

CHAPTER III

RESEARCH METHODOLOGY

This research utilized a mixed-method approach — a combination of a positivistic, product-oriented approach and a naturalistic, process-oriented one. Even though this study is rather qualitative in nature, the data gathering procedure was done by using qualitative and quantitative measures. A case study was carried out to investigate naturalistic qualitative data. A one-group pretest-posttest design was the measure of collecting quantitative data. The data —both ‘process’ and ‘product’— were then analyzed descriptively and statistically.

3.1 Population and Samples

The population of this study were first-year students at King Mongkut’s University of Technology Thonburi. Since the setting of this research is an autonomous public university, like other public universities in Thailand, the institutional policy did not allow researchers to use simple random sampling. According to Hatch and Lazaraton (1991: 85), neither random sampling nor random assignment is possible in classroom research as it involves the use of classes where students have already been assigned on the basis of some principles such as test scores, timetables or by self-selection (the students decide which course to take). In this research, it is impossible to select students randomly to begin with. The researcher, therefore, decide to work with an established class of students, an intact group. Choosing intact groups from the whole population of first-year students at KMUTT is not feasible because there are different types of undergraduate programs: regular programs, bilingual programs and international programs provided by four faculties: the Faculty of Engineering, the Faculty of Science, the Faculty of Industrial Education, and the Faculty of Architecture. The Faculty of Engineering provides regular and bilingual programs. The Faculties of Science and of Industrial Education offer only regular programs. The Faculty of Architecture provides only international programs. Therefore, to minimize threats to validity and reliability of the intact design, the researcher chose multi-stage sampling which cluster sampling and stratified random sampling are combined. Cluster sampling involves a multistage sampling in which sampling large clusters is primarily done, and smaller clusters are then sampled. A sample of participants is, finally, selected (Leary, 2004: 122). Stratified random sampling (also known as proportional or quota random) involves

dividing the population into homogeneous subgroups and then taking a simple random sample from each subgroup. If the strata or groups are homogenous, it is certain that the variability within a group is lower than the variability for the population as a whole (<http://www.Socialresearchmethods.net/kb/sampprob.htm>).

In this study, cluster sampling was done primarily to select the desired sample of the regular program students. Architecture students were excluded as they are international students. Their English proficiency might be higher than students in other programs. The next stage was separating regular program students from bilingual students. In semester 2 of the academic 2006, there were 46 classes available for students taking the LNG 102 course to choose according to their timetables organized by the KMUTT Registration Office. These 46 classes were divided into 4 clusters. The first cluster (10 sessions) consisted of industrial education students majoring in Mechanical Engineering, Civil Engineering, Industrial Engineering and Electrical Engineering. The second cluster (15 sessions) included science students together with industrial education students majoring in Printing Technology and Multimedia. The third cluster (17 sessions) was composed of engineering students from every department. The fourth cluster (4 sessions) was engineering students in the bilingual programs which were finally excluded from the sample. The industrial education students, who are regarded as low proficiency students, were also discarded to prevent any bias in terms of the students' English proficiency. Therefore, the total population of this study consisting of the second and the third clusters was 1179 students. The optimal sample size is around 285 (Krejcie and Morgan, 1970 cited in Isaac and Michael, 1983: 193). To reach the samples that can represent the whole population, six groups – three from cluster two and another three from cluster three were selected for quantitative measure. Initially, each subgroup was supposed to contain around 40-45 students, but after the first two weeks had passed, it was found that there were only 202 students enrolling in the selected sections. However, a number of students also withdrew from the course; as a result, 189 students were retained in the study. For naturalistic inquiry, stratified random sampling was applied. Subjects for interviews included 18 students who were randomly selected from cluster two and cluster three. Also, 10 students from cluster two, and 10 from cluster three were chosen as cases for a case study using portfolio assessment as a measure.

