CHAPTER II

RESEARCH DESIGN

The description of the status of language in terms of political, social and historical factors is a research tradition in sociolinguistics, which can be dated as early as 1950 (Weinreich: 1953). In sociolinguistic research, language data has been obtained either through, (a) Language Census or (b) Sociolinguistic Survey.

Data on language is obtained most frequently through the population censuses, wherein information is sought about mothertongues and languages known by the respondents. But the dynamics of language use and language choice have never been justifiably dealt with in language census. The language census due to absence of data on "language use patterns", at best provides a broad geographic distribution of reported mothertongue and languages known, a more description without commenting on the social processes involved in the use of language in different spheres of activity.
As against language census surveys, the sociolinguistic surveys of multilingual populations are not only designed to obtain a broad profile of the social distribution of language but also to elicit self reported data on patterns of language use and attitudes of individuals. They also attempt to observe or map how interactions are perceived by those involved, how interactions manifest attitudes of participants and how their attitudes towards other languages, their own language and other speech communities in general influence their interactions. The belief and behavior mechanisms may perpetuate the social reality in the particular context.

The basic distinction between a language census and a sociolinguistic survey is that while census aims at a complete coverage of a population and are limited in number of questions that can be asked, a survey involves more detailed questioning of representative samples within a given population and are designed to collect linguistic and language behaviour data.

The results obtained are intended to form the basis for generalization.
Systematically undertaken sociolinguistic surveys can serve as a guide to plan the changes in society. The surveys bring out clearly that how strikingly complex is the linguistic and social knowledge of the members of a given society. With such understanding, it becomes impossible to maintain a position of linguistic prejudice and intolerance. Planning based on widespread ignorance of socio-cultural and linguistic diversity could be avoided if the systematic information is available through sociolinguistic surveys.

Further, sociolinguistic surveys can serve as an important pre-requisite to more indepth studies in the areas of sociolinguistics like, language in contact, language convergence, language maintenance and shift, language identity, language variations, and language teaching in general and education of linguistic minorities in particular.

When the present research was designed, it was decided to have a broad scope as initially broad scope is more likely to produce the comprehensive description. It is also more likely to generate a wide variety of hypotheses which could be refined and/or verified later in the study or in subsequent research.
However, during the discussions, it was felt that there is a need to stay focussed on a specific setting and seek answers for precise questions. The focus had to be decided by which sociolinguistic perspective had to be adopted: whether the focus of study is sociology of language or it is Labovian sociolinguistics. The sociology of language as expounded by Fishman (1974), deals with the social or the societal aspects of language behaviour. It examines into the co-variation of diversity and of patterns in both the language and society. While sociolinguistics as advocated by Labovian is concerned with how the structure of language functions in society as a means of communication or how certain linguistics features serve characterise particular social arrangements. It was decided to undertake study of multilingual speech community covering language demography, degree of multilingualism, patterns of language use and attitudes of different language speakers in the first phase and the study of structure of language in terms of social context would be taken in the second phase. The present research deals with the first phase of the study. Therefore, the focus of the present study falls in the realm of sociology of language.
2.1 Methodological Approach

Large-scale sociolinguistic surveys based on questionnaires such as postal, telephone or computer mail surveys do not permit the use of sensitive instruments and refined quantifications, such as precise and detailed questions or actual language behaviour in a great variety of situations.

Further, large scale sociolinguistic surveys where emphasis is on group usage, individual responses get submerged in statistical frequencies. The surveys to study sociolinguistic environments may reveal gross differences but fail to indicate the specific social processes that take place therein. Tabulating frequencies of specific variables can overlook how context relates to activities.

On the other hand, sociolinguistic surveys undertaken with ethno methodological approach make use of participant observation and emphasise personal interaction in data collection which is evident from anecdotal observations. It allows use of more sophisticated measures and analysis but because of the complexity of their use, they are limited to small samples. Such sociolinguistic surveys have their limitations, as
it is often impossible to gain data that confirm to strict rules of empirical social research, in terms of representativeness and generalisability.

In the present study, both approaches, statistical and ethno-methodological are used to collect data which complement each other, the ethno-methodological approach uses both interactive and non interactive methods for collection of data.

Among the interactive methods, participant observation is used primarily for data collection. Interviews are also used for the purpose of triangulation of data. Further, interviews with prominent personalities of Medan gave the oral history of Medan town apart from insight into Medan’s multilingual ethos. Interaction with informants allowed multiple variables to interact simultaneously, and brought out the issues which people value rather than being determined statistically. Informal conversations with subjects helped to gain more of an emic perspective. This approach being an open-ended approach, it allowed analysis to begin immediately, further, whenever, data failed to provide sufficient detail, immediate changes in data collection strategies could be made.
Non-participant observation is used on occasions, in which the resource person took on a detached uninvolved role. These observations were of great importance at the time of verification and refinement of findings.

