

## CHAPTER 3

### RESEARCH DESIGN AND METHODOLOGY

#### 3.0 Introduction

This chapter will explain the procedures of data collection and the methodology used to analyse the research questions. The research questions and a brief explanation of the research sample will be given.

#### 3.1. Research Questions

As explained in Chapter 1 and Chapter 2, the use of English on the Internet is an inherent feature due to the technological ease and the origins of Internet having roots in an English speaking world. In such an environment, how do non-native speakers of English communicate with each other on the Internet? To address this question, a website organised by and for non-native speakers of English was chosen to analyse the language used. The website chosen for this study is a website on Tamil Language at <http://www.tamil.net>. As mentioned in chapter 1, the following research questions will be asked :

a) Research Question 1 : How much English has been used in this website on Tamil Language <sup>was</sup> <sup>in a month</sup> over the course of a month?

PERPUSTAKAAN UNIVERSITI MALAYA

b) Research Question 2 : What are the types of English and Tamil words used in this website on Tamil Language?

c) Research Question 3 : Is there a pattern of code-switching in the emails of the chosen website?

### **3.2. Research Sample**

Being a frequent user of the Internet using local and foreign websites, it was observed that English was quite frequently used in the non-English websites and also among web-users who are non-native speakers of English. Most times, a combination of both English and the vernacular languages were used.

There are about 13,800 websites on Tamil language as mentioned in Chapter 1. Among them, Tamil.Net was chosen because it is one of the oldest and most outstanding of all links in terms of quality and academically-inclined projects. It started a few years ago and is still an actively used site.

The Tamil.Net is a Tamil cyber community aimed at bringing together Tamilians from all over the world to get information and to communicate with each other. It is an online community which has forums and discussions, projects and frequently asked questions. Facilities, such as appropriate software, discussions and contributions, both in Tamil and English are available. The discussion group has members who speak and understand the Tamil Language. They come from various parts of the world,

such as, Canada, United States, Australia, India, Mauritius and Malaysia. Users of this website are encouraged to join the cyber community to be able to listen to other participants ideas and views on prevailing topics under discussion.

### **3.3. Collection of data**

Data was collected from the homepage of Tamil.net and from the archived emails available from the discussion group of this website. The Tamil.Net website has other projects which were too large to be included in this study. The method of collecting data based on the research questions is discussed below :

#### **3.3.1. Research Question 1 : How much English was used in this website in a month?**

For the first research question, both the homepage and the emails from the month of November 2001 was used. From the archived emails in this website, November 2001 was the first month listed and so the emails from this month were chosen for the study. The method of data collection from the homepage and the emails are as follows :

## Homepage

The entire homepage was printed out and a word count was made manually to count the number of words in English and in Tamil. The percentage of English words and Tamil words was calculated.

Number of English words / Total number of words x 100 = Percentage of English words used in the homepage. The same method was used for Tamil words.

Tamil words in the homepage include Tamil words in the Tamil script and in the Roman script (transliterated). Transliterated words relate letters to sounds and are written in Roman alphabet making the words look like English words.. For example, the word 'Internet' is written as 'Inaiyam', which means Internet in Tamil. This enables someone who understands Tamil but does not know how to read and write in Tamil script, to communicate in a variety that can be understood by Tamil speakers. For the purpose of analysis, names of individuals were excluded. All abbreviated words were considered as one word; for example words such as 'you'll'.

Since Tamil.net has the properties of a hypertext as was mentioned in Shortis (2001), the words representing links which are to the left column of the homepage were counted.

## **Emails**

The 'archives' in the homepage which links to the emails in the discussion group was used to choose the emails for the study. The archives in the website was arranged according to months beginning from November 2001. A total count of the emails in the month of November, 2001 was made.

Emails solely in Tamil, solely in English and with both languages were counted separately. Where there is some form of interaction between sender and receiver, there is a tendency at times to combine languages especially when there is a language choice available. It ranges from a simple combination of the title in Tamil and the reply in English or a reply in Tamil to a previous mail in English. These emails with a combination of both languages in some form were considered as emails with mixed code and were used for this study.

The percentage of English emails and Tamil emails was calculated as follows :

Number of emails in English / Total number of Emails x 100

= % of emails in English

Number of emails in Tamil / Total number of Emails x 100

= % of emails in Tamil

Number of emails in both languages / Total number of Emails x 100

= % of emails in both languages

### **3.3.2. Research Question 2 : What are the types of English and Tamil words used in this website?**

For the second research question, only the homepage was used for the analysis. This was because homepages represent texts which have a longer life span unlike emails. The type of English words used in the homepage were analysed using Shortis (2001) properties of technology texts and Hymes (1997) fundamentals of communication as explained in Chapter 2.

The homepage was analysed for the type of words used in the tags which appear on the left margin of the homepage. The universal properties of technology texts as formulated by Shortis was observed. The general layout of the homepage, the Internet jargon words, and the abbreviated words were analysed.

Using Hymes' components of speech, the relationship of the addressor and addressee was analysed by looking at the language used by the producer of Tamil.Net to communicate to the participants. Personal statements, reference to names of people, and the style used to communicate was observed. The ways of expressing the message contents, the types of clauses, spoken versus written text were also analysed.

The Tamil words used in place of English words in the homepage were categorised according to noun, verbs, adverbs, adjectives and proper nouns. Words which appear more than once were counted and the number of

times they were used in the homepage appear by the side of the word. These words were then arranged together by broad topics such as food, fashion or coined words, words referring to the Internet and expressions of belongings in order to find out the reasons for using these words.

**3.3.3. Research Question 3 : Is there a pattern of code-switching in the emails of this website?**

The emails which had both English and Tamil words were chosen and copied. However, only the original email texts were included. Portions of the emails which were forwarded or copied were excluded. The font was changed to New Roman Times and the size was changed to size 9. This was necessary to suit the text of the research.

In each email, sentences which were solely in English and Tamil were omitted. The English words in the Tamil sentences and the Tamil words in the English sentences were placed in bold to emphasise that these words are used for the analysis of this study. This also included English and Tamil words in the openings and closings of the emails. Long sentences in Tamil and English were replaced with '.....'.

These words, both English and Tamil were then categorised according to broad topics such as pronouns, technical words, phrases or expressions, religious terms, food, names of places, words relating to Internet

and words relating to music and sound as was similarly done for the Tamil words found in the homepage for Research Question 2.

The reasons for using English words in Tamil sentences and Tamil words in English sentences were then analysed using Valdes-Fallis' strategies of code-switching as mentioned in Chapter 2. Referring to Tables 2.1 and 2.2, the external and internal factors were looked at to analyse the code-switching patterns in the following manner :

- a. Role relations – the openings and closings of the emails were observed closely to look into the pattern of code-switching in relation to role relations among the participants.
- b. Settings – the language choice in the titles of the emails and in the text message of the emails was observed.
- c. Style – the personal style of the participants in relation to the language choice was observed.
- d. Proficiency in both languages – the use of both languages in the sentences was observed to study the proficiency of the participants in both English and Tamil.

### **3.4. Conclusion**

The procedures of data collection for this study have been outlined in this chapter. In the following chapter, these procedures will be used to analyse the homepage and selected emails of Tamil.Net.