1 shows the theoretical framework that indicates all those factors involved in the sociolinguistic analysis done in this study. The results (including observations and findings) are derived based on the factors presented below.



**Figure 3.1: Factors involved in the discourse analysis done**