CHAPTER 4

METHODOLOGY OF DATA COLLECTION

4.1 INTRODUCTION

This chapter discusses the methodology of data collection, specifically pilot testing, formulation of questionnaires and interviews, general data collection, recording, transcription, analysis and generalization of the data and data presentation.

4.2 THE PILOT TEST

The pilot test was conducted between February and May 2006 on 12 selected informants. The purpose was to ensure the effectiveness and validity of the questionnaire, the word-list, the passage-reading and all the interview questions as part of the research instrument. All the answers and feedback were not considered as primary part of the research. The idea was to see if the questionnaires needed be modified in order to be read and understood easily by the selected informants, ensuring that the informants would be able to give answers in a manner that the researcher could utilise.

4.3 THE QUESTIONNAIRE

Based on the pilot test, most of the questions had to be reformulated in order to enhance their effectiveness and validity. Many irrelevant and unproductive questions were amended. Some new questions were added, while others were altered. The order of the questions was also rearranged to achieve the desired result.

The research instrument comprises five parts: Part I-Particulars, Part II-Interview, Part III-Word-list reading, Part IV-Passage-reading and lastly, Part V-Conversation. The ideal method would be for all the informants to participate in all the five parts of the entire questionnaire. However, only 97 informants were able to participate in all parts of the entire questionnaire (Part I to Part V). Of the 97 sets of data collected, only 78 of the entire questionnaires of the informants' were used. In the other 13 questionnaires, only Part I-Particulars, Part III-Word-list Reading and Part IV-Passage-reading could be utilised for the study, and encoded with an (R) for readings. This is because these informants either spoke too fast or too slow, paused too much or too long, or code-switched too often. For the other remaining six (6) questionnaires, only Part I-Particulars, Part II- Interview and Part V-Conversations could be utilised and encoded with a (C) for conversation. This is because some parts of the reading material were not read out by the informants.

The remaining 23 informants only participated partially in the questionnaire. They fell into the older age group of 56-65 year olds (3.4.1.2) and the ethnic group of Other Non Bumiputera (ONB) (3.4.1.3). These informants were only required to complete Part I-Particular, Part II- Interview and Part V-Conversation and they were also encoded as (C). It was not compulsory for them to read the Part III-Word-list Reading and Part IV-Passage-Reading. This is because most of the older informants were illiterate or unable to read Malay. Reading of the word-list was difficult and reading the passage was quite impossible for them. Therefore, informal and formal questions were asked regarding the family, health or even the 'good old days' of these groups of informants. The same rules applied to the ONB, as most of them were immigrants and reading a foreign language was also a difficult task for them. This could be associated with Gardette (1968) and Francis (1983) who highlighted in some

studies that several informants were interviewed for different parts of a questionnaire, on the grounds that not every speaker is equally competent in all parts of the subject matter covered in the questionnaire. The distribution of questionnaires and the number of questionnaires accounted for the in study is shown in Table 4.1.

Table 4.1: Distribution of the Questionnaires and
Number of the Questionnaires Utilised by informants of the study

Questionnaires Distributed	**No. of Informants**	**Questionnaires Utilised**	**No. of Informants**
Part I-V	97	Part I-V	78
		Part I, III and IV (R)	13
		Part I, II and V (C)	6
Part I, II and V	23	Part I, II and V (C)	23
Total	**120**	**Total**	**120**

In summary, Part I-Particulars was completed by all 120 informants. Part II-Interview and Part V-Conversation were completed by 107 informants, while Part III-Wordlist Reading and Part IV- Passage Reading were completed by 91 informants as shown in Table 4.2.

Table 4.2: Number of Each Part of the Questionnaires
Utilised in the study

Parts of the Questionnaire	Part I	Part II	Part III	Part IV	Part V
Questionnaire Utilised	120	107	91	91	107

4.3.1 PART I: PARTICULARS

First of all, the researcher explained in brief the purpose of the study to individual informants, before asking them to fill in Part I of the questionnaire (see Appendix Bi). Part I consists of two sections, namely Sections A and B. Section A contains 12 questions, which were set out to obtain the information on the social background of the informants (3.3.2.1-3.3.2.2). Section B included a table, which was set in order to inquire about the informant's language choice and use (3.3.2.3). The questionnaire was

set in two languages, namely English and Malay, to minimize misunderstanding by the informants.

