CHAPTER III

RESEARCH METHODOLOGY

Introduction

This chapter aims to describe the research design, population and samples, stages of research, data collection, treatment of data, and data analysis respectively. However, due to the complex nature of investigating both product and process of second language learning, the outline of this chapter is provided as follows:
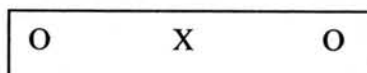
3.1. Research design

3.2. Population and samples

3.3. Stages of research

    3.3.1. Preparations before main study

        3.3.1.1. Course descriptions

        3.3.1.2. Teaching materials development

            3.3.1.2.1. Pre-course needs assessment

            3.3.1.2.2. Long-range lesson plans

            3.3.1.2.3. Lesson plans and materials development

        3.3.1.3. Research instruments development

            3.3.1.3.1. Oral Proficiency Test and Band Scales

            3.3.1.3.2. Teacher's Observation Form

            3.3.1.3.3. Student's Self-Evaluation Form

            3.3.1.3.4. Interview Questions for Co-construction Process

        3.3.1.4. Summary of research instruments

    3.3.2. Pilot studies

    3.3.3. Final revisions

3.4. Data collection: Main study

    3.4.1. Preparation

        3.4.1.1. Class management: Groups and roles

        3.4.1.2. Getting consents

    3.4.2. Conducting Pretests

    3.4.3. Conducting non-participant observations

3.1. Research design

The research design of this study was the One-Group Pretest-Posttest Design.

Figure 3.1: Research design

O	X	O

O means a Pretest and Posttest which was the same form of test
X means a treatment which was narrative task repetition

The One-Group Pretest-Posttest Design was manipulated due to the researcher's need to investigate the development of the participants of this study. Another reason was because the Socio-Cultural or Social Constructivism Theory, which was one of the frameworks of this study, normally paid close attention to the settings and participants in interactions and focused on the intensive nature of transcription and analysis in order to show a robust picture of SLA processes through the small group of study (Foster & Ohta, 2005: 2).

One single form of the Oral Proficiency Test was used as the Pretest and Posttest (O) in order to measure the gained score of the participants as well as to able capture and explain the participants' improvement on the same situations or topics of the test tasks.

The treatment (X) which was used with the participants was the narrative task repetitions based on five different topics obtained from the Needs Assessment.

The thematic contents or topics were Snack Attack, Shopping, Free Time, Travel, and Study Time respectively. The treatment was planned to be conducted as follows:

Figure 3.2: Treatments planning

O	X	O
	X1	
	X2	
	X3	
	X4	
	X5	

X1 = Narrative task repetition of Topic 1/Week 6
X2 = Narrative task repetition of Topic 2/Week 8
X3 = Narrative task repetition of Topic 3/Week 10
X4 = Narrative task repetition of Topic 4/Week 12
X5 = Narrative task repetition of Topic 5/Week 14

3.1.1. Quantitative and qualitative data

This research used both quantitative and qualitative data. The justifications of using both types of data were not only for the triangulation of the data (Denzin, 1978, cited in Brown, 2002: 243), but also for more understanding of the data of which quantification may not be able to produce. Foster and Ohta (2005: 2) stated that from a Social Constructivism Theory perspective, quantification risked sacrificing the richness of the interaction that occurred from the participants being investigated.

The quantification of data was used to measure the gained score of the Posttest from that of the Pretest, and to compare the oral language performances of the participants' narrative repetitions with those of new narrative tasks on the three language aspects which were fluency, accuracy, and complexity.

In terms of the qualitative data, it was collected through observations, interviews, and from the students' report forms to ensure the methodological triangulation (Denzin, 1978, cited in Brown, 2002: 244). The focus point of using qualitative data was to analyze how the students had co-constructed their knowledge, and what they had learned from that co-construction process. The "theorizing" technique of LeCompte and Preissle (1993:239) which was the cognitive process of discovering or manipulating abstract categories and the relationships among those categories was used for the qualitative data analysis.

3.1.1.1. Treatment of quantitative data

The recordings of participants' Pretest and Posttest were played back and rated using the descriptive band scales ranging from band 1 to 5 on seven focused areas for each interview and narrative test. For the interview test, it was evaluated on the areas of complexity, accuracy, fluency, pronunciation, vocabulary, flexibility, and confidence to initiate speech respectively. On the other hand, the narrative test was rated on the areas of complexity, accuracy, fluency, pronunciation, vocabulary, thematic development, and coherence and cohesion respectively.

Through the use of the T-unit, the transcription data of the new narrative task (NNT) and narrative task repetition (NTR) that participants performed in the classrooms was measured on three aspects of language: fluency, accuracy, and complexity. However, the measurement was based on one of the two topics for each Low and Medium score achievers since either one of the two topics was a representative between themselves, and the last topic for High score achievers. T-unit was defined as a single clause or a dependent clause with one or more independent clauses attached to it. The suggestions made by Polio (1997: 138-140) in counting clauses or phrases for T-units, word counts and errors had been consulted as a guideline in preparing the speech data.

3.1.1.2. Treatment of qualitative data

Referring to the Assessment Glossary (Allen, Noel, & Rienzi, 2001: 5, cited in Diaz-Rico, 2004: 78), qualitative assessment were results of assessment that were described verbally and may involve counts of categories such as those for scoring rubrics or rating scales. The recorded data of participants' performances for NNT and NTR were described for their language qualities using the descriptors of the band scales on each seven focused areas that had been mentioned above.

The observation, interview, audio-taped and VTR data of the focus group were analyzed, synthesized, and categorized using the "theorizing" technique (LeCompte and Preissle, 1993). Its process was to record the phenomena, classify the data into categories, establish linkages or relationships, and make inferences in order to study the co-construction process of the participants and what they have learned from that process.

3.2. Population and samples

The population of this study was the third year Undergraduate English major students studying at the Faculty of Education. They were trained to be teachers of English language after graduation.

The samples were twenty students purposively selected from the third year English major students studying at the Faculty of Education, Chulalongkorn University. They enrolled the course named "Effective English Speaking" of which this study was conducted in the first semester of academic year 2006.

All twenty subjects, or so called "participants" were randomly assigned into four groups. Each group was comprised of five members with different levels of English oral proficiency determined from the Pretest in order to promote the scaffolding among the learners. There were 2 Low score achievers, 2 Medium score achievers, and 1 High score achiever in each group. All of the twenty samples were also posttested by the end of the semester to measure the gained score in order to answer the research question 1.

Figure 3.3: Groups dividing for all twenty participants
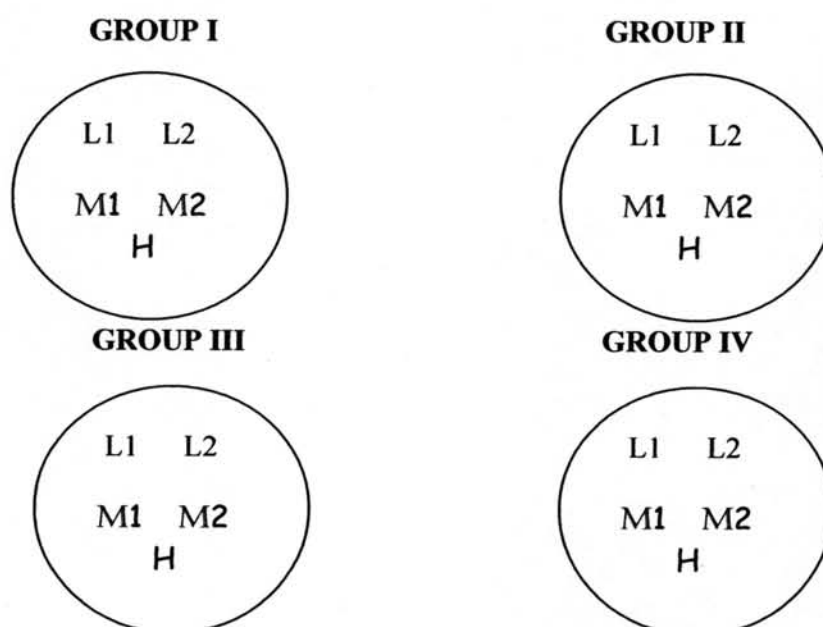
The oral narrative performances of the participants L2, M1, and H of each group were selected to be analyzed to answer the research questions 2 to 5 as shown in Figure 3.4. The total of the selected participants were 12.

Figure 3.4: Twelve selected participants

GROUP I

L2
    M1

H

GROUP II

L2
    M1

H

GROUP III

L2
    M1

H

GROUP IV

L2
    M1

H
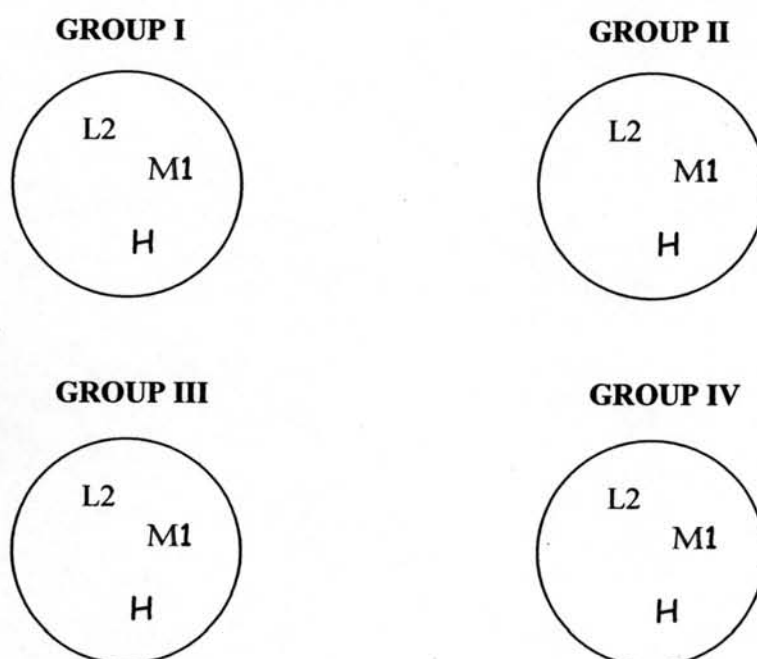
Regarding the investigation of co-construction process in the research questions 6 and 7, one group was selected to be the focus group for the investigation. The criteria of choosing the focus group were it must be a representative of all the four groups in term of the numbers of members, roles, and mixed-ability levels. Additionally, the members of the focus group should reflect positive interdependence, individual accountability, and responsibility for the success of all group members which were the key characteristics of collaborative or cooperative learning (Putnam, 1993: 16-21). Considering the above mentioned criteria, Group I was selected to be the focus group because it matches all the criteria. There were five members in Group I with mixed level of English oral proficiency as shown in Figure 3.5.

