

Chapter Three

Method

This chapter presents the method of this study. It consists of four sections. The first section is a brief description of the participants. The second and third sections describe the research design, including the instruments and the procedures of data collection. The final section presents the methods of data analysis.

3.1 Participants

In this study, sixty-nine female students were recruited from two classes of Chayi Girls' Senior High School in Chayi City. Among them, twenty-nine were from the second-year Mathematically-Gifted Class and the other forty from a normal third-year class. The minimal English learning experience of the participants was five years, and the age range was from 16 to 18. Additionally, based on their raw scores on the reading section of a General English Proficiency Test—Intermediate Level, the participants were further divided into three levels: High, Mid, and Low English proficiency levels (see Table 3.1).

3.2 Treatment

The materials for treatment are as follows: (1) materials for instruction of text structure, (2) a recorded think-aloud modeling on Article 1 of the pre-test, (3) materials for the think-aloud technique.

The materials for instruction of text structure were based on the literature review in the previous section (see Appendix A), on the rationale that text structure, especially cohesive relations among sentences in the text, should be explicitly taught to students to enhance their reading comprehension. Also, since the “Discourse Structure” test aims to test the understanding and command of text structure, the two most essential elements of a well-structured text—unity and coherence—were

introduced to the participants. Some extra practices that required the students to recognize topic sentences and delete irrelevant ones were presented to help consolidate the concept of paragraph unity. Besides, the concrete manifestations of coherence in text—cohesive devices—were also taught to assist students in better understanding of coherence, followed by practices to reinforce the target concepts. Halliday and Hasan's (1976) model of categories of cohesive devices and examples to demonstrate them were adopted as the teaching material. Besides, de Beaugrande & Dressler's (1981) idea of regarding Tense/Aspect as a tie signaling relationships was included, too. Furthermore, a reference list of grammatical cohesion mainly borrowed from Chang (1997) and Liang (1997) was also included in the teaching materials. The instruction session took two 50-minute class periods.

After the instruction session, a recorded think-aloud modeling of taking the "Discourse Structure" test was played to the participants. The recorded think-aloud modeling was designed by the researcher herself for this study on the rationale that mental modeling through think-aloud can be critical in making explicit the invisible mental process during reading (see Appendix B for the script). Article 1 of the pre-test was used as material for think-aloud modeling since the participants would have just taken the pre-test and had fresh memory of how they took the DS test while listening to the recorded modeling, and therefore, were supposed to pick up the strategies demonstrated in the modeling more readily. In the modeling, the targeted strategies of taking a DS test were explicitly verbalized. That is, the unity of a paragraph and cohesive relations among sentences were pointed out either as clues for choosing an answer or as evidence to confirm an answer. To make a rounded modeling for reading comprehension, other reading strategies such as guessing, prediction and inference were used spontaneously but not emphasized. Besides, since the think-aloud technique is reader-based, the prior knowledge of the modeler was spontaneously

activated as well, as the Schema Theory suggests what reading is like (see 2.1). The speed of the modeling was deliberately made slower than that of ordinary speech, and the participants were encouraged to learn from the modeling what the think-aloud was and how the concept of unity and cohesive devices were utilized as clues to take a DS test. The recording was played only once.

Subsequently, the participants practiced the think-aloud technique with the DS test. The materials for the think-aloud technique contained four articles in the “Discourse Structure” test mode (see Appendix C). In total, four DS tests for think-aloud practice were administered in four class periods extending over two weeks. The present thesis, aiming to explore how the participants took the “Discourse Structure” test by detecting cohesive relations in the text, adopted the focused instructions. That is, the participants were asked to report what they were thinking about the cohesive clues they were coming across or searching for while reading. Also, since the criterion of “unity” in paragraph development was emphasized in this study, the participants were encouraged to read through a paragraph first and then commented on what they were thinking about at any time or when they reread each sentence of the paragraph. Therefore, the procedure this present study adopted for the think-aloud technique is focused instructions but not strictly sentence-by-sentence: The participants were told to say out loud what they were thinking about after reading aloud whatever length of the passage. Literally, no time limit was set; however, most participants finished each DS test through think-aloud within twenty minutes. For the regular third-year class, the researcher remained in the classroom during each practice session without interrupting the procedure. For the Gifted class, since four participants were asked to have their think-aloud protocols recorded with mini tape recorders in an adjacent classroom (see 3.4), the researcher monitored both classrooms to see if there were any mechanical problems.

The selection of the articles and the design of the tests were based on Yeh's (2002) description of the criteria for creating tests of "Discourse Structure." Specifically, each article contained approximately fifteen sentences, ranging from fourteen to eighteen, with the average article length being from 196 to 226 words. In each article, five sentences, exclusive of the first and the last sentences, were deleted from the text as items for alternative. Besides, the articles were of moderate difficulty level (see Table 3.1). First, word difficulty level was controlled with less than six words beyond the 4,000 word frequency level in each article. Second, the complexity of sentence structures was carefully analyzed so that each article contained mostly simple sentences, with no more than one compound sentence, six complex sentences, and one compound-complex. As for the sources of the articles, Articles 2 and 4 were adopted from the 2002 Department Required English Test (DRET), and the other two were adapted from articles in mock examinations for DRET takers³. Most importantly, two native speakers were asked to examine the coherence and logic of the articles so that no ambiguity would arise.