In section A, the informants were requested to give their particulars either orally or by filling in a form, such as their hometown, resident address and residential duration, ethnic membership, gender, age, highest level of education, occupation, income and type of housing. Information such as name, contact numbers, venue and time of interviews were recorded for future reference. Nevertheless, the names of these informants were not revealed in the research report or findings. Informants were referred to by using a code or reference number and not by name.

After filling in the particulars, the researcher made an analysis and asked the informant some extra questions. This enabled the researcher to get a clearer picture of the social background of the informant and to clarify any uncertainties regarding answers to some items in the questionnaire that could be confusing and needed clarification such as Ethnic Membership (3.4.1.3), Occupation (3.4.1.4.1.1), Income (3.4.1.4.1.2) and Type of Housing (3.4.1.4.1.4).

In section B, the informants were asked questions like “What language do they speak?”, “Who do they speak it to?” and “Where do they speak it?”. From the answers given, information such as the mother tongue, first language and home language of each informant was gathered. Besides that, the researcher also had a clearer picture of the functions of the languages used, specifically the SMD. By analysing sections A and B, the researcher could obtain a clearer picture of the social factors that have contributed to social variations and language choices and use.

4.3.2 PART II: INTERVIEW

Part II was formulated in order to obtain formal speech (FS) of the informant (3.4.2.3). In this part, the informants were interviewed by the researcher. Formal questions were asked regarding the city of Kota Kinabalu. For example, their whereabouts when Kota Kinabalu achieved its status as a metropolitan city; their opinions on the city council and its activities and suggestions to improve the city of Kota Kinabalu. Questions regarding their houses and workplace or schools were also asked (see Appendix Bii).

4.3.3 PART III: WORD-LIST

Part III was designed in order to obtain the Word-list Style (WLS) of the informant (see 3.4.2.1). In this part, the informants were asked to read aloud a word-list of 150 words in their most natural way. A compilation of 150 word-list from the Swadesh word-list (Nothofer 1991) and the word-list for use in Borneo (Prentice 1969) was used. The words list comprised of basic nouns such as body parts, numbers, nature, basic verbs and adjectives which are familiar to the people, native to the place and originating from proto-Malay, not loan words. Some additional words were used for a clearer and more precise phonological variations (see Appendix Biii).

4.3.4 PART IV: PASSAGE-READING

The researcher formulated Part IV in order to obtain the Reading Passage Style (RPS) of the informants (3.4.2.2). In this part, the informants were asked to read two passages taken from a local newspaper on a current issue. The first passage was an easily read text with 299 day-to-day words about a Tsunami alert in Sabah, adapted from *Harian Ekspress* on Monday, February 7, 2005, page 1, column 1. The second passage was also an easy-reading text with 283 day-to-day words on a Storm in Kota Kinabalu. This article was adapted from *Harian Ekspres* on Saturday, July 1, 2006, page 1,

column 4 (see Appendix Biv). Here, the researcher corrected some of the typographic mistakes in the articles, added some words such as the conjunctions, *ekoran* and *serta*, and changed some of the words such as *tidak mahu* to *enggan* in order to get more linguistic variations.

4.3.5 PART V: CONVERSATION

In order to acquire the casual speech (CS) of the informants, Part V was designed (3.4.2.4). In this final part of the data collection, the informants were interviewed once again on topics used in Part IV. This was more a casual conversation rather than a formal interview. For this study, the informants were asked about their feelings and experiences during the recent Tsunami alert in Sabah. In addition, questions on their near-death experiences were also asked. Informants were also asked about the hottest topics of the day such as *Akademi Fantasia* and other popular television programmes (see Appendix Bv).

4.4 DATA COLLECTION

Once the questionnaire for this study was prepared and tested (as in 4.3), and the number and types of informants had been finalised (as in 3.3), the actual data collection began. Data collection was the most important and the most difficult phase of the research. Hence, it was done with careful planning. For this study, the researcher decided to use the direct method[3] of data collection. This method involves the use of questionnaires by fieldworkers in the actual field.