Figure 3.5: Focus group

GROUP I

L1   L2

M1   M2

H
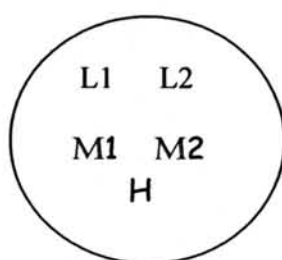
3.3. Stages of research

The stages of research are as follows:

3.3.1. Preparation before main study

Initially, the preparation had to be taken concerning the course descriptions, and the development of teaching materials including other necessary instruments.

3.3.1.1. Course descriptions

When the English major students of the Faculty of Education at Chulalongkorn University became the third-year students, they were required to take one course which was called "Effective English Speaking" (2701358) in the first semester. The course descriptions of this course (Faculty of Education, Chulalongkorn University, 2004) were to teach the principles of speaking in front of the audience in various occasions, especially in the classrooms and meeting rooms; and the fundamental knowledge of speaking. Students should know how to use appropriate lexicons or phrases, idioms, and rhetorical metaphors as well as how to listen and speak in the academic settings; and should be able to summarize what had been listened to.

After looking at the course descriptions, the key concept of this course was to train the students of how to speak in public in various situations. After reviewing the literatures on public speaking, it was found that there were many genres of public speaking, and narrative genre was one of them. Since there were some empirical studies (Gass & Varonis, 1985; Shimojima, 1997; Gass et al., 1999; Lynch & Maclean, 2001; Bygate, 2001a) showing the positive effects of using the narrative and narrative task repetition to improve the English oral language skills of the learners

such as they became more fluent and expressed themselves more clearly, the researcher then decided to focus the teaching and learning of this course on the practice of this recurring task. Hence, the narrative genre and its repetition were selected to be the core learning task for the students taking this course. The next step was to develop the teaching materials through the needs assessment conducted with the prospective learners.

3.3.1.2. Teaching materials development

Before developing the teaching materials, the needs assessment was conducted to gather the information necessary to meet the learning needs of the target group of subjects. It was conducted with all the to-be third-year English major students studying at the Faculty of Education because it was not known which group would be selected to be the participants in the main study. All of them were required to enroll the course named "Effective English Speaking" when they became third-year students in the following semester.

3.3.1.2.1. Pre-course needs assessment

The Pre-course needs assessment was conducted by using the questionnaire with a purpose of gathering information for developing the teaching and learning materials and lesson plans, particularly, concerning the topics or thematic contents of the participants' preference for enhancing their learning involvement.

The Pre-course Needs Assessment Questionnaire was prepared in English and Thai, and it consisted of four parts. The first part asked for the general data of the participants such as their personal information, English educational background, and language skills. The second part was the self-assessment of their own English abilities on various situations or topics. The third part was a list of twelve thematic contents or main topics for the participants to select 4 topics and rank them from 1 as the most preferred topic to 4 as the least preferred one. The last section of this part was an open-ended question asking the informants to suggest other thematic contents or topics that they liked. Finally, the fourth part contained some questions about the participants' expectations from the English course focusing on the speaking skills and their plans to achieve their goals arisen from their expectations. Details of each part of the Pre-course Needs Assessment Questionnaire were provided in the Appendix 1.

To ensure the content validity of the Pre-course Needs Assessment Questionnaire, the technique of expert validity was used. The criteria of selecting the three experts were one person was from the English language teaching, the second one was an expert in second language acquisition, and the last one was from the curriculum development and instruction area. All of them were also experienced English language teachers of more than five years. The content validity value of the Index of Item Objective Congruence (IOC) of the Pre-course needs assessment questionnaire was high at 0.84. The content validity value of IOC was normally accepted at 0.50. More information of the content validity evaluation was provided in Appendix 2.

3.3.1.2.1.1. Revision

Table 3.1 was the suggestions made by the experts for revising of the Pre-course Needs Assessment Questionnaire.

Table 3.1: Suggestions for Needs Assessment Questionnaire

Parts	Suggestions
Instruction	Add a sentence, "Please respond to the Questionnaire as truthfully as possible." And a few statements should be added to reassure that the participants' real names would be concealed, and the data would be reported as group data only.
Part 1	Add a question asking whether they have any extra English learning/practicing outside their classes.
Part 2	Add following questions: - Are you satisfied with your English speaking ability? - If not, what areas of speaking problems do you want to improve? - How are you going to improve your English speaking skills given that you have the opportunity to do whatever you want to? - In your opinion, what factors can help you improve your English speaking skills? - If you have to tell a story about something, what would you do to make it interesting to the listeners? - If you have to repeat telling the same story, what would you do about it? Why would you do it that way?
Part 3	The instruction was modified to ask the participants to not only select the topic they liked best, but to rank from the one they liked best which was number 1 to the one they liked least which was number 4.
Part 4	Add a question, "What do you think the teacher can do to help you achieve your expectation?"

3.3.1.2.1.2. Data from Needs Assessment

The data obtained from the Needs Assessment was summarized in Table 3.2.

Table 3.2: Data from Pre-course Needs Assessment

General info	Participants' ages ranged from 19 to 24 years old. 69% never been abroad and had no extra English classes 89% spoke Thai at home
Part 2: Sub part 1	Around 50% of the informants considered their speaking ability good to excellent on general topics and fair on abstract topic.
Sub part 2 Note: Q) = question A) = answer	Q) What are you good at when speaking English? A) 85% thought they were good at speaking English on general conversation, something about themselves or their daily lives Q) What do you find difficult when speaking English? A) the organization of ideas when speaking; to speak spontaneously without planning; to use appropriate words and correct sentence structures; and to speak smoothly and fluently when they were panic. Q) Are you satisfied with your English speaking ability? A) 69% were not satisfied; 18% were satisfied; and 15% were somewhat satisfied. For those who were not satisfied, they wanted to improve their accent, intonation, vocabulary, fluency, grammar, organization of thoughts and words, and self-confidence respectively. Q) What are you going to do to improve your speaking skill if you have the opportunity to do whatever you want to? A) 82% would try to speak and practice more; 10% would take extra English courses; and 8% thought that going abroad would improve their English. Q) What factors can help you improve your English speaking skill? A) 62% mentioned regular English practice and use; 18% courage and self-confidence; 17% teacher and students including their English knowledge and experience of learning English; 12% English-speaking environment; 12% getting more exposures in English such as watching soundtrack movies, etc.; 4% paying more attention and concentration in English classes; and 3% getting a boyfriend who is an English native speaker. Q) If you have to tell a story about something, what would you do to make it interesting to the listeners? A) 72% stated pitch, tone of voice, voice clarity, gestures, acting, eye-contact, and expressions of feelings; 22% using realias, pictures, or real artifacts; 3% using interesting introduction; and 3% using jokes or examples.

	Q) If you have to repeat telling the same story, what would you do about it? A) Their ideas were changing the delivery styles such as singing songs, using different sets of pictures, or adding feelings into their repeated stories. Their reasons were to make the stories become more interesting, to avoid boredom, and to get the audience's attention.
Part 3:	Theme/topic selections from the most favorite ones to the least favorite ones were: Free Time; Travel; My Family; Myself; Food; Shopping; and Education.

After that, the researchers had a meeting with the participants again and requested them to re-select only five themes/topics. And the five thematic contents were Free Time, Travel, Food, Shopping, and Education respectively.

3.3.1.2.2. Long-range lesson plans

After reviewing the course descriptions of "Effective English Speaking" and the data gained from the Needs Assessment, the rationale of designing the lesson plans was set to be to develop and design the learning and teaching materials including the lesson plans where the theory of Social Constructivism and Task-based Language Learning and Teaching supported the course descriptions of the Faculty of Education, Chulalongkorn University (2004).

Therefore, the learning objectives of the overall course were:

- Students would be able to speak in front of the classrooms.
- Students would be able to speak in different genres.
- Students would be able to get the fundamental knowledge of speaking, especially, that of oral narrative.
- Students would be able to know how to use appropriate lexicons, or phrases, idioms, and rhetorical metaphors.
- Students would be able to know how to listen and speak in the academic settings.
- Students would be able to reflect on what had been listened.

The long-ranged lesson plans of Effective English Speaking Course were developed as demonstrated in Table 3.3.

Table 3.3: Long-range lesson plans

PRETEST
Week 1: Introduction to public speaking Week 2: Effective speech planning and effective listening Week 3: Informative and Persuasive speaking Week 4: Planning a narrative speech Week 5: New narrative task practice (# 1: Snack Attack) Week 6: Narrative task repetition (# 1: Snack Attack) Week 7: New narrative task practice (# 2: Shopping) Week 8: Narrative task repetition (# 2: Shopping) Week 9: New narrative task practice (# 3: Free Time) Week 10: Narrative task repetition (# 3: Free Time) Week 11: New narrative task practice (# 4: Travel) Week 12: Narrative task repetition (# 4: Travel) Week 13: New narrative task practice (# 5: Study Time) Week 14: Narrative task repetition (# 5: Study Time)
POSTTEST

3.3.1.2.3. Lesson plans and materials development

The lesson plans for this study were focused on week 5 to week 14 where the data of participants' English oral performances and observation data during their planning time were collected and analyzed. Five thematic contents/topics taken from the Needs Assessment were classified into three different levels: Low, Medium, and High levels according to the cognitive familiarity suggested by Estaire and Zanon (1994). This classifications of topics and pictures used for narration were then validated by three the same group of experts who validated the Pre-course Needs Assessment Questionnaire. The experts all agreed with the topic classification and pictures with the IOC value of content validity at 1.00. The topics of Snack Attack and Shopping were for the Low proficiency of students, Free Time and Travel were for Medium level, and Study Time was for the High level.

