CHAPTER III
METHODOLOGY

This research was designed to investigate the effects of synchronous and asynchronous Web-based learning in task-based instruction on English language learning achievement of undergraduate students at Kasetsart University. The design was based on a repeated measures factorial design in which two independent variables (IVs) were studied to determine the main effects and interactive effects on the achievement (DV). Two sets of instruments were used. One was for the treatments and the others were for data collection. The main effects and interaction effects were used to test the hypotheses set for this study. This chapter starts with the population, subjects of the study, research instruments, instrument validation, treatment procedures, data collection, and data analyses.

Population

The population of this study was undergraduate students at Kasetsart University during the academic year 2005.There were 26,867 undergraduate students consisting of 14,965 students at Bangkhen Campus, 3,108 students at Sakon Nakhon Campus, 3,908 students at Kamphaeng Saen Campus, 739 students at Suphan Buri Campus, 4,048 students at Si Racha Campus, and 99 students at Krabi Campus. These students were required to take at least 9 credits in English from the English language syllabus to fulfill their study. One of the courses required was Foundation English III (FE3) course of which the approximate number of students taking this course was 2,000.

Subjects

One English course (355254 'English for Ground and In-flight Attendants') was selected from the English language syllabus by purposive sampling. All the students registered in the course in the 355254 'English for Ground and In-flight Attendants' in the first academic semester of 2005 at Kasetsart University: Bangkhen and Kamphaeng Saen Campus, were used in this study. The reason was that all of the 102 subjects represented students both from the main campus in Bangkok and the up-country campus. There were undergraduate students who came from various

fields of study i.e. Engineering, Business Administration, Economics, Social Science, Humanities, Agro Industry, Education, Science, and Liberal Arts and Science. They consisted of both males and females with various years of study.

Next, the subjects from both campuses were assigned to two experimental groups according to their English language ability which derived from the grades from Foundation English III (FE3) course. The grades from FE3 achievement test were managed by the English Foundation course committee, the Department of Foreign Languages, Faculty of Humanities, Kasetsart University. One of the characteristics of FE3 test was that the scores were distributed in the shape of 'normal' which contained certain statistical characteristics that were known and constant (Bachman, 1990:72-3). Also FE3 test shared some characteristics of standardized test in that it was based on a fixed content. The administration of the test did not vary from one administration of the test to the next. Therefore, the grades derived from FE3 were constant and did not vary. Specifically, the measurement scale was provided. The score distribution norms had been established with groups of individuals and if there were alternate forms of the test, they were equated statistically to assure that reported scores on each test indicated the same level of ability, regardless of the particular form of the test being used.

Then, a matched pair technique was used to divide students into two experimental groups by their grades A (=4.00), B+ (3.50), B (=3.00), C+ (2.50), C (2.00), D+ (1.50) or D (=1.00) from FE3 course. The number of subjects was 102. There were 51 in each treatment group equally. Next, an independent samples t-test was conducted to test the significant difference of their English language ability. Finally, they were randomly assigned to two treatment groups: Convergent Group (convergent tasks) and Divergent Group (divergent tasks).

Means and standard deviations of English language ability of the two treatment groups were shown in Table 1.

Table 1. Mean and Standard Deviation of English Language Ability of Group One and Group Two

	Minimum	Maximum	Mean	Std. Deviation
GROUP1	1.00	4.00	2.8627	.73538
GROUP2	1.00	4.00	2.7745	.84447

n = 51 for each group

Table 1 shows the mean scores from previous English course of the two experimental groups. The mean scores of Group One = 2.8627 and Group Two = 2.7745. It is seen that the mean scores of Group One are slightly higher than Group Two.

To test the significant difference between the mean scores of the two experimental groups, independent samples t-test was conducted at the 95 percent of confidence levels as can be seen in Table 2.

Table 2. t-test of English Language Ability of the Two Experimental Groups

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
									Lower Upper
PREVIOUS SCORES	Equal variances assumed	2.333	.143	.575	19	.572	.1618	.28128	-.42696 .75049

$p < 0.05$ n = 51 for each group

Table 2 shows the t-test value of mean scores of the two groups $t = .575$, $p = .143$ and yields no significant difference at the 0.05 level. This reveals that the English language ability of the two experimental groups is not different.

Next, one group was randomly assigned for convergent tasks and the other was assigned for divergent tasks. The first treatment was given in synchronous learning environment to both groups whereas the second treatment was given in asynchronous learning environment.

Table 3. Frequency Distributions of Fields of Study of the Subjects in Convergent and Divergent Group

Fields of study		Convergent Frequency	Group Percent	Divergent Frequency	Group Percent
Valid	Economic	0	0	1	.6
	Agriculture	2	1.3	0	0
	Humanities	15	9.6	16	10.2
	Engineering	6	3.8	7	4.5
	Agro-Industry	1	.6	1	.6
	Education	10	6.4	7	4.5
	Business Administration	7	4.5	9	5.7
	Science	2	1.3	1	.6
	Social Science	2	1.3	3	1.9
	Liberal Art	6	3.8	6	3.8
	Total	51	32.5	51	32.5
Missing	System	106	67.5	106	67.5
Total		157	100.0	157	100

n = 51 for each group

Table 3 shows the frequencies of fields of study of the subjects in Convergent and Divergent Group. The subjects from both groups were from various fields of study. The number of subjects from each field of study was slightly different. The highest percentage of subjects was from the Faculty of Humanities, there were 9.6 percent in Convergent Group and 10.2 percent in Divergent Group. No subject was from the Faculty of Economic in Convergent Group while there was no subject from the Faculty of Agriculture in Divergent Group.

Research Instruments

Two sets of instruments were constructed, one was used for the experiment and the others were used for data collection. The instrument for the experiment was a Web-based courseware while the instruments for data collection consisted of two language achievement tests, and a questionnaire asking for student opinions concerning Web-based learning environments at the end of the treatments.

1. Instrument for the experiment

The Web-based courseware was constructed in two forms. The first one used convergent tasks (making conclusion) and the other used divergent tasks (discussing). They were given in two Web learning environments: synchronous (SL) and asynchronous (ASL). The design was as follows:

The first form of the courseware was constructed using convergent tasks. It consisted of four learning modules. The first two modules were given in synchronous environment and the others were given in asynchronous environment.

