

CHAPTER 3

RESEARCH METHODOLOGY

This study is aimed at investigating the effects of the pre-listening question and the post-listening question techniques on the listening comprehension of the third year English major students at Prince of Songkla University, Pattani Campus, and examining students' attitudes towards the two approaches. This chapter describes the design and procedures of the study. It is divided into four sections: the subjects, the research instruments, the testing procedures, and the statistical devices.

3.1 Subjects

3.1.1 Subject Selection

Seventy-four of the third year English major students, studying in the second semester of the academic year 2006 at Prince of Songkla University, Pattani Campus, were intentionally selected as the subjects in this study for 3 reasons. First, they are considered having a high degree of English listening proficiency when compared with students in other programs since they had studied English for approximately ten years. Second, they have taken several English courses, including *Listening and Speaking 1 (417-237)* and *Listening and Speaking 2 (417-238)* which are courses for students at an intermediate listening level in the English program. Therefore, it could be assumed that the subjects having passed these two courses were at an intermediate listening level. Third, since this research study is aimed at providing useful information to the teaching EFL in the Thai context, it is appropriate to use Thai native speakers as subjects. Having considered all these components, the researcher was confident about the effective measurement of their listening comprehension tests.

3.1.2 Subject Grouping

3.1.2.1 Comparable Groups

After the subject selection procedure, a week prior to the experiment the subjects were asked to undertake a listening proficiency test (see the research instrument) of which the scores were used to rank the subjects, from the highest to the lowest. Then, the subjects with comparable scores were paired. A subject in each pair was assigned to Group A, and the other to Group B. The subjects were then equally split into two groups with similar listening ability. However, only 29 pairs could be matched (see Appendix E) because some of the 74 students could not be paired and some missed the listening proficiency test. Totally, only 58 students (29 students in each group) with comparable listening ability as measured by the listening proficiency test were chosen to participate in this experimental study.

3.1.2.2 High and Low Groups

To investigate further into the next step, the researcher subdivided each group of twenty-nine students into high and low listening proficiency groups based on their scores in the listening proficiency test. The scores were arranged from the highest to the lowest. Eight subjects who got scores of 25-30 were considered as the high listening proficiency group and eight subjects who got 8-12 were assigned to the low listening proficiency group based on the 27% technique (Hughes, 1989). To restate, 8 subjects constitute a group. The distribution of the subjects is presented in Table 3.1.2 below.

Table 3.1.2

The Distribution of the Subjects

Group	No. of the Subjects	High (25-30 marks)	Low (8-12 marks)
A	29	8	8
B	29	8	8

Table 3.1.2 above shows the two main groups formed from the subjects' comparable listening ability of 29 subjects each. The high and low listening ability groups each consisted of 8 subjects regarded as *the good listeners and the poor listeners*, respectively.

Upon completing the grouping, both groups A and B were requested to undertake the same listening tests but they were asked to adopt different approaches. Group A was assigned to do the listening tests and complete the questionnaire using the pre-listening question technique while Group B was assigned to do the listening tests and complete the questionnaire using the post-listening question technique.

3.1.3 Subject Test

To ensure that the subjects in both Groups, A and B, had a similar level of listening ability, the means of the two groups were compared using a t-test, as shown in Table 3.1.3.

Table 3.1.3

The Means of the Subjects' Listening Abilities

Group	No. of the Students	Mean	S.D	T-value	Two-tailed
A	29	18.14	6.66	- .571	.573
B	29	18.21	6.68		

As shown by the data presented in Table 3.1.3, the mean scores of listening ability measured by the listening proficiency test stated above are 18.14 for Group A and 18.21 for Group B. The result of the t-test was .573 which did not indicate a statistically significant difference at a level of 0.05 ($t = -.571$, $p = .573$). Thus, it could be assumed that their average listening ability prior to the experiment was at approximately the same level.

Furthermore, a comparison of the difference between the mean scores of the high and low ability students within the two groups was not statistically significant at a level of 0.05 (for the high groups, $t = -.552$, $p = .598$; for the low groups, $t = -.552$, $p = .598$). The mean scores of the high listening ability groups were 26.75 and 26.88, and the mean scores of the low listening ability groups were 10.00 and 10.13 for Groups A and B respectively. At this point the researcher assumed that each high and low listening ability group was at approximately the same level of listening ability.

3.2 Research Instruments

To investigate English listening proficiency, listening comprehension levels under the use of the pre-listening question and the post-listening question techniques, and the attitudes towards the two techniques of the subjects, the researcher developed a set of research instruments to elicit this information. They were the three parts of *a listening proficiency test, four listening comprehension tests and the two questionnaires* surveying the subjects' attitudes.