The subjects in this study were all KMUTT students who took the English foundation course LNG 102 in the second semester of the academic year 2006. The demographic characteristics of the subjects were as follows:

- A. Age: 17-22 years of age
- B. Gender: male and female
- C. Education: undergraduate level
- D. Language proficiency: pre-intermediate
- E. Others : students who can pass the foundation English course I (LNG 101), and take a foundation English course II (LNG 102) in the second semester of the academic year 2006.

According to the task-based EFL curriculum at KMUTT, all of the students have to take at least three English courses and further optional courses, depending on the requirements of their faculties. An English placement test is used to place students into two groups based on proficiency. Those who can pass the English Placement Test are regarded as *more proficient students* whereas those who could not pass are regarded as *less proficient* (see figure 3.1 for more details).

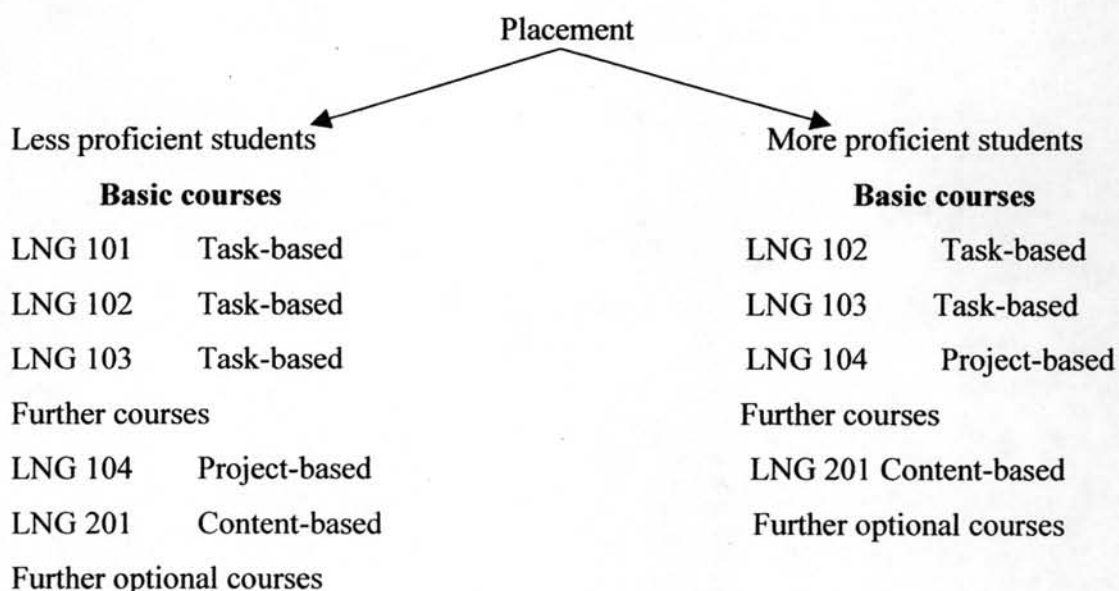


Figure 3.1: Overview of course type in the English curriculum at KMUTT

(Source: Watson Todd, R. ,2001: 5).

The samples of this study were those *less proficient students* who had to take three basic courses starting from LNG 101-LNG 103. They are regarded by most of the English teachers as the 'weak' students who are not ready for task-based instruction (Watson Todd, 2006). However, in piloting task-based language teaching in the initial stage of curriculum renewal at KMUTT, Srimavin and Watson Todd (2000) reported that tasks can lead to beneficial learning for low-level students in a Thai situation. The reasons for choosing the LNG 102, a second course in a series, as an evaluated course are:

(1) Having been modified several times, the current LNG 101 course is now a remedial course for the less proficient students to prepare them for the more task-based courses, LNG 102 and LNG 103. Its content emphasizes grammar and learning strategies rather than tasks.

(2) LNG 102 is the most appropriate course to be studied. It places a strong emphasis on developing the cognitive skills of the students which are dealt with through a series of English medium tasks that include dictionary tasks, resourcing tasks and portfolio tasks. The goals of this course are to provide strategy training and to enhance autonomous learning skills. Students are also encouraged to learn independently using a Self-Access Learning Center as well as the Internet and library resources (see Appendix A). Therefore, the task-related materials used in this study consist of teaching materials (e.g. course handouts, course-related document); learning materials -- authentic materials-- such as newspaper, journals, magazines, e-books, e-magazines, etc.) as well as self-access learning materials available at KMUTT's Self-Access Learning Center.