The pictures used for the narrative tasks were taken from the website of Kate Singleton available from http://www.cal.org/caela/health and a book called Action English Pictures by Frauman-Prickel (1985). These pictures were then modified and redrawn to suit the levels of student's score achievements in terms of cognitive load from the increasing number of picture frames. There were many reasons why pictures were used as stimulus materials or realias to stimulate their oral narratives. First of all, pictures on pieces of paper were easier to be modified and duplicated than pictures on CD or cartoon animations. Secondly, accordingly to Schwartz (2000), pictures

were a great incentive for language production. Lastly and most importantly, pictures contributed to interest and motivation, a sum of the context of the language, and a specific reference point or stimulus (Wright, 1989: 19). The selection and modification of pictures had been done by considering the following lesson plan objectives:

- Students should be able to describe the setting, character(s), time, and situation.
- Students should be able to describe what the character(s) is doing in the pictures.
- Students should be able to tell the problem/turning point/crisis of the story.
- Students should be able to tell what finally happened by using their own imagination/creativity.
- Students should be able to give their own opinion(s) to highlight/evaluate the story.

The teaching materials were the new narrative task and its repetition materials based on five previously selected topics. The materials included lesson plans and stimulus materials which were pictures, newspapers, VCD and audio-tape. The teaching materials of new and repeated narratives vary according to the frameworks and procedures of the new narrative work and narrative repetition (Figures 3.6 and 3.7) which were developed from the proposed framework of task-based lessons in Table 2.5.

Figure 3.6: Framework and procedures for new narrative task

Pre-task stage

- Warm- up

- Useful words and phrases provision

- Topic and objectives introduction

During-task stage

- Groups dividing
  I II III IV

- Roles assignment for each group
  Chair
  Monitor
  Time-keeper
  Note-taker
  Reporter

- 10-min. planning time

- Task performance (5 mins)

Post-task stage

- Discussion
  Grammatical Structures,
  Expressions,
  Vocabulary

- Reflection
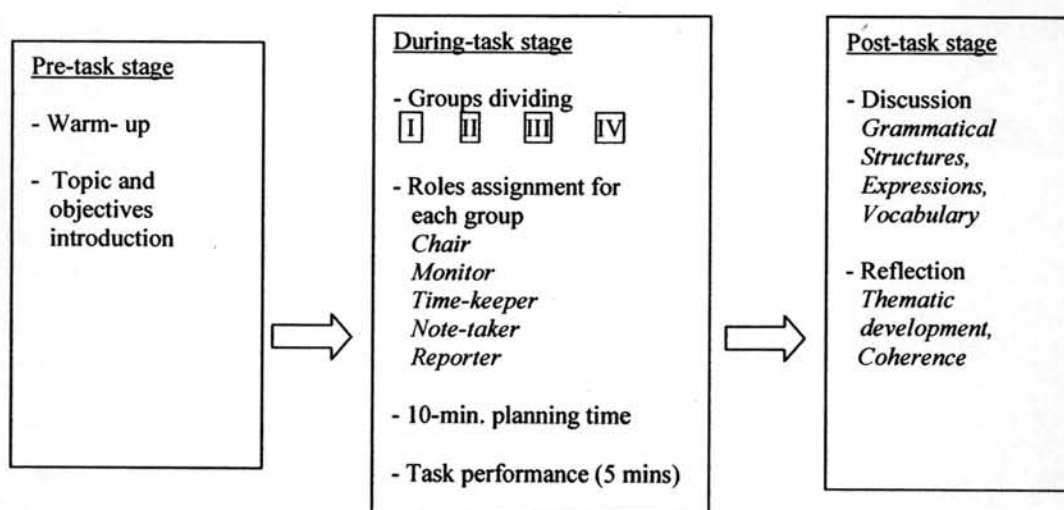 *Thematic development,*
  Coherence

In the Pre-task stage, the materials for the warm-up activities were prepared which were real artifacts, slides of pictures, games, reading and listening passages, or films. A set of questions relevant to that particular topic were prepared in advance on slides of PowerPoint with a purpose of prompting the learners to think about that topic. After that, the learners would be informed about the topic of the class and the objectives of the task. Some useful vocabulary and phrases were also prepared in advance and shown on the PowerPoint.

In the During-task stage, a blank sheet of paper with the name of the topic was provided for each group of the participants to make some notes for their narratives. Other materials were the cassette recorders which were placed for each group. A camera was also set up for recording the performances of their narratives.

Before the Post-task stage, every participant was given another sheet of paper to jot down any grammatical errors that they could notice from their peers' narratives, and any comments that they would like to make on their peers' narratives. This would help the audience practice the noticing of language errors. The main purpose of doing this was to try to make the participants who were the audience to pay attention to their narrators. This will be done in a friendly atmosphere to avoid a feeling of being humiliated which can harm the self-confidence of the speakers.

After that, the participants will be asked to redo the same topic of the narrative task in the following week. The framework of NTR was provided in Figure 3.7.

Figure 3.7: Framework and procedures for the narrative task repetition

Pre-task stage	During-task stage	Post-task stage

- Warm- up
 	
- Topic and objectives introduction
	- Groups dividing
	 I II III IV
	
	- Roles assignment for each group
	  Chair
	  Monitor
	  Time-keeper
	  Note-taker
	  Reporter
	
	- 10-min. planning time
	
	- Task performance (5 mins)
		- Discussion
		  Grammatical Structures,
		  Expressions,
		  Vocabulary
		
		- Reflection
		  Thematic development,
		  Coherence

In the Pre-task stage, though the NTR was based on the same topic as that of the previous week, a new set of warm up activities which were different from the new narrative task were prepared in order to lead them into the theme of the task. The difference between the NNT and NTR procedures was in the Pre-task stage in which there was no interaction with the students through questions and answers and words and phrases were not provided to the class in NTR.

In the During-task stage, a blank sheet of paper with the name of the topic was provided for each group of the participants to make some notes for their narratives. A cassette recorder was placed for each group, and a camera was also set up for recording the performances of their narratives.

Before the Post-task stage, every participant was given another sheet of paper to jot down any grammatical error that they could notice from their peers' narratives, and any comment that they would like to make on their peers' narratives. Again, the purpose of doing this was to try to make the participants who were the audience to pay attention to their narrators.

Concerning the objectives of accomplishing the narrative task, a framework of oral narrative in Table 2.2 (Labov, 1972) was modified to be used as a guideline for analyzing narrative structures shown in Table 3.4.

Table 3.4: Modified framework of oral narrative

Orientation	Give details of time, persons, place, situation.
Complication	Give the main event sequences and show a crisis, problem, or turning point.
Evaluation	Highlight the point and reveal the teller's attitude by emphasizing parts of the narrative. [Optional]
Result	Show a resolution to crisis.
Coda	[Optional way of finishing by returning listeners to present]

3.3.1.2.3.1. Revision

There was a comment on three questions which were used during the Pre-task stage of the new narrative task under the topic of Snack Attack. These questions should be addressed to the participants. Therefore, the questions were corrected to be:

- When you were children, how many kinds of vegetables did you know?
- Could you come up with a lot of answers?
- Why could/couldn't you come up with a lot of answers?

Details of the content validity evaluation of the teaching materials were given in Appendix 4.
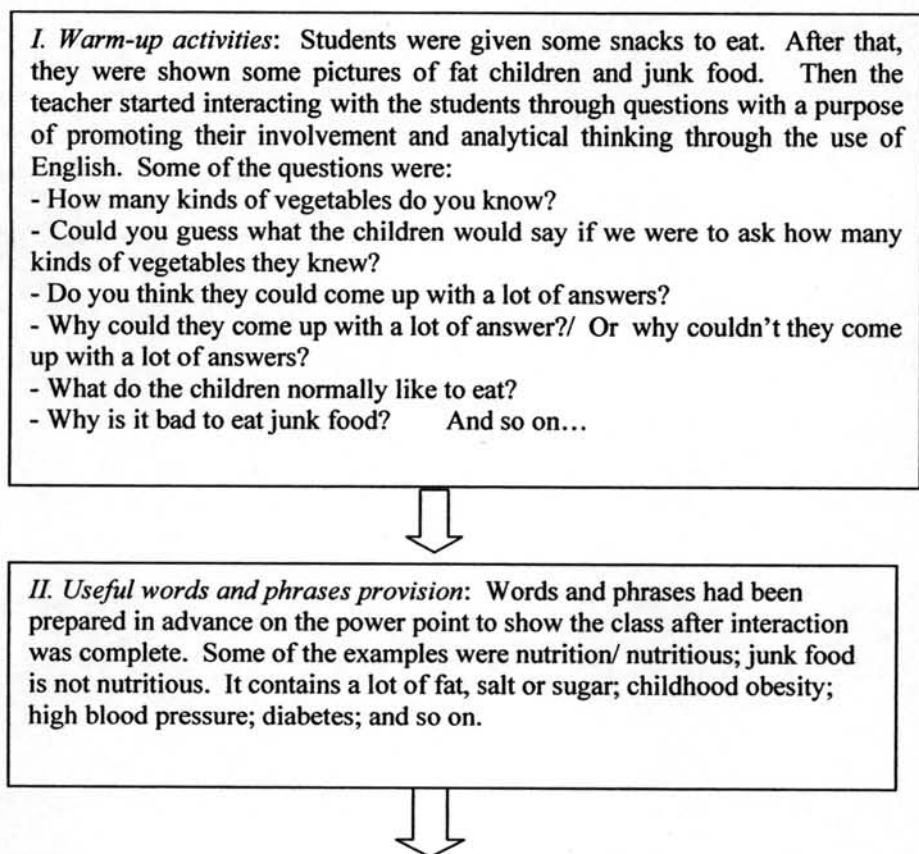
3.3.1.2.3.2. Lesson plan sample

One sample of the lesson plan for each new narrative and its repetition for the topic of Snack Attack was provided in Figures 3.3 and 3.4. (See Appendix 3 for all teaching materials and lesson plans.)

