Chapter 3

Research method

3.1 Introduction

This chapter describes the research methods and procedures applied in the present study. It provides information on the subjects and how they were selected, and on the research instruments and materials, data collection procedures and methods of data analysis.

This study is concerned with how different immersion experiences – in terms of length of time spent in immersion and the quality of the immersion environment – will affect the vocabulary size of non-mother tongue speakers of English in Grade 7 at certain types of South African schools. More precisely, it is concerned with how much, and what kind of immersion in English is necessary for immersion subjects to develop a vocabulary size comparable to that of mother tongue speakers of English at the same academic level (§1.2).

The study is based on data collected from a group of Grade 7 school students, between the ages of 12 and 14. In contrast to the Kiel study (§2.2.2), which was a major stimulus for the present research, and which examined both spoken and written data, the present study focused primarily on direct vocabulary tests and secondarily on written exercises. The instruments and procedures were tested in an initial pilot study conducted at School S. In this pilot, spoken data was collected but it was found that certain children monopolised the conversation, while others made almost no contributions at all. In order to counter this and to provide richer productive data in the main study, all subjects completed the questionnaire designed to elicit information on language background and language environment, the three vocabulary tests to assess their receptive vocabulary and a writing exercise, rather than a spoken discussion, to assess their productive vocabulary (§3.4.4). In the light of the experience of the pilot study, I felt that a written exercise would be more appropriate to the aims of the study and would provide richer productive data, facilitating a comparison of subjects' productive and receptive

vocabulary.

3.2 Hypotheses: a brief rationale and formulation

Two different classifications of immersion are used in this study: the terms early immersion (EI) and late immersion (LI) refer to the number of years spent in an immersion situation, while deep immersion (DI) and shallow immersion (SI) refer to the quality of the immersion experience deriving from the language environment in which the subjects found themselves (§1.3.1).

Early immersion subjects are those children who, at the time of the study, had been immersed in an English environment at school for four years or longer. Late immersion includes those subjects who were fairly new to immersion, having spent less than four years in an immersion situation at the time of data collection. The choice of this cutoff point between early and late immersion was motivated essentially by pragmatic factors: making the division at this point meant that the number of students in each group was large enough to comprise a viable sample.

Subjects were classified as having undergone deep or shallow immersion depending on the school they attended. Both schools used in this study were what is known as 'ex-Model C' schools (see §2.2.3). In the previous political dispensation these would have been reserved exclusively for white students, and would have been staffed by white teachers only; at the time this study took place, however, their pupil body made them very different from each other. Both schools had changed considerably in make-up over the last decade: numbers of mother tongue speakers of English had shrunk and numbers of African language speaking children had increased. At the time of data collection in 2001, School D (deep immersion environment) had a pupil ratio of African language speakers to English speakers of about 80:20, while the ratio at School S was 95:5. The teaching population had remained more or less the same, however, comprising a majority of white English or Afrikaans mother tongue speakers at both schools. All teachers (21 in total) at School D were white; teachers at School S (41) were predominantly white with two coloured teachers speaking Afrikaans as mother tongue, and three black teachers. Thus, through the opening of schools to all races a situation had arisen in which many

teachers and students did not share common languages or cultural backgrounds. Nonetheless, despite this similarity, subjects at School D were classified as deep immersion subjects because the English language environment of the school was comparatively rich: at the time of data collection, teachers of the Grade 7 group were predominantly MT English speakers while, despite the high ratio of African to English mother tongue speakers at the school, students in the Grade 7 classes were made up of half MT English speakers (29) and half African language L1 speakers (29). African language speaking children were concentrated in the lower grades. As the school had a large component of foreign students (there was a total of 14 foreign language speakers in the Grade 7 classes tested, for instance) and a wide range of languages other than indigenous South African languages was spoken by students, the language spoken both formally and informally in the classroom and in the playground tended to be English (private communication with teachers and own observations).