The published and unpublished documents about Medan were examined as a data source as well as for adding precision to other data in the present survey, which used both questionnaire and participant observation as tools for data collection, the data on group usage in terms of statistical frequencies is used and at the same time care has been taken that valuable individual responses are preserved. This amalgamation of both methodological approaches gives a better insight to the multilingual situation of Medan. In short, the sociolinguistic scene for the present study was appraised on the basis of three types of data:

1. Main data obtained from the specially designed questionnaires.

2. Auxiliary data obtained from the interviews of some residents of Medan who are knowledgeable about the socio-cultural, historical developments of Medan; also from the participant observations.
3. Secondary sources such as census, gazetteers and other published and unpublished material available on Medan.

2.2 Tools for Data Collection

The interactive or non-interactive methods, qualitative method or quantitative method, whichever method is chosen for data collection, the tools for data collection are equally important in any survey as it ensures appropriateness of information in relation to the objectives of the study. There are several tools, which can be designed depending upon the target population and its composition. It is possible that more than one tool will be chosen for the same survey, which would compliment each other to arrive at a holistic picture. Therefore, in any research design preparation of tools for data collection demands careful understanding of the nature of the population and the objectives of the study.

The emphasis of the present survey was to seek social distribution of responses, so, questionnaire was felt as a proper tool. The population under study consist of both non-literate and educated as well as from varying occupations,
hence self administered questionnaire is expected to bring out the required information.

The format of the questionnaire has four parts viz.,

1. Personal information
2. Multilingualism/language competence
3. Language use in different domains
4. Attitudes

2.2.1 Personal Information

Twelve questions pertaining to personal information are structured at the beginning of the questionnaire. Each question represents an independent variable in terms of which the responses are analyzed later. The variables are:

1. Sex
2. Religion
3. Marital Status
4. Residence
5. Age
6. Place of birth
7. Ethnic
8. Father’s ethnic
9. Mother's ethnic
10. Education
11. Occupation
12. Period of stay in Medan

2.2.2 Multilingualism/language competence

Information on language competence in mother tongue and other languages known by the informant is obtained in this section. This information is obtained in terms of four skills namely, understanding, speaking, reading, and writing. In addition, information on the mode of acquisition of these languages is also collected. This information was sought to test the hypothesis that ‘multilingualism setting in Medan is more of a natural state of behavior than a learned languages behavior.’

2.2.3 Language use in different domains

In order to study the languages use patterns, different domains were chosen covering both formal and informal contexts. The domains for study are as follows.
1. Home
2. Neighborhood
3. Market
4. Work place
5. Administration
6. Religion

2.2.3.1 Home

Information pertaining to use of languages at home, choice of languages for listening, speaking over telephone, reading newspapers/magazine, writing letters and viewing television of self and family members are collected in home domain. In this domain it was expected that as mother tongues are more conservatively preserved at home the use of mother tongues would be reported maximum as compared to other domains.

2.2.3.2 Neighborhood

The hypothesis is that the neighborhoods in Medan are multilingual and acquisition of neighbor's languages with noticeable linguistic towards the other group is norm of the people of Medan. In order to test these findings, information on languages of neighborhood and what language/s are used to speak with neighbor's was sought in the questionnaire.
Information is also elicited about the language/s used among the friends belonging to different linguistic groups.

2.2.3.3 Market

In the domain of market, which is domain of more transactions and interactions and the place where almost every individual will participate the information is elicited regarding the use of language/s in market. It was expected that Medan being multilingual, the domain of market would show more use of different than a single lingua franca.

2.2.3.4 Work place

Work place of an individual would depend on the occupation she/he is engaged in. The occupation covered in present survey being varied. The population was expected to have a wide choice of language at work place. Because of geographical location of Medan, most of the work places have persons speaking Java, and Batak. It was thus expected that along with Java and Batak, other languages might be used at work place. Information is also obtained on language used interaction with superiors, subordinates equals and outsiders,
this information was expected to bring out the power and prestige accorded to different languages and the functionality of the languages involved.

2.2.3.5 Administration

Domain of administration is unique in case of Medan because of its multilingual composition to administrative staff of the Government. Even Bahasa Indonesia has been constituted, but the use of ethnic languages among administrative staffs of the Government flourished very much. It was therefore, decided to elicit the information about language use in administration.

