

I. Introduction

Background and Purpose of the Study

Interpretation, an activity undertaken as early as peoples of different tongues first met and communicated, has “just faced its professionalization and standardization in these decades” (Zhou & Chen, 1987, p. 9). It was not put into frequent use until the end of World War I, when interpretation has become an indispensable assistance in diplomatic affairs between countries (Liu, 1993, p. 2). Today, convenient transportation tools and cutting-edge technology greatly facilitate interactions between countries; yet language barrier still exists. Interpretation, therefore, is in great demand by people from all walks of life who want to communicate with their counterparts speaking different languages. Consequently, more and more people are attracted to join this industry; nevertheless, their abilities differ. Most of them can handle lowest level of performance requirements, “but as requirements increase, performance problems arise: comprehension problems, reformulation problems, technical problems” (Gile, 1995, p. 2). Hence, training programs and schools mushroom all over the world. The thriving of training programs has inspired many studies on the efficacy of interpretation training by “practisearchers” (Gile, 1992a, p. 150), who work as interpreters and researchers at the same time.

A great number of previous studies have been dedicated to consecutive interpretation, and the focus has been put on the relationship between note taking and the interpreter’s performance. As Tsui (2004) put it:

Over the past literature about research on note taking, Gile (1995) mentioned the relationship between note taking and other events pertaining to consecutive interpretation. And Liu (1993) studied students’ notes from a pedagogy perspective to seek the way improving note-taking skills. Li (2000) probed into the relationship between the accuracy of note taking and interpretation performance. (p. 1)

Tsui, himself, further investigated the relationship between student interpreters' interpretation output, notes taken and the duration of the training received. From above, it can be seen that lots of efforts have been made to research on interpretation performance and note taking. Although previous studies have done extensive exploration, this study would take a step further by taking into consideration the difficulty degree of speeches in an attempt to examine its relationship with duration of interpretation training.

Literature Review

Definition of Interpretation

Interpreting is an “interlingual explanatory act” (Liu, 2004, p. 2). To put it in a simple way, interpretation refers to orally translate one language (source language) into another language (target language). Interpreters are different from translators because they do not use written texts to convey messages from the source text; neither do they have rather plenty of time to ponder the correctness of every word choice they make. Instead, interpreters convey messages by speaking, either simultaneously or consecutively (Mahmoozadeh, 1992, p. 231; Paneth, 2002, p. 31). Besides, the occupations that need interpreters most are those who have to frequently get in touch with people speaking different languages, e.g., businessmen, politicians or scientists (Renfer, 1992, p. 173). They need interpreters' assistance to communicate with foreigners for some purposes in some occasions (business dealings, military negotiations, lawsuits, academic interactions, etc). They always negotiate with each other vis-à-vis and these occasions make much of instantaneity. Interpreters, therefore, are expected to do the interpretation properly on the spot, regardless of the environment of the occasion or other challenges. And time and space, thus, become the two key constraints for interpreters, which constitute the greatest differences between interpretation and translation (Padilla & Martin, 1992, p. 195; Renfer, p. 174). Due to the two constraints, what interpreters pursue is the

ability to get used to all kinds of environment and to convey the information correctly and properly, instead of paying attention to the correctness of word choice. So, interpretation is an art which displays different beauty from translation, and it is more like an “act of communication” (Cheng, 1994, p. 11; Jones, 1998, p. 3). Additionally, since interpreters cannot require speakers to speak with the accent they are familiar with and speakers might come from any social level with a variety of diction habits, they also have to learn to listen to and understand all kinds of dialects and accents (Harris, 1981, p. 159).

Then, what constitutes a good interpreter? This question has been discussed by many scholars. And the researcher compiled scholars’ perceptions into the following list:

1. Interpreters should be able to well handle the source language and target language.
2. Interpreters should have up-to-date general knowledge and domain knowledge to deal with all kinds of subjects, participants and world or local events.
3. Interpreters should have rapid decision making ability and be quick-witted.
4. Interpreters should be good at emotion-controlling, pressure-counteracting and public speaking.
5. Interpreters have to not only be versed in the source language and target language, but also know well about source culture and target culture, so as to convey the information in the most proper way and arouse expected reaction.
6. Interpreters should have good memory.
7. Interpreters should master interpretation skills, such as note-taking, sight translation, consecutive interpretation and simultaneous interpretation, etc.
8. Interpreters should always make beforehand preparations and be

responsible.

(Bowen, 1984, p. 6; Jones, 1998, p. 4; Liu, 1993, p. 45; Mamoodzadeh, 1991, p. 233; Roy, 2002, p. 348-352; Seleskovitch, p. 62-63; Setton, 1994, 60-61; Yang, 2000, p. 200; Zhou & Chen, 1995, p. 46-59)

Translation and Interpretation Training in Taiwan

Due to the business opportunities and the lack of professional translators and interpreters in the market, translation and interpretation programs have mushroomed these decades. From established BA or MA programs, electives in universities and colleges to non-credit courses in continuing education, more and more channels are available for Taiwanese to learn the skills. Under such circumstances, the efficacy of interpretation (and translation) training turns out to be a popular issue. And studies about it have been conducted by a great number of scholars. The predecessors have concluded several findings though, this study tries to make further investigation on the issue. And the details shall be lighted upon in the following sections.

Methodology

Research Questions:

1. Does school training enhance learners' consecutive interpretation abilities and help them interpret more accurately?
2. Will there be any remarkable difference between these subjects in their note-taking style?
3. Speaking of Notes-Output Conversion Rate, is it true that subjects receiving longer training tend to have higher conversion rate?
4. From three five-minute speeches (in three degrees of difficulty), how do student interpreters receiving training with difference duration perform?
5. In addition to the accuracy rate, the influence from the difficulty degree of

speeches and note-taking style, is there any other difference towards the overall performance among these subjects?

Participants

Nine students from two Graduate Institutes of Translation and Interpretation were invited to take part in the experiment. All of them speak Mandarin Chinese as their first language and obtained TOEFL CBT scores greater than 250.

These subjects were divided into three groups by the duration of interpretation training they have received. And the groups are:

Group 1: Subjects who received training for less than one year. This group includes subject A, B and C.

Group 2: Subjects who received training for one to two years. This group includes subject D, E and F.

Group 3: Subjects who received training for more than two years. This group includes subject G, H and I.

Additionally, the researcher views subjects in Group 1 as novice student interpreters, whereas those in Group 2 and 3 as senior student interpreters.

Materials

Three speeches with different difficulty degree were used to investigate the subjects' performance. The topic of these speeches is: 1) the history of European Union and its fifth enlargement; 2) a speech in fund-raising banquet in Fu-Jen University; 3) J. F. Kennedy's Inaugural Address. The difficulty degree is: speech 1 < speech 2 < speech 3. And the ranking is determined by four criteria:

1. Technical terms or proper nouns: Technical terms refer to special words which describe the detail of a specialized activity (qtd. in Tsui, 2004). Proper nouns are must-noted ideas and usually cause heavy burden (Gile, 2002, p. 170). It requires more mental efforts to deal with the two kinds of words. In the experiment, the

density increases from speech 1 to 3.