³ Articles 1 and 3 were adapted from two passages in Lungteng versions of DRET mock examinations intended for third-year senior high school students who would take the DRET.

Table 3.1

The Lengths and Difficulty Levels of the Articles for think-aloud practice

	No. of words Beyond the first 4000 most frequently used	No. of sentences	No. of simple sentences	No. of compound sentences	No. of complex sentences	No. of compound complex sentences	No. of words	topic
Article 1	2	14	7	1	6	0	196	history
Article 2	0	18	11	0	6	1	207	fiction
Article 3	6	14	9	0	5	0	189	health
Article 4	3	15	12	0	3	0	226	invention

3.3 Instruments

Four instruments were used in this study to collect data: (1) the reading section of a General English Proficiency Test: Intermediate Level, (2) pre-tests and post-tests in the Discourse Structure mode, (3) a metacognitive awareness questionnaire prior to and after the integrated treatment, (4) a questionnaire on each individual participant's responses to the instruction of text structure, think-aloud modeling and the think-aloud technique.

The General English Proficiency Test (the GEPT), developed by the Language Testing and Training Center, is the only test specifically designed to evaluate the English proficiency of the Taiwanese people. Its intermediate level targets mainly at people who have completed their senior high school education. It was therefore adopted as the placement test for the participants, who were supposed to be more or

less at this English proficiency level. The GEPT, in fact, contains four subtests, listening, reading, speaking and writing. However, only the reading subtest was administered since issues regarding reading are the main concern in this study.

Table 3.2 is the result of the participants' performance on the reading section of a GEPT mock test, showing that the sixty-nine participants were roughly evenly divided into three groups of different proficiencies, 22 for High, 24 for Mid and 23 for Low. The mean scores are 98.1, 74.3, and 54.9 for High, Mid and Low respectively, when the total score is 120.

Table 3.2

The results of the GEPT Reading Test (N=69)

High				Mid				Low			
ID	GEPT Reading	ID	GEPT Reading	ID	GEPT Reading	ID	GEPT reading	ID	GEPT reading	ID	GEPT reading
1	111	12	99	23	84	35	72	47	63	59	57
2	108	13	96	24	84	36	72	48	63	60	54
3	108	14	96	25	84	37	72	49	63	61	54
4	108	15	93	26	84	38	69	50	63	62	51
5	105	16	93	27	81	39	69	51	63	63	51
6	102	17	93	28	78	40	69	52	60	64	51
7	102	18	90	29	78	41	69	53	60	65	48
8	102	19	90	30	78	42	69	54	60	66	48
9	99	20	90	31	78	43	69	55	60	67	45
10	99	21	87	32	75	44	69	56	60	68	42
11	99	22	87	33	75	45	66	57	57	69	33
				34	72	46	66	58	57		
Mean:			98.1	Mean:			74.3	Mean:			54.9

As for the pre-test and the post-test, each included four articles in the Discourse Structure test mode (see Appendix D). All the articles were selected, like those for think-aloud practice, according to the criteria for designing the Discourse Structure test specified by Yeh's (2002). As shown in Table 3.3, both tests were designed in such a way as being comparable in difficulty in terms of article length, word difficulty and complexity of sentence structures. In addition, the difficulty level of the articles in each test varied to avoid the ceiling effect, so that actual individual difference could be reflected in the test results. Like the articles for think-aloud practice, the articles for both tests were examined by two native speakers, based on whose suggestions a few changes were made so that coherence and logic were ensured for all the articles. As for the sources of the articles, Article 4 of the pre-test and Article 3 of the post-test were adopted from Kletzien (1992), Article 3 of the pre-test from *Guided Practice for Reading Comprehension* (2002), and the others were adapted from articles in different versions of DRET mock examinations. With regard to the scoring, a correct answer was given one point, with a possible maximum total score of 20 for each test. The pre-test and post-test scores were compared to probe the effects of the integrated treatment on the performance of the DS test among the participants with different reading abilities.