On the other hand, subjects from School S were classified as shallow immersion subjects in that, although at the time of data collection the majority of teachers were English or Afrikaans mother-tongue speakers, the students were almost exclusively non-mother-tongue speakers of English: 54 of the 56 students in the class spoke an African L1. The remaining two children were foreign language speakers. Playground as well as classroom interaction tended to be in the students' mother tongue rather than in English as the shallow immersion subjects were a more homogenous language group than immersion subjects at School D (see Table 3.1), with 47 of the 54 subjects speaking North Sotho, South Sotho or Tswana. These languages are closely related as they all belong to the Sotho group. Languages represented at School D were spread more evenly over a number of indigenous South African languages (see Table 3.1).

Although there were differences in the quality of immersion at these two schools, it must be pointed out that as far as the nature of English immersion in South African schools is concerned, they were in fact quite close together on the continuum of school types. Both were staffed mainly by English MT or fluent English speaking teachers who were not fluent in any African languages, making little or no code-switching possible; both schools were urban and both had the same type of resources, such as a library, sports facilities, educational aids and so on. School business was conducted through the medium of English in both cases. Compared to poor rural schools, which lack even the most basic

resources such as water and electricity and which are staffed by teachers who are not fluent in English themselves, these represented relatively rich English language environments.

Table 3.1 (§3.3) indicates the home languages spoken by the Grade 7 groups at the two schools. These are representative of the school as a whole in the case of School S, but not of School D, where the overall ratio of non-mother-tongue speakers of English to mother-tongue speakers of English in the Grade 7 classroom was lower than that of the school as a whole.

Against this background, two hypotheses were formulated. These are expressed generally as the length of immersion hypothesis and the quality of immersion hypothesis, but each is tested in terms of receptive vocabulary and productive vocabulary (§1.6). The length of immersion hypothesis is expressed as a directional hypothesis and predicts that early immersion subjects will have a larger vocabulary than late immersion subjects, and that the vocabulary size of mother-tongue speakers of English will be larger than that of both groups of immersion subjects. This is supported by research conducted in the Kiel study (Daniel and Nerlich, 1998; Kickler, 1995; Wode, 1995)(§2.2.2).

The quality of immersion hypothesis (§1.6), expressed as a non-directional hypothesis, states simply that there will be a difference in vocabulary size between the two immersion groups, deep and shallow immersion, and between these groups and mother-tongue speakers of English, without specifying the direction of this difference. The reason for the non-directionality of the hypothesis is that there is very little research evidence in this area to support the idea that the quality of the immersion environment, as distinguished in this study, has a positive effect on vocabulary size (Hacquebord, 1994).

3.3 Subjects

As was described above (§3.2), subjects were drawn from the Grade 7 classes of two Pretoria primary schools, referred to here as School D and School S. School D's Grade 7 group comprised 79 students, while School S had 59 Grade 7 students. Speakers of foreign languages (see §3.2) were excluded as I wished to focus on the experience of the majority of the subjects, that is, those who spoke either English or an African language. Also excluded were those subjects who missed one or

more of the tests. This brought totals to 58 subjects from School D and 54 from School S. All subjects were in Grade 7 at the time of the study: this is the last year of primary schooling in South Africa but is regarded as the first year of the secondary phase or General Education and Training Certificate, which learners achieve at the end of Grade 9.

Once speakers of foreign languages and subjects with missing data had been eliminated from the data set, half the subjects (29) from School D were MT English speakers and, of the immersion subjects, early immersion subjects were in the majority (22 as opposed to 7 late immersion subjects). School S's student population, on the other hand, comprised entirely mother-tongue speakers of South African indigenous languages. Early immersion subjects (33) outnumbered late immersion subjects (21). Participants in the study represented several South African language groups (see Table 3.1).