[3] As opposed to the indirect method which usually involves delivering the questionnaire by post either directly to a previously identified respondent, or, more often, to some intermediary, who in turn finds an appropriate respondent who fills it in and returns it (Francis1983:79).

4.4.1 FIELDWORK

This study relies greatly on information gained first-hand from informants, so fieldwork was an essential part of the study. The fieldwork of this research took about six months to reach completion, beginning from June until December 2006, with the assistance of two research assistants were specifically employed as fieldworkers. This was not only because each interview took about 20 to 40 minutes but time was needed between interviews for reviewing records, sorting out notes made rapidly in the course of the interview, listening to the tape-recording to make sure that it was technically adequate, noting gaps to be filled in, and planning strategies for the next interview session (Francis 1983). The researcher constantly had to hold discussions with the fieldworkers after every interview session to review problems and avoid further obstruction.

4.4.2 FIELDWORKERS

All interviews and data collection were conducted by the researcher herself together with two other fieldworkers. The fieldworkers were trained graduate teachers from Universiti Malaysia Sabah, who were waiting for their posting to a government teaching position. Unlike the researcher, they had no linguistic background. Hence, the researcher felt the need to give them a few sessions of briefing and training on issues which might arise such as:

> …how to 'enter' a community sensitively enough to gain the confidence of its members; how to find relevant people to study, how to persuade them to be part of the investigation, including being recorded; how to minimize the disruption caused to their lives by the data collection process; and how to somehow recompense the community for its co-operation (Britian 2005:5).

The researcher and both the fieldworkers are all natives of Sabah. They are Sino-Kadazan, Dusun and Kimaragang respectively, who have first-hand experience and knowledge of the day-to-day way of life of the informants that they interviewed. They are all speakers of the Sabah Malay dialect (SMD), and they speak SMD on a daily basis at home with family and friends. Being native speakers has its advantage, as Gardette puts it:

> A fieldworker who speaks his dialect fluently, who can therefore carry on the interview without speaking the standard language, has several advantages; he puts the informant at ease from the beginning, he distracts him as little as possible from the dialect, he does not impose on him the wearying task of hearing a question in standard language and answering in dialect (Gardette (1968:45) as cited in Francis (1983:81)).

The two fieldworkers are also trained teachers who know the Standard Malay (STM) well. They are able to distinguish between the standard variety and SMD, and know the appropriate place and time to use one language rather than the other. What is most important is that they are good conversationalists, flexible, friendly and easygoing, who are able to enlist the enthusiastic cooperation of the informants.

4.5 DATA RECORDING AND ACCESS

Data collection techniques in this study relied more on capturing the speaker's speech as naturally as possible. A portable high-fidelity tape recorder was used for this purpose. The question and answer sessions with the informant in Part I were recorded for future reference in case of any uncertainties regarding their social background and language choice and use. All interviews conducted in Part II and Part V were between the researcher or fieldworkers and the informants; word-list reading by informants in Part III and passage reading by informants in Part IV were recorded for later

transcription and analysis. This helped the researcher to obtain a clearer phonological variations used by the informants.

All interviews were recorded using a hi-technology Sony Cassette-Recorder, TCM-200DV/150 onto TDK90-minute or TDK60-minute cassettes. Each side of the cassette was allocated for one informant's recording only. It was not to be used for another informant's recording even though it was not fully used up. For example, Side A of a cassette was used for informant No. 01. Therefore, the next informant, who is informant No. 02, was allocated Side B of the cassette. All cassettes were then carefully coded with the reference number of each informant. This enabled the researcher to identify which recording belonged to which informant.

After all the recordings were done, they were typed out on the computer, word by word on wide space specifically with a double-spaced line and triple-spaced character broad transcription, which is the phonemic transcription. All the phonemic transcriptions were then printed out. These phonemic transcriptions print-out are later used for narrow transcription, which is the phonetic transcription. Due to time constraint, the phonetic transcription was be done on papers rather on computer. Data of transcriptions were bound into six volumes (20 sets of informant each) for easy access.