(3) In the first semester of the academic year 2006, all of the first-year students were required to enroll in the so called "12-day intensive courses": LNG 101 or LNG 102 that were newly designed during the summer and used in June 2006. The course content and methodology were changed dramatically. Also, the two courses were taught within a fairly short time frame. Due to such constraints, the researcher decided not to evaluate those courses.

Other participants of this study included: 3 English language teachers; 3 subject teachers from three different faculties: the Faculties of Engineering, of Science and of Industrial Education, 5 ex-students and 2 audiences-- the faculty dean and the department head of the School of Liberal Arts, King Mongkut's University of Technology Thonburi.

3.2 Research Instruments

As this research was a mixed-method evaluation research, the research methodology was based on both quantitative and qualitative methods. The research instruments used in this study consisted of:

1. Curriculum-based achievement tests
2. Student portfolios
3. Self-assessment checklist
4. Semi-structured interviews, and
5. Classroom observations, and
6. Materials evaluation.

3.3 Research Instrument Construction and Validation

1. Curriculum-based achievement tests

A curriculum-based achievement test was developed by the researcher of this study to determine the degree of the students' attainment of the course objectives (see Appendix B). There were two equivalent forms of the test: Form A and Form B. Form A was administered at the beginning of the course (a pretest). Form B was administered at the end of the course (a posttest). The content of the test was consistent with the course objectives identified by the researcher of this study (see Appendix C). The test was composed of 3 main parts: Dictionary, Getting main ideas and related details, and Note-taking and summary writing. It included a variety of test types e.g., gap-filling, short response, multiple choice, and matching (see test manual in Appendix D for more details).

With regard to the quality of the test, three experts were asked to validate the test using the Index of Item-Objective Congruence. It was found that its content validity index was 0.85. The test was also field-tested by selecting 40 students who have similar characteristics with the subjects of the main study to participate in the pilot study. Additionally, the Cornbach alpha was applied to measure internal consistency of the test. The reliability of the overall test was 0.82. The test was then applied in the main study.

2. Student portfolios

A portfolio task is an adjunct of the LNG 102 course. All students are required to do the portfolio task to learn how to self-direct their learning. Twenty students were randomly selected to be cases for a case study. Individual students

handed in four pieces of portfolio: 2 product portfolios and 2 process portfolios. The scoring was done by three raters using two types of rubrics as scoring criteria for assessing students' learning product (see Appendix E) and process (see Appendix F).

The content validity of this research tool consisting of rubrics for assessing product and process portfolios was approved by three experts using the Index of Item-Objective Congruence. The index of the product portfolios was 0.78, and that of the process portfolios was 0.77.

3. Self-assessment checklist

Students' learner autonomy in language learning was assessed by using a self-assessment checklist adapted from Barnett's Attitudes questionnaire for self-access in Wenden (1991). The self-assessment checklist assessed student autonomy by looking at two main components of autonomy: *ability* and *willingness*. Ability included *knowledge* about alternatives from which choices have to be made (items 7, 8, 11, 12, 13, 14, 15), and *skills* for carrying out whatever choices seem most appropriate (items 6, 9, 10, 27). Willingness included *motivation* (item 30), and the *confidence* to take responsibility for the choices required (items 1, 2, 3, 4, 5, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29). The content validity of the self-assessment checklist was examined by three experts using the Index of Item-Objective Congruence, and it was found that the index was 0.76. About 40 students who had similar characteristics with the subjects of the main study participated in the pilot study. To measure its internal consistency, the Cronbach alpha was applied. The reliability coefficient of the self-assessment checklists was 0.79. Before using it in the main study, the self-assessment checklist was translated into Thai (Appendix G).