Topic 1: Snack Attack (Low score achievers)/ New narrative task

The stimulus materials used for narrating a story of this topic was a set of six picture frames about a boy who spent his time watching TV and playing computer games while eating snacks.

Figure 3.8: Lesson plan of Topic 1 (NNT)

A. Pre-task stage (about 45 minutes)

I. Warm-up activities: Students were given some snacks to eat. After that, they were shown some pictures of fat children and junk food. Then the teacher started interacting with the students through questions with a purpose of promoting their involvement and analytical thinking through the use of English. Some of the questions were:
- How many kinds of vegetables do you know?
- Could you guess what the children would say if we were to ask how many kinds of vegetables they knew?
- Do you think they could come up with a lot of answers?
- Why could they come up with a lot of answer?/ Or why couldn't they come up with a lot of answers?
- What do the children normally like to eat?
- Why is it bad to eat junk food?   And so on...

II. Useful words and phrases provision: Words and phrases had been prepared in advance on the power point to show the class after interaction was complete. Some of the examples were nutrition/ nutritious; junk food is not nutritious. It contains a lot of fat, salt or sugar; childhood obesity; high blood pressure; diabetes; and so on.

III. Topic and objectives introduction: Students were introduced to the topic and objectives of the lesson.

B. During-task stage (about 45 minutes)

I. Groups dividing: Students were divided into four small groups with five students of mixed abilities per group.

II. Roles and stimulus materials assignment: Assign students in an individual group a different role and explain to them what their job responsibilities were. Then give each group the same set of pictures to plan for their narrative tasks, and tell them the instructions.

III. Planning time: Students were given ten minutes to plan for their narrative tasks. They could take notes of what they were going to say. They were encouraged to discuss in English as much as possible.

IV. Task performance: Students who had been assigned a role of a "reporter" came to narrate a story in front of the class. They were also reminded to refer, but not to read from their notes.

C. Post-task stage

> *I. Discussion*: Any language problems were discussed with the class in a relaxing atmosphere.

⇩

> *II. Reflection*: Students were encouraged to reflect their ideas on their peers' narrative performances on how they built up the stories and related the events together.

Topic 1: Snack Attack (Low score achievers)/ Narrative task repetition

Figure 3.9: Lesson plan of Topic 1 (NTR)

A. Pre-task stage (about 30 minutes)

> *I.* Warm-up activities: Students were asked to answer some questions on the handouts about "You Are What You Eat" followed by a discussion. Then there were some funny quizzes for them to answer such as, "Are French Fries from France?", and so on.

⇩

> *II. Topic and objectives introduction*: Students were introduced to the topic and objectives of the lesson.

B. During-task stage (about 45 minutes)

> *I. Groups dividing*: Students were divided into four small groups with five students of mixed abilities per group.

⇩

> *II. Roles and stimulus materials assignment*: Assign students in an individual group a different role and explain to them what their job responsibilities were. Then give each group the same set of pictures to plan for their narrative tasks, and tell them the

⇩

> *III. Planning time*: Students were given ten minutes to plan for their narrative tasks. They could take notes of what they were going to say. They were encouraged to discuss in English as much as possible.

⇩

> *IV. Task performance*: Students who had been assigned a role of a "reporter" came to narrate a story in front of the class. They were also reminded to refer, but not to read from their notes.

C. Post-task stage (about 30 minutes)

> *I. Discussion*: Any language problems were discussed with the class in a relaxing atmosphere.

⇩

> *II. Reflection*: Students were encouraged to reflect their ideas on their peers' narrative performances on how they built up the stories and related the events together.

3.3.1.3. Research instruments development

Despite the teaching materials, there were additional four research instruments to be developed for this study.

3.3.1.3.1. Oral Proficiency Test and Band Scales

The Oral Proficiency Test (OPT) was used as both a Pretest and Posttest of the study. The Test and descriptive band scales were developed in order to assess the English oral proficiency of the participants before the beginning of the study to divide them into three levels of Low, Medium, and High score achievers. Another objective of developing the Oral Proficiency Test was to be able to measure the gained score of the English oral proficiency of the participants at the end of the study to see whether the narrative task repetition could improve their English oral language performance.

The Test consisted of two speaking tasks which were the interview and narrative tasks. Besides the narrative test, the interview test was included in the Oral Proficiency Test because it was the most common procedure for assessing spoken language proficiency (Burns & Joyce, 1997: 104). The interview test was constructed using Fulcher's framework of characterizing interview test which had been shown in Table 2.6. For the narrative test, it was also constructed using Fulcher's framework for characterizing narrative test as provided in Table 2.7.

The procedure of an oral interview had followed the suggestion made by Manidis and Prescott (1994: 27, cited in Burns & Joyce, 1997: 106). There were three stages for the procedure as shown in Table 3.5.

Table 3.5: Procedure of an oral interview

Stages	Procedure	Assessment
Exploratory	Welcome learner and introduce self. Explain purpose of interview. Initiate conversation.	Make an intuitive assessment
Analytical probing and extending	Phase I: Elicit language samples by directing questions that allow the learner to display features of language behavior described in the level you have selected.	Check level
	Phase II: Extend depth of questions, use visual stimuli and change topics to seek extension of the language behavior.	Adjust level

	Phase III: Take the learner to the point at which she/he can no longer function comfortably. Compare with previous level of interaction for confirmation of level.	Confirm level
Concluding, winding down	Return to comfortable level of interaction for the learner. Return to general conversation and proceed with a reminder of assessment.	Record assessment

Descriptive band scales which were a series of short descriptions and band scales of different levels of language ability (from 1 which was the lowest score to 5 which was the highest one) were developed for assessing the English oral proficiency of the participants. The descriptive band scales were developed by reviewing the theoretical bases of Communicative Approaches to Second Language Teaching and Testing of Canale and Swain (1980), the Common European Framework of Reference for Language: Learning, Teaching, Assessment (2001), and other literature on speaking testing and assessment (Burns & Joyce, 1997; Fulcher, 2003; Hughes, 2003; Luoma, 2004) as well as the measurements suggested by Chalhoub-Deville (1995: 21) for each part of the test.

For the interview test, its measurement depended on the following performance criteria:

- *Complexity of the language structure* involved the use of fragments and/or choppy sentences to the ability to use complex sentence structures.
- *Accuracy of the language structure* referred to the grammatical errors in sentences to no detection of grammatical error.
- *Fluency of the speech* meant the production of short utterances with evident pauses and false starts to smooth and effortless flow of expression.
- *Pronunciation* covered from mispronunciation and being difficult to understand to almost native-like intonation of English speakers.
- *Appropriacy of register or vocabulary* reflected from a little repertoire of vocabulary to a good command of broad lexical repertoire and idiomatic expressions.

- *Flexibility to adapt to the change of topic or task* covered from little flexibility in using basic sentence patterns to a native-like flexibility in reformulating ideas to express herself/himself on all topics of conversation.

- *Confidence to initiate speech* involved an ability to answer very simple questions about personal details, but with repetition, rephrasing, and repair to an ability to interact with ease, confidence, and skill.

For the narrative test, the measurement was based on the following criteria:

- *Complexity of the language structure* involved the use of fragments and/or choppy sentences to the ability to use complex sentence structures.

- *Accuracy of the language structure* referred to the grammatical errors in sentences to no detection of grammatical error.

- *Fluency of the speech* meant the production of short utterances with evident pauses and false starts to smooth and effortless flow of expression.

- *Pronunciation* covered from mispronunciation and being difficult to understand to almost native-like intonation of English speakers.

- *Appropriacy of register or vocabulary* reflected from a little repertoire of vocabulary to a good command of broad lexical repertoire and idiomatic expressions.

- *Thematic development* referred to a little ability to describe about the scene, character(s), and/or event from the prompt to elaborate descriptions with sub-themes, turning point or problem and appropriate conclusion or solution.

- *Coherence and cohesion* referred to no apparent use of any simple connectors to the smooth and well-structured speech with the use of connectors, transition expressions, and other cohesive devices.

More details of the Oral Proficiency Test and Descriptive Band Scales were provided in Appendix 5.

Both of the Oral Proficiency Test and Descriptive Band Scales were validated by three experts. The first one was an expert in English teaching and instruction, the second expert was in English education, and the last one was an expert in English instruction and assessment. All of the experts were also experienced English instructors. The overall IOC value of the content validity of the Oral Proficiency Test was 0.934, and the Descriptive Band Scales was 1.00. (See

Appendix 6 for the evaluation for content validity of Oral Proficiency Test and its Descriptive Band Scales.)

For the reliability of the instrument, the researcher's assistant who acted as the inter-rater was trained before rating the recorded oral performance of the participants. The results of the inter-raters reliability of the Pretest was .844 and .925 for the Posttest.

3.3.1.3.1.1. Revision

After reviewing the comments of the experts, the implementation was done as follows:

- Regarding the interview question in the Exploratory Stage, question 4 was changed to be "How do you like studying at…..(name of test-taker's university)…?"

- Regarding the prompts card A used in the Analytical Probing and Extending Stage in Phase II and III, the prompt card A with questions on the topic of technology was taken away. The test-taker was asked to respond to questions of prompt card B only.