4.6 DATA IDENTIFICATION

All phonemic transcriptions are needed in order to identify the place and number of the possible positions. It is required before the frequency usage of each linguistic variable can be calculated and tabulated. For all the interviews of each informant in Part II and V the researcher identified the possible positions of all nine linguistic variables

(3.4.3.1-3.4.3.9) by circling the vocalic variables, underlining the consonant variables and placing a question mark at all open vowel words (see Appendix Bvi and Bvii). Then, all the possible positions of each variable were calculated and tabulated for each informant. This information was filled in the Possible Position/Frequency Count Form for each speech style (see Appendix Bviii and Bix). The process of identifying the nine variables from the collected data took about three months, from January- March 2007.

However, as all informants were reading from the same word-list in Part III and reading-passage in Part IV (Appendix Bix and Bx), all the possible positions of these two readings were identified and tabulated beforehand, and recorded in the Possible Position/Frequency Count Form (see Appendix Bxi and Bxii).

4.7 DATA TRANSCRIPTION

After all the data from the interviews in Part II, word-list reading in Part III, passage reading in Part IV and casual interviews in Part V were phonemically transcribed and identified, the next step was to transcribe them into phonetic transcriptions. As data transcribing, especially phonetic transcription is a long process, it took about 10 months to reach completion, from April 2007- January 2008.

For each identified possible position, a phonetic symbol was placed. All these phonetic transcriptions were then used to investigate the linguistic variations, particularly the phonological differentiations made by the informants. For the interviews in part II and V, the number and place of possible positions of linguistic variables were different from one informant to another. The researcher played back the recordings and marked each possible position with the phonetic symbol as it was demonstrated by each informant in these two speeches (see Appendix Bxiii and Bxiv).

For example, for informant SM11, there are 9 occurrences of variable (ε) in formal speech and 15 instances of variable (ε) in casual speech. Of these, informant SM11 had demonstrated variable (ε) as [ε] variant 7 times, twice as [i] variant and none as [a] and [↔] variants in word-medial position of Formal speech. In casual speech, variable (ε) was realized as [ε] variant once, as [↔] variant 14 times and none as [a] or [↔] variants in word-medial position. The total frequency of [ε], [i], [a] and [↔] variants used was then calculated and inserted into a specially designed form used for possible position count and frequency count, which the researcher referred to as Possible Position/Frequency Count Form (see Appendix Bxv and Bxvi).

Next, the researcher processed the word-list reading and passage reading using the same procedure. There were 9 and 15 possible positions of the variable (α) respectively, which were predetermined and calculated in advance, as everyone was reading from the same word-list and passages. With the phonemic transcription ready in hand, the researcher transcribed each identified possible position of the variable (ε) with either [ε], [i], [a] or [↔] variant following the pronunciation of each informant in these two speech styles (see Appendix Bxvii and Bxviii). The total frequency of [ε], [i], [a] and [↔] variants used was calculated and filled in the Possible Position/Frequency Count Form (see Appendix Bxix and Bxx). For example, informant SM11 used the [ε] variant 5 times, the [i] variant twice, and none of [a] or [↔] variants, in the 7 identified word-initial positions of variable (ε). SM11 also used the [ε] variant 7 times, the [i] variant twice and none of [a] or [↔] variants, in the 9 identified word-medial positions of variable (ε) in the word-list reading. As for the passage-reading, informant SM11 used the [ε] variant once, the [i] variant twice and none of [a] or [↔] variants, in the 3 identified word-initial positions of variable (ε). On

the other hand, she used the [ɛ] variant 14 times, [i] variant 8 times and none of [a] or [↔] variants, in the 22 identified word-medial positions of variable (ɛ). This whole process was repeated for each of the 120 informants.