2.2.3.6 Religion

The composition of Medan in terms of religion is Muslim, Christians, Hindu and Budhis. The information about language use in the domain of Religion was elicited on two levels.

1. Personal prayers
2. By priests for performing religious activities in mosque and church services
This information was elicited to find out whether different languages are used in this domain by different religious groups or the same language is used across religions. It was also important to find out whether same religious group uses different languages in the domain of religion. The choice of language for reading religious books was also sought in order to find out whether the language differs among personal prayers religious formal/informal discourses and the religious texts. It was observed that many a time religious communities under study use different languages than that of the mother tongue of the communities. For example, Christians with Batak Toba language as mother tongue often reported to use English in their religious texts when it comes to reading as the education level goes up, the preferred language for formal discourse thus is English. In case of Muslim, irrespective of mother tongue, language of religious text is Arabic. It was thus, decided to find out whether these observations are substantiated by the data (both by elicitation and observation) or the language choice in reality is different in these three situations namely informal/formal discourse, personal prayers and religious text.
2.3 Attitudes

Most studies on attitudes about language and linguistic communities and their effect on the use of language or choice of a particular language in multilingual situations have shown that attitudes play a very important role in the choice of the language. Most communities show reference for the magical powers of a particular language or adverse value judgments about certain language. The value judgments that underlie these attitudes may be ethical or aesthetic, or they may rest on a pragmatic appeal to efficiency. The degree of social sanction they carry according to the language community, whether the judgments and attitudes are social or individual, the individual expounding them frequently claims objectivity for his opinions. The far reaching value judgments passed on whole language that some languages are intrinsically better than others, cannot be dismissed as ignorance or prejudice, as it is a view held by people who are both intelligent and serious and can bring forward evidence to support it. It also reflects in their choice of that particular language in certain contexts in a multilingual situation and hence has to be taken note of.
The attitudes of the informants are elicited in two areas namely, symbolic/emotional use of language and functional use of language. For the symbolic use it was expected that most of the language communities in Medan would rate their own mother tongue high for uniting both family and speech community, for selecting friends and spouses. It was also expected that for cultural expression mother tongue would be the preference language. As far as functional preferences of language at different levels of education, as for securing job or business and as an official language the preferred would either be Bahasa Indonesia or English. The information was sought on the hypothesis that higher the education more the preference for Bahasa Indonesia and though people prefer mother tongue for integrative purpose, for instrumental purpose Bahasa Indonesia is preferred.

Medan being a language contact situation where many languages are spoken each language would have influence of other language. It was expected that respondents would have the perception of this influence and being exposed to the same language spoken elsewhere in nearby towns they may be
able to bring out the difference in their variety and the standard variety.

Another question asked in this section is, ‘which language is used more often in daily oral communication in Medan?’ Through this question it was expected to find out which is the lingua franca of multilingual Medan.

2.4 The Setting

Next important aspect of research design is the selection of a site for a survey. Selection of a place or a group to undertake survey may be done on the basis of initial definition of the group, theoretical direction, curiosity, or other influences such as convenience. For the present research, Medan was selected as the area for investigation because:

1. It is a capital city of North Sumatra province

Being a capital city, Medan attracts native speakers of different ethnic languages besides Melayu to come looking for a better life.
2. **Due to historical reasons majority of language communities residing in Medan i.e. Melayu, Batak, Minang, Java and Chinese, are natives of Medan and not the migrants.**

The different language speakers residing in a place over generations would provide a different perspective in terms of patterns of language use, language choices and attitudes than a situation where migrants or immigrant linguistic groups come in contact with a dominant language group.

3. **It is a place where many language communities exist, providing a multilingual contact situation where any sociolinguistic hypothesis could be examined.**

4. **It is an area which is developing at a phenomenal rate of progress and area with high political awareness.**

There is a definite relationship between political/economic development and linguistic pluralism. At a low level of development, masses are normally inert and pluralism is not a threat. At a transitional stage where masses are ‘mobilized’, the divisive force exerted by multiple language groups is greatest. At a high level of political/economic development movements of masses across language areas are more, mass
media minimize the remoteness and above all, upwardly mobile individuals have opportunity to learn one or more languages. Medan falls under third category where linguistic pluralism is an asset and with proper language policies Medan could be developed much faster.

2.5 Selecting Respondents:

The basic purpose of using sampling techniques to select respondents is to ensure the representativeness of the population under study. The other important aspect of sampling is also fixing the size of a sample in relation to the nature and magnitude of the population. This also involves the cost, time and manpower available to undertake the study. The sample size thus obtained is further allocated to sub-classes or sub-categories of the variables selected for the study. For example, the variable sex has two-sub categories while the variable like occupation or education may have large number of subcategories depending on the purpose of study. The allocation of sample size to those sub categories need not be strictly proportionate to the population of these subcategories but certain degree of age weight has to be given to these sub
categories depending on the goals and objectives of the study.