- Regarding the validating question asking whether the descriptors were easy to use, it was recommended not to rely only on the live rating. Therefore, the rating was conducted by listening to each test-taker's recorded production.

3.3.1.3.2. Teacher's Observation Form

Observation was an important method of collecting the data because it enabled researchers to document and reflect upon classroom interactions and events as they actually occurred rather than as we had thought they occurred (Burns, 1990: 80). Therefore, it should be systematic and precise. Observation would become meaningful if researchers or teachers knew what to look for and had a framework or structure of what to be observed (Borich, 1994: vii).

Regarding the framework of observation, Genesse and Upshur (1996: 76-95) had proposed a framework of classroom observation as in Table 3.6 which was used as a guideline for the observation process of this study.

Table 3.6: Genesse and Upshur's framework of observation

Genesse and Upshur's Framework of Observation	THIS STUDY
1. Identify the <u>purpose</u> for which the observational information will be used.	<u>Purposes:</u> - To explain the co-construction English language learning process of the participants in the focus group. - To explore what the participants in the focus group had learned from that co-construction process.
2. Identify the kinds of <u>observational information</u> that would be useful for the purpose.	<u>Observational information:</u> The participants' interactions to their peers in the group including their reactions to instructional materials, tasks, and topics of learning.
3. Decide <u>how</u> you will observe, <u>whom</u> to observe, <u>how often</u> and <u>when</u> to observe in order to collect the desired information.	<u>How:</u> Using non-participant observation by the researcher and researcher's assistant. <u>Whom:</u> The focus group who was a representative of the other three groups. <u>How often:</u> Every class of narrative tasks and their repetitions. <u>When:</u> During the focus group's planning time.
4. Select a <u>method</u>(s) for recording your observation.	<u>Methods:</u> Using the Teacher's Observation Form and VTR recorder.
5. Prepare the necessary reporting <u>forms</u>.	<u>Form:</u> See Appendix 9

The Teacher's Observation Form was developed in order to collect the data of the co-construction process of the focus group during the planning-time. It consisted of two parts: checklist items and three open-ended questions. The non-participant observation was conducted by both the researcher and researcher's assistant. Additionally, a cassette recorder was set for each group to record its participants' voice, and a camera was also turned on for recording the focus group during the planning-time to enable the researcher to get as much data as possible.

This Observation Form was validated by three experts. The criteria of selecting the experts were that one expert had an expertise in Social-Constructivism Theory, the second one in the field of Second Language Acquisition, and the last one was in the Qualitative Assessment. The first expert was a highly experienced instructor at the University of Pittsburgh, and the last two experts were also experienced English language instructors at the Language Institute of Chulalongkorn University. The IOC value for the content validity of this Teacher's Observation Form was 1.00 which showed an agreement of the experts towards the contents of the

Form. However, there were suggestions and comments from the experts to reshape this Form.

3.3.1.3.2.1. Revision

The revisions of the Teacher's Observation Form were mainly on adding some more open-ended questions such as "How?", "In what ways", and so on, in order to direct the researcher to be more skeptical to the observation data. A blank space was also added after each question to enable the researcher to make field note. The name of this form should be changed to be Teacher's Observation Form and Field Notes. This Observation Form was provided in Appendix 7, and its content validity evaluation was in Appendix 8.

3.3.1.3.3. Student's Self-Evaluation Form

The information gained from the participants themselves was very helpful in terms of triangulation of data sources because it was viewed from different perspective which could prevent the subjectiveness from the observation. As Reed and Bergemann (2005: 19) had mentioned, "If observation was to be effective in the learning process, it must be objective."

Therefore, the main objective of using this instrument was to ask them to report on what they had learned from working with their peers and how they had learned it. Another purpose was to emphasize a sense of peers relationship and a learning community that to accomplish a group task, they had to collaboratively work towards the goal.

The Student's Self-Evaluation Form consisted of open-ended questions. (See Appendix 9 for more details of Student's Form.) Its content validity value measured by the IOC Index was high at 1.00. However, there were some suggestions made by the experts for further revision.

3.3.1.3.3.1. Revision

The revisions were mostly done on the questions to be more specific and clear. For example: Instead of asking only "Why?", it was changed to be "Could you explain why this happened?" Some questions were extended to probe more on the participants, not just to ask only for "yes/no", e.g., "Did you listen and pay attention to each others?" "Yes/No". "If YES, why did you listen and pay attention to each others?", "If NO, why didn't you listen and pay attention to each others?" Lastly, the name of this instrument should be Student's Report Form. (See Appendix 10 for the content validity evaluation.)

3.3.1.3.4. Interview Questions for Co-construction Process

An additional instrument that was used with the members in the focus group, and other participants who were randomly selected from the other three groups were also interviewed after the classes were over. The objective of having this instrument was to ensure the methodological triangulation for the investigation of the co-construction process and its product.

The Interview Questions was a list of preplanned and semi-structured questions in a fixed order for greater consistency and reliability (Burns, 1999: 119). The questions were mainly emphasized on the interviewee's points of view on the problem or trouble while working with her/his group and its solution, the method they used for planning the task, the language knowledge they learned from that collaboration, the way the group worked together, and , finally, the suggested idea to improve the group's task. (See Appendix 11 for more details of the Interview Questions.) The content validity value from the IOC was very high at 1.00 with some suggestions for further revision.

3.3.1.3.4.1. Revision

The revisions were made on question 1 to 3 to elicit more replies from the participants. Details of content validity evaluation and revisions were provided in Appendix 12.)

3.3.1.4. Summary of research instruments

The summary of research instruments was shown in Figure 3.10.

Figure 3.10: Summary of research instruments

Types	Objectives	Characteristics	Validation	Means of analysis
1. Pre-course Needs Assessment	To gather information for developing learning materials & lesson plans, and to get more information on participants (Ps)	- Eng & Thai - Four parts: 1. Ask general data 2. Self-assessment 3. Choose topics 4. Ps expectations from Eng course	Content validity value of IOC was at 0.84.	Qualitative method: Content analysis and Frequency counting
2. Teaching materials	To enable Ps to practice NNT & NTR	Lesson plans & stimulus materials on 5 topics obtained from Needs Assessment	Content validity value of IOC was at 1.00.	Oral performances were analyzed by using T-unit: FLUENCY, ACCURACY & COMPLEXITY; and Descriptive Band Scales on 7 aspects.
3. Oral Proficiency Test and Descriptive Band Scales	To divide Ps into levels, and to measure Ps gained scores	2 speaking tests: 1. Interview 2. Narrative Descriptive Band Scales: 5 scales (1=lowest to 5=highest) with descriptors on 7 aspects of language learning for each test	Content validity value of IOC: 1. Oral Proficiency Test = 0.934 2. Band Scales = 1.00 Inter-rater reliability: 1. Pretest = .844 2. Posttest = .925	Descriptive Band Scales: 1. Interview test = 7 aspects[1] 2. Narrative test = 7 aspects[2]
4. Teacher's Observation Form and Field Notes	To collect data of co-construction process of focus group, and ensure methodological triangulation	- Non-participant - Two parts: 1. Checklist items 2. Open-ended Qs	Content validity value of IOC was at 1.00	Content analysis
5. Student's Report Form	To get Ps reflection on what they had learned and how they had learned it.	Open-ended Qs	Content validity value of IOC was at 1.00	Content analysis
6. Interview Qs for Co-construction Process	To investigate co-construction process and ensure methodological triangulation	Semi-structured Qs	Content validity value of IOC was at 1.00	Content analysis

---

[1] COMPLEXITY, ACCURACY, FLUENCY, PRONUNCIATION, VOCABULARY, FLEXIBILITY, and CONFIDENCE TO INITIATE SPEECH

[2] COMPLEXITY, ACCURACY, FLUENCY, PRONUNCIATION, VOCABULARY, THEMATIC DEVELOPMENT, and COHERENCE AND COHESION

3.3.2. Pilot studies

The pilot studies were conducted with the Oral Proficiency Test and one topic of each level from the teaching materials due to the following objectives:

- To identify potential problems necessary for implementation before the main study.
- To identify areas where there was a chance that Murphy's Law, which stated that what could go wrong would go wrong, might happen.
- To prepare the researcher, researcher's assistant including researcher's inter-raters the practice of how to collect and analyze the data which lead to the reliability of the study.

3.3.2.1. Oral Proficiency Test

The pilot of the OPT was conducted with twenty-one third-year English major students studying at the Faculty of Education, Chulalongkorn University. They shared the same characteristics in terms of their major, faculty, year of study as well as the university they were studying in. The reason of using the OPT in the pilot study was to see the applicability of the test questions and its rubrics to the target group of the participants for the final revision.

Each test-taker was asked to take the OPT for about fifteen minutes. Their oral production was tape-recorded for rating on a later occasion. The rating process could not be handled live because the researcher had to operate the tape-recording equipment as well as monitor the time. The test started with the interview test followed by the narrative test. For the rating process of the test, the inter-raters were trained and explained of how to use the descriptive rating scale. And if there was any disagreement in any of the test-takers' oral performance, the third inter-rater, who was a native speaker of English and instructor of English language, would come to give the final rating.

The rating process was conducted by listening to the participants' audio-recordings several times before giving the scores because there were seven aspects of language learning for each test. For the interview test, the rating had to be done on COMPLEXITY, ACCURACY, FLUENCY, PRONUNCIATION, VOCABULARY, FLEXIBILITY, and CONFIDENCE TO INITIATE SPEECH respectively. On the other hand, there were also other seven aspects for the narrative test which were COMPLEXITY, ACCURACY, FLUENCY, PRONUNCIATION, VOCABULARY, THEMATIC DEVELOPMENT, and COHERENCE & COHESION.

3.3.2.1.1. Data from the pilot study of OPT

The results of Oral Proficiency Test of the students showed that the maximum score that they could get was 4 out of 5 bands. Their mean score from rater one was 3.02 whereas it was 3.23 from the rater two. The inter-rater reliability of rating was significant at .758.