Table 3.1: Subjects' home languages

		School D	School S	Total
Home language	English	29	0	29
	Tswana	7	26	33
	South Sotho	8	14	22
	North Sotho	2	7	9
	Zulu	8	1	9
	Xhosa	2	4	6
	Tsonga	1	2	3
	Swazi	1	0	1
Total		58	54	112

3.4 Instruments

The instruments and materials for this study consisted of a questionnaire, two vocabulary tests which were compiled by the researcher, the Vocabulary Levels Test (VLT) designed by Nation (1990), and a writing exercise. This section describes the design and application of these instruments.

3.4.1 Questionnaire

The first step in the study was the design of a questionnaire to identify the language background of the

subjects. This questionnaire elicited information which was useful also in determining the richness of the language environment which subjects experienced outside the school environment. It requested information on home language, years spent in immersion, language/s spoken at home, to both parents and siblings, and languages spoken in the playground and outside the confines of the school (see Appendix A). It was on the basis of the information provided by subjects in this questionnaire that they were assigned to the early or late immersion group, or designated as MT English speakers or foreign language speakers.

3.4.2 Tests A and B

These instruments (see Appendix B) were designed in the following way: two chapters of a Grade 7 history textbook (Marzo and Stokes, 1997) were taken as the corpus for the vocabulary tests. These two chapters were chosen because they were authentic texts (in use in schools) and dealt with content which was accessible and fairly familiar to Grade 7 learners (one chapter covered life at the Cape in the time of Jan van Riebeeck, and the other dealt with everyday life in Medieval Europe). This similarity of passages, it was hoped, would add reliability to the test results. These chapters were tested for readability by using the Grammatik function in WordPerfect. This function uses the Flesch Readability Index, which is a measurement of the ease with which a document can be read. Table 3.2 below shows typical Flesch RES (reading ease score) values for some common reading material:

Table 3.2 Flesch RES

Material	Flesch Index
Comics	95
Consumer Adverts	82
Sports Illustrated	65
Time	57
New York Times	39

Flesch readability indexes are also often translated into the educational level that is generally necessary to understand a document, as indicated in Table 3.3 below. The row indicating Grade 7 has been

highlighted, as this is the level of texts which subjects in this study would be expected to read with ease.

Table 3.3

Flesch Index	Educational Level
91-100	5th grade (grade 5)
81-90	6th grade (grade 6)
71-80	**7th grade (grade 7)**
66-70	8th grade (grade 8)
61-66	9th grade (grade 9)
51-60	High School
31-50	Some Colleges
0-30	College Graduate

The Grammatik function also provides statistics on a number of features that have been linked to readability, such as the use of the passive voice, sentence complexity and number of words per sentence. The scores for the two chapters used in this study are tabled below (Table 3.4)**:**

Table 3.4: Readability features

	Medieval times	The Cape Colony
Flesch Reading Ease Score (100 = very readable)	69	69
Passive voice (%)	9	10
Sentence complexity (100 = very complex)	22	40
Vocabulary complexity (100 = very complex)	13	13
Words	1985	1552
Sentences	184	96
Simple sentences	96	32
Syllables per word	1.5	1.43
Words per sentence	10.78	16.16
Sentences per paragraph	2.35	4.17

As it happened, these two passages were identical in readability according to the RES, both having a score of 69. This places the texts at a level which is marginally higher than US Grade 7 material intended for mother-tongue speakers of English. Other values which are very similar if not identical are the vocabulary complexity and percentage of passive verbs. There is a marked difference in sentence length between the two passages, however, with Passage 2 having a much higher average sentence length than Passage 1.