Table 3.3

The Lengths and Difficulty Levels of the Articles for Pretests and Posttests

Pre-test articles								
	No. of words Beyond the 4000 frequency level	No. of sentences	No. of simple sentences	No. of compound sentences	No. of complex sentences	No. of compound complex sentences	No. of words	topic
Article 1	1	16	13	1	2	0	181	theater
Article 2	4	13	6	1	5	1	211	health
Article 3	6	19	10	0	8	1	251	science
Article 4	5	16	9	1	6	0	264	value
Total	16	64	38	3	21	2	907	
Post-test articles								
	No. of words Beyond the 4000 frequency level	No. of sentences	No. of simple sentences	No. of compound sentences	No. of complex sentences	No. of compound complex sentences	No. of words	topic
Article 1	0	15	11	1	3	0	172	animal
Article 2	5	13	6	1	6	0	226	health
Article 3	6	17	7	0	10	0	241	popula- tion
Article 4	6	18	11	0	7	0	271	collection
Total	17	63	35	2	26	0	910	

In this study, the same questionnaire was used to tap the participants' metacognitive awareness in taking a DS test prior to and after the treatment to see if there was any change caused by the treatment. The questionnaire employed in this study made reference to Carrell (1989), Yang (1996) and Hsu (2002) for questions on reading strategies. In addition, questions concerning the role of cohesive devices in solving or causing difficulties in taking the DS test were added. The questionnaire contained two sections: the first section aimed at probing the participants' awareness of the strategies they used to take a DS test; the second section tried to tap the participants' awareness of the difficulties they encountered in taking a DS test. All the answers in the questionnaire were constructed in a five-point scale ranging from "strongly agree," to "strongly disagree." All the statements in the questionnaires were presented to the participants in their first language, Mandarin Chinese (see Appendix E).

Another questionnaire designed by the researcher was used to explore the participants' responses to the three stages of treatment, namely, the instruction of text structure, think-aloud modeling, and think-aloud practice (see Appendix F). Items 1 to 5 contained both a yes-no question concerning the effect of the treatment and an open question probing the areas of the effect. The answers of the open questions were categorized for discussion. Items 6 to 10 were constructed in a five-point scale ranging from "strongly agree," to "strongly disagree." And the last question aimed at probing the degree of the participants' perceptual effect of the three stages of treatment mentioned above.

3.4 Data-collection Procedures

The procedures for data collection of this study are as follows: (1) a placement test, (2) the pre-test and the metacognitive awareness questionnaire prior to the treatment, (3) the recording of think-aloud protocols, (4) the post-test and the

metacognitive awareness questionnaire after the treatment, and (5) students' response questionnaires.

Since the participants came from two classes, the data were collected separately, and the interval for the collection of equivalent data from the two classes was within a week for some practical and unexpected reasons. Besides, the researcher, who was also the English teacher for the two classes, did the instruction and collected all the data herself.

In early October 2002, all the participants did the reading section of a mock GEPT-Intermediate Level to have their English reading proficiency determined. Then, within the same week, the pre-test and the metacognitive awareness questionnaire were administered to attain basal data of the participants' performance on the DS test and their metacognitive awareness of the strategies they used and the difficulties they encountered in taking the DS test. Afterwards, the integrated treatment of instruction of text structure and think-aloud modeling followed by think-aloud practice was administered (see 3.2). During think-aloud practices, four participants from the Gifted class were recruited for the recording of think-aloud protocols. One of them got the highest score on the placement test (111) and will be referred to as Highest in the following discussion. The others were from each level of reading proficiency—High, Mid, and Low, and will be referred to as such in the following discussion. The scores on the placement test for High, Mid and Low were 96, 78, and 60 respectively, resulting in intervals of 15, 18 and 18 for the scores obtained by these four participants. These three participants were selected from those with moderate performance on the placement test in their groups so that they could be better representative of their groups. Also, they were more verbal than their counterparts, so that intended information from the think-aloud protocols could be expected. Their recordings were done in an adjacent classroom from their regular classroom, where

their classmates were doing thinking-aloud for the same DS test. The researcher monitored both classrooms to see if there were any mechanical problems. The participants from the Gifted class were recruited for the recording mainly because of administrative convenience and their familiarity with the researcher.

Immediately after the four think-aloud practice sessions, all the participants were given a post-test in the DS test mode, which, as mentioned above, was comparable to the pre-test in passage length and difficulty index.

After the post-test, all the participants were asked to fill out two questionnaires. One, the same as was administered prior to all treatments, focused on the change of the participants' metacognitive awareness of how they took a DS test. The other questionnaire was to reveal the participants' response to the instruction, think-aloud modeling and think-aloud practice.

3.5 Data analysis

Quantitatively, a statistical program, SPSS for Window t-test, was applied to analyze the collected data, including the results of the pre-and post- tests, and those of the pre-and post-activity questionnaires.

Besides, the tape-recorded think-aloud protocols by four participants, sixteen passages in total, were transcribed verbatim for analysis (see Appendix G). Mainly, the frequencies of each category of cohesive ties that each participant pronounced to use as clues in doing the DS test were calculated for comparison. Moreover, to answer the third research question, data collected from the participants' response questionnaire were sorted and categorized to explore the participants' responses to the instruction of text structure, think-aloud modeling and think-aloud practice with the DS test⁴.

⁴ All the codings, both for the transcripts and for the response questionnaires, were done by the researcher and her sister Ms. Li-wen Chen, who is an instructor at the Department of Foreign