2. Repetition of word or information: The repeated word or information would become old information for the interpreter, and repetition helps memorize the information (Cheng, 1995, p. 89; Hu, 2005, p. 68). Therefore, it requires less mental efforts to process.

3. Lexical Density: It divides words into content words (n, v, adjective and adverb) and function words (preposition, conjunction, article and pronoun); more content words make a text more difficult to understand. This concept was introduced into interpretation by Gile, who consented that greater lexical density could make a speech harder to understand (p. 170). According to Text Content Analysis Tool (2007), the lexical density of speech 1 is 49.36%, speech 2 is 39.67%, and speech 3 is 52.73%.

4. Text familiarity: Gagné (1984) illustrated an example to show how the text familiarity would influence the reader's (or the listener's) comprehension:

- A. Smith's father was a jewelry storekeeper.
Smith's favorite sister likes jewelry very much.
- B. Hollingshead's sire was a traveling cooper.
Hollingshead's esteemed sibling appreciates sonnets considerably.

Both A and B consist of two simple sentences, but A is easier to understand than B. Here, we can perceive the influence of text familiarity. From the lexical density, speech 2 is easier than speech 1. In order to make speech 1 easier, the study tried to enhance subjects' familiarity to speech 1 by offering two news articles which covered plentiful information and background knowledge about speech 1. Meanwhile, subjects were informed of less information about speech 2 (only a brief introduction to the speech), and further fewer ideas towards speech 3 (the speech title was the only thing to be mentioned).

Procedures

Procedures of the experiment involve: 1) recruiting subjects; 2) pre-experiment communications with subjects; 3) conducting the experiment. The third step could be further divided into: 3.1) filling out a questionnaire; 3.2) introduction of the experiment process; 3.3) testing the machine; 3.4) playing the speech part by part; 3.5) subjects' interpreting part by part; 3.6) post-experiment interview (an interview is conducted to probe into each subject's perceptions about consecutive interpretation, note taking and their performance in the experiment). We could dub step 1 and 2 as the pre-experiment phase, and step 3 as experiment phase.

Evaluation Method

In the second-stage final report of Taiwan's National Institute for Compilation and Translation (2005), accuracy and completeness are the most mentioned concerns from the interpretation clients (p. 66-67). Additionally, Kurz's compilation (2001) also leads to similar conclusion (p. 398-403). Besides, in Wang's survey (2006), respondents were asked how they distinguish whether an interpretation performance is good or not. The result shows that up to 59% of the respondents chose accuracy as one of the ways to tell, whereas fluency, standing in second place, only accounts for 15% (p. 6).

From the preceded paragraph, the study aimed to evaluate the subjects' performance by the accuracy and completeness of their speech production. But, in this study, accuracy concerned meaning-based accuracy, which means the subjects' interpretation would be counted correct as long as the message that the speaker wanted to convey has been processed and produced in the output.

The evaluation format of this study is a modification from Tsui's (2004) format. There are two formats in the study, evaluating the accuracy rate, the note-taking style and Notes-Output Conversion Rate. Table 1 and 3 are the samples

of the two formats.

Table 1
Evaluation format towards the accuracy of the subjects output

Speech 1		
Information	Word(s)	Output
	During	
	the	
	WWII	
	the	
	economy	
	of	
	the	
	western	
	Europe	
	got	
	into	
	serious	
	recession	

Table 1 involves three columns—information, word(s), and output. Word(s) shows the transcription of the source speech. As for the output column, it is used to examine if the subject has produced the output. If the subject produced it, then there would be a “O” sign in the corresponding blank; contrarily, “X” is marked if the subject did not produce it or produced incorrectly. Finally, information column segments the speech into information units, which are the smallest units of ideas. According to Tsui (2004), pauses separate information units (p. 51), which could offer the researcher a more precise way to investigate if the subjects interpret all messages. Take subject G’s performance in speech 1 for instance:

Speaker: In April 1951, the European Coal and Steel Community was
founded by six countries...

Subject G: 在 1961 年四月的時候，一共有六個國家組成了歐洲煤鋼
共同體...

And evaluation would go like Table 2:

Table 2
Sample for evaluation

Speech 1		
Information	Word(s)	Output
	In	
	April	O
	1951	X
◎	the	
	European	O
	Coal	O
	and	
	Steel	O
	Community	O
	was	
	founded	O
	by	
	six	O
	countries	O

Note: O = produced correctly; X = did not produce / produced incorrectly; ◎ = correct information

Here, four things need to be noted:

1. Checkpoints, referring to words that are checked in the table, light upon content words (nouns, verbs, adjectives and adverbs). Function words (prepositions, conjunctions, articles and pronouns) are not considered. Besides content words, the researcher further take into consideration particular prepositions and conjunctions for evaluation because some of them also contain important information, e.g., those signifying the linkage of sentences (e.g., not only...but also...), the transition of mood (e.g., although), the cause and effect (e.g., because), or the setting of the context in a given period of time (e.g., during). Without interpreting them, listeners would be confused.
2. As long as the meaning is conveyed, it can be counted correct.
3. ◎ mark is filled only when all the checkpoints within the information unit are correct. And the tallied amount of ◎ would be the indicator for each

subject's performance in output.

4. The subjects' performance in output would be shown on the basis of:

$$\frac{\text{total number of correct info. units}}{\text{total number of info. units}} \times 100\%$$

As for the total number of information units in the three speeches, they are:

Speech 1: 64; speech 2: 97; speech 3: 91.

And then, we can see the sample of the second evaluation format in Table 3, which consists of three columns—word(s), notes taken and notes interpreted. The function of word(s) column is like that in Table 1. As to notes taken column, the corresponding blank would be marked with a code if the subject took down note in his / her own sheets. And it would leave empty if the subject did not take down the notes. In this column, coding scheme was adopted and six codes were worked out, encompassing S, A, AN, F, W and M. S represents symbols, such as “□” (symbolizing a country), or “☺” (indicating gladness, goodness, etc.); A refers to abbreviations, such as “EU” (European Union); AN stands for Arabic numerals, such as 33, 1998, etc.; F means fragmental words or BoPoMoFo (Chinese phonetic symbols), such as “euro” (signifying Europe), and “又” (symbolizing the Chinese character 歐 “ou”); W signifies words, such as “西” (xi, the Chinese character, which means west), “recession,” “electronics,” etc. Besides, M would be marked if the subject did not take down the notes, but produced corresponding output correctly. And then, X is given if there is no corresponding note in the sheets and no relevant speech production in the interpretation.