A lexical analysis was then performed on the two passages, using the VocabProfile computer program. This is a freeware computer program developed by Nation (1990) and is available at several URLs, such as *www.jbauman.com*. This program analyses any text against three different word lists. Baselist 1 contains the 1000 most frequent words in English and baselist 2 the second thousand most frequent: according to Nation (1990:19), the words from these two lists account for around 87 percent of the running words in a text. The third baselist is made up of Xue and Nation's University Word List (1984) which contains 836 word families (§2.3.3.1) which are particularly common in academic as opposed to general texts, that is, what is sometimes referred to as 'subtechnical vocabulary', and is referred to in this study as the 3000-word level. According to Nation (1990) the words from this category account for approximately eight percent of the running words in a university academic text, irrespective of the discipline from which it was taken. It must be kept in mind that this study dealt with children who were on the point of entering secondary schooling and as such were not expected to be completely familiar with this third category. However, a developing knowledge of these words was expected from the MT English speakers and from the early immersion and deep immersion subjects at least. The remainder of the words which are not found in any of the first three lists make up the fourth category, the above-3000-word level, low frequency words. The program divides the words in any given text into these four categories according to their frequency.

Several research studies have used this program. Laufer and Nation (1995) found that the VocabProfile score correlated with an independent measure of vocabulary knowledge, the Vocabulary Levels Test, which was used in the present study. They also found that the VocabProfile

predicted broader language proficiency measures, in that learners at three proficiency levels – ranging from low intermediate foreign learners of English in New Zealand to graduates of Israeli high schools who were in their first semester at university, to a third group who had completed two semesters at the same university – had significantly different VocabProfile scores.

In the present study the Lexical Frequency Profile (LFP) of the passages was established by using the VocabProfile to analyse the vocabulary levels of each text. In both Passage 1 (*Life in Medieval times*) and Passage 2 (*Life in the Cape Colony*) well over half the words came from the 1000-word level (62.6% and 75.5% respectively), and just over ten percent from 2000-word level (12.3% and 11.5% respectively). Less than five percent (4.6% and 2.2%) came from the UWL or 3000-word level. The words that were not on any of these three lists constituted 20.5 percent of the total words in Passage 1 and 10.8 percent of Passage 2. This last category consisted mostly of proper nouns and subject-specific terminology. From the Grammatik readability features it appears that although a greater percentage of words in Passage 1 came from the 2000-, 3000- and above-3000-word levels, the vocabulary complexity scores were identical at 13. Passage 2 had a higher sentence complexity rating than Passage 1, which may have compensated for the higher number of high frequency words and the fact that Passage 1 may have been marginally more difficult. Based on these RES scores, however, it seems valid to make an objective comparison of the two passages, and to say with some confidence that they were comparable, and that tests set on their vocabulary would be of equal difficulty.

Once the LFP of the passages had been established, two vocabulary tests were devised by selecting a total of 60 words, 30 from each text, on the basis of the LFP data – 20 words from the 1000-word level, 20 words from the 2000-word level and 20 words from the 3000-word/UWL level. Selection included the whole spectrum of function words in an attempt to cover as wide a vocabulary as possible: nouns, adjectives and verbs, prepositions and linking words such as conjunctions and adverbs. Two parallel tests were compiled using these 60 words: Test A comprised a total of 30 questions, ten words from each of the three levels, taken from the passage on medieval life (Passage 1). Test B comprised a total of 30 questions, ten words from each of the three levels, taken from the

passage on the Cape Colony (Passage 2). Tests were made up of discrete item, multiple-choice cloze-type questions. Test and distractor items in each of the three levels of words per test were phonologically or orthographically similar and all from the appropriate frequency level, the aim being to standardise as far as possible the level of difficulty of each item in each section, and across both tests. Below are two examples of items from Test A and two from Test B. Correct answers are in bold type.