The second form of the courseware was constructed using divergent tasks. It consisted of four learning modules. Similar to the first form, the first two modules were given in synchronous environment and the others were given in asynchronous environment.

Each learning module lasted six hours. SL and ASL were differentiated by time and types of communication technologies. The contents in this Web-based courseware consisted of four topics designed for four learning modules. They were: 1) flight history, 2) flight crew, 3) ground services, and 4) in-flight services. The main task outcomes at the end of each learning module focused on performing two different types of tasks: convergent and divergent tasks. The four modules and the tasks were designed under the following framework.

1.1 Characteristics of a Web-based courseware

This Web-based courseware is a hyper-media-based instruction which utilizes the attributes and resources of the World Wide Web to create a meaningful learning environment where learning is fostered and supported. The design involved multi-media i.e. text, graphics, video, and audio which were integrated into a single delivery system under computer control (CMC). Several software programs were used for the development such as Macromedia Dreamweaver, Flash, Photoshop, Movie Maker, Wavelab etc. Technologies enhanced TBI by providing collaborative functional learning environments and interactive skill-based programs. The system typically consisted of the following basic components: course design features, course management features, communication tools, and administration. The overall design attempted to map student learning onto the three main stages: pre-task, during task and post-task. The concepts of convergent and divergent tasks were involved in all levels

of designing learning activities. Various instructional methods such as presentation, problem solving, drill and practice and also cooperative learning were included. These methods were used in different objectives. The video was used at the presentation stage to give information and explanation at the beginning and wherever explanation was required. Exercises consisted of True/ False, filling in the blank, writing short answers and matching whereby students could check the answers by themselves. Navigation bars were designed for students to click to the next page in linear ways or go back to the previous pages.

For collaborative work, a ChatRoom and a live Webboard were used for communication in SL. Webboard and e-mail were communication tools used in ASL. In SL environment, the subjects were required to attend a computer lab on a fixed schedule. There was a teacher assistant facilitating their learning and solving any problems which might occur during the class. They could communicate with their peers and teacher via ChatRoom and live Webboard not via a face-to-face communication. In ASL environment, students learned at their own place and pace. They did not attend class. They could communicate with their peers and teacher via Webboard or e-mail. Both environments required students to work in small groups and submit their works within the time limit. They submitted their work at the task menu created on the left side of the Webpage at the end of each learning module. Feedback was given via live Webboard for SL environment at the real time and via Webboard for ASL environment at the delayed time.

The local area network (LAN) used in this study was the Nontri network which was developed by Kasetsart University. Two servers: the Maxlearn server and the e-course server were used for delivering the WBI of which user-friendly hardware and software systems had already been set up for the students. The servers provided tools for teachers to organize teaching and learning on the Web. To access Web-based classes, Kasetsart University provided account and password for each student. Students had to log in the URL http://course.ku.ac.th and selected the course 355254 from the course lists.

There was a time limit for students to access the Web-based course. Students were allowed to access each learning module within six hours. After six hours, the learning materials were deleted from the site to give way to the new module. After

each module, students had to perform a task assigned for their group. They were required to create a PowerPoint presentation and submit on the Web. The university server provided tools for teachers to create a place for students to submit their works either in an electronic file or a link to other URL. The works were kept in the database which could be retrieved anytime by the course organizer.

Assessments at the end of the treatments were given in traditional forms to ensure that the students did the test at the same time. Students sat in the exam room and finished the test except a part in direct speaking which required a face-to-face testing.

1.2 The development of a Web-based instruction

The stages in developing this Web-based instruction followed Andrews and Goodson's concepts (1980 cited in Jonassen ed., 1988). They concluded from their study that the general instructional design of a WBI mainly consisted of three phases namely analysis, development and evaluation.

1.2.1 Analysis

1.2.1.1 The audience

The audience was undergraduate students from various fields of study. They were studying in their second, third or fourth year at Kasetsart University, Bangkhen and Kamphaeng Saen Campus, registering in the course 355254 'English for Ground and In-flight Attendants' in the first semester of the academic year 2005. They had completed all English foundation courses. Their English proficiency levels were ranging from lower intermediate to intermediate level.

1.2.1.2 Content

The content was developed from the course 'English for Ground and In-flight Attendants' which was English for specific purposes (ESP). It was an elective course in English language syllabus provided for students in all fields of study. The course objectives aimed to present English concerning the tasks of ground and in-flight attendants. The topics: 1) flight history, 2) flight crew, 3) ground services, and 4) in-flight services were constructed into four Web-based learning modules. Each module presented vocabulary, expressions, reading materials and language functions in English concerning the tasks that the airline personnel had to perform under each topic

in various situations. The learning materials were presented in four language skills: listening, reading, speaking and writing respectively.

1.2.2 Development

1.2.2.1 Tasks and learning activities

The development focused on two types of tasks namely convergent and divergent tasks in two types of WBI (SL & ASL) (see Appendix A). Convergent tasks focused on making conclusion whereas divergent tasks focused on discussing.

Characteristics of 'convergent' and 'divergent' tasks were shown in the following chart.

Characteristics	**Convergent tasks**	**Divergent tasks**
knowledge-based	• require true justified knowledge • abstract conceptualization and active experimentation	• require new significant knowledge •various outcome options with more than one goal possible
activity	• allow for collaboration in meaning negotiation where collaboration leads to a single goal	• allow independent work of where individuals can perform the tasks differently according to their cognitive styles • might lead to different outcomes
question	• require only one correct answer • allow collaborative work with short answers which are not highly cognitively demanding • no reference making	• encourage students to generate questions with more than one goal possible • cognitively demanding, such as making inferences

Chart 9. Characteristics of Convergent and Divergent Tasks

1.2.2.1.1 The pre-task stage

Activities were designed to review and develop the appropriate language repertoire, leading learners to the target language outcomes. Explanations were given in the form of texts and recorded voice. Types of learning activities were designed relating to the types of tasks learners had to perform in each topic. For example, the learners in convergent group had multiple-choice, filling in the blank, True/False questions and matching activities. The learners in divergent group had discussion, giving opinions and writing short answer activities. The software programs used in this stage allowed learners to write their answers, drag and drop for matching, click to

play the recorded sound, video or check the correct answers. The menu bars on the left side of the frame provided links to activities in listening, reading, speaking and writing skills. The pre-task activities were not absolute convergent or divergent characteristics since they must contain equal difficulty levels. Thus, the types of activities were limited as can be seen in the following chart.