4. Semi-structured interviews

Different groups of participants were interviewed for different purposes. Teachers were asked about the rationale, scope, and activities of the program (Appendix H). Audiences were inquired about the potential purposes of the study (Appendix I). Subject teachers were asked to identify students' target or real-world needs (Appendix J). Students were asked to report their perceptions of the course as well as their own performance (Appendix K). The four sets of questions were translated into Thai and its content validity was validated by experts using the Index of Item-Objective Congruence. It was found that its index was 0.77.

5. Classroom observations

During the period of the main study, the researcher observed what was going

on in the natural classroom setting, particularly teacher performance. A classroom observation checklist was designed using Willis (1996)'s model for task-based learning framework as a guideline (Appendix L). The researcher observed three classes of three teachers, and three lessons of each class were video-taped: (1) the first lesson when the task was introduced, (2) the during-task lesson, and (3) the final lesson when students presented the task outcome. Each lesson lasted between 20 minutes and 2 hours depending on its purpose. The content validity of the checklist was examined by the experts using the Index of Item-Objective Congruence, and its index was 0.76.

6. Materials evaluation

Course materials were reviewed, and materials evaluation was carried out to find out whether each set of materials is relevant to the course objectives. The objective grid was designed for material evaluation. Its content validity was examined by the experts using the Index of Item-Objective Congruence. The index was 0.75.

3.4 Criteria for the Selection of Experts

The experts who were invited to validate the research instruments of this study had the following characteristics (see the expert profiles in Appendix M):

1. They had some experience in task-based instruction, curriculum development, self-access learning, and/or learner autonomy.
2. They had taught the LNG 102 course, the evaluated course, for at least one semester.
3. They had been involved in any academic research in applied linguistics concerning task-based instruction, curriculum development, self-access learning, and/or learner autonomy, in particular.

3.5 Data Collection

The data collection consisted of two main stages: the pilot study and the main study.

1. The pilot study

- 1.1 Discussions with stakeholders and audiences were done to confirm or disconfirm the issues or problems specified by the researcher.
- 1.2 All research tools were selected and developed.

1.3 All research tools underwent a validating process before being used in the main study.

2. The main study

The data collection was done throughout the 15-week course. The procedures were as follows:

- 2.1 The achievement test (Form A) was administered to the subjects as a pretest in week 1.
- 2.2 Classroom observations, materials evaluation and portfolio assessment were carried out during weeks 2-14.
- 2.3 Semi-structured interviews were conducted, recorded, and analyzed in week 14.
- 2.4 The subjects did the post-test (Form B) in week 15.
- 2.5 The subjects were asked to fill out the self-assessment checklist to find out how they perceived self-directed learning and learner autonomy in week 15.
- 2.6 The scores on the pre-test and the post-test were compared to investigate any differences between the scores obtained from the two tests.
- 2.7 The scores on the self-assessment checklists were also compared, and correlated with the post-test scores to examine correlation between the subjects' gains and their perceptions toward self-directed learning and learner autonomy.

3.6 Data Analysis

In this design, the data obtained were analyzed qualitatively and quantitatively. The analysis were divided into 3 phases:

Phase 1: Analysis of Context

Phase 1 was concerned with a qualitative analysis of the first feature: context. To answer research question 1, research questions 1.1 and 1.2 were involved.

Research question 1: What is the context within which the curriculum is working?

Research question 1.1: Do the goals and objectives of the course meet the needs of the stakeholders?

The data obtained from interviews with stakeholders were conceptualized, coded, quantified, tallied and presented in the form of a frequency distribution. The Chi-square test was applied to find out whether or not there was a significance difference in the responses of the stakeholders.

Research question 1.2 Are the goals and objectives appropriate for the specified groups of students?

The data obtained from interviews with teachers and stakeholders were conceptualized, coded, quantified, tallied and presented in the form of a frequency distribution. The Chi-square test was applied to find out whether or not there was a significance difference in their responses.

Phase 2: Analysis of Implementation

The second phase involved qualitative analysis of the second feature, implementation. To answer research question 2, research questions 2.1-2.5 were included.