Test A

In Medieval times explorers sailed to the East _____ they wanted to trade for spices, silk and jewellery.

a *although*

b *yet*

c ***because***

d *still*

People like Leonardo da Vinci made wonderful _____ which still amaze us today.

a ***inventions***

b *institutions*

c *intentions*

d *impressions*

Test B

There were four different groups at the Cape: the Company officials, the _____ burghers, the Khoikhoi and the slaves.

a *tree*

b ***free***

c *fear*

d *fine*

Ryk Tulbagh was _____ by the Dutch East India Company and was sent to the Cape as Governor in 1751.

a ***employed***

b *entertained*

c *escaped*

d *explained*

3.4.3 Nation's (1990) Vocabulary Levels Test

In addition to these two vocabulary tests (Test A and Test B), a third instrument was used (see Appendix B). Each subject completed the Vocabulary Levels Test, or VLT (Nation, 1990). This test was included to add to the richness of the data on receptive vocabulary size and also to act as a benchmark for the other tests, as it is an internationally recognised and widely used test which is freely available to researchers. This test was devised as a measure of vocabulary size, for use by teachers who wished to develop suitable vocabulary learning and teaching programmes for their students. It has been widely used in New Zealand and other countries as a diagnostic vocabulary test for immigrant students when they arrive in an English-speaking country. It is generally accepted as an appropriate measure of second language (English) vocabulary size (Laufer, 1992a; Yu, 1996, in Qian, 2002) and was designed specifically to establish where learners should be given help with vocabulary learning.

The VLT is made up of five sections, each representative of a different vocabulary level in English relating to specific vocabulary learning objectives, namely the 2000-word level, the 3000-word level, the 5000-word level, the university word level (UWL) and the 10 000-word level. Nation's (1990:261) theory is that all learners should know the high frequency words contained in the 2000- and 3000-word level lists to be able to function effectively in English. The 5000-word level is on the boundary between high and low frequency words. The University Word List (UWL) represents one type of specialised vocabulary and should assist learners in reading academic material. The 10 000-word list comprises low frequency words (§2.3.3.1).

According to Read, 'scholars work on the assumption that, in order to read independently, learners

should know at least 95 percent of the running words of a text. This means that only one word in 20 will be unfamiliar to them' (2000:83). According to both Nation (1990:24) and Laufer (1992; 1997), to achieve this kind of coverage, learners must have a vocabulary of at least 3000 word families (this issue is dealt with in more detail in §2.3.3.1). Such estimates of the vocabulary size of MT English speakers can be used as a benchmark for the acquisition of vocabulary of non-mother-tongue speakers who enter a school system which uses English as the language of learning and teaching (LOLT). Scores of MT English speakers were used in a similar way in the present study as a foil against which to compare immersion subjects' vocabulary size.

The Vocabulary Levels Test comprises word-definition matching type items, with six words and three meanings in each question. Each section comprises 18 possible correct choices. Each word was chosen 'so that they would be representative of all words at that level' (Nation, 1983, in Qian 1998:54). Subjects match words to definitions, and one point is awarded for each correct choice. This format was used to avoid guessing correctly as far as possible. In each group, all words belong to the same word class, thus reducing the chance of a grammatical clue to the right definition. The test constitutes a broad measure of word knowledge and, as stated above, was included in this study to provide a third set of results that should raise the overall validity of the findings on receptive vocabulary. An example of an item from the first section (2000-word level) of the VLT is provided below:

1.	*original*	
2.	*private*	
3.	*royal*	_____ *complete*
4.	*slow*	_____ *first*
5.	*sorry*	_____ *not public*
6.	*total*	

3.4.4 Writing exercise

To be proficient in a language learners need not only to know or recognise words, but also to demonstrate their knowledge of words appropriately in their own speech and writing (Read, 2000:3). Researchers have adopted a communicative approach in which learners complete tasks simulating real life situations. Bachman and Palmer (1996, in Read, 2000:5) believe that there is a place for vocabulary assessment within task-based language testing because language ability has two components – language knowledge and strategic competence. Learners need to know about vocabulary, grammar, the sound system and so on, but they should also be able to draw on that knowledge effectively for communication purposes under normal, everyday conditions. Although the main focus of this study was the size of receptive vocabulary among immersion subjects, it was decided to include an activity which would reveal something about the subjects' productive vocabulary and their ability to use the language to perform communicative tasks as well. Although these data were not substantial enough to form a basis for making any strong claims about subjects' productive vocabulary, the inclusion of the exercise did allow the researcher to link the present study more closely to the Kiel study (e.g. Wode, 1995, see §2.2.2) and allowed for some comparability between the studies as Wode used the same passage 'A Desperate Decision' (from Klippel, 1987) as his prompt in eliciting productive data.