2.6 Analytical Procedures and Processing of Data

The third aspect of any research design is the analytical procedures. The degree of sophistication in the analytical procedures depends mainly on the type of data collected and the quantum of information gathered for the study. The techniques of data processing and analysis are determined by the formulation of the different hypothesis which in turn are based on the objectives of the study.

The quantitative data collected through questionnaires and qualitative data collected through participant observations and through secondary sources were processed through different techniques.

Formal analysis and reanalysis of qualitative data and writing report was done simultaneously while conducting fieldwork. A description of the group or language behavior could be revised to take into account the accuracy of responses and the increasing knowledge about Medan situation.
In the final stage of analysis of the qualitative data, however, all notes and papers are classified to synthesize ideas which lead to useful insights. Scanning the data from field notes and supplementary sources was scanned and arranged which resulted in the emergence of patterns and categories were formed. This process involved comparing, contrasting, the development of linkages and relationship in the categories and constructs, and making speculative inferences which are tested as the research continues. Two principal methods were: constant comparison and analytic induction. The first facilitated the coding of tentative categories at the beginning of the study, while the latter helped produce recursive rules. The notes from the observation were examined and more investigations were set for more observations to provide verification and detail. Further, the qualitative data was used to support the finding arrived at from the analysis of quantitative data or explain the apparent contradiction/disparities found through statistical analysis.

As far as analysis of quantitative data prior to analysis, it is important to examine the data collected through questionnaires for its correctness in all respect as the
inferences drawn are based on the obtained data. Therefore, each filled-in questionnaire was examined before accepting for tabulation. The total number of questionnaires examined are 870 covering seven mother tongue groups and no questionnaire was rejected. The number of questionnaires collected were also checked for its conformity to sample allocation in terms of ward, age, sex, education and mother tongue. The tabulation was checked for its correctness by cross checking the totals with the sub-totals obtained over different sub-categories of the variables.

The purpose of the research is to study the use of different languages in various societal domains like home, neighborhood, market, work-place, education, administration and religion. These patterns are studied in terms of social variables like sex, age, education and mother tongue. Since the study aims at obtaining broad patterns of language use in different domains in terms of different variables, no specific hypotheses were formulated in statistical terms. The analysis consists of obtaining bi-variate and multivariate frequency tables. The data is analyzed in terms of such variables which are expected to yield significant correlations. Data obtained
through bi-variate and multivariate tables were further computed to get exact numbers so that comparisons across mother tongue, sex, age, and education in each domain could be made. In addition to these basic bi/multivariate tables many more correlational tables were computed to explain certain phenomenon like language use across generations and degree of multilingualism.

The data obtained from survey was analyzed either on nominal or ordinal scale of measurement. Therefore, very limited statistical test could be applied for finding out significant in statistical terms. The information on migration was obtained to study whether there is any difference in language use patterns between the migrants and native.

The information about the language choice for reading purpose was elicited for both respondent and his/her family members and different type of reading materials such as newspaper, magazines and fiction. These distinctions in terms of generations and reading materials were made to capture the differences in language use across generations due to exposure and education. The examination of filled - in questionnaires brought out the languages all the respondents
use to read any material for both the self and family members. Hence, the different categories in terms of reading materials and family members were not further analyzed.

The following is the list of bi-/multivariate table formats, designed for the analysis of the data obtained through questionnaires:

- Distribution of respondents in terms of age, sex, mother tongue and religion
- Distribution of respondents in terms of districts
- Languages known in terms of age and sex
- Language use at home in terms of mother tongue
- Language used for writing, reading and speaking.
- Language of neighborhood in terms of mother tongue
- Language choice for speaking with neighborhood and friends having different mother tongue in terms of mother tongue
- Language use in market in terms of sex and mother tongue
- Language use in workplace in terms of mother tongue
- Language use in administration by self in terms of mother tongue
- Language use in administration by officials in terms of mother tongue
- Language use in religious practices in terms of mother tongue and religion
- Attitudes in terms of sex, age and ethnic group

From these basic tables, several tables were computed and data was interpreted.

2.7 Presentation of Report

Finally, the important aspect of any survey is the presentation of the results of findings in proper and comprehensive manner. The presentation should take into consideration the objectives of the study, the co-relational aspect in terms of variables, the level of abstraction, and the target users thus, the effective presentation of results decides the impact of the study and its goals to be achieved.