Research question 2: How well is the curriculum implemented?

Research question 2.1: Is the teaching method relevant to the prespecified objectives?

The data obtained from classroom observation were analyzed and presented in the form of a frequency distribution. The criteria were as follows:

1 = Observed

0 = Not observed

NA = Not applicable

The Chi-square test was applied to find out whether or not there is a significant difference in the observed teachers' uses of teaching techniques in relation to task-based teaching methodology ($p = 0.05$). The following criteria were utilized to determine how much task-based teaching methodology was used in the observed classrooms:

0.67-1.00 = a lot

0.34-0.66 = moderate

0-0.33 = little

Research question 2.2: Are the teacher skillful in task-based instruction?

The data obtained from classroom observation were analyzed and presented in the form of a frequency distribution. The Chi-square test was applied to find out whether or not there was a significance difference in the observed teachers' skill in teaching a task-based English course ($p = 0.05$). To determine whether the observed teachers were skillful in task-based instruction, the following criteria were used:

56-82 = skillful

28-55 = moderately skillful

0-27 = not skillful

Research question 2.3: Are the tasks consistent with the course objectives?

The data obtained from interviews with teachers were analyzed using content analysis. They were conceptualized, coded, quantified, tallied and presented in the form of a frequency distribution.

Research question 2.4: Are the teaching materials relevant to the prespecified objectives?

The data obtained from interviews with teachers were conceptualized, coded, quantified, tallied and presented in the form of a frequency distribution. The following criteria were used to determine the quality of teaching materials of each task:

More than 75% = perfectly matched

51%-75% = matched but not need to be supplemented

26%-50% = need to be revised

0%-25% = mismatched

Research question 2.5: Are the resources adequate?

The data obtained from interviews with teachers and students were conceptualized, coded, quantified, tallied and presented in the form of a frequency distribution.

Research question 2.6: Are the assessment procedures appropriate to the prespecified objectives?

The data obtained from interviews with teachers were conceptualized, coded, quantified, tallied and presented in the form of a frequency distribution. The Chi-square test was applied to find out whether or not there was a significant difference in the responses of the stakeholders.

Phase 3: Analysis of Student Outcomes

In Phase 3, both quantitative and qualitative analyses of the third feature, student outcomes, were carried out to examine 'product' and 'process' of their learning. Research questions 3.1-3.3 replied to research question 3.

Research question 3: Are student outcomes due to the effects of what is happening within the curriculum (context and implementation)?

Research question 3.1: Do students achieve the prespecified objectives?

Research question 3.1.1: Do the students make significant gains in their language abilities after taking this course?

In the pilot study, the Cronbach alpha was used to investigate the reliability of the test. The reliability coefficient was 0.82. *Item analysis* was also conducted to find out difficulty indexes and discrimination indexes of the test (Appendix N). The criteria for measuring these indices were as follows (Sukamolson, 1995:31):

For the difficulty index (p):

$p < 0.20$	means the item was difficult.
$p = 0.20-0.80$	means the item was good in terms of its difficulty.
$p = 0.81-0.94$	means the item was easy.
$p \geq 0.95$	means the item was very easy.

For the discrimination index (r):

$r = 0$	means the item had no discrimination ability.
$r \leq 0.19$	means the items had a low discrimination ability.
$r = 0.20-0.29$	means the item had a fair discrimination ability.
$r = 0.30-0.39$	means the item had a high discrimination ability.
$r \geq 0.40$	means the item had a very high discrimination ability.

Based on the criteria set, the test items of which difficulty indices ranged between 0.20 and 0.80, and of which discrimination indices were equal or higher than 0.20 were chosen for the main study.

Descriptive statistics were used to measure central tendency and dispersion of the test scores on the curriculum-based achievement tests. Students' scores of the 6 intact groups from pre- and post- tests were analysed using One-Way ANOVA ($p = 0.05$) to test whether or not there is a significant difference of their mean scores before and after course implementation. As the resulting F values are significant, LSD, a type of post hoc analysis that can isolate which groups were significantly different, was, then, applied because the six intact groups were unequal in size but had similar variance (note: the LSD test is more sensitive than the Scheffe test). (http://staff.harrisonburg.k12.va.us/~gcorder/test_post_hocs.html).