4.8 DATA PROCESSING

The main objective of the study was to observe the relationship between the social and phonological variations in different stylistic differentiation. For this study, social variations (3.4.1) comprises four variables, which are age, gender, ethnicity, and one's standing in society, also known as social stratification. Phonological variations (3.4.3), on the other hand, emphasize linguistic variables of some distinctive vowel and consonant variations. The stylistic differentiation (3.4.2) consists of speech contexts, or more specifically, the formality of speech such as the reading word-list style (WLS), reading-passage style (RPS), formal speech (FS) and casual speech (CS). Thus, this study not only highlights the linguistic variations between individuals of different social backgrounds, but also within the speech patterns of individual informants in different contexts.

This study is a combination of qualitative and quantitative research. Therefore, the researcher used SPSS 15.0 (Statistical Package for Social Sciences, Version 15) to measure the frequency of linguistic variations. Each variable of the social differentiation, stylistic differentiation and phonological differentiation was encoded so that the data could be entered into and processed by the SPSS (see Appendix Bxxi).

4.9 DATA ANALYSIS

Data that has been collected, transcribed and processed was then combined or grouped according to their social groupings such as age, group, gender, ethnicity and social

stratification. For example, informant SM11 was grouped under the age group: 36-45 year olds; gender: Female; ethnic membership: Kadazandusun; and social stratification: Middle Working Class, in order to be further analyzed. The duration taken by the researcher to process and complete analyse the data was about three months, from February-June 2008.

There are two ways of measuring the level of use of a linguistic variable; first by the percentage mean (4.8.1) and second by index score (4.8.2). Here, the researcher used percentage means to calculate and to show the frequency use and percentage differences of a certain variant as compared to another variant(s) of the same variable in different stylistic differentiation, which are reflected in real figures. Real figures in themselves are not sufficient to confirm whether or not the percentage differences within a social variation and between stylistic variations are significant. Hence, these real figures are further verified by significant tests, namely Independent-Samples T-test or One–Way ANOVA test (4.1.1.1) and Paired-Samples T-test (4.1.1.2). On the other hand, the researcher used the index score as a discussion to determine and conclude whether or not social and stylistic variations play any role in shaping linguistic variables.

4.9.1 PERCENTAGE MEAN

In order to calculate a percentage mean, each informant's data of linguistic variable frequency usage will be identified and tabulated. For example, informant SM11 used the [ɛ] variant 5 times, the [i] variant twice, and none of [a] or [↔] variants, in the 7 identified word-initial positions of variable (ɛ) in WLS. The percentage means of variable (a) by SM11 are as follow:

$$[e]: \quad 5/7 \quad \times 100 \quad = 71.42\%$$

[i]: 2/7 x 100 = 28.58%
[a]: 0/7 x 100 = 0.00%
[↔]: 0/7 x 100 = 0.00%

The percentage use of each variant stated above of the informant SM11 was combined with the other informants in each social grouping and stylistic differentiation in order to get the percentage mean. For example, the percentage mean of variable (a) for gender is equivalent to the total percentages of 68 informants who took part in WLS divided by 68. For example, the total percentages for each variant of variable (e) in WLS by 68 females are [e]=11199.60, [i]=427.20, [a]=0, and [↔]=373.20, then multiply by 100. Finally, the percentage means for females are as follow:

[ɛ]: 11199.60 / 11664 x 100 = 96.02%
[ɪ]: 427.20 / 11664 x 100 = 3.66%
[ɑ]: 0.00 / 11664 x 100 = 0.00%
[↔]: 373.20/ 11664 x 100 = 0.32%

In order to investigate the frequency of the linguistic variants used by different social groups, the percentage mean of the combined data for each social group was calculated and then compared with different stylistic differentiation. An example of this would be to measure the use of variable (ɑ) (e.g. [ɑ] variant and [↔] variant) in an identified word position (e.g. word-final position) by social context (e.g. Female) in a different stylistic context (e.g. Word-list Style). Through these processes, any mean percentage of a linguistic variable of any position produced by a particular social group in any stylistic context can be calculated and tabulated, and seen in a percentage mean table.