Convergent activity	Divergent activity
True or false	Write short answer with can be either YES or NO
Matching	Complete the statements. OR Find the word or words from the passage for the following definitions.
Complete the following sentences.	Write short answers to these questions.
Fill in the blanks.	Write short answer to the following questions.

Chart 10. Convergent and Divergent Pre-task Activities

1.2.2.1.2 During-task stage

Target tasks were identified relating to the topic of each learning module. The tasks were designed differently between convergent and divergent tasks. The subjects performed different types of tasks namely convergent (making conclusion) or divergent (discussing) in small groups to achieve the goals of the tasks. This process dealt with using language in the required situations. Students had to use all they had learnt in order to fulfill the tasks. Tasks were designed following the definitions of tasks given by Ellis (2003), Bygate, Skehan and Swain (2001). They were as follows:

- *A task is a work plan. A task constitutes a plan for learner activity.*
- *A task involves a primary focus on meaning. A task seeks to engage learners in using language pragmatically rather than displaying language. It seeks to develop L2 proficiency through communicating. Thus, it requires a primary focus on meaning.*
- *A task involves real-world processes of language use. The work plan may require learners to engage in a language activity such as that found in the real world.*
- *A task can involve any of the four language skills.*

- *A task engages cognitive processes. The work plan requires learners to employ cognitive processes such as selecting, classifying, ordering, reasoning, and evaluating information in order to carry out the task.*
- *A task has a clearly defined communicative outcome. The stated outcome of a task serves as the means for determining when participants have completed a task.*

It was this stage that the differences between convergent and divergent tasks were seen clearly as shown in the following chart.

Module	Convergent tasks (making conclusion)	Divergent tasks (discussing)
1.Aircraft History	Use the information from the websites to compare and contrast the aircraft in these aspects: body, capacity, type of engine, altitude, size, speed, and flight routes.	Use the information from the websites to make decision in the following situation: you have to assign an airplane to Japan. What type of aircraft will you use among B 747-400, B 737, A 300, DC 10? The number of passengers is 180. Give reasons to support your decisions.
2. Job Descriptions	From the airlines websites, compare the requirements for 'passenger service personnel'. Are there any differences among airlines?	Study the requirements from the airline websites. If you are an interviewer, what criteria you'll use to select an applicant for 'passenger service personnel'. Give reasons to support your decisions.
3. Ground Services	Use the information from the airline websites provided or you can search more online to describe in what situations the airlines allow ticket cancellation and in what situations the airlines do not allow cancellation without penalty.	Explain what you think an airliner should do if a passenger loses his/her ticket. Then use the information from the airline websites provided or you can search more online to find out what the airlines do. Then compare that with what you plan to do.
4. In-flight Services	Using information from the airline websites provided, write the tasks that a cabin crew has to perform from the beginning until the end of the flight.	Plan the procedures for serving meals to 400 passengers within 30 minutes, using 6 cabin crews. Then compare the information from the airline Websites provided with your plan.

Chart 11. Convergent and Divergent Tasks

Different cognitive strategies were required when doing these two types of tasks. Students collected information from different sites and used them as conclusions while they were working with convergent tasks. Along the same lines, they had to discuss of what they could do in each situation when they worked with divergent tasks. The tasks also required students to create their work online. This stage increased the highest level of self-directedness and student control of their learning. Design features of this stage included flexible time scales and opportunities for students to set outcomes and goals for self-managed learning. The task instructions set the directions, provided a scope for the students to explore and enabled them to use the information they found to fulfill the tasks. In addition, links to other sites were constructed for students to submit their work online. Students worked in small groups creating PowerPoint presentation. All groups had to submit their work within the time limit. They had about one week to finish their work. After the dead line, the pass was closed.

1.2.2.1.3 The post-task stage

The feedback was given on the Webboard. Comments and suggestions were given via the live Webboard in SL and via a Webboard in ASL.

1.2.3 Administration

The Internet technologies and university central repository offered learners-oriented administrative activities from registration to reporting their grades. Technologies also maintained student information (names, user IDs, passwords, e-mail addresses), course information, course schedules, discussion boards, ChatRooms. The student login assisted the researcher in managing class activities for different experiment groups and checking the frequency of their attendance. Students of each experimental group could only log in the materials provided for their group. The system did not allow them to log in the materials of the other group. The central repository helped set file preferences, such as availability date, expiry date, and the software application to open the file. It also helped the researcher and students participate in group discussion using communication technologies such as ChatRoom or Webboard to communicate synchronously or asynchronously. The main servers used for course organization was Maxlearn server which linked to the e-course server where the learning materials were stored. The e-course server provided software that supplied multimedia, whereas the Maxlearn

server provided software that supported various tools for class logistics such as tools for creating or editing a new course, bulletin board to post an announcement, a database that stored student profile, and communication tools such as a ChatRoom, Webboard, e-mail and work-group software.

1.2.4 Access

When students came to the site (http://course.ku.ac.th) they had to log in with their user names and passwords. These user IDs helped the researcher to group students with the type of tasks they were assigned to perform. It also controlled which students accessed to the materials. After their login, students could see a list of courses they registered for on the left frame. Once again, they had to select the course for this experiment. The record showed the frequency of their login. Before accessing the course materials, students were required to fill out the personal information which was kept as learner profile in the electronic database. With user ID and password, the researcher could access all groups and reach every student. Hi-speed modems or the faster ISDN connections were recommended for students who accessed the course at their place.

1.2.5 Human interaction

In order to achieve the goal of this course, students were provided with technologies that supported various projects through communication, collaboration and coordination. These technologies were used separately for SL and ASL. This could avoid the feeling of loneliness since students could discuss ideas and talked to their peers and got immediate responses from their teacher in SL using ChatRoom and live Webboard by clicking and logging into the ChatRoom where they found their peers and teacher. Webboard and e-mail were provided for them to communicate in ASL.