The passage in question tells the following story: a group of children and their teacher go hiking in the Scottish Highlands. An accident occurs and the children are faced with a situation in which they have to find solutions to several problems. In the Kiel study, this scenario was used as the starting point for group discussions and written exercises. In the group discussions, subjects were arranged in groups of three and asked to discuss together how they might get themselves out of their plight. Their discussions were recorded and analysed. To elicit free written (productive) data, subjects were asked to imagine themselves in this situation and to write down their responses. In the same way, in the present study subjects were required to complete several tasks in writing, which were adapted from the prompts used in the Kiel study (Wode, 1994): to provide an interpretation of a map, to explain what they would have done under the circumstances described in the passage 'A difficult decision', to write a letter to a friend relating the event as if they themselves had experienced it, and to describe a similar predicament in which they had found themselves in real-life (see Appendix C). This was a

free active/productive writing exercise, and students were at liberty to write what and how they pleased. It was not a controlled-active vocabulary exercise which would have demanded, by means of prompts, that subjects used specific vocabulary.

In the present study the written data received from each subject was transcribed onto the computer and analysed, using VocabProfile, to establish the LFP of each piece of writing. This established the total number of words (tokens and types) used by each subject and the number of words (types) from each level (1000-, 2000-, 3000-word levels, and the words not on any list). Those word types which were not on any of these three lists were individually analysed. Generally, these comprised spelling errors and proper names, as well as any words from the above-3000-word level, such as 'ambulance' and 'dehydrated'. These lists were then edited to reflect only those words which came from the above-3000-word level. The totals of word types on each of the four lists for each subject were then converted to a percentage of the total number of words (types, not tokens) used by each subject. In this way it was possible to compare the number of word types used by each subject at each level.

3.5 Procedure

I visited School D on two consecutive days to administer the tests and to facilitate the writing exercise. Subjects came from the three Grade 7 classes at the school and each group was dealt with separately, during their English-as-subject periods. This procedure worked quite successfully as children did not get too tired or bored with the activities; it did have disadvantages though, as children took a long time getting to the classroom and some valuable time was wasted. Also, there were children who were absent on one of the two days: these subjects were subsequently excluded from the data (§3.3).

Data collection in the vocabulary size tests (Tests A and B and the VLT) took the form of paper-and-pencil testing. The testing session at School D commenced with subjects filling in the questionnaire (see §3.4.1) and completing Test A. Roughly 30 minutes was allowed for the completion of the test, and together with completing the questionnaire this took up one double period (about 50 minutes). The

children were given a short break and then returned to the classroom where they completed Test B, also in about 30 minutes. On the following day, the VLT was administered and after a break, the writing exercise. After reading the story aloud, with children following a transparency on which the passage had been printed, I gave a brief introduction to the exercise with the aid of a transparency featuring a map of the imagined area in the Highlands where the children in the story find themselves. This map featured landmarks such as a river, a range of hills, a fisherman's hut, a lake and a road. A third transparency listed the activities required of the subjects (see Appendix C). After a very brief class discussion as to how the problem in the story could be solved, children were given about 40 minutes to complete the writing exercise. I moved around the classroom and where necessary tried to keep students focused on the task. This procedure was repeated for the other two classes.

At School S, the English subject teacher arranged for all the Grade 7 students to sit together in one venue for an entire day and this made it easy for me to administer the tests and elicit the written data on the same day. This had the advantage of ensuring that all subjects completed all the activities. The session started with the completion of the questionnaire. The order and method of testing followed that of School D, and tests were collected in each case after about 30 minutes. Short breaks were allowed between the tests and the writing exercise.