3.3.2.2. Teaching materials

One lesson of each Low and Medium score achiever level including another one for the High level were piloted in order to identify any implementation regarding the stimulus materials, planning time, class management, and framework of the narrative tasks including all other necessary instruments. Altogether, there were five thematic contents or topics that were employed for each new and repeated narrative task: two topics/lessons were for Low score achievers, another two for the Medium level, and only one topic for High level. The topics for the Low score achievers were Snack Attack and Shopping. The topics of Free Time and Travel were prepared for the Medium level. Finally, the topic of Study Time was for the High score achievers.

According to the IOC value of content validity of the teaching materials, it was shown that the topics indicated above were appropriate for that particular level and that anyone of the two topics for low or medium level was a representative of the two. Three topics that had been randomly selected to be tried out in a pilot study were Snack Attack, Travel, and Study Time respectively. The pilot studies for these three topics had not been conducted in a sequence. There was a span of time from topic 1 to topic 2 and to topic 3 since the researcher needed some times to get the results of the Oral Proficiency Test back in order to determine who would be the participants for topic 2 and 3. Moreover, during the time of conducting the OPT and pilot studies, it was the summer school break. Therefore, it was not that easy to get full cooperation from the target group. The details of the pilot study of each topic are provided separately below:

3.3.2.2.1. Topic for Low level: Snack Attack

There were five participants for the pilot study of topic 1 which was Snack Attack. Four of them were females, and one was a male student. They were third-year English major students, and they were studying at the Faculty of Education. During this time, it was not clear what score level these participants were.

The procedures were taken step-by-step as indicated in the lesson plan. After they had been informed about the topic and objectives of the lesson, they were given a

set of pictures which was shown below and ten-minute planning time. They were allowed to discuss with their friends. They were able to make some notes on a piece of paper for their individual narrative. After the planning time was over, each one of them had to narrate her/his story for about three minutes, and it was tape-recorded.

Picture 3.1: Picture for Topic 1 (Snack Attack)

One week later, the lesson plan on the topic of Snack Attack for the narrative task repetition was introduced. The procedures were followed according to the lesson plan. After the warm-up activities were over, they were informed about the topic and objectives. The same pictures used last week were distributed to each one of them. Ten-minute planning time was provided as their planning time. They were able to interact with their friends and write some notes for their repeated narrative. Each one of them was tape-recorded during their three-minute performance of the narrative repetition.

The audio-recordings of the participants' performances of both new and repeated narratives were transcribed using the modified guidelines of the transcription conventions of Allwright and Bailey (1991) and convention analysis of Perakyla (2005). After that, the transcriptions were counted for the number of T-unit, which was defined as an independent clause with or without any dependent clauses attached to it.

In counting clauses and phrases for T-units, some guidelines of Polio (1997), which were provided in Appendix 13, for errors, words, and phrases were used. Then

each T-unit was inspected for several times to look for pauses, and grammatical errors, and count for number of words per T-unit. The purpose of looking for pauses, finding grammatical errors, and counting the words was to measure the fluency, accuracy, and complexity of the participants' English oral performances. The measurement for three aspects of language performance had followed the suggestions made by Bygate (2001) which stated that:

- FLUENCY was measured in terms of the number of unfilled pauses per T-unit, that is, the higher the number of unfilled pause there was, the less fluent the talk.

- ACCURACY was calculated by the incidence of errors per T-unit, that is, the higher the number there was, the less accurate the language. The construct of measurement for the accuracy extensively covered word order, subject-verb agreements, prepositions, sentence structures, and tenses.

- COMPLEXITY or lexical complexity was measured in terms of number of words per T-unit, that is, the higher the number there was, the complex the language.

After counting the pauses, errors, and words using the T-unit, the total numbers of pauses, errors and words were used to find the means to explain a probable tendency whether there was a change in the fluency, accuracy, complexity.

3.3.2.2.1.1. Data from the pilot study of Topic 1

From the Figure 3.11, it showed that the number of pauses found in the new narrative task (NNT) decreased from 5.60 to 2.40 which was the mean of pauses in narrative task repetition (NTR). On the other hand, the mean of errors in NNT slightly increased from 14.60 to 17.40 in NTR. Finally, the mean of the number of words in NTR showed a slight increase.

Figure 3.11: Total numbers of pauses/errors/and words of Topic 1

Pauses	Mean	Std. Deviation
Pauses in NNT	5.60	5.03
Pauses in NTR	2.40	3.36
Errors		
Errors in NNT	16.40	7.20
Errors in NTR	17.40	7.99
Number of words		
Number of words in NNT	185.60	106.37
Number of words in NTR	188.40	80.91

The figures obtained from Figure 3.11 showed that there was a potential that the narrative task repetition had an effect on the fluency of the English oral language performance of the participants.

One of the participants was interviewed as to how she thought about the pilot study. She said that overall it was interesting, but the ten-minute planning time was not enough for her to plan for the narrative. Concerning the picture, it was suggested that the words and phrases showing time sequences in the picture frames should be deleted so that the narrator was able to sequence the story on their own. The pilot studies of the other two topics were conducted to see how their results would come about.

3.3.2.2.2. Topic for Medium level: Travel

Though the researcher could identify what levels the participants who had got the OPT were, it was still quite difficult to ask them to attend the pilot study since it was the school break. What the researcher could do was to be able to get one participant for each pilot study of Medium and High score achiever levels. However, the information gained from them was satisfactory relating to the purposes of conducting the pilot studies.

The participant who attended this pilot study was a Medium score achiever. She was one of the participants who participated in the pilot study of Topic 1. She was third-year English major student in the Faculty of Education, Chulalongkorn University.

The lesson plans and teaching materials of both NNT and NTR for the topic of Travel were introduced to the participant. There were eleven frames for the picture used for this level. This time the planning time was extended to be fifteen minutes

instead of ten minutes. The participant's audio-recordings were transcribed. Then it was analyzed to separate into the T-units and counted for pauses, errors, and number of words.

Picture 3.2: Picture for Topic 4 (Travel)
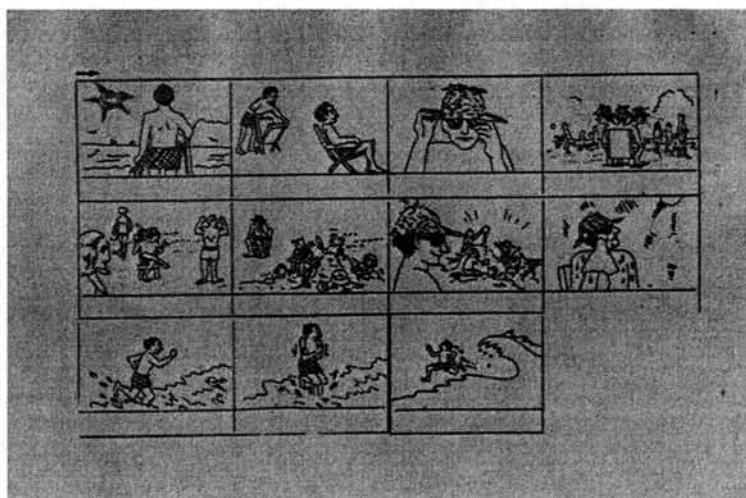
3.3.2.2.2.1. Data from the pilot study of Topic 4

The numbers of pauses, errors, and words of both new and repeated tasks were summarized in Figure 3.12. It showed that the pauses in her performance of the repeated narrative decreased from 8 to 3. Moreover, the grammatical errors in her NTR dropped from 27 to 11. On the other hand, the total number of words in her repeated narrative did not increase.

Figure 3.12: Total numbers of pauses/errors/and words of Topic 4

Types of narratives	Total pauses	Total errors	Total number of words
New	8	27	228
Repetition	3	11	215

After the pilot study, the participant had been interviewed. She expressed her opinion that the task was more complicated than last time because there were more picture frames for her to look through and came up with a narrative about the pictures. However, she was fond of the topic because she liked traveling. Her opinion corresponded with the experts' agreement on the content validity of the teaching materials, particularly, the increasing number of picture frames to put more cognitive load for medium level of score achievers. When she was asked about the

planning time that had been extended from ten minutes to be fifteen minutes, she was contented with it because ten-minute planning time was too short.

3.3.2.2.3. Topic for High level: Study Time

The only topic for the high score achievers was Study Time, and it was also piloted with one participant who got high score from the OPT. The purpose of trying out this topic was to see the reaction of the participant towards the pictures that had been increased to be sixteen frames.

The participant who attended this pilot study was a male third-year English major student studying in the Faculty of Education. He was a high score achiever of the OPT. The lesson plan and teaching materials of both NNT and NTR for the topic of Study Time were tried out with him. He was allowed fifteen minutes to plan for his narratives. There were sixteen frames of the pictures for this topic.

Picture 3.3: Picture for Topic 5 (Study Time)

3.3.2.2.3.1. Data from the pilot study of Topic 5

The participant's audio-recording was transcribed and calculated for pauses, errors, and number of words which were shown in Figure 3.13. The total number of pauses and errors in the repeated narrative decreased which was the same as the participant in the pilot study of Travel.

Figure 3.13: Total numbers of pauses/errors/and words of Topic 5

Types of narratives	Total pauses	Total errors	Total number of words
New	24	35	383
Repetition	4	19	322

The participant was also interviewed after the pilot study was over. When he was asked about the picture, he said there were so many frames that he had to pay a lot of attention to make sure that he covered all contents from the pictures when he was planning for a story.

3.3.3. Final revision

The information gained from the pilot studies not only implied preparedness, but also a final revision of the instrument. The first thing to be prepared was the tape-recorder. It was mal-functioned when it was in need to be used. A spare tape-recorder should be prepared in case of the Murphy's Law. The second thing was that it was better to prepare a planning sheet for the participants to plan for their narratives. Last but not least, the researcher's assistant was needed in the classroom at all time for the observation because the researcher could not manage a dual role of both the teacher and observer well enough to ensure a thorough observation.