Research question 3.1.2: Do the students develop their autonomy in language learning during and after taking this course?

Students' learner autonomy was investigated using the self-assessment checklist and the portfolio task. In the pilot study, the Cronbach alpha was applied to measure the internal consistency of the self-assessment checklist. The reliability coefficient was 0.79. The following criteria were utilized to determine the students' attitudes towards learner autonomy:

2.51 –3.00 = Yes

1.51-2.50 = Undecided

1-1.50 = No

To complete the portfolio task, individual students needed to do 4 pieces of written assignment. 20 subjects from the 2 main groups: Faculty of Science, Faculty of Industrial Education, and Faculty of Engineering were randomly selected, and their work (80 pieces) was rated by 3 raters. The following are the criteria for interpreting scores on portfolio tasks:

21-30	High
11-20	Average
0-10	Low

To investigate the interrater reliability of the rubrics for the portfolio task, the Pearson Product Moment correlation was used.

Research question 3.2: Do the learners think the course is appropriate?

Frequency distribution was used to illustrate the nature of the subjects. The Chi-square test was applied to find out whether or not there is a significant difference in the stakeholders' opinions.

Research question 3.3: Is the student learning the result of instruction or extraneous factors?

The data obtained from interviews with teachers and students were recorded, coded, quantified, tallied and presented in the form of a frequency distribution. The Chi-square test was applied to find out whether or not there is a significant difference in the proportions of the responses.

3.7 Validity and Reliability of the Research Instruments

The Pearson Product Moment correlation coefficients were used to examine the relationships between three pairs of scores: scores on the self-assessment checklist were compared with scores from achievement tests (posttests) to examine the relationship of students' language abilities and their perceptions as autonomous learners. The scores on the ratings of product portfolios were correlated with scores on the curriculum-based achievement tests to measure convergent or concurrent validity of the two measures and the model itself. Scores on process portfolios were also correlated with those on self-assessment checklists to examine the students' learning process.

To assure the reliability of the rubrics for the portfolio tasks, generalizability coefficient was applied to investigate the sources of error in the measure. The generalizability coefficient of the analytic scale for assessing product portfolios was 0.89, and that of the holistic scale for assessing process portfolios was 0.73.

3.8 Statistical Tools

For quantitative data, different statistical analyses were utilized with a particular instrument.

1. SPSS/PC Version 14 was utilized to compute frequencies, means, mode, median, SD, t-test, One-Way ANOVA, Chi-square test, reliability coefficient, Pearson product-moment correlation coefficient, post-hoc analysis and item analysis.

2. GENOVA, a computer program for Generalizability Theory, designed by Robert Brennan, Director of Center for Advanced Studies in Measurement and Assessment at the University of Iowa, USA was employed to calculate generalizability coefficient (<http://www.education.uiowa.edu/casma/GenovaPrograms.htm>).

The following table illustrates the detailed information of evaluation dimensions, sources of data, types of instruments, data analysis, and criteria for judging the success of the course (see Table 3.1).

Table 3.1: Evaluative dimensions

Dimensions	Evaluative questions	Sources of information	Instruments	Data analysis	Criteria for making judgements
Context	1. Do the goals and objectives of the course meet the needs of the stakeholders?	Stakeholders	Interviews	Frequencies Content analysis	More than 50% of the interviewees agree with this statement.
	2. Are the goals and objectives appropriate for the specified groups of students?	Teachers, and stakeholders	Interviews	Frequencies Content analysis	More than 50% of the interviewees agree with this statement.
Implementation	1. Is the teaching method relevant to the prespecified objectives?	Teachers	Classroom observation	Frequencies	$\chi^2_{0.05} < 14.07$
	2. Are the teachers skillful in task-based instruction?	Teachers	Classroom observation	Frequencies	More than 50% of the teaching techniques used are consistent with task-based teaching methodology.
	3. Are the tasks consistent with the course objectives?	Teachers	Interviews	Frequencies Content analysis	$\chi^2_{0.05} < 5.99$
	4. Are the teaching materials relevant to the prespecified objectives?	Course materials	Material evaluation	Frequencies	More than 50% of each set of materials is consistent with the course objectives.
	5. Are the resources adequate?	Teachers Students	Interviews	Frequencies Content analysis	More than 50% of the interviewees agree with this statement.