Table 3
Evaluation format towards the subjects' note-taking style

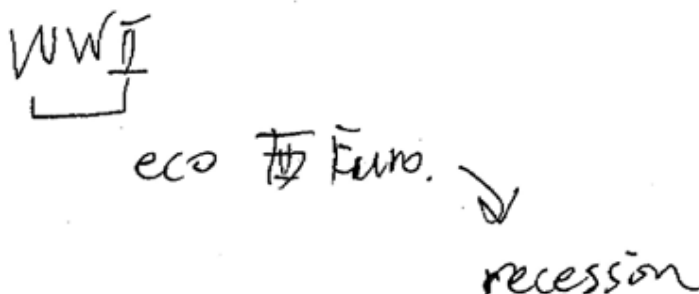
Speech 1		
Word(s)	Notes Taken	Notes Interpreted
During		
the		
WWII		
the		

economy		
of		
the		
western		
Europe		
got		
into		
serious		
recession		

Take subject E for example:

Speaker: During the WWII, the economy of the Western Europe got into serious recession.

Subject E's notes:



WWII
eco W Euro.
↓
recession

Subject E's output: 在第二次世界大戰的時候，西歐的經濟已經逐漸衰退。

And evaluation would go like Table 4.

Table 4
Sample for evaluation

Speech 1		
Word(s)	Notes Taken	Notes Interpreted
During	S	O
the		
WWII	A	O
the		
economy	F	O
of		
the		
western	W	O

Europe	F	O
got	S	O
into		
serious	X	X
recession	W	O

Note: S = symbols; A = abbreviations; AN = Arabic numbers; F = fragmental words; W = words; M = memory; X = Did not produce / Produced incorrectly

Finally, the percentage of S, A, AN, P, W and M would reveal the subjects' note-taking style, and we can see if there are any similarities or dissimilarities between the three groups. And the calculation method is as follows:

$$\text{S percentage} = \frac{\text{total number of S}}{\text{total number of checkpoints}} \times 100\%$$

The total number of checkpoints of the three speeches is: speech 1 (247); speech 2 (315); speech 3 (278). By the way, the number of checkpoints will not influence the difficulty degree of a speech.

Speaking of notes interpreted column, it is a column to examine to what extent the subjects interpreted their notes which were taken down on the sheets. Notes-Output Conversion Rate lights upon the rate the subjects interpreted their notes. In this study, the researcher assumed: senior student interpreters have higher conversion rate than novice ones do. And from the evaluation, we can see if the assumption makes sense. Additionally, since we only dealt with notes taken down on the sheets, the total number of strategy M would be ruled out. Then, the conversion rate could be counted through the following equation:

$$\frac{\text{total number of correctly processed notes}}{\text{total number of correct notes taken down on the sheets}} \times 100\%$$

Only correctly processed notes are considered because Notes-Output Conversion Rate talks about the extent the subjects interpreted from the notes taken.

Results

Accuracy Rate of Each Subject's Interpretation Output

After evaluating the performance of each subject and calculating, the results were compiled into a table.

Table 5
Accuracy rate of subjects' interpretation output in three speeches

	Speech 1	Speech 2	Speech 3
Subject A	65.62%	49.48%	19.78%
Subject B	56.25%	50.52%	--
Subject C	73.44%	48.45%	29.67%
Subject D	84.38%	68.04%	39.56%
Subject E	84.38%	67.01%	74.73%
Subject F	84.38%	75.26%	49.45%
Subject G	89.06%	87.63%	49.45%
Subject H	85.94%	64.95%	49.45%
Subject I	92.19%	81.44%	54.95%

Note: Subject B's performance in speech 3 is ignored because he practiced the same speech at school before.

In speech 1, subject I got the highest accuracy rate (92.19%), whereas subject B performed worst (56.25%). Speaking of speech 2, the highest accuracy rate is 87.63% (subject G) and the lowest one is 48.45% (subject C). Finally, in speech 3, subject E got 74.73%, which is the highest, while subject A, the worst, got 19.78%.

And it could be found from the table that each subject's performance was evidently influenced by the difficulty degree of three speeches. Generally, every subject, no matter how long he / she received interpretation training, would have lower accuracy rate as speeches became harder. The reason might be attributed to the fact that they did not understand speech 2 and 3 as much as they did in speech 1.

Moreover, as Gile's Effort Model in CI (1992b, p. 191-192; 1995, p. 178-183; 2002, p. 167-168) puts:

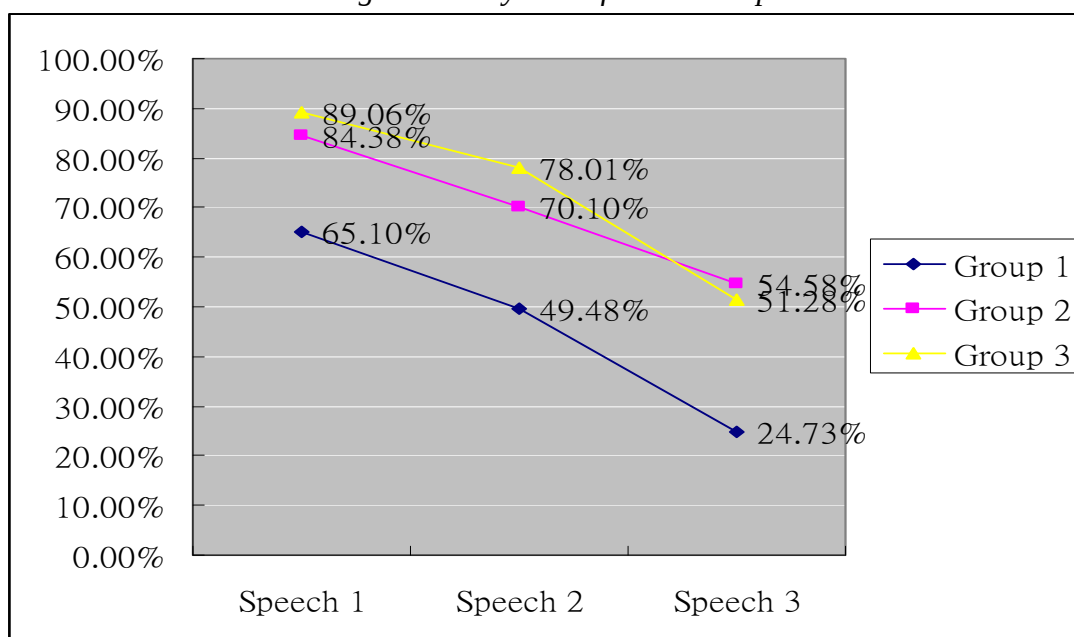
$$\begin{aligned} \text{CI (listening)} &= L + M + N + C \\ \text{CI (reformulation)} &= \text{Rem} + \text{Read} + P \end{aligned}$$

L: listen + analyze; M: memory; N: note-taking; C: coordination; Rem: memory recall; Read: note-reading; P: output production

Effort Model examines the allocation of “mental energy” (1995, p. 161) in the interpreter’s mind. In Gile’s opinion, the total effort capacity is of fixed quantity. As the speeches get harder, the subjects would have to pay more efforts to L for understanding speeches. Once the total efforts exceeded the fixed quantity, he / she may not be able to put enough efforts in M, N and C. Consequently, troubles take place. On the other hand, if an interpreter chooses to keep the balance among L, M, N and C, he / she is under the risk of not being able to comprehend what the speaker said. Then, as the comprehension drops, the performance would be deteriorated. Therefore, whether the subject pays more efforts to L or not, his / her performance would be negatively influenced by higher difficulty degree.