4.9.1.1. INDEPENDENT-SAMPLES T-TEST AND ONE-WAY ANOVA TEST

It is important to determine whether or not social contexts, such as gender, age, ethnicity and social stratification play a role in influencing the use of linguistic variables and how significant they are. For example, it can be determined whether or not males and females (or any social group) use a significantly different amount of [α] variant (and [↔] variant) in any of the stylistic contexts. Significant differences in figures, specifically in mean percentages were difficult to establish. Hence, a statistically significant test was needed. In this study, either the Independent-Samples T-test, for moderator with two variants, such as Gender with only two variants-male and female (see Appendix Bxxii), or the One–Way ANOVA test, for moderator with three or more variants such as Age (5 variants), Ethnicity (7 variants) and Social Stratification (5 variants) was utilized (see Appendix Bxxiii). The significant level is at 5 % (p<0.05) (2-tailed).

As the One–Way ANOVA test is to search for significant difference between particular social groups such as age group, ethnic group or social stratification group, a Post Hoc test, namely *Scheffé* was further used to find any significance by any possible pair of comparison to be made within a group. This *Scheffé* post hoc test was especially useful for those social variables with more than two variants such as age, ethnicity and social stratification (see Appendix Bxxiv). One-Way ANOVA is able to tell that there are significant percentage differences by ethnicity, but in order to know which ethnic group differs from which one, a post hoc test *Scheffé* was needed to be employed (see Appendix Bxxv). For example, in the case of variable (tΣ), there are percentage differences of [tΣ] or [s] variants between ethnic groups. By using a post hoc test, a more comprehensive result could be obtained, such as differences between CHN and MLY, CHN and KDZ, or CHN and BJU, but there is no difference between MLY, KDZ and BJU. Moreover, when there are significant differences, the post hoc

test *Scheffé* can detail which percentage is bigger or smaller, for instance CHN use more of the [s] variant than MLY, KDZ and BJU.

4.9.1.2 PAIRED-SAMPLE T-TEST

To determine whether or not stylistic context plays a role in influencing linguistic variables by a particular social group and whether or not the difference is significant or otherwise, a statistically significant test, namely, the Paired-Samples T-test, was employed. This paired-Sample T-Test can determine whether or not males or females (or any social group) use a different percentage of [α] variant (or [$\leftrightarrow$] variant) in different stylistic variations, which are the WLS, RPS, FS and CS (see Appendix Bxxvi). Although there are four stylistic variations, this test was used to test the significance of variations produced in social context by pair-wise comparison of different stylistic contexts. The pairings are grouped according to their formality procession, for example, Pair I: WLS-RPS; Pair II: RPS-FS; and Pair III: FS-CS. The significant level is at 5 % (p<0.05) with 2-tailed.

4.9.2 INDEX SCORE

No one social group used just one single variant. Rather all groups used all forms which are variants of one another, in different proportions. Thus, an index score was established in order to measure the tendencies and probabilities of use, as it measures the pronunciation movement of a variable for example from a standard variant to the less standard variant or vice versa. First, the percentage mean of the combined data for each social group mentioned above has to be converted into index scores. These indices range from 100 to 200 or more depending on the number of variants of a particular linguistic variety. For example, variable (ε) has 4 variants: (ε)-$_1$ = [ε], (ε)-$_2$ = [ι], (ε)-$_3$ = [α] and (ε)-$_4$ = [$\leftrightarrow$], thus, index scores would range from 100 to 400. The

consistent use of $(\varepsilon)\text{-}_4$ will produce a score of 400, while the consistent use of $(\varepsilon)\text{-}_1$ will produce a score of 100. To compute the index score, first, the percentage mean of all the variants has to be multiplied by the number of $_X$ in a variant such as Mean x $(e)_{\text{-}X}$ = Index Score. For instance, the percentage means of variable (ε) for the particular social group MMC are 98.86 %, 0.86 %, zero % and 0.29 % respectively. The index score was calculated as follows:

98.86	x	$(e)\text{-}_1$	= 98.86
0.86	x	$(e)\text{-}_2$	= 1.72
0.00	x	$(e)\text{-}_3$	= 0.00
0.29	x	$(e)\text{-}_4$	= 1.16
Total Index Score			= **101.74**

This means the tendencies and probabilities of variable (ε) used by this particular social group is almost consistent with the use of [a] variant as the total index score of 101.74 is very close to 100 (see Appendix Bxxvii for summary index scores for all linguistic variables).