1.2.6 Timing

Each Web learning module lasted 6 learning hours. The first experiment was in SL environment; students were required to attend class on a fixed schedule. The researcher also stayed online with them at the real-time. The second experiment was in ASL environment; students had 24-hour-a-day accessing the Web-based course, completing and submitting assignments and exercises and interacting in Web-based forum using Webboards or other Web-based components such as e-mail.

In this phase, students did not attend class on a fixed schedule but their attendance was recorded by the university systems.

1.2.7 Feedback

Most activities provided immediate feedback after finishing the task by clicking the navigation bar at the bottom of the page. Since answers varied in divergent group, only suggested answers were given. Feedback for their task presentation was also written on the Webboard.

1.3 Courseware validation

1.3.1 Validation of the lesson plans

Three experts (see Appendix B) in language teaching from universities in Thailand and USA were used to validate the lesson plans. The experts suggested revising the learning activities and the language in the pre-task stage to be of equally difficulty levels while the differences between convergent and divergent tasks should be given in the task stage. Additionally, they suggested revising the proportion of the time required to perform the task.

1.3.2 Validation of the courseware

To validate the courseware, three experts in language teaching and technology were consulted for validation (see Appendix C). Additionally, the courseware was also uploaded to the Web and allowed the members of TESLCAL (Computer Assisted: Teachers of English as a Second Language-Learning) all over the world to give comments before it was used in the experiment. The comments received from the experts focused on the good characteristics of a well-designed courseware including how contents were presented to the learners. The experts suggested using tables instead of textboxes in each frame. Additionally, they suggested improving the quality of the video presentation since the visions were not clear enough. The only one comment from a member in the TESLCAL group focused on the contents of the teaching materials because she had experience in training the airline staff in various positions. The comments given stated that the materials did not provide enough information and tasks for the job of the airline staff. More specific details were needed to provide knowledge and understanding of each position. However, the technological skill used in designing the Web-based courseware was admirable. Anyway, all the comments were used to improve the

Web-based courseware relating to the criteria for evaluating the characteristics of a well-designed courseware given by Boling and Soo (1999: 442-458) as follows:

1. Consistency

To allow the users to concentrate on mastering the content, the design concerned that graphics did not distract users from what the lessons wanted to present.

2. Good use of space

To ensure that each section of content fit onto one screen and different parts of the information are distinguished from one another spatially.

3. Legibility

To ensure that the text on the screen is legible and readable, the font size is not smaller than Time New Roman 11.

4. Contrast, repetition, alignment, proximity

To ensure the screen design exhibits good use of contrast, repetition, alignment, and proximity.

5. Ease of navigation and recovery

To ensure the design prevents users from making serious errors with the interface of the software. The navigation bars were clearly marked and tested for the effectiveness.

6. High quality playback

To ensure the audio and video playback is of sufficiently high technical quality. Small screen in the main window was created for video presentation. The video files were minimized to decrease time in downloading when the subjects used dial-up modems by telephone lines.

7. High interface control

To ensure that learners can control over pacing and sequence. The presentation was designed using linear presentation.

The courseware was pilot tested for the first time in the second semester of the academic year 2003. The network was connected using Kasetsart University LAN; therefore, there was no problem downloading the video or audio files. Additionally, the communication tools worked effectively. Students did not have problem with communicating via a ChatRoom or a Webboard. After that, the courseware was fully

developed by adding more exercises and learning materials. It was pilot tested again in the first semester of the year 2004 with two groups of students taking the same course used in this study in order to improve its effectiveness. The outcome showed that students had positive attitude toward learning on the Web. They also paid more attention on their learning comparing to the normal classroom setting.

2. Instruments for data collection

2.1 Language achievement test

A language achievement test consisting of two parts was constructed to assess learning achievement in four language skills namely listening, speaking, reading, and writing (see Appendix D). According to the course objectives, the test was constructed into four language skill sub-tests consisting of listening, speaking, reading, and writing skills following Bachman (1990), Bachman and Palmer (1996), Alderson and others (1995), and Douglas' frameworks (2000). Skills emphasized in these tests were listening, speaking and reading with lower scores required on writing sub-test. The cut-off score was set in advance (50 percent) and those who did not meet the criterion are considered 'not pass.' Students were not evaluated by comparison with the achievement of other students, but instead their achievement was measured with respect to the degree of their learning or mastery of the specified content domain.

In order to construct the four-skill sub-tests, the analysis of target language use (TLU) tasks to the test task was carried out. The quality of test usefulness was concerned to adapt TLU tasks to the test situations or eliminated some of them. The students needed to communicate in various aspects in airline business relating to the topics of the learning modules. They were expected to use appropriate strategic competence to engage an appropriate language discourse in executing a complex performance, interacting with the passengers or colleagues to achieve the specific communicative goals in various language functions.

2.1.1 The construction of the test

The constructs of the four sub-tests followed the five steps in the ideal development of a CR test (Roid and Haldyna, 1982: 30) as follows:

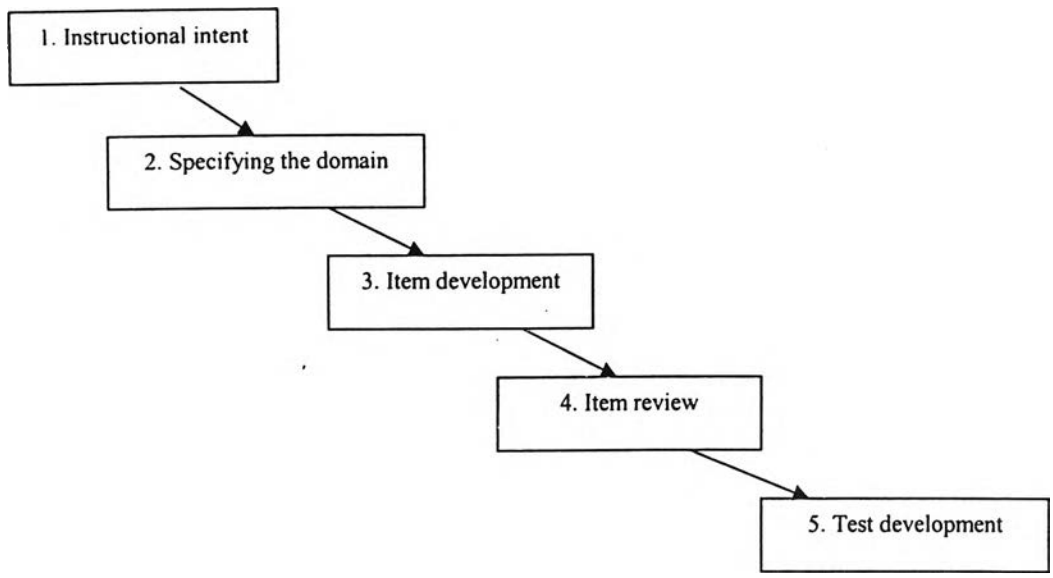

Figure 7. Five Steps in the Ideal Development of a CR Test

2.1.1.1 Listening sub-test

The rubrics of the listening skill sub-test were provided at the time of the test. The time allotted for listening test was thirty minutes. The test consisted of 20 items in each test. The students were given a few minutes to complete their personal information in the answer sheet before listening. Then they heard communications between airline staff and passengers, and an announcement twice and they had to write their answers on the answer sheet.

The input of the audio passages used in the listening subtest was scripted, but it contained many elements of natural production, including false starts, self-correction, and other hesitation phenomena. The passage ranged in length from about a minute to five minutes. The topics of the test were: requesting a passenger to store his belonging, asking for traveling purposes or asking to see travel documents, making a welcome announcement or making a landing announcement, and making reservations.

The test tasks required the subjects to select one best answer in a multiple-choice task, and wrote not more than three words in each blank for a sentence completion task.

The test rubrics can be seen in the following chart.

Objectives	To measure listening ability required to comprehend spoken English in airline business in various situations.
Time	30 minutes
Construct	Hughes' (1989) macro-skill listener functions: 1) listen for specific information, 2) obtaining gist of what is being said, 3) following directions, and 4) following instructions.
Functional knowledge	Expressions and vocabulary in real-life English required to perform the tasks of ground and in-flight attendants in these situations: requesting a passenger to store his belonging, asking for traveling purposes or asking to see travel documents, making a welcome announcement/ landing announcement, and making reservations.
Grammatical knowledge	1) Y/N & wh-questions, use of relative clauses, modifiers/ adjectives, indirect statements/ questions, modals, passive constructs, phrases.
Test tasks	1) choose the correct answer: 10 items / 10 points. 2) fill in the blanks with the information from listening (Not more than three words are needed): 10 items/ 10 points.

Chart 12. Test Rubrics for Listening Sub-test

2.1.1.2 Reading sub-test

The rubrics contained information about time allotment and the number of questions. The time allotted was 60 minutes. The test consisted of 2 reading passages with 15 items in each test. The reading passages were constructed in different topics accompanying with a variety of tasks for each passage i.e. selecting choices, and completing summaries. The reading passages were taken from the real-life resources i.e. the Internet, airline documents, newspaper, and magazines. Each separate task in this section contained its own rubrics in addition to the general one above. The questions were constructed based on Alderson and Lukmani's taxonomies of reading comprehension (1989 cited in Cohen, 1994) to specify what the students should be able to do relating to course objectives. Test tasks were constructed using various techniques i.e. multiple choice, and completion which requires minimal writing. The rubrics provided some informative information about evaluation criteria and the procedures for responding as follows:

Objectives	To measure reading ability required to comprehend written texts in real-life English required to perform the tasks of ground and in-flight attendants.
Time	60 minutes
Construct	Taxonomies for assessing reading comprehension skill (Alderson & Lukmani, 1989 cited in Cohen, 1994): 1) The recognition of words and phrases of similar or opposing meaning. 2) The identifying or locating of information. 3) The discriminating of elements or features within context. 4) The interpreting of complex ideas, actions, events and relationships. 5) Inferring the deriving of conclusions and predicting the continuation.
Functional knowledge	Knowledge of vocabulary and rhetorical organization in reading authentic passages from written texts and airline's online materials.
Grammatical knowledge	Use of relative clauses, modifiers/ adjectives, indirect statements/ questions, modals, passive, adverbs/ phrases and tenses.
Test tasks	1) Choose the one best answer (A), (B), (C) or (D) 15 items / 15points. 2) Select the words from the box below to complete the following outline 15 items/ 15 points

Chart 13. Test Rubrics for Reading Sub-test

2.1.1.3 Speaking sub-test

The test intended to assess the necessary features of quality airline personnel in speaking skill. Test tasks were constructed into a multiple choice test consisting of 20 items. The time allotted was 20 minutes. The construct followed the course contents and topics, three dimensions of Cohen's (1994) socio-cultural ability were used to specify what the students should be able to perform according to the course objectives. Test items consisted of short dialogues in various situations. The assessment criteria encompassed both in linguistic and task fulfillment features. The linguistic criteria included resources of grammar and expressions, fluency, vocabulary, and use of polite forms. Some criteria were more salient in specific tasks; for example use of polite language in answering the passenger's questions.

Objectives	To measure speaking ability required to understand the speaker, and to express language functions in different situations performed by airline personnel.
Time	30 minutes
Construct	The construct involves these functions: asking and giving information, describing things/ jobs/ a sequence of events, indicating roles someone has, talking about work, dealing with complaints/ apologies, making requests/ offers, expressing opinions, thanks, inquiring about distance/ time, talking about ability/ possibility, prohibition.
Functional knowledge	Vocabulary and syntactic knowledge of real-life English which is required to perform the tasks of ground and in-flight attendants. Socio-cultural ability: Cohen's (1994) three dimensions: 1) socio-cultural appropriateness. 2) situational relevance and accuracy. 3) proper amount of information.
Grammatical knowledge	Y/N & Wh-questions, relative clauses modifiers/ adjectives, indirect statements/ questions, conditional sentences, modals, passive constructs, adverbs/ adverbial phrases.
Test tasks	Choose the most appropriate expression used by airline personnel: 20 items /20 points.

Chart 14. Test Rubrics for Speaking Sub-test

2.1.1.4 Writing sub-test

The construct of the writing sub-test was a subjective test consisting of two items in each test. The test rubric provided the instructions for performing the test. The time allotted for writing was 30 minutes. The students were expected to employ a high level of background information and language knowledge in order to complete the task. Test items were derived from the course content in four areas namely types of aircraft, job descriptions, making announcement, and in-flight services.