3.2.1 Listening Proficiency Test

In order to place the subjects in comparable groups, high groups and low groups mentioned earlier, a listening proficiency test (see Appendix A) adopted from the TOEFL test (Test of English as a Foreign Language) was used as an instrument for this process. This test was originally designed to test English listening proficiency of non-native English speakers and was a standardized test used to determine proficiency level. Therefore, the listening proficiency test from the TOEFL test was a suitable instrument to measure the listening proficiency of the subjects in this study. In

addition, the TOEFL test listening section has been specially designed to test listening ability. Forty minutes are allowed for the subjects to complete the test.

This test is composed of 50 multiple-choice items divided into three main parts: restatement, short conversations and long conversations. In the restatement section, subjects hear 20 brief statements and they have to select the written restatement nearest in meaning to what they have heard. In the second part, 15 short conversations between two speakers are presented and a third voice asks a question about what was said in the conversation. Subjects choose the correct written answer to the spoken question based on each short spoken dialogue. In the last part, subjects answer 15 questions based on long spoken dialogues between two speakers or talks given by one speaker. They are asked several questions about each text and have to choose the correct written answer.

3.2.2 Listening Comprehension Test

Four English listening comprehension tests adapted from a commercial textbook *Intermediate Matters 1996* (Bell and Gower, 1996) were used as the instrument to measure the subjects' listening comprehension. This commercial book was found to provide a good source of material as it was carefully written, graded for each difficulty level, and based on suitable topics. The four texts which were used for both groups were taken from the original cassette accompanying the textbook. The length of the texts was between 375-530 words.

The transcript and the test items of the four texts are given in Appendix B. The four texts are *Reception Service*, *Ann Web interview*, *Wilar the crocodile man*, and *Sue's story* as described in Table 3.2.2.1.

Table 3.2.2.1

Text Types

No.	Text	Text Type			Length (words)	Time (minute)
		Reservation	Interview	Narrative		
1	Reception service	✓			530	03.15
2	<i>Ann Web</i> interview		✓		520	03.27
3	<i>Wilar</i> the crocodile man			✓	520	03.37
4	<i>Sue's</i> story			✓	375	02.27

Using a commercial textbook had many beneficial advantages. First, all the listening passages used in this study were carefully chosen so that they would well suit the English listening proficiency level of the subjects in this study because they were designed for students at an intermediate level of English proficiency. Second, all the four texts had the same level of difficulty in both content and vocabulary. However, they were different in content providing the subjects with an opportunity to listen to a variety of listening texts which in turn could motivate them to concentrate on the listening texts. Third, the words and phrases in each text were known to the subjects. Some unfamiliar words were pre-taught in the introduction stage. Thus, vocabulary was not a linguistic problem for the subjects. Fourth, each text was recorded on a cassette by a different native speaker in order to prevent the subjects from becoming familiar with any individual native speaker's voice and accent. Fifth, the four texts used in this study were all authentic spoken texts and it is generally accepted that learners need to listen to real spoken language with all the characteristics of natural language use in order to develop listening skills (Buck, 2001).

To reach the goal of testing listening comprehension of the subjects, forty short-answer questions were constructed. Some were taken from the original questions in the textbook and some were written by the researcher to tap two levels of local information and global information (see Appendix B). The local questions required the subjects to locate specific details or information from the text. The global questions required the subjects to synthesize or infer information from the text. These

two types of questions were used to assess subjects' listening comprehension when the subjects were tested by different methods of the pre-listening question technique and the post-listening question technique. Totally, there were 40 questions allocated to the 4 texts and each text had different numbers of the two question types as shown in Table 3.2.2.2 below.

Table 3.2.2.2

Number of Test Items at the Two Comprehension Levels

No.	Text	Type of listening question		No. of questions
		Local questions	Global questions	
1	Reception service	5	2	7
2	Ann Web interview	7	3	10
3	Wilar the crocodile man	7	4	11
4	Sue's story	8	4	12
Total		27	13	40

3.2.3 Attitude Questionnaire

The two attitude questionnaires relating to the pre-listening question technique and the post-listening question technique were used to elicit the subjects' attitude towards the use of the pre-listening question technique and the post-listening question technique. The questionnaire relating to the pre-listening question technique was constructed to be used on the subjects in Group A while Group B answered the questionnaire relating to the post-listening question technique. The content of the questionnaires was designed to elicit what subjects thought about the pre-listening question technique and the post-listening question technique and also to find out whether the subjects' attitude affects their test taking.