In the preceded paragraphs, we tried to analyze the results from the overview (Table 5), which detailedly describes each subject’s performance and is filled with numbers. To simplify the overview, the average accuracy rate of each group was worked out and converted into Figure 1:

Figure 1
Average Accuracy rate of three Groups



Speaking of the accuracy rate of speech 1, Group 2 and 3 performed in similar league, whereas Group 1's performance is around 20 percent lower than the other two groups. As to the accuracy rate of speech 2, there is an 8% difference between Group 2 and 3, while Group 1 is still 20% lower than Group 2 and 28% lower than Group 3. Finally, speech 3 shows a bit different situation, in which Group 2's accuracy rate is a bit higher (about 3%) than Group 3. As for Group 1, it further drops to 24.73%, which constitutes around 27% difference against Group 2 and 3. If we decipher the three groups' performance in three speeches, we can find that difficulty degree of speeches would influence subjects' performance and make them perform worse and worse.

It can be indicated from the previous paragraph that Group 1's performance remains around 20% lower than Group 2 and 3 in speech 1 and 2. In terms of the third speech, the gap enlarges to 27%. The accuracy rate difference in three speeches constitutes a great gap between Group 1 and the other two groups. However, if we turn to look at the performance difference between Group 2 and 3, the difference becomes much smaller (5% in speech 1, 8% in speech 2, -3% in speech 3). Generally, Group 3's performance is slightly better than Group 2 despite the minus-three-percent difference against Group 2 in speech 3.

A similar study has been conducted by Tsui in 2004, and he obtained alike research results. With similar results from two studies, it might be persuasive to say that training does help improve student interpreters' abilities. The enhancement, however, is more remarkable during the first year of training. As training continues, the enhancement scale would become smaller.

Furthermore, the researcher used the one-way ANOVA to investigate whether the difficulty degree of the three speeches influenced all the subjects and made them perform worse and worse from speech 1 to 3. In this part, difficulty degree was the only thing discussed, so, during calculation, all subjects were put into one group.

Table 6
ANOVA of three speeches about accuracy rate

	n	M	SD	F	Sig.
Speech 1	9	79.52	11.91	11.816	.000
Speech 2	9	65.86	14.22		
Speech 3	8	45.88	16.63		

* $p < 0.001$

Generally, a result would be called significant if p value is less than 0.05. From Table 6, we can see that p value is even less than 0.001, which means the performance difference between the three groups is very significant.

Each Subject's Note-Taking Style

The second evaluation is about each subject's note-taking style. The researcher made use of the concept of coding scheme and worked out six codes (S, A, AN, F, W and M) to sort subjects' notes. First of all, tables of compilation should be shown:

Table 7
Compilation of each subject's note-taking style in speech 1

	S	A	AN	F	W	M	X
Subject A	4.86%	6.07%	5.67%	18.55%	11.34%	36.29%	17.00%
Subject B	8.10%	6.48%	5.67%	21.86%	9.72%	28.75%	19.43%
Subject C	14.17%	6.48%	4.05%	12.55%	0.81%	50.61%	11.34%
Subject D	23.48%	6.48%	5.26%	16.19%	7.69%	30.77%	10.12%
Subject E	14.98%	8.50%	5.67%	28.75%	11.34%	26.32%	4.45%
Subject F	14.58%	6.48%	4.86%	28.75%	11.74%	27.53%	6.07%
Subject G	17.41%	8.91%	5.26%	31.58%	17.00%	16.60%	3.24%
Subject H	6.48%	4.45%	5.26%	28.34%	12.96%	35.22%	7.29%
Subject I	8.10%	3.64%	5.67%	14.98%	22.27%	41.70%	3.64%

Table 8
Compilation of each subject's note-taking style in speech 2

	S	A	AN	F	W	M	X
Subject A	11.43%	1.27%	2.54%	9.21%	12.06%	29.52%	33.97%
Subject B	12.70%	0.32%	1.95%	15.56%	11.11%	25.40%	33.02%
Subject C	7.30%	0.32%	2.22%	11.43%	2.54%	34.60%	41.59%
Subject D	18.10%	0.32%	2.22%	8.89%	10.16%	36.19%	24.13%
Subject E	16.83%	3.17%	2.54%	13.33%	12.06%	27.30%	24.44%
Subject F	15.87%	0.32%	3.49%	19.37%	6.67%	40.32%	13.97%
Subject G	11.43%	0.32%	3.49%	23.81%	18.10%	38.10%	4.76%
Subject H	7.62%	0%	3.81%	16.51%	13.97%	32.70%	25.40%
Subject I	6.35%	0%	2.22%	6.98%	19.68%	52.70%	12.06%

Table 9
Compilation of each subject's note-taking style in speech 3

	S	A	AN	F	W	M	X
Subject A	5.76%	0.72%	0.72%	9.00%	14.39%	11.87%	57.55%
Subject B	--	--	--	--	--	--	--
Subject C	7.19%	0.36%	2.16%	12.59%	0.72%	20.86%	56.12%
Subject D	8.99%	0%	0.36%	7.19%	6.48%	25.54%	51.44%
Subject E	14.03%	0%	1.44%	21.94%	16.19%	35.61%	11.15%
Subject F	8.63%	0.72%	0%	14.75%	7.19%	29.50%	39.21%
Subject G	9.71%	0.36%	0.72%	17.99%	20.50%	25.54%	25.18%
Subject H	5.76%	0.72%	0.72%	16.19%	12.95%	27.34%	36.33%
Subject I	3.24%	0%	0.36%	6.48%	22.30%	35.25%	32.37%

Note 1: S = Symbols; A = Abbreviations; AN = Arabic numbers; F = Fragmental word(s); W = Word(s); M = Memory; X = Did not produce / Produced incorrectly

Note 2: Subject B's performance in speech 3 is ignored because he practiced the same speech at school before.

What needs to be mentioned first is that X is not viewed as a note-taking strategy; it is used to signify subjects' not producing the corresponding output or producing it incorrectly. Therefore, it is more like a representation of incorrectness.