Cohen's holistic assessment and multi-trait scoring scales (Cohen, 1994:331) were used as criteria for scoring the writing ability. They assessed three aspects in writing ability namely main idea/ opinion, rhetorical features, and language control.

Objectives	To measure English writing ability required to perform the tasks of airline personnel.
Time	30 minutes
Construct	The construct is a subjective test.
Functional knowledge	Vocabulary and syntactic knowledge of real-life English which is required to perform the tasks of airline personnel in writing skill.
Grammatical knowledge	Use of relative clauses, modifiers/ adjectives, indirect statements/ questions, modals, passive, adverbs/ phrases and tenses.
Test tasks	First test 1) Describe these aircraft in at least 80 words. Select one of the following: balloon, glider, propeller plane, jumbo jet, concord, Boeing 747, ornithopter (5 points). 2) Write a welcome announcement of at least 80 words (5 points). Second test 1) Write at least 80 words to describe a job description from the following: reservation staff, passengers service agent, pilot, flight crew check-in staff (5 points). 2) Describe in-flight services provided by any airlines in at least 80 words (5 points).

Chart 15. Test Rubrics for Writing Sub-test

The multi-trait scoring scales for writing are as follows:

Scores -Main idea/ opinion	**Rhetorical features**	**Language control**
5 - The main idea in each of the two articles is stated very clearly, and there is a clear statement of change of opinion.	A well-balanced and unified essay, with excellent use of transitions.	Excellent language control, grammatical structures and vocabulary are well chosen.
4 - The main idea in each articles is fairly clear, and change of opinion is evident	Moderately well-balanced and unified essay, relatively good use of transitions.	Good language control and reads relatively well, structures and vocabulary generally well chosen.
3 - The main idea in each articles and a change of opinion are indicated but not so clearly.	Not so well-balanced or unified essay, somewhat inadequate use of transitions.	Acceptable language control but lacks fluidity, structures and vocabulary express ideas but are limited.
2 - The main idea in each article and/ or change of opinion is hard to identify in the essay or is lacking.	Lack of balance and unity in essay, poor use of transitions.	Rather weak language control, readers aware of limited choice of language structures and vocabulary.
1 - The main idea of each article and change of opinion are lacking from the essay.	Total lack of balance and unity in essay, very poor use of transitions.	Little language control, readers are seriously distracted by language errors and restricted choice of forms.

Chart 16. Cohen's Scales for Scoring Writing

2.1.2 Test validation

Three experts in language teaching and testing were consulted for test validation (see Appendix E). They suggested to test speaking skill directly and also to revise the constructs of the test items and the language used. According to the comments, the tests were revised in the speaking part. The direct speaking sub-test was added for testing speaking skill. Direct speaking test consisting of 5 situations was used to assess the socio-cultural knowledge and language performance. The students selected one situation at random. The tasks included topics, level of formality, the relative status of the participants and familiarity. The time allotted for the test was 12 minutes. Students had 3 minutes to prepare and 9 minutes to speak. Cohen and Olshtain's rating scales (Cohen, 1994: 285) were used for scoring the appropriateness of socio-cultural strategies for speaking in a given context. The students must display competence in communicative language ability to fulfill the task in a manner consistent with qualified experienced airline personnel believed to be of effective communication in the target language use situations.

Cohen and Olshtain's communicative ability scales (Cohen, 1994: 285) for assessing the appropriateness of the socio-cultural strategies are as follows:

Score	Communicative Ability
5	The message is socio-culturally appropriate, reflects the situation clearly, and contains the proper amount of information.
4	The message is for the most part socio-culturally appropriate, reflects the situation relatively clearly, and/or generally contains the proper amount of information.
3	The message is somewhat lacking in socio-cultural appropriateness, calls for some interpretation to be understand, and/ or contains too much or too little information.
2	The message is mostly lacking in socio-cultural appropriateness, calls for much interpretation to be understand, and/or contains too much or too little information.
1	The message is completely lacking socio-cultural appropriateness, is not clear, and/ or contains far too much or too little information.

Chart 17. Scales for Scoring Speaking

The test was pilot tested with students taking the course 355254 'English for Ground and In-flight Attendants' in the second semester of academic year 2003 during the Midterm and the Final Exam.

Brown's formula (1996: 66-67) for item discrimination and item facility index were used to analyze the facility values and the discrimination index. The discrimination index and the facility values of 0.20-not over 0.80 were set to select

the test items for the experiment (Pattiyatanee, 1998). Test items which were not in the range between 0.20 – 0.80 were revised and pilot tested again. K-R20 was used to calculate the reliability value of the objective test items while SPSS program was used to calculate the Cronbach alpha values for the subjective test.

2.1.2.1 The first pilot-test

The test was constructed following afore mentioned test rubrics. The first and the second part of the test were used in the second semester of the year 2003. They were administered with 61 students taking the course 355254 'English for Ground and In-flight Attendants' at Kasetsart University during the Midterm and Final Exam respectively. The test items were calculated for the IF and ID index to select the items which were in the range between 0.20-0.80. The numbers of items with appropriate difficulty and discrimination values were shown in Chart 18.

Subtest Total time = 60 minutes	**Part One Number of Items**	**Items between 0.20 -0.80**	**Part Two Number of Items**	**Items between 0.20 -0.80**
Listening	20	7	20	3
Speaking (indirect)	20	9	20	7
Reading	20	6	20	11
Reliability values of Part One = 0.69			Reliability values of Part Two = 0.65	

Chart 18. Numbers of Test Items in the First Pilot-Test

The test items which were too difficult or too easy were revised, and increased some test items in the reading and listening parts, on the other hand some items in the speaking sub-test were decreased since it could be tested in the direct speaking test. The Cronbach alpha was calculated for reliability values of the writing and speaking tests. The Cronbach alpha values of the writing test were = 0.51, 0.72 and the values for the direct speaking test were = 0.68.

2.1.2.2 The second pilot-test

The test was improved and used again in the first semester of the year 2004 with 138 students taking the course 355254 'English for Ground and In-flight Attendants' at Kasetsart University during the Midterm and Final Exam respectively. The test items were calculated for the IF and ID index and reliability values to select

the test items which were in the range 0.20-0.80. The numbers of test items selected were shown in Chart 19.