The index score of each linguistic variable is then divided according to stylistic context. A line-graph was formed from index scores. From the index line-graph, it is possible to determine whether or not stylistic context plays a role in influencing the relationship between social context and phonological differentiation.

In order to observe the correlation of social context and phonological differentiation, the social lines and the gaps between them were examined in the index line-graph. First, the more distinct the social lines are from each other and not overlapping or crossing each other, the more the social variation is correlated to the

phonological differentiation. Second, the greater the gap between social lines, the greater be the role of the social groups be in determining the phonological differentiation. This is further supported by the significant percentage differences of the linguistic variable between one social group and another in different stylistic context, which are determined by to the Independent-Samples T-test or One-way ANOVA test (see 4.9.1.1). For example, if the index line-graph shows social lines are well distinguished, gaps are wide and the percentage differences between the lines are tested as significant at 0.05 %, then there is correlation between social variations and phonological differentiations. But if the social lines are overlapping or crossing each other and the significance test indicate insignificant results, then the phonological differentiation is not correlated to the social context. If the social lines are all well lined up and well distinguished but not well spread out and supported by the insignificant test result, then the phonological differentiation is also not correlated to the social context (see Appendix Bxxviii for summary significance of social variations).

On the other hand, to observe the correlation of stylistic and phonological differentiations, the movements of the social lines from one stylistic differentiation to another are examined in the index line-graph. First, the more gradually the social lines are moving to one direction and not leveling at same point, the more the stylistic differentiations are involved in determining the relationship between social contexts and phonological variations. This is then supported by the significant percentage differences of the linguistic variable realisation by social context between one speech style and another, which are determined by the Paired-Sample T-Test (see 4.9.1.2). For example, if the index line-graph shows that social lines are consistently rising or dropping in the less formal stylistic context and the percentage differences of the

movement between one stylistic context and another are tested as significant at 0.05 %, then there is a correlation between stylistic context and phonological differentiation. But if the social lines are almost level or moving up and down drastically and the significant-test indicated insignificant results, then the phonological differentiation is not correlated to stylistic context (see Appendix Bxxix for summary significance of stylistic variations).

Some linguistic variables are subject to stylistic variations as well as social variations such as gender, age, ethnicity and social stratifications, and are referred to as a MARKER. A marker plays some, if not a major significant role in marking a social difference. It marks the speech differences between male and female, different agae groups, different ethnic groups and different social strata. However, if the linguistic variable is subject only to social variation but not involved in sytematic stylistic variations, it is referred to as an INDICATOR. If a variable is merely an indicator in a particular speech community it plays a less significant role in marking a social difference in that community than does a variable which is a marker. A variable that is subject only to stylistic variations but not social variations or subjected to neither one, is neither a marker nor an indicator, and plays no significant role in marking a social difference in the speech community (see Appendix Bxxx for summary social functions of the linguistic variables in the speech community)

The findings will be generalized and discussed under the following headings in the findings chapters:

i) Social and Stylistic Variations of Variable (α)

ii) Social and Stylistic Variations of Variable (ε)

iii) Social and Stylistic Variations of Variable ($\leftrightarrow$)

iv) Social and Stylistic Variations of Variable (ο)

v) Social and Stylistic Variations of Variable (η)

vi) Social and Stylistic Variations of Variable (ρ)

vii) Social and Stylistic Variations of Variable (κ)

viii) Social and Stylistic Variations of Variable (τΣ)

ix) Social and Stylistic Variations of Variable (?)

4.10 DATA PRESENTATION

The data for this study are presented in tables and not in its raw form. Data collected for Questionnaire-Part I are presented in tables as shown in Table 3.1-3.14 (see 3.4). However, the results from data collection of Questionnaire-Part II, Part III and Part IV are presented statistically, summarized in tables and graphs, and are presented in the findings chapters.

4.11 CONCLUSION

The methodology of carrying out this study is discussed in this chapter and the previous one, Chapter III. The following Chapters V-XIV will discuss the findings of this study, namely the correlation between phonological differentiation of the SMD and both social and stylistic variations.