From the tables, it could be found that the difference of the note-taking style among these subjects is not remarkable. Speaking of A and AN, the use frequency is found similar among all subjects. This reveals that student interpreters, either novice or senior, would choose the best and most familiar way to take notes. For example, when it comes to information such as July, December, 1951, United States, United Nations, etc., everyone knows to take these kinds of messages down in Arabic numbers or abbreviations.

As to other note-taking strategies (S, F, W and M), their percentage is not as even as A and AN; there is a bigger percentage difference in these strategies between subjects. Take speech 2 strategy S for example, it accounts for 18.10% in subject D's notes, while it is only 6.35% in subject I's. There is an around-12-percent difference between them. Moreover, in speech 1 strategy M, it is 50.61% for subject C, whereas 16.60% for subject G. In this case, the difference grows to 34 percent or so. From the situation, we could reason that longer training would not make student interpreters use certain type of strategies more frequently

for easy note-taking. Instead, it all depends on each interpreter's preferences and habits towards linguistic style. Therefore, training is just a process for student interpreters to learn to wield note-taking and find the way best comfortable for them to take notes.

Finally, X column manifests a tendency that is not found in other columns: the X percentage in all these speeches decreases from subject A to subject I. To see it more clearly, the researcher worked out a table:

Table 10
Average percentage of X among three groups in three speeches

	Speech 1	Speech 2	Speech 3
Group 1	15.92%	36.19%	56.84%
Group 2	6.88%	20.85%	33.93%
Group 3	4.72%	14.07%	31.29%

It can be figured out that the percentage of X diminishes from Group 1 to 3 in every speech. Similar to the results of accuracy rate analysis, the percentage difference between Group 1 and 2 is bigger than that of Group 2 and 3. Since X stands for a tally on the percentage of output that was wrongly interpreted or was not produced, the decreasing tendency could imply that training helps improve novice student interpreters' ability in comprehending and interpreting what the speaker said.

Besides, from the above-mentioned tables, we could work out each group's average note-taking tendency among three speeches. The ranking is as follows:

Group 1: 1) M (28.25%); 2) F (13.51%); 3) S (8.66%)

Group 2: 1) M (31.03%); 2) F (17.68%); 3) S (15.05%)

Group 3: 1) M (33.97%); 2) F (18.06%); 3) W (17.75%)

It is indicated that M is the most used strategy by all the three groups. If we look at the average percentage of M, it is about 30% in the three groups. This proves the argument that interpreters only take down key ideas of a speech and

retain secondary messages by heart.

Then, F ranks second in all the groups. This implies that all the subjects preferred to take notes with the most fundamental strategy—to take down notes by words. As it takes time to be familiar with certain symbols and it takes chances to make use of Arabic numerals and abbreviations, to take down notes by words turns out to be a choice. However, because all subjects were aware that it should be time-consuming to take notes by words, they chose to use fragmental words. Moreover, since every Taiwanese is educated with the system of BoPoMoFo (Chinese phonetic symbols), the subjects also adopted it as one way to take notes.

Speaking of the third place, it shows a different situation. While S ranks third in Group 1 and 2, it is W that occupies the third place in Group 3. This might be because, according to follow-up investigation, subjects in Group 3 had not done frequent practices since the end of their professional test, held in June, 2006. In other words, they had not received frequent practices for months at the time the experiment was conducted, which was on November 11 and 12, 2006. Thus, it is possible that Group 3 subjects used less strategy S because they had not practiced for a long time and were not that familiar with symbols.

Other Findings from Subjects' Notes

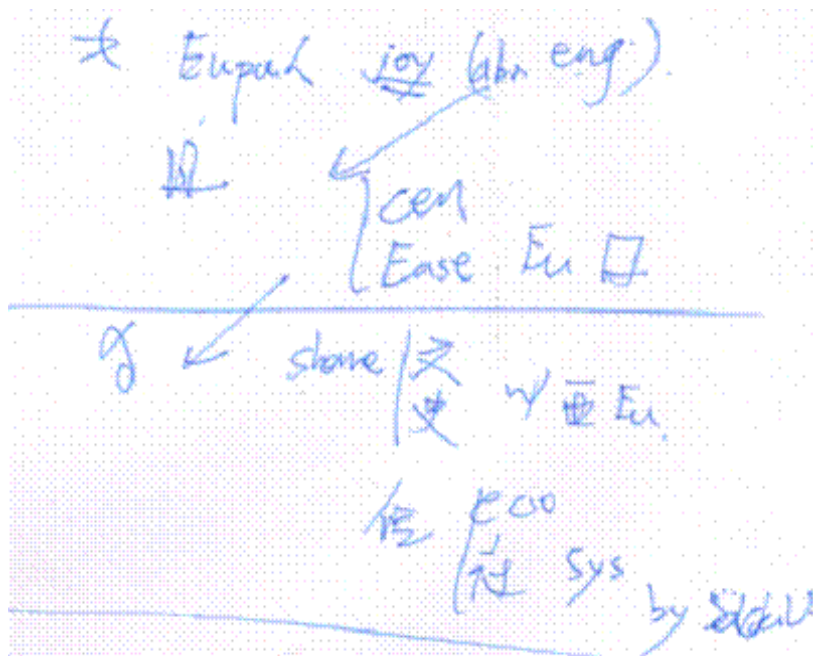
Besides the analyses from the previous compilation tables, the researcher also gained two findings during evaluating subjects' notes:

1. Smart use of obtained information: Interpreters use arrows to represent the idea of directions, process of transformation, linkage, increase / decrease, etc. But some of them extend the use of arrows to other way. Take subject E's performance for example:

Speaker: We can see from TV that Europeans are so overjoyed about the **enlargement**. Besides, this **enlargement** also benefits *central and*

east European countries. Although **central and east Europe** share the same culture and history with west Europe....

And subject E's notes are:



Subject E's output: 我們可以看到，那些國家的人民，爲了，因爲加入歐盟而感到很高興，因爲這一次的擴大行動，多半都是中歐和東歐的國家，雖然中歐和東歐的國家和西歐的國家擁有相同的歷史和文化傳統....

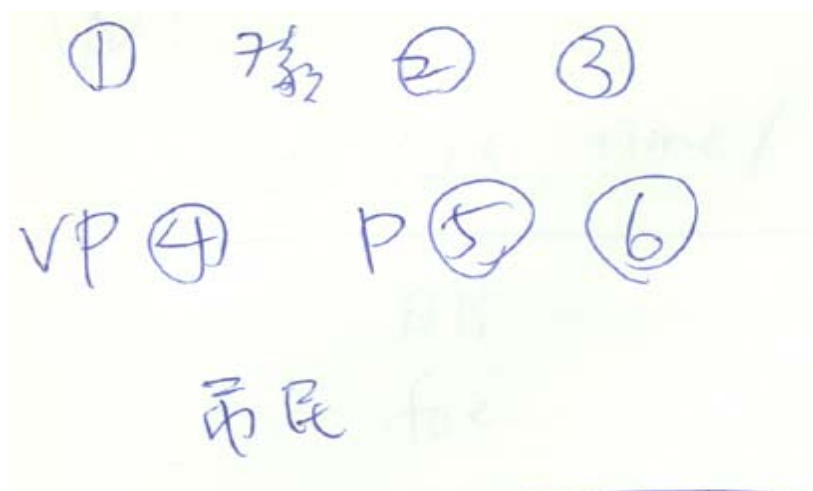
The speaker repeated the idea of “enlargement” and “central and east Europe.” From notes, we can see that subject E used arrows to first represent the repeated idea of enlargement. And then, it stood for central and east European countries. With arrows, he did not have to spend time on taking same notes again.