For the revision, the first one was about the planning time which was extended to be fifteen minutes instead of ten minutes. Moreover, the criteria of counting the T-units for the fluency, accuracy, and complexity as well as the conventions of transcription were revised. The revised criteria of counting the T-units were as follows:

FLUENCY was calculated in terms of the number of unfilled pauses (for every one second) indicating by a sign of (+) per T-unit. The highest the number of pause it was, the less fluent the speech. However, fillers and false starts such as er, and um were not counted.

ACCURACY was calculated in terms of incidence of error per T-unit. The higher the number of error it was, the less accurate the speech. Errors included the incorrect use of articles, tenses, subject-verb agreements, phrasal verbs, prepositions, and vocabulary. The repetitions of words, grammatical errors, or slips of tongue before immediate self-corrections were not counted as errors.

COMPLEXITY was calculated in terms of the number of grammatically correct words per T-unit. The higher the number of words it was, the more complex the speech. However, the repetition of words or grammatical errors made before the self-corrections were not counted as well as words that were used to interact with the audience, e.g., ok, ah ha, and so on.

For the conventions of transcription, they were provided in Appendix 13.

To conclude, the data from the pilot studies showed that there was a tendency that the narrative task repetition would have effects on the English oral language performance of Thai learners, but there were many questions needed to be investigated to confirm these effects. Therefore, it was necessary to conduct a main study to carefully investigate as to what effects the narrative task repetition had on the English oral language performance of Thai learners. However, there were also many other things to prepare before the data collection of the main study which were the handling of the participants in heterogeneous roles and groups, the participants' consents as well as the long-range lesson plans for the whole study.

3.4. Data collection: Main study

Before the data collection for the main study was conducted, some preparations concerning the class management and ethical issue had to be handled.

3.4.1. Preparations

3.4.1.1. Class management: Groups and roles

The reason of grouping all twenty participants into small groups (see Figure 3.3) was because the research into the effectiveness of solutions developed by small groups had shown that groups tended to do much better when several alternative solutions for the problem were possible (Adams & Galanes, 2006: 7).

Concerning the roles assignment, an individual member in each group was assigned a certain role as a responsible member (Woodward, 2001: 219-220): a *chair*, *time-keeper*, *monitor*, *note-taker*, and *reporter* respectively. Each role was rotated for each new topic, or every two weeks. The responsibility for each role was as follows (Gray, 2000: 164):

- Chair: This person was responsible for leading group discussion. She or he should make sure that everyone was included in the discussion.
- Time-keeper: This person reported the time left for the finishing the task.
- Monitor: This person kept the group on the task and controlled the use of English as much as possible.
- Note-taker: This person took notes of the details from the discussion.
- Reporter: This person narrated to the class.

3.4.1.2. Getting consents

Since the study was conducted in the classroom of "Effective English Speaking" which involved classroom experience and discourse of the learners, the consent forms together with the explanatory letters were distributed to all of the participants on the first day of class. All twenty of them had signed and returned the consent forms to the researcher. Below is a summary of information given in the consent forms: (Conducting research [Online])

- A description, purpose, and duration of the study
- A statement of how to protect their confidentiality
- The name of a contact person who is responsible for the study
- A list of benefits that this study could contribute to the English education

The consent form and an explanatory letter were provided in Appendix 21.

3.4.2. Conducting Pretest

The Oral Proficiency Test (OPT) was used to pretest all twenty participants before the beginning of the classes. The Pretest was conducted in the recording studio of the Faculty of Education. The test took about fifteen minutes for each person. The purpose of pretesting them was to be able to group them in four groups of five mixed English oral proficiency which were Low, Medium, and High score achievers. The second important purpose was to have this individual Pretest score to be a baseline to compare the gained score from their Posttest when the course was over.

Concerning the purpose of dividing them into different levels, their scores were arranged and the Percentile was used to separate the low and high score achievers. Tirakanun (2003: 63) had suggested that when the curve was skewed, it is more appropriate for the researchers to use the Percentile to divide the data into groups, e.g., to use the ranks of 30th Percentile and 70th Percentile. The data below

30th Percentile was in the low group while the data above 70th Percentile was in the high group, and the data in between the two was in the medium group. More information regarding the participants' Pretest scores, ranks, and levels computed by the Percentiles were provided in Chapter 4.

3.4.3. Conducting non-participant observations

While the participants were working in their groups for their narratives, the researcher's assistant would observe only the focus group, complete the checklist, and write down the field notes in the Observation Form. The researcher had to try to avoid a threat to internal validity by observing every group so the other three groups would think that they were not neglected. The fact that the focus group was given more attention for the observation was not disclosed to its members, and they did not know that they were the selected focus group.

3.4.4. Having the planning discussions audio- and video-taped

The audio recording was done with all groups while they were planning for their narratives with their peer members. Each group was given a small tape-recorder, and a "monitor" was asked to take care of the tape-recorder to make sure that the volume and recording buttons were at the right positions.

Besides the audio-recording, the discussion during the planning time of the focus group was also VDO taped to make sure that the data could be collected as much as possible. By doing this, it could ensure the methodological triangulation. And for the VDO recorder, it had been left since the very first day of the class though the narrative lessons were not introduced yet to let them get used to having a camera in the classroom.

3.4.5. Having the participants' narrative performances audio- and video-taped

When the participants who had been assigned a role of "reporter" performed their new narratives or repeated them, their performances were both audio-taped and VDO taped.

3.4.6. Having the Student's Report Forms filled in

After the discussion and reflection sessions in the Post-task stage had been completed, each participant was asked to fill in the Student's Report Form and submitted it before leaving the class.

3.4.7. Interviewing some participants

While the participants were filling in the Student's Report Forms, some participants were randomly selected to be interviewed in order to obtain the data of the co-construction process. The interviews were conducted in every class.

3.4.8. Conducting Posttests

The same Oral Proficiency Test which had been used as a Pretest was reshuffled to be used as a Posttest in order to avoid the retention of memory of the participants. The Posttest was sequenced to begin with the narrative test task and followed by the interview test. It was conducted after the completion of the classes by the end of the semester. The purpose of conducting the Posttest was to measure the gained score from that of the Pretest to see the improvement of the participants' English oral performance.

3.5. Treatment of data

Before further investigation, the raw data had to be prepared and treated. Those were the audio-recordings and VDO recordings of the participants who performed the new narrative tasks and narrative task repetitions including the audio-recordings and VDO recordings of the focus group during their planning-time.

3.5.1. Data of narrative performances

The audio-recordings and VDO recordings of the English oral performance of new and repeated narratives of the participants were transcribed using the following conventions as shown in Figure 3.14.

Figure 3.14: Convention of Transcription

/	indicated the cutting point for each T-unit
+	indicated a pause of one second
X	indicated an incomprehensible or inaudible word
XXX	indicated an incomprehensible or inaudible phrase length
( )	indicated a translation of a word or a phrase spoken
{ }	indicated a description or an explanation of an action
(!)	indicated an interruption
w0rd	indicated an emphasis
wo:rd	indicated a prolongation of sound
£word£	indicated smile voice
wo(h)rd	indicated laugh particle inserted within a word
((word))	indicated transcriber's comments

The transcriptions were then being treated in two dimensions: quantifiable and qualifiable dimensions.

3.5.1.1. Quantifiable dimension

According to Scholfield (1995: 4), quantification was a term for any way in which aspects of language or speaker are observed and turned into numbers by measurement or categorization or any other means. Hence, the data of new and repeated narratives were measured by using the T-unit on three different measurements which were fluency, accuracy, and complexity.

The T-unit was defined as an independent clause with or without any dependent clauses attached to it. After the T-unit had been used to break the transcriptions into clauses, the following measurement would be taken to investigate the FLUENCY, ACCURACY, and COMPLEXITY of the English oral performance of the participants' narratives. The operationalized definitions for each aspect of language learning had been modified from those of Bygate (2001) and results from the researcher's pilot studies.

-   FLUENCY was calculated in terms of the number of unfilled pauses (for every one second) indicating by a sign of (+) per T-unit. The higher the number of pause it was, the less fluent the speech. However, fillers and false starts such er, um, and so on were not counted. Though other Second Language Acquisition researchers had thought that pauses may not represent dysfluency (Fillmore, 1979; Fathman, 1980, cited in Lennon,

1990: 393), the researcher viewed unfilled pauses as the indicators of dysfluency because there was fifteen-minute Pre-task planning time provided for the participants to mentally rehearse and word collaboratively with their peers before their narration. Therefore, the ability to produce the target language under time pressure with smoothness should be applicable, especially, in the case of topic or content familiarity to the speakers.

- ACCURACY was calculated in terms of incidence of error per T-unit. The higher the number of error it was, the less accurate the speech. Errors included the incorrect use of articles, tenses, subject-verb agreements, phrasal verbs, prepositions, and vocabulary. The repetitions of words, grammatical errors, or slips of tongue before immediate self-corrections were not counted as errors.

- COMPLEXITY was calculated in terms of the number of grammatically correct words per T-unit. The higher the number of words it was, the more complex the speech. However, the repetition of words or grammatical errors made before the self-corrections were not counted as well as words that were used to interact with the audience, e.g., ok, ah ha, and so on.

3.5.1.2. Qualifiable dimension

Though the quantification dimension of language improvement analysis was popular among the scholars and researchers, there were other interesting aspects that figures or statistics could not explain clearly. Therefore, to measure the language data in the qualifiable dimension could explain what was happening in the narrative performance of the participants from another perspective.

The transcriptions and audio-recordings of the new and repeated narratives were analyzed by using the same descriptors which were used for the narrative test in the OPT. These descriptors were arranged into seven categories:

- COMPLEXITY or SYNTACTIC COMPLEXITY involved a variety use of sentence structures.

- ACCURACY involved a correct use of word order, subject-verb agreements, preposition, and tenses.

- FLUENCY involved a natural smooth flow of oral language performance.