Subtest Total time = 60 minutes	Part One Number of Items	Items in 0.20 -0.80	Part Two Number of Items	Items in 0.20 -0.80
Listening	25	18	25	14
Speaking (indirect)	15	6	15	5
Reading	25	20	25	16

Chart 19. Numbers of Test Items in the Second Pilot-Test

Chart 19 shows the numbers of test items after the second pilot-test. It is obvious that there were more items in Part One which were in the range 0.20 -0.80. To make the equal numbers of test items between the two parts and to meet the time allotted, the items which asked for less important details and a small reading passage for a cloze test in Part One were deleted.

The items which were selected for the experiment were concluded as follows (see details in Appendix F):

a) Part One consisted of 14 items in the listening part, 16 items in the reading part, 5 items in the speaking part, and 2 items in the writing part. These scores were added with another 10 scores from a direct speaking test. The total scores were 55 and the time allotted for the whole test was 2 hours plus 12 minutes for the direct speaking.

b) Part Two consisted of 14 items in the listening part, 16 items in the reading part, 5 items in the speaking part, and 2 items in the writing part. These scores were added with another 10 scores from a direct speaking test. Similar to the first part, the total scores were 55 and the time allotted was 2 hours plus 12 minutes.

The numbers of the test items and reliability values after the try-outs can be seen in the following chart.

Part 1	Reliability	Part 2	Reliability
Listening 14 items	0.80	Listening 14 items	0.76
Speaking (indirect) 5 items	089	Speaking (indirect) 5 items	0.80
Speaking (direct) select 1 from 5 situations	0.81	Speaking (direct) select 1 from 5 situations	0.81
Reading 16 items	0.86	Reading 16 items	0.85
Writing 2 items	0.85	Writing 2 items	0.72

Chart 20. Test Items and Test Reliability

2.2 Questionnaire asking opinions after the treatments

Since this study was conducted in technologically mediated environments, two critical dimensions of Web learning environments –time and place were in concerns when constructing the questionnaire. The opinions of the subjects were investigated relating to the Web learning environments that might affect learning on the following topics:

1. The environment that faciiitates learning.
2. The environment that enhances learning achievement.
3. The environment that enhances motivation.
4. The environment that enhances attitude toward learning English.

The students' responses were categorized by key words and analyzed by means of frequency and percentage for describing the data.

Treatment Procedures

The treatments were conducted in the first academic semester of 2005 at two campuses of Kasetsart University: Bangkhen and Kamphaeng Saen Campus with 102 subjects enrolling in the course 355254 'English for Ground and In-flight Attendants'. The subjects from both campuses were divided into two groups according to their grades from the previous English course (FE3). Next, the subjects from each grade were equally randomly assigned to two experiment groups. One group was assigned for convergent tasks while the other was assigned for divergent tasks. Both groups received 2 treatments in synchronous and asynchronous learning in repeated measures design. The repeated measures design was used within 2 reasons. First, to ensure that every subject received both Web learning environments in SL and ASL. Second, to

repeat types of tasks the subjects received in the treatments and that the time to perform those tasks was long enough to allow the effect between the 2 IVs to occur. The treatment procedures for each experiment group were as follows:

Convergent Group

1. Treatment One

The subjects from both campuses were given the first and the second learning modules with SL convergent tasks. They had to attend class on a fixed schedule not via face-to-face communication with the teacher. They sat in a large computer lab, with each student seated at a separate computer terminal. They were dispersed throughout the lab to prevent verbal communication, Chat CMC was used to simulate the same time/ different place communication environment. A teacher assistant was with them in the computer room giving instructions on how to access the program and facilitated the classroom learning. Prior to the experiment, the subjects were explained how to learn during the experiment and told to communicate with the teacher or their peers via Chat CMC in SL. Thus, during the class periods the researcher stayed online at the same time with the subjects. When they finished their work, they submitted their work on the Web. The simultaneous feedback and answers were given via the live Webboard. As the feedback comments were entered, they were displayed sequentially in real time for each group and were automatically saved.

The first learning module started on June 23rd, 2005 at Bangkhen Campus, whereas the experiment started later on July 18th, 2005 at Kamphaeng Saen Campus due to the problem in registration. After six hours, the subjects submitted the first task on the Web. The pause between each module was the time before the next learning module was uploaded to the Web at the beginning of the week. Then they continued learning the second module in SL for another six hours. At the end of the second module, they were given the first part of the test in a traditional form of which the scores were analyzed for further statistical analyses. The direct speaking test was given individually. The subjects had 3 minutes to prepare and 9 minutes to speak and their speaking was recorded for the co-raters. The other parts such as listening, reading, writing and a part in in-direct speaking test were given in the exam room during the Midterm Exam between July-August 2005.

2. Treatment Two

The subjects from both campuses were given the third and the fourth learning modules with ASL convergent tasks. They did not attend class on a fixed schedule. They learned at their own pace/ place/ time. In ASL environment, subjects communicated with the teacher or peers through e-mail or a Webboard. When they finished their work, they submitted their work on the Web. The feedback and answers were given via the Webboard at the delayed time. At the end of the fourth module, they were given the second part of the test of which the scores were analyzed for further statistical analyses. The testing procedures were administered exactly like the first testing. This test was scheduled during the Final Exam in September 2005.

Divergent Group

1. Treatment One

The subjects from both campuses were given the first and the second learning modules with SL divergent tasks. Similar to the first group, they attended class in a computer lab on a fixed schedule not via a face-to-face communication with the teacher. The teacher stayed online at the same time but in a different place while a teacher assistant was with them giving instructions on how to access the program and facilitated the classroom learning. Prior to the experiment, the subjects were given instructions to communicate with the teacher or peers through ChatRoom or a live Webboard. When they finished their work, they submitted their work on the Web. The simultaneous feedback and answers were returned via the live Webboard. Similar to the convergent group, they were given the first part of the test at the end of the second module during the Midterm Exam between July-August 2005. The scores were analyzed for further statistical analyses.