Besides, subject B made smart use of the glossary. Before the experiment, glossary about the three speeches was given. Subject B made beforehand preparation and wrote a new list of glossary, in which he listed unfamiliar

vocabulary and proper nouns and numbered each of them. Then, in the experiment, every time when the words in the new glossary were spoken, he used numbers instead of other kind of strategies to jot down the messages. Take speech 3 for example:

Speaker: Vice President Johnson, Mr. Speaker, Mr. Chief Justice, President Eisenhower, Vice President Nixon, President Truman, Reverend Clergy, fellow citizens...

And subject B's notes are:



Subject B's output: 強森總統、發言人、總統艾森豪、副總統尼克森、總統杜魯門跟令人敬愛的宗教領袖跟我國國民...

At the beginning of the Inaugural Address, Kennedy mentioned a list of important figures who were present at the ceremony. When other subjects were busy taking down all the names, subject B, with beforehand numbering, only had to write down numbers and paid attention to each person's title. Because subject B only had to write down numbers, he could save a lot of time from writing notes. This manifests the effect of numbering.

2. Overlapped use in note-taking strategies: The argument that symbols should

be used unequivocally was proposed by many scholars, such as Jones (1998, p. 57) and Liu (1993, p. 72). However, this concept does not make sense only in strategy S; it can be extended to other kinds of strategies, too. Take speech 1 for example:

Speaker: ...Under such circumstances, they will again become part of Europe...

Subject G's notes:



Subject G's output: 而在這樣子的狀況，我們可以看到部分的歐盟會經歷....

In subject G's notes, EU was mostly used to stand for the abbreviation of European Union. In this part, he used EU to represent the idea of Europe. And this led to the misinterpretation. In subject G's output, we can see that he wrongly interpreted EU into *oumeng2* (歐盟, means European Union), which is a different idea from Europe. And this is because he used the symbol for European Union to stand for the idea of Europe. And the overlapped use made him misinterpret it.

Notes-Output Conversion Rate of Each Subject

Notes-Output Conversion Rate refers to the extent (on percentage base) to which the subjects produced their output from the notes they took down, i.e., how much the subjects interpreted from their notes. In this study, the researcher made an assumption that those receiving longer training (Group 2 and 3) would have higher conversion rate than those receiving shorter training (Group 1). Since the

conversion rate probes into the extent the subjects interpreted from the notes they took, what we should pay attention to is the total number of notes that subjects took down correctly and the total number of notes correctly processed. Here, the total number of strategy M is ruled out because it stands for memory in mind, not written notes in sheets. Then, after evaluating the subjects' performance and calculating, a table was worked out:

Table 11
Compilation of each subject's notes-output conversion rate in three speeches

	Speech 1	Speech 2	Speech 3
Subject A	93.04%	86.96%	54.12%
Subject B	89.06%	87.02%	--
Subject C	94.68%	78.67%	84.38%
Subject D	98.63%	91.20%	85.94%
Subject E	97.08%	92.05%	93.29%
Subject F	96.95%	95.14%	94.25%
Subject G	96.47%	96.11%	75.18%
Subject H	98.59%	92.42%	91.09%
Subject I	99.26%	95.50%	90.00%

Note: Subject B's performance in speech 3 is ignored because he practiced the same speech at school before.

It can be seen that, in speech 1, the three groups were able to interpret almost everything they took down while listening. Moreover, almost all the subjects achieved a conversion rate over 90 percent in speech 1 (subject B is the only one whose conversion rate did not reach 90%, but he almost made it). Moreover, subjects in Group 2 and 3 got conversion rate higher than 95%; subject I even got a rate quite close to 100%. In speech 2, the conversion rate of subjects in all groups dropped. Subject C is the one whose performance was not beyond 80%. Besides, those in Group 2 and 3 were still able to maintain a rate beyond 90%; some of them even gained 95%-above conversion rate. As to subjects in Group 1, they faced a drop which was more serious than those in Group 2 and 3. Finally, in speech 3, the conversion rate dropped further. We can see that subjects in Group 2 performed best among the three groups, whereas those in Group 1 performed worst. In group 1, subject A's performance even dropped to 54.12%.

Table 12
Average notes-output conversion rate among three groups in three speeches

	Speech 1	Speech 2	Speech 3
Group 1	92.26%	84.22%	69.25%
Group 2	97.55%	92.80%	91.16%
Group 3	98.11%	94.68%	85.42%

After dissecting each subject's conversion rate, we could further examine the difference of the conversion rate amongst three groups. Table 12 sheds a light on the average notes-output conversion rate among three groups in three speeches. From the table, it can be found that the notes-output conversion rate of Group 2 and 3 is always higher than that of Group 1.

If we compare the conversion rate of Group 1 with that of the second highest Group in three speeches, we can find that the rate difference between the third place and the second place is 5.29% (speech 1), 8.58% (speech 2) and 16.17% (speech 3). This could indicate that as the speeches get harder, the performance gap between novice student interpreters (Group 1) and senior student interpreters (Group 2 and 3) could become bigger. Speaking of senior student interpreters, their conversion rate is rather stable; the rate among three speeches is generally above or nearly 90%.

Generally, all the subjects performed better in easier speeches and worse in harder ones, and the tendency is even more evident in Group 1. Again, it can be explained by Gile's Effort Model (1992b, p. 191-192; 1995, p. 178-183; 2002, p. 167-168). As the speeches get harder, the subject has to pay more efforts to tackle. Once the total effort requirements exceeded the fixed quantity, the subject's performance would be hampered.

In addition to Gile's model, the researcher figured out some other possible assumptions to answer why the subjects did not process some notes:

1. The subject had forgotten the corresponding message while interpreting.
2. In the listening phase, the subject just heard the word and took it down

without understanding its meaning in the context.

3. The subject did not tackle the note because he neglected it.
4. The subject used same symbol to represent more than two messages, so he was confused while interpreting.

To Notes-Output Conversion Rate, the one-way ANOVA can also be adopted to examine whether the speech difficulty degree influences every subject's conversion rate significantly.