- PRONUNCIATION involved a clear, natural, and intelligible stress and intonation.

- VOCABULARY involved sufficient lexical repertoire including idiomatic expressions and colloquialism.
- THEMATIC DEVELOPMENT involved a clear description of the story and its details with an appropriate conclusion.
- COHERENCE and COHESION involved a well-controlled use of organizational patterns, connectors, and cohesive devices.

3.5.2. Data during planning-time

The audio-recordings and VDO recordings of the planning discussion of the focus group for Topics 2, 3, and 5 (Shopping, Free Time, and Study Time) were transcribed according to their verbiage without making any correction for their grammatical errors.

After the completion of transcriptions, this fifteen-minute planning time was compared against the "Punctuated Equilibrium Model", which was a model of group development suggested by Gersick (1988, cited in Burn, 2004: 183-185). It was the development of a small group working on a task with a specific deadline. The reason of comparing the focus group's development with Gersick's Model was because they shared some characteristics in common. Both of them were small group work, and they were assigned to work on a task within a time deadline. However, a modification had been made in order to fit the working cycle of the focus group within its planning time. The modified model of the group development of the focus group was provided in Table 3.7.

Table 3.7: Modified model of group development

Phase 1: Start (5 minutes)	Phase 2: Midpoint (5 mins – 7 mins)	Phase 3: Get Ready (5 mins - 3 mins)
A period where the group members start discussing and planning for their task after being informed of the task, its topic, its framework, and the members' roles.	A period where the group members are sharing their ideas to build up the content of the narrative task according to its framework and making an agreement about how to complete the task. The sense of urgency to finish planning the task on time is shared among the members.	A period where the members are putting pieces of information to prepare for the final narrative presentation.

The transcription of each of the focus group's planning discussions for its new and repeated narratives was divided into three Phases: Start, Midpoint, and Get Ready Phases. As suggested by LeCompte and Preissle (1993) on the "theorizing" technique and by McGrath and Altermatt (2003) on the group interaction coding systems, the VDO recordings were reviewed and the audio-recordings were played back several times to find the patterns which displayed the focus group's co-construction process. The data from the Teacher's Observation Forms and Field Notes as well as from the Student's Report Forms were also used as the supplementary data for exploring any specific patterns occurring in the co-construction process. After that, the literatures and empirical studies on peer or social interaction (Lee, 2004; Foster & Ohta, 2005: Cao & Philp, 2006; Pinter, 2006) and negotiation of meaning or content (Branden, 1997; Rulon & McCreary, 1986) were reviewed in order to find any resemblance between the phenomena of this study and other studies as well as to support the coded categories.

These categories were then validated by three experts for the content validity. The first expert was the famous lecturer and well-known author in a field of Social Constructivism Theory. The second and the third ones were the experts in an area of Second Language Acquisition. The overall content validity (IOC) value was .88. Details of the content validation were provided in Appendix 15. The suggestions given from the experts were used to implement these coded categories.

There were altogether nine categories which were employed to code the co-construction process of the focus group. Their names and definitions were as follows:

Category 1: *Continuers* (modified from Foster and Ohta, 2005) referred to non-verbal or verbal expressions a speaker used to *encourage* his/her peers *to contribute more ideas* in their collaborative work. For example: "yes", "ah ha", a complimentary encouragement like "that's a good one", or a phrase such as "go ahead" which is sometimes accompanied by a body gesture to indicate his/her peers to continue speaking. It was coded as **CON**.

Category 2: *Other-initiated suggestion/comment* (modified from Lee, 2004) referred to expressions in which a speaker used to *suggest or comment on her/his own initiative* on what other members had been discussing in order to indicate her/his opinion. For example: "I think...", or the speaker's own expressions of her/his own suggestion/comment on that context. For example: "How about...?" It was coded as **OISC**.

Category 3: *Other-initiated correction* (modified from Lee, 2004) referred to expressions in which the listener provided more *accurate pieces of information on her/his own initiative,* in terms of content, vocabulary, and grammar for the speaker. It was coded as **OIC**.

Category 4: *Appeal for ideas or assistance* (modified from Lee, 2004) referred to expressions when *a speaker asked for the contribution of content,* characterized by "Wh- questions" or expressions such as "What about you?", "What do you think?", and so on, or *for the correct word or spelling,* characterized by expressions such as "What is it called?" or "How do you spell it?", and so on. It was coded as **AIA**.

Category 5: *Helping out* (modified from Lee, 2004) referred to *expressions* which can be a word, phrase or whole sentence; or an idea or suggestion *provided by a listener after being asked by one peer member to help contribute a content or solve a problem of a word choice or spelling as well as to help complete another member's sentence.* It was coded as **HO**.

Category 6: *Content confirmation checks* (modified from Rulon & McCreary, 1986) referred to expressions made to *confirm the speaker's previous utterance* which were characterized by the rising intonation or the questioned-word form. It was coded as **CONC**.

Category 7: *Content clarification checks* (modified from Rulon & McCreary, 1986; Foster & Ohta, 2005) referred to expressions made to *elicit clarification of the speaker's previous utterance* which were characterized by "Wh-questions", "Yes-No questions", tag questions, repetitions of all or part of the speaker's preceding utterance(s) with rising intonation or other utterances such as "Do you understand?", "All right?", and so on. It was coded as **CONCLA**.

Category 8: *Content confirmation or clarification* (modified from Rulon & McCreary, 1986; Foster & Ohta, 2005) referred to non-verbal or verbal *expressions* which can be a body gesture, word, phrase or whole sentence provided by a speaker *to confirm or clarify the question being asked* by one peer member, or to *confirm the previous utterance of another peer speaker as to show support or agreement* on what s/he had said. It was coded as **CCC**.

Category 9: *Other-initiated paraphrase* (modified from Lee, 2004) referred to words, phrases or sentences that a *listener, on her/his own initiative, repeated or expanded upon another speaker's utterance.* It was coded as **OIP**.

3.6. Data analysis

After the data had been treated, they were then analyzed according to the probing questions. In order to explain clearly of how the data were analyzed in order to answer their related research questions, each question was restated before the explanation of the data analysis.

1. <u>Research question 1</u>: Would the Posttest scores of the subjects be significantly higher than those of the Pretest?

The ratings of the Pretest and Posttest were conducted by the researcher listening to their audio-recordings. After that, the ratings were re-rated by the researcher's inter-rater, and the inter-rater reliability was calculated. Since the subjects in this study were small in number, the scores of the Pretest and Posttest were tested to see whether the assumption of normality or equality of variance was met. If it violated the assumption of normality, the Wilcoxon signed-rank test, a non-parametric test, would be used to measure the difference between the Posttest and the Pretest. On the contrary, if it did not violate the assumption, the *t*-test for dependent means would be used for the measurement. The accepted level of significance for the result of measurement would be ≤ 0.05 (one-tailed).

2. <u>Research question 2</u>: Would there be a significant difference in the fluency of the subjects' English oral language performances between new narrative task and narrative task repetition?

3. <u>Research question 3</u>: Would there be a significant difference in the accuracy of the subject's English oral language performances between new narrative task and narrative task repetition?

4. <u>Research question 4</u>: Would there be a significant difference in the complexity of the subject's English oral language performances between new narrative task and narrative task repetition?

Since any one of the two topics for each level was the same in terms of content validity value, the new narrative tasks (NNT's) and narrative task repetitions (NTR's) of Topics 2, 3 and 5 were selected for the data analysis. Previously, it had been planned to use the data collected from the Topics 2, 4 and 5 to give a chance to the participants to try out one topic of different level, but the data from Topic 4 had to replaced by Topic 3 because they were contaminated by one participant who had planned her narrative task repetition in advance before the class.

After the transcribed data of NNT's and NTR's of the three topics had been treated, the pauses, grammatical errors and number of grammatical correct words were counted for fluency, accuracy and complexity. The counting of pauses, grammatical errors and number of words was randomly checked again by an inter-

rater who was a native speaker of English and an experience English language teacher.

Chi-Square was used to test the differences of the numbers of pauses, grammatical errors, and words between the new and repeated narrative tasks. As mentioned by Connor-Linton (2007), that the hypothesis tested with Chi-Square was whether or not two different samples (of people, texts, whatever) were different enough in some characteristic or aspect of their behavior that we could generalize from our samples that the populations from which our samples had been drawn were also different in the behavior or characteristic.

5. Research question 5: What are the descriptive qualities of English oral language performances of the subjects performing new narrative task and narrative task repetition?

The audio- and/or video recordings of the selected participants in all four groups (L2 + M1 + H x 4 groups) performing the NNT and NTR were played back and listened to several times and compared with the transcriptions. The purpose was to analyze them against the descriptive band scales to measure their English oral performances on seven areas which were COMPLEXITY, ACCURACY, FLUENCY, PRONUNCIATION, VOCABULARY, THEMATIC DEVELOPMENT, and COHERENCE AND COHESION respectively.

6. Research question 6: How do the subjects in the focus group co-construct their English language knowledge when they were doing their new narrative task and narrative task repetition?

The transcriptions of the focus group's planning time for the NNT and NTR of the Topics 2, 3, and 5 were coded using the implemented categories. After that, a few of the participants in the focus group were interviewed to confirm whether what she or he had said during the planning-time was really a helping out for the group's members, a content confirmation check, and so on. Then the coded categories were compared while watching the VTRs and listening to the audio-recordings for several times in order to explore the whole co-construction process in each phase of group development. The purpose of exploring the whole co-construction process was to explain any dominant pattern in each phase of the group development.

7. <u>Research question 7</u>: What do the subjects in the focus group learn from the co-construction process while they were planning the narrative tasks and narrative task repetitions?

The transcriptions of the focus group's planning time were reviewed and analyzed. Additionally, the data from the Student's Report Forms, Teacher's Observation Forms and Field Notes, and the interviews were inspected to see what had emerged from the process of co-construction in order to find out what the participants had learned from working together in their group trying to accomplish the narrative tasks.