Each questionnaire (see Appendix C) was constructed based on the features which were reviewed and compiled by many scholars (Underwood, 1989; Buck, 1990; Boonyakarn and Syananondh, 1991; Weir, 1993; Shohamy, 1997, and Lingzhu, 2003). After that the questionnaires were revised based on comments and suggestions made

by lecturers in the Department of Language and Linguistics, Faculty of Liberal Art, Prince of Songkla University, Hat Yai Campus. All items in the two questionnaires were weighted following Likert's five point rating scales. The levels of agreement varied from 5 (strongly agree), 4 (agree), 3 neutral, 2 (disagree), to 1 (strongly disagree). There are 11 items in the questionnaire relating to the pre-listening question technique and 10 items in the questionnaire relating to the post-listening question technique. These two questionnaires were constructed in Thai for the subjects in order to avoid confusion and misinterpretation. However, English versions were also designed so that the interested readers other than Thai native speakers can understand the content of the questionnaires and make use of this study in the future.

3.2.4 Pilot Study

The process of pilot testing was done before conducting the main study. The main purpose of the pilot study was to test the validity and reliability of the four listening tests and the two questionnaires so that they could be improved and revised to be used in the main research study. In addition, the pilot study enabled the researcher to designate an appropriate testing time for the main study.

The pilot study was conducted with a group of 30 third-years English major students at Thaksin University, Songkla. Their learning backgrounds were similar to the subjects in the main study - third year English major students. They were asked to do the four listening tests and answer the two questionnaires. After carrying out the pilot test, the researcher conducted an item analysis of the four tests. *The difficulty of the test (p) and the discrimination power of the test (d)* were examined. According to Nuttal & Skurnik (1969), the level of difficulty should be between 0.20-0.80 and the discrimination power should be equal to or more than 0.20. Hence, the items which did not accord with the criteria above were deleted or revised. Forty items of an acceptable level of difficulty and discrimination power were retained (see Appendix D). Based on Kuder-Richardson Formula 20 (cited in Harris, 1969), the researcher calculated the reliability of the test which should be at least 0.70. The analysis result showed that the reliability of the test was 0.78 which was acceptable, as shown in Table 3.2.4 below.

Table 3.2.4

Pilot Study Results of the Tests

No. of Test Takers	No. of Test Items	Highest Score	Lowest Score	\bar{X}	S.D.	KR 20 (R ^{tt})
30	40	33	14	19.53	4.91	0.78

Following the pilot of the four tests, the two questionnaires were tried out with the students who had taken the tests. The Cronbach Alpha method was used to analyze the reliability of the questionnaires. The alpha of the pre-listening question and the post-listening question questionnaires were 0.79 and 0.86 respectively. The total alpha of both questionnaires was 0.82 which was acceptable. Based upon the alpha scores, the researcher was confident in using the two questionnaires with the main research subjects.

3.3 Testing Procedure

In order to collect the data in this study, two stages of testing procedure were set. The details of the two stages are described as follow:

3.3.1 Introduction Stage

During the introduction stage, the researcher first provided the subjects with appropriate background information. This included giving information about the topic of the text by stating its title. Next, the researcher explained the purpose and nature of the listening tasks so that subjects would know what they had to concentrate on when listening to the text and how to answer each question type. The researcher explained that the subjects would be required to answer two types of questions: local questions which required the subjects to locate specific details or information from the listening text, and global questions requiring subjects to synthesize or infer information from the listening text. Finally, the meanings of any unfamiliar words which the subjects might hear in the listening text were explained. The unfamiliar words from the all 4 texts are shown in Table 3.3.1 below.

Table 3.3.1

Lists of Unfamiliar Words

Text	Unfamiliar Word
1. Reception service	Single, twin, double, reserve, registration
2. Ann Web interview	Tarantulas, collection, passionate, crickets, locusts, cockroaches, awe, shuddering, terrified, aggressive, equivalent.
3. Wilar the crocodile man	Grab, scream, belly, mud, squeeze, chest, squat down, club.
4. Sue' s story	Rebuilding, mask, blindfolds, remote, kidnap, extremist, autonomy, release, nightmare, and security-conscious.

3.3.2 Testing Stage

Following the introduction stage, the researcher gave instructions to each group. Subjects in Group A, using the pre-listening question technique, were given a handout containing the questions prepared specifically for this research. They had 5 minutes to study those questions before returning the handout to the researcher. Next, the researcher had the subjects listen to the text twice and they were allowed to take notes while listening. After that, the researcher gave the question handout back to the subjects and they were allowed 15 minutes to write down their answers on the answer sheet.