Table 13
ANOVA of three speeches about conversion rate

	n	M	SD	F	Sig.
Speech 1	9	95.97	3.26	4.766	.019
Speech 2	9	90.56	5.58		
Speech 3	8	83.53	13.37		

* $p < 0.05$

We can find that p value is 0.019, which is less than 0.05. Hence, the difficulty degree of the three speeches poses significantly influence.

Findings from the Interview

The first finding is about the speech difficulty degree. In the interview, most subjects agreed that the ranking was: speech 1 < speech 2 < speech 3. And the most-mentioned factor for determining the difficulty degree was background knowledge. To all subjects, it was the main reason why speech 1 was the easiest one and speech 3 the hardest one. In addition to background knowledge, the subjects also mentioned wording, lexical density, speech delivery speed and the speaker's accent. But these factors were not as frequently mentioned as background knowledge.

The second finding is about the way these subjects saw note taking. All the subjects agreed that note taking is helpful, but their reliance on note taking differed. Some subjects (A, B, D, E and G) relied very much on it, while some (subject C, F, H and I) regarded it as a kind of assistance and they did not depend too much on it.

Moreover, subject I even mentioned that he would choose to forsake certain notes if those notes could not fit in the corresponding context. To subject I, it was the coherence that should be made much of, not whether all the notes are processed.

When it comes to other methods to help memorize the source speech, converting messages into images was the most mentioned strategy, followed by purely memorizing and paying some attention to the speaker's body language or intonation.

The third finding is about the way to tackle negligence or not understanding some words. The most adopted method was to use guessing strategy. When some words were neglected or not understood, subjects would guess from the former information and the context and then choose safer target words to fill in the gap. Additionally, similar result could also be found in Chang's study on listening strategy (2006).

Besides guessing from former information and the context, most subjects also mentioned that they would choose to tackle the part they are not sure of with a vague interpretation. For example, if they do not know how to interpret "crane" into Chinese, they would interpret it with the hypernym of crane (such as bird).

The final finding is about the subjects' definition of short-term and long-term consecutive interpretation. From the interview, we can find that every subject held different perceptions to it (see Table 14). More uniform perception was found in Group 2 and 3. Their understanding towards CI was clearer and they better grasped what short-term and long-term CI should be because they had received more training and practices than Group 1 subjects. Additionally, from Liu's work (1993, p. 7) and the interview, it is indicated that short-term CI is generally a speech within 30 seconds, while long-term CI is a speech that lasts for more than one minute.

Table 14
Compilation of each subject's definition about short / long-term CI

	Short-Term	Long-Term
Subject A	shorter than 1 min.	longer than 1 min.
Subject B	shorter than 3 min.	longer than 3 min.
Subject C	shorter than 1~15 sec.	longer than 1 min.
Subject D	shorter than 3 min.	longer than 5 min.
Subject E	shorter than 30 sec.	longer than 30 sec.
Subject F	shorter than 30 sec.	longer than 2 min.
Subject G	shorter than 30 sec.	longer than 30 sec.
Subject H	shorter than 30 sec.	longer than 30 sec.
Subject I	shorter than 30 sec.	longer than 4 min.

Conclusions and Discussions

Summary of Major Findings

In the experiment, subjects' performance in accuracy rate, note-taking style and Notes-Output Conversion Rate were examined. In accuracy rate (see Table 5) and Notes-Output Conversion Rate (see Table 11), we can find that those with longer training (more than one year) achieved higher accuracy rate and notes-output conversion rate than those with shorter training (shorter than one year). It could be concluded that consecutive interpretation training does help improve student interpreters' abilities. The enhancement, however, is more remarkable during the first year of training. Then, as training continues, the enhancement scale becomes smaller. Such an outcome is similar to that in former studies conducted by graduate students in National Taiwan Normal University. Firstly, in Pey-chich Lee's thesis (1999), she argued that interpretation training would help enhance students' ability in interpreting messages more correctly. Besides, Ho-ching Chang (2001) designed an experiment to examine the sight translation performance difference between three kinds of subjects: 1) subjects possessing high language proficiency and ST training; 2) subjects possessing high language proficiency and without ST training; 3) subjects possessing weaker language proficiency and ST training. In Chang's results, training was also regarded as an element boosting students' performance. Finally, Chien-chang Tsui (2004) pointed out that training

would help improve students' interpretation abilities.

As to subjects' note-taking-strategy-using tendency in each speech, which had been compiled into Table 7, 8, 9, remarkable difference or tendency could not be found. Instead, each subject had his own note-taking habit, which matched former scholars' argument (Bowen, 1984; Chang, 2006; Cheng, 1994) that every interpreter has his / her own note-taking style, deriving from personal linguistic habit and preferences. Therefore, we might say that training would not alter one's note-taking style and make him / her use certain type of note-taking strategies more frequently. What training can do is to help students get familiar with note taking, learn to well wield note-taking, find out the way best comfortable for themselves to take notes down and strike a balance between the efforts made on note taking and listening.

Speaking of the influence of difficulty degree, it could be noticed that each subject, novice or senior, performed worse and worse as the difficulty degree increases from speech 1 to 3. Nevertheless, those receiving longer training would have smaller performance decay than those receiving shorter training do. Moreover, Table 6 and Table 13 offered us statistical evidences that difficulty degree would significantly influence subjects' performance in accuracy rate and conversion rate.

Finally, senior student interpreters (Group 2 and 3) tend to interpret idea by idea, whereas novice ones (Group 1) would interpret from the notes, which might cause incoherence and fragmentation of ideas from time to time. In other words, senior student interpreters' interpretation is more comprehensible than novice ones'. The reason why novice student interpreters' interpretation is less comprehensible might be that they incline to interpret every note as possible as they can. Additionally, note taking beginners always want to take every idea down (Cheng, 1994, p. 13). With the desire to note everything, beginners would get into the trouble that they take down ideas without understanding what they mean in that context. Consequently, their interpretation would become less comprehensible.

Implications and Suggestions

From the results of the experiment, it can be noticed that interpretation training would help improve students' performance in accuracy and notes-output conversion rate. However, what we should pay attention to is that training may not be able to make students use certain type of note-taking strategies more frequently. From the compilation of note-taking strategies in the study, we can find that each subject has his / her own note-taking style. Therefore, for a teacher to instruct students, he / she should not force students to learn to use certain kinds of strategies more frequently; instead, general note-taking principles can be given for students to refer to. Students can adopt ideas they like from these general principles, from which they would develop their own note-taking method

Additionally, during instructing, the teacher should care more about the percentage of X (signifying wrong interpretation output or no interpretation output), shown in the Table 7, 8, 9 and 10. From the tables, we could see that it is the indicator that decreases from Group 1 to 3, which means it can be improved through training. To make it, the teacher ought to ask students not to try to take every idea down without understanding the meaning of each idea in corresponding context. It is comprehension towards the source speech that counts. Without comprehension, students will not be able to produce good interpretation. Besides, students should do more practices to well wield note taking and know how to strike a balance on the efforts paid to listening and note taking. Under the teacher's instruction and students' diligent practices, the percentage of X could be reduced.