Dimensions	Evaluative questions	Sources of information	Instruments	Data analysis	Criteria for making judgements
	6. Are the assessment procedures appropriate to the prespecified objectives?	Teachers	Interviews	Frequencies Content analysis	More than 50% of the interviewees agree with this Statement.
Student outcomes	1. Do the students make significant gains in their language abilities after taking this course?	Students	Objectives-based achievement test	Paired-sample T test	$t_{0.05} < 1.96$
	2. Do the students develop their autonomy in language learning during and after taking this course?	Current students & Ex-students	Self-assessment checklist Portfolio assessment Interviews	Frequencies Descriptive statistics G-coefficient Frequencies Content analysis	$\chi^2_{0.05} < 43.77$ Mean scores on each type of portfolios should be greater than 30. ($r > 0.75$) $\chi^2_{0.05} < 7.81$
	3. Do the students think the course is appropriate?	Students	Interviews	Frequencies Content analysis	$\chi^2_{0.05} < 3.84$
	4. Is the student learning the result of instruction or extraneous factors?	Students	Interviews	Descriptive Content analysis	More than 50% of the interviewees perceived that their learning outcome is the result of in-class instruction.

3.9 Stages of Research

The research procedure of this study was conducted according to the following time line.

Table 3.2: Stages of research procedure

Time	Procedure
April 2006	Interview stakeholders and audiences, overview program scope and activities, to identify purposes and issues Identify data needs and select instruments.
May- June 2006	Develop instruments
July 2006	Validate the instruments by experts Revise the instruments Conduct pilot study
August 2006	Revise the instruments for the 2 nd time. Apply analytical procedures.
September 2006	Review course-related documents. Conduct task analysis and material evaluation. Analyze and interpret data.
October 2006	Prepare informal reports on the first part of the study and present them to the stakeholders and audience to react to the data.
November 2006- February 2007	Main study
March- August 2007	Analyze and interpret data. Prepare informal reports on the second part of the study and present them to the stakeholders and audiences for their reactions to the findings.
September 2007	Assemble reports and prepare a final report Present the final report to the audiences

Chapter Summary

This chapter concerns the research methodology of the present study. The population were 1179 undergraduate students at King Mongkut's University of Technology Thonburi who took the LNG 102 courses, the evaluated course, in semester 2 of the academic year 2006. Samples were six intact groups selected by multi-stage sampling to assure that they could represent the whole population. Each subgroup was supposed to contain around 40-45 students, but, indeed, only 189 were students retained in the study due to some unforeseeable problems: (1) Numbers of student registering in each session were unequal; (2) Students from different faculties had to take the same classes; and (3) A large number of students dropped the course before the semester ended. Therefore, the subjects of this study were 189 students classified into two main clusters: engineering students and science students combined with industrial education students. The two clusters were subdivided into 6 groups (three from each) which were the subjects for quantitative inquiry. For naturalistic inquiry, stratified random sampling was applied. Also, 18 students randomly selected from each cluster were assigned as subjects for interviews. In addition, 20 students (10 from each cluster) were chosen as cases for a case study using portfolio assessment as a measure.

There were 6 research instruments used in this study: (1) curriculum-based achievement tests, (2) student portfolios, (3) self-assessment checklist, (4) semi-structured interviews, (5) classroom observations, and (6) materials evaluation. A number of statistical analyses were utilized: descriptive statistics, t-test, one-way ANOVA, Chi-square test, Cronbach alpha, Pearson product-moment correlation, test item analysis, and GENOVA.