Subjects in Group B, using the post-listening question technique, on the other hand, listened to the text twice and were allowed to take notes while listening, without seeing the questions before listening. The researcher gave the question handout to them only after they had listened to the text and they then had 15 minutes to answer the questions on the answer sheet. The administration procedure of the pre-listening question technique and the post-listening question technique is presented in Table 3.3.2.1 and 3.3.2.2 respectively.

Table 3.3.2.1

Test Procedure of the Pre-Listening Question Technique

No.	Group A (N=29) Pre-listening question technique	Time (minute)
1	The subjects read the questions and give the question sheets back to the researcher.	5
2	The subjects listen to the listening texts twice and take notes while listening.	10
3	The subjects answer the given questions.	15

Table 3.3.2.2

Test Procedure of the Post-Listening Question Technique

No.	Group B (N=29) Post-listening question technique	Time (minute)
1	The subjects listen to the listening texts twice and take notes while listening.	10
2	The subjects read the questions.	5
3	The subjects answer the given questions.	15

Since the duration of each test was 30 minutes, each group took two tests at a time which made a total of one hour test session. In the first week, each group took the first two tests (*Reception service* and *Ann Web interview*) and in the second week they took the last two tests (*Wilar the crocodile man* and *Sue's story*). In addition, since the two subject groups took the tests at different times, the subjects in Group A were kept in the class for 20 minutes after they finished the tests to prevent them from discussing the test with Group B subjects.

Immediately after the subjects finished the 4 tests, the attitude questionnaires relating to the pre-listening question technique and the post-listening question technique were administered. The subjects in Group A answered the attitude questionnaire relate to the pre-listening question technique while the subjects of Group

B answered the attitude questionnaire relate to the post-listening question technique. The time taken to complete the questionnaires was 10 minutes.

3.4 Data Analysis and Statistical Techniques

The data in this study consisted of the scores from the four listening tests and the information on attitudes elicited through the use of the two questionnaires. The scores were used to interpret the effects of the pre-listening question technique and the post-listening question technique on the subjects' ability to understand listening passages and the means scores obtained from the questionnaires were used for interpreting the level of agreement. Independent and dependent variables as described below were used in the statistical analysis.

Independent variables

The independent variables were the two listening testing techniques employing the pre-listening question technique and the post-listening question technique.

Dependent variables

There were two main dependent variables: the scores on the tests of Group A and B, and the means scores on a rating scale obtained from each item in the questionnaires on the pre-listening question technique and the post-listening question technique.

The data were analyzed using the following methods:

1. To answer the first research question, concerning how much the pre-listening question and the post-listening question techniques affect the listening ability of all the participating subjects in responding to questions in listening comprehension tests, the tests were first marked following which the total scores were calculated. The mean scores and standard deviations (SD) of both Group A and B were then calculated by using the SPSS/PC program and a t-test was run to determine the differences between the mean scores.
2. To answer the second and the third research questions, concerning how much the pre-listening question and the post-listening question techniques affect the listening ability of the high and low listening proficiency groups, the total scores of the high and low groups were calculated. The mean scores and standard

deviations (SD) of each group were then calculated by using the SPSS/PC program and a t-test was run to determine the differences between the mean scores.

3. To answer the fourth research question, concerning the degree the pre-listening question and the post-listening question techniques affect subjects' listening ability in responding to local questions and global questions, the total scores of each type of question were calculated. The mean scores and standard deviations (SD) of each group were then calculated by using the SPSS/PC program and a t-test was used to determine the differences between the mean scores.
4. To answer the fifth research question, exploring the subjects' attitude towards the pre-listening question and the post-listening question techniques, the data derived from the five-point rating scale were analyzed to find the mean scores of the subjects' responses on each item using the SPSS/PC program. The ranges of the mean scores of each level were used to interpret the level of agreement. The criteria for the rating scale interpretation (Som-in, 1988:34) are shown in Table 3.4 below.

Table 3.4

Criteria of Rating Scale Interpretation

Ranges of the Total Mean Value (\bar{X})	Interpretation of Preference	Level of Agreement
4.21-5.00	Most	Strongly agree
3.41-4.20	Much	Agree
2.61-3.40	Moderate	Uncertain
1.81-2.60	Slightly	Disagree
1.00-1.80	Least	Strongly disagree