2. Treatment Two

The subjects from both campuses were given the third and the fourth learning modules with ASL divergent tasks. They did not attend class on a fixed schedule. They learned at their own pace/ place/ time. In ASL, subjects communicated with the teacher or peers through e-mail or a Webboard. When they finished their work, they submitted their work on the Web via Maxlearn server. After each six-hour time, they submitted the task on the Web. The feedback and answers were given via the Webboard at the delayed time. At the end of the fourth module, they were given the

second part of the test which the scores were analyzed for further statistical analyses. Similar to the convergent group, the test was administered during the Final Exam in September 2005.

Data Collection

Data in this study are scores from the achievement tests, and frequencies from the questionnaire asking students' opinions after the treatments.

1. Learning achievement

Subjects were given two parts of the test: the first part was given after the second learning module and the second part was given after the fourth learning module. The test consisting of four language-skill sub-tests was scored using the 1 score for correct answer criteria for the objective test items. Cohen's scoring scales (1994) were used as criteria for scoring speaking and writing skill (see page 98-99). For scoring consistency in the subjective test, inter-rater technique was used. One foreign teacher was assigned to be a co-rater rating speaking or writing subtests together with the researcher. The inter-rater reliability between the researcher and the three raters were $\alpha = 0.8163$, $\alpha = 0.9517$, and $\alpha = 0.8528$. The speaking was rated from the recorded voice. Mid point of the scores from two raters were used for scoring writing and speaking tests. Finally, means and standard deviations of the total scores were used in the factorial ANOVA analysis with repeated measures for the main effects and the interaction effects.

2. Opinions from the questionnaire

The subjects were given an open-ended questionnaire at the end of the treatments to obtain attitudes of the subjects toward WBI. They were asked to respond to the questions and to express their opinions about the Web-based learning environments between SL (which requires class presence) and ASL (which students can learn at their convenient time and does not require class presence) as follows:

Question one: Which environment facilitated learning?

Questions two: Which environment enhanced learning achievement?

Question three: Which environment enhanced motivation?

Question four: Which environment enhanced attitude toward learning English?

The data were collected from 102 participants at the end of the treatments. There were only 88 participants who responded in the questionnaire. The data obtained from

the questionnaire were used as background explanatory variables and were not included in the statistical analysis.

Data Analysis

1. Analysis of learning achievement

To test the three hypotheses set for this study. The 2 × 2 two-way ANOVA with repeated measures was conducted to explore the main effects with learning achievements as the dependent variable and the WBI and TBI were the two independent variables. Two factors of WBI were SL and ASL whereas two factors of TBI were convergent and divergent tasks. The confidence level was set at the 0.05 significance level. The methods of statistical analysis were used to answer the research questions shown in Chart 20. Along with the research questions, main effects and interaction effects were explained in Table 5-7.

The values of the effect sizes were used for the interpretation in terms of the correlation between the independent variable classification and the individual scores on the dependent variable. Cohen (1988: 21-23) defined the effect size (ES) as the interpretation of the percent of non-overlap of the treated group's scores with those of the untreated group. Measures of effect sizes in ANOVA are measures of the degree of association between an effect (e.g. a main effect, an interaction effect, a linear contrast) and the dependent variable. They can be thought of as the correlation between an effect and the dependent variable that is attributable to each effect. Eta squared and partial Eta squared are estimates of the degree of association for the sample. The rationale underlying the computation of effect sizes as given by Cortina and Nouri (1999:1-8) is: "*Due to the interpretational limitations, we often want information relating not only to the question of whether or not an effect exists, but also to the magnitude of the effect. .....The resulting value is a standardized index of magnitude of effect and can be used to compare the contributions of different variables within a given study of the contributions of the same variable across different studies. It also relates to the underlying population not to data from a sample.*"

Research Questions/ Hypotheses	Interpretation of Data Analysis
Research Question 1 Are there any differences between the English language learning achievement of students learning through convergent and divergent tasks? **Hypothesis 1** The English language learning achievement of students learning through divergent tasks is significantly higher than that of student learning through convergent tasks.	1. The main effect is used to determine the effect of two types of tasks on English learning achievement. 2. If the hypothesis is accepted, it means that there is a significant difference between the outcomes of convergent and divergent tasks and that the achievement of divergent tasks is higher. 3. If the hypothesis is rejected, it means that there is no significant difference between the outcomes of convergent and divergent tasks. Or the achievement of divergent task is not higher than that of convergent tasks.
Research Question 2 Are there any differences between the English language learning achievement of students learning in SL and ASL? **Hypothesis 2** The English language learning achievement of students learning through ASL is significantly higher than that of student learning through SL.	1. The main effect is used to determine the effect of SL and ASL environments on TBI learning achievement. 2. If the hypothesis is accepted, it means that ASL can enhance higher learning achievement. 3. If the hypothesis is rejected, it means that ASL does not enhance higher learning achievement. Or the achievement of ASL is not higher than SL.
Research Question 3 Are there any interactive effects among convergent and divergent tasks and SL and ASL WBI environments on English language learning achievement? **Hypothesis 3** There is an interactive effect between types of tasks (convergent & divergent) and SL and ASL WBI on English language learning achievement.	1. The interaction effect is used to determine the effect of WBI and TBI on learning achievement. 2. If the hypothesis is accepted, it would mean that the interaction between SL and ASL WBI environments and types of tasks will have effect on learning achievement. 3. If the hypothesis is rejected, it would mean that there is no interaction between WBI environments and types of tasks. Accordingly, they do not have neither negative nor positive effect on learning achievement.

Chart 21. The Interpretation of Data Analysis

The main effect referred to the impact of IVs on DV. In this study, the main effect of WBI showed whether WBI environments had influence on language learning achievement or whether there was a significant difference among the mean scores of SL and ASL on learning achievement. Similarly, the main effect of TBI showed whether convergent and divergent tasks had any influences on learning achievement or whether there was a significant difference between the mean scores of students who performed convergent and divergent task in WBI environments.

The interaction effect referred to the impact that one IV had on another. In this study, it was determined whether there was an interaction effect between WBI and TBI on learning achievement. If there was an interaction, the effect that WBI had on TBI would differ significantly on the achievement of TBI. If there was no interaction effect, there would be no difference on learning achievement no matter what types of tasks students performed in whatever WBI environments.

2. Analysis of responses from the questionnaire

The responses from the questionnaire were categorized by key words. Then, they were analyzed for frequencies and percentage which were used to describe the data.