As we can see from the interview, background knowledge is the factor which was most mentioned by research subjects as the criterion to determine a speech is difficult or easy to interpret. Thus, in addition to the instruction of CI skills, the teacher should also try to enrich students' knowledge about various domains. Having basic knowledge about various domains can help students deal with various

topics. To enrich students' knowledge about various domains, the teacher can prepare interpretation exercises of different topics in class for students to practice or encourage them to dabble in all kinds of topics as possible as they can. Moreover, to read all kinds of topics will also help them get familiar with all kinds of registers.

Finally, due to the word limit, this study mainly discussed the interpretation training efficacy from the perspective of output and notes. For future researchers, it is suggested that investigations be made from some other perspectives, e.g. the relationship between training and memory capacity, which, to certain extent, also poses impacts on to interpreters' performance. It is hoped that some new light can be shed on the field of interpretation study in the future.

References

Chinese References:

- Cheng, Li-yu. (1995). *Cognitive psychology*. Taipei: Wu-nan. 【鄭麗玉。1995。《認知心理學》。台北：五南圖書。】
- Hu, Zhi-wei. (2005). Human's memory and memory training. *Cosmic light*, 32, 66-75. Taipei: Christian Cosmic Light Holistic Care Organization. 【胡志偉。2005。從人類記憶的本質談記憶訓練。《宇宙光》，第32期，66-75。台北：基督教宇宙光全人關懷機構。】
- Liu, Min-hua. (1993). *Consecutive interpreting and notes: Theory, practice and teaching*. Taipei: Fu-Jen University. 【劉敏華。1993。《逐步口譯與筆記：理論、實踐與教學》。台北：輔仁大學出版社。】
- Liu, Mi-qing. (2004). *Studies on theories of translation and interpreting*. Beijing: China Translation and Publishing Co. 【劉宓慶。2004。《口筆譯理論研究》。北京：中國對外翻譯出版公司。】
- National Institute for Compilation and Translation. (2005). *Final report for establishing the translation and interpretation accreditation system*. 【國立編譯館。2005。《建立國家翻譯人才評鑑基準之研究：期末報告》。】
- Tsui, Chien-chang. (2004). Note-taking and speech production in consecutive interpretation from English into Chinese. Unpublished MA Thesis. 【崔建章。2004。《英譯中逐步口譯筆記選擇內容與語言產出之關係》。國立台灣師範大學翻譯研究所碩士論文。】
- Wang, En-mian. (2006). Interpreting into the B language: the East Asian experience. 2nd Chinese symposium for translation studies, 1-7. 【王恩冕。2006。從母語譯入外語：東亞三國的經驗對比。《第二屆兩岸三地中華譯學論壇會前論文集》，1-7。】
- Yang, Cheng-shu. (2000). The value and price of interpreting. *Studies of*

interpretation and translation, 5, 197-200. Taipei: Taiwan Association of Translation and Interpretation. 【楊承淑。2000b。論口譯的價值與價格。翻譯學研究集刊，第五輯，197-209。台北：台灣翻譯學學會。】

Zhou, Zhau-siang. & Chen, Yu-zhan. (1987). *The theory and practice of interpreting*. Taipei: The Commercial Press. 【周兆祥、陳育沾。1987。□譯理論與實踐。台北：臺灣商務印書館。】

English References:

- Bowen, D. (1984). *Steps to consecutive interpretation*. Washington D.C.: Pen & Booth.
- Chang, H. F. (2006). Listening strategy use as applied by graduate student of translation and interpretation. *Selected papers from May and June 2006 conferences on language teaching, literature, linguistics, translation and interpretation*, 147-171. Taipei, Crane.
- Cheng, Y. P. (1994). Consecutive interpretation: How to use your symbols intelligently. *Translation and interpreting: Bridging east and west*, 11-17. Hawaii: University of Hawaii Press.
- Gagné, E. D. (1984). The effects of text familiarity and cohesion on retrieval of information learned from text. *Journal of experimental education*, 52 (4), 207-213.
- Gile, D. (1992a). Opening up in interpretation studies. *Translation studies: An interdisciplinary*, 149-158. Amsterdam: John Benjamins.
- Gile, D. (1992b). Basic theoretical components in interpreter and translator training. *Teaching translation and interpreting: Training, talent and experience*, 185-193. Amsterdam: John Benjamins.
- Gile, D. (1995). *Basic concepts and models for interpreter and translator training*. Amsterdam: John Benjamins.

- Gile, D. (2002). Conference interpreting as a cognitive management problem. *The interpreting studies reader*, 162-176. London & New York: Routledge.
- Harris, B. (1981). *Language at crossroads, proceedings of the 29th annual conference of the American Translators Association*. Medford: Learned Information Inc.
- Jones, R. (1998). *Conference interpreting: Explained*. Manchester: St. Jerome.
- Kurz, I. (2001). Conference interpreting: Quality in the ears of the user. *Meta*, XLVI, 2, 394-409.
- Mahmoodzadeh, K. (1992). Consecutive interpreting: Its principles and techniques. *Teaching translation and interpreting: Training, talent and experience*, 231-236. Amsterdam: John Benjamins.
- Padilla, P. & Martin, A. (1992). Similarities and differences between interpreting and translation: Implications for teaching. *Teaching translation and interpreting: Training, talent and experience*, 195-203. Amsterdam: John Benjamins.
- Paneth, E. (2002). An investigation into conference interpreting. *The interpreting studies reader*, 30-40. London & New York: Routledge.
- Renfer, C. (1992). Translator and interpreter training: A case for a two-tier system. *Teaching translation and interpreting: Training, talent and experience*, 173-184. Amsterdam: John Benjamins.
- Roy, C. B. (2002). The problem with definitions, descriptions, and the role Metaphors of interpreters. *The interpreting studies reader*, 344-353. London & New York: Routledge.
- Seleskovitch, D. (1978). *Interpreting for international conferences*. Washington D.C.: Pen & Booth.
- Setton, R. (1994). Training conference interpreters with Chinese—Problems and prospects. *Translation and interpreting: Bridging east and west*, 55-66. Hawaii: University of Hawaii Press.

On-line References

- National Institute for Compilation and Translation. (2005). Final report for 2nd stage study on establishing the translation and interpretation accreditation system. Retrieved April 22, 2007, from <http://www.nict.gov.tw/tc/learning/032/941030.pdf> 【國立編譯館。2005。建立國家翻譯人才評鑑標準第二期研究期末報告。2007 年 4 月 22 日取自 <http://www.nict.gov.tw/tc/learning/032/941030.pdf>】
- Text Content Analysis Tool. (2007). Retrieved April 17, 2007, from <http://www.Usingenglish.com/resources/text-statistics.php>