3.6 Data analysis

As already indicated in the discussion of the hypotheses (§1.6), there are two aspects to the data analysis. The first is the measurement of receptive vocabulary size by means of scores on the vocabulary tests, and the second is the measurement of productive vocabulary size, by means of a writing exercise.

3.6.1 Receptive vocabulary size

Marking of the two parallel tests, Tests A and B, and the Vocabulary Levels Test was done manually. Each correct answer was awarded one point and the results entered using the SPSS statistical program. For Tests A and B, results for each item were entered (in binary code, 1 for a correct answer, 0 for an incorrect answer). Each subject's choice on each item was also entered for use in a possible later item analysis (options were listed as 1,2,3 and 4 and the subject's choice was indicated by entering the corresponding number). In the case of the VLT, subjects' correct or incorrect responses were simply recorded, using 1 for a correct answer and 0 for an incorrect answer. The sum of correct answers at each level (in Tests A and B this was out of 10 for each level, and in the VLT it was out of 18 at each of the five levels) for each subject was then calculated. Means, score ranges and standard deviations were computed for each level of each test.

3.6.2 Productive vocabulary size

The results from the writing exercise were entered in the following way, using the SPSS program: total number of tokens used; total number of lexical types used; total number of lexical types at each of the three levels (1000-, 2000- and 3000-word levels); total number of types from the advanced list (those words not appearing in any of the three lists and determined to be from a higher level). Once the totals at each level had been computed they were reduced to percentages of the total number of word types each child had written. These percentages were reflected in the data as 1000-word level, 2000-word level, 3000-word level and above-3000-word level. Means, score ranges and standard deviations were computed for each level.

3.7 Statistical analysis

The chief statistical test used in this study is the analysis of variance, or ANOVA. The one-way ANOVA allows researchers to handle data from experiments that have designs involving more than two conditions (factors, or variables). This single test informs the researcher whether the change in the independent variable (IV) has affected the scores, that is, whether the different conditions have

resulted in significantly different scores between groups, and whether these are larger than the differences within the groups.

In this study the dependent variable is vocabulary size, operationalised by testing both receptive and productive vocabulary size. The independent variables are length of immersion and quality of immersion. These independent variables each have three levels, or conditions: in the case of length of immersion, MT English speakers, early and late immersion subjects make up the three levels. The three levels of the quality of immersion variable are MT English speakers, deep and shallow immersion subjects. As explained in the sections above (§3.6.1 and §3.6.2), means and standard deviations of the scores on the tests and the writing exercise data were computed. Once the one-way ANOVAs had been done, and where significant differences were found, *post hoc* Scheffé tests (multiple comparisons) allowed me to determine where exactly the greatest differences in the scores lay – that is, how the levels of one independent variable differed in their influence on subjects' performance on the dependent variable.

Subsequently, two-way between-subjects ANOVAs were performed to check for any interaction effects. Interaction effects are the effects of a combination of the independent variables on the dependent variable. The interaction effect reveals whether there were cases where the effect of one independent variable was moderated by the effects of the other. The advantage of factorial designs over conducting multiple one-way ANOVAs lies precisely in this capacity of ANOVA to look at the interaction effect of the combination of variables. Whenever there is a strong interaction effect, the main effect is considered to be less important and the interaction effect overrides the main effect. The interpretation must then focus on the interaction effect. If the interaction is not significant then much stronger claims can be made about the effects of the independent variables (Hatch and Farhady, 1982).

3.8 Conclusion

This chapter presented the methodology of the study and described instruments, procedure and subjects. It briefly outlined the hypotheses of length of immersion and quality of immersion: these are discussed in more detail in Chapter 4. An explanation of the methods of data analysis was provided. A detailed discussion of the results is presented in Chapter 4.