5W1H Training Effectiveness for Information Extraction: Interpreting Summarized Chinese Indictments into English

Karen Chung-chien Chang

In the past decade, court translation/interpretation has attracted much research interest in Taiwan. In 2008, relevant guidelines and details were established, and court translation and interpretation became formal practices. However, some problems have become more noticeable and have drawn the attention of the authorities and researchers (Chang, 2013, 2016; Y. L. Chen, 2018; Chen & Chen, 2013; Y. T. Chen, 2018; Tu, 2019). One problematic issue is the lack of training for certified court interpreters. Because the threshold of language competence for becoming a court-certified interpreter is not very high, the assumption that a certified interpreter can effectively assist with a court case has received much criticism. This difficulty stems from the fact that most legal documents are not reader-friendly, a feature further compounding the comprehension problem when all messages are conveyed orally. As the syntactic structures of Chinese and English differ substantially, novice court interpreters, when hearing a paragraph of condensed expressions, tend to become baffled and unsure of where to start. Consequently, much time is required for information processing. This situation is especially evident in court interpreters' handling of summarized indictments. The present study adopts a training method to help student interpreters to parse a summarized indictment and extract key information. Employing the 5W1H strategy, this study examines the effect of training on students' ability to process the given indictments in Chinese and to interpret them accurately into English. The participants processed 13 indictments in total, and their renditions were evaluated based on the criteria of information completeness and grammatical correctness. The results have indicated the 5W1H strategy greatly helped the participants filter through the layers of information more effectively and produce English interpreting renditions more accurately.

Keywords: 5W1H strategy, information extraction, summarized indictment, indictment interpreting

Karen Chung-chien Chang, Associate Professor, Department of Foreign Languages and Applied Linguistics, National Taipei University, E-mail: changcc@mail.ntpu.edu.tw

檢視運用 5W1H 資料擷取策略於口譯訓練之成效：以簡易起訴書中譯英為例

張中倩

　　過去 10 年，法庭口譯成為新的研究議題。於 2008 年開始，臺灣設立、實施法庭通譯之相關規定及細節，但有關當局、學界均注意到一些主要問題（Chang, 2013, 2016; Y. L. Chen, 2018; Chen & Chen, 2013; Y. T. Chen, 2018; Tu, 2019）。其中，法庭通譯訓練的缺乏為主要議題之一，由於成為法庭通譯的語言能力門檻不高，在特約通譯取得證書後，是否有能力協助處理法庭案件，儼然引起批評及憂慮。由於多數法律文件內容艱澀，在法庭訊問中，以閱讀方式表達時，更容易造成通譯人員在聆聽訊息後，無法迅速處理訊息及翻譯；再加上中、英語言句法結構上存在許多差異，對於新手法庭通譯來說，當聽到一大段資料濃縮的訊息時，常常不知如何著手進行翻譯。在案件處理時，此情況常發生於簡易起訴書的翻譯。本研究採用 5W1H 資料擷取策略訓練學習「法庭口譯」的學生，檢視此策略是否可以有效地幫助學習者擷取簡易起訴書中的重要訊息，在檢視學習成效時，重點為中譯英的「信息完整度」及「文法正確性」，學生共處理了 13 件簡易起訴書，分析顯示：5W1H 策略對於訓練訊息擷取非常有成效，學習者能擷取、處理案件所提之多重細節（人物、地點、內容、時間、緣由、過程），在 13 週訓練之後，學生在斷句、擷取細節、中譯英的表達及文法正確度方面，皆有明顯進步。

關鍵詞：5W1H 策略、訊息擷取、簡易起訴書、起訴書翻譯

張中倩，國立臺北大學應用外語系副教授，E-mail: changcc@mail.ntpu.edu.tw。

Introduction

Court interpreting has been a burgeoning field in Taiwan since, in 2008, the Judicial Yuan formally established the guidelines and regulations of recruiting, training, and making use of interpreters to assist the communication of court cases involving a foreign party. Such a result can be attributed to the increasing presence of foreign workers, the developing tourism industry, and interracial marriages in Taiwan. Facing the needs of enabling the involved foreign parties to express themselves fully in court, the Judicial Yuan realized the importance of forming a talent pool from which court interpreters working in different language combinations could be located. In the past decade, more efforts have been invested into studying different aspects of court interpreting as a practice. Among different issues, how to train certified interpreters to develop adequate competence to assist court cases has become a major topic. Researchers (Chang, 2013, 2016; Y. L. Chen, 2018; Chen & Chen, 2013; Chen & Liao, 2016; Y. T. Chen, 2018; Tu, 2019) have directed their research attention to details, such as training material development, needs analysis, training effectiveness, case performance evaluation, and interpreters' self needs assessment.

When evaluating a court interpreter's competence in assisting a court case, legal professionals and interpreters themselves have named inadequate training as a major issue. Consequently, efforts have been made to investigate suitable training approaches as well as materials. Previously, Chang's (2013) study has pointed out such training can be provided at two levels: the training for certified interpreters and the

training for student interpreters. The former may be held as part of the in-service training, whereas the latter should start with course training at the university level. This study covers the case simulation training student interpreters received in the course of Court Interpreting, especially focusing on the training results from their handling of summarized indictments. The goal is to investigate the training effectiveness of the 5W1H information extraction strategy, with the aim of finding a feasible and effective way of training future court interpreters.

Literature Review

This section of literature review consists of two parts. The first part is a short review on the practice of court interpreting, especially on the challenges facing those in this profession and field practice. Second, as this study sets the goal to find an effective training approach for student interpreters in handling summarized indictments, a fixed feature in all criminal cases, the background information and past relevant studies carried out in the framework of 5W1H, a strategy for information extraction, are provided.

Challenges Facing Court Interpreters

Court interpreting refers to the practice in which an oral interpreting activity is performed by an interpreter who works in a courtroom. Such an interpreter bears the responsibility to present information from a source language (SL) to a target language (TL)

faithfully (Li & Zhang, 2006). However, different from other types of interpreting, court interpreting encompasses a high level of complexity and demands much from interpreters; consequently, court interpreters often encounter many difficulties and challenges in processing information presented in court. Moreover, a court interpreter faces the issue of "trust." According to Hale (2004), court interpreters frequently have to:

> prove themselves amidst constant suspicions of infidelity to the original text, the extremely high demands placed on them, the inherent complexities of the interpreting process, the inadequacies of the system they are to work in, the misunderstanding of their role by lawyers and witnesses alike, the poor working conditions and the low remuneration. (p. 2)

With these issues, the training of court interpreters has always been a complicated task.

Similar concerns, especially those related to the lack of interpreting training and linguistic competence among court-certified interpreters for providing quality interpretation, are shared by those involved in the field of court interpreting in Taiwan (Chang, 2013, 2016; Y. T. Chen, 2018). On the issue of quality interpreting, Alvarez and Vidal (1996), Berk-Seligson (1999), Hale (1999), and Morris (1999) highlighted that good interpreting must take into account not only linguistic content but also the involved social, cultural, and psychological variables. Moreover, for an utterance to convey

a speaker's intended meaning successfully, the accurate interpreting rendition must reflect both the semantic and pragmatic views of that utterance (Yule, 1996). González, Vásquez, and Mikkelson (1991) have stressed that it is required that court interpreters must "interpret the original source material without editing, summarizing, deleting, or adding while conserving the language level, style, tone, and intent of the speaker or to render what may be termed the legal equivalence of the source message" (p. 16). With these principles in mind, researchers have examined the quality of court interpreters' renditions from different angles. While emphasizing quality assurance, some have placed much focus on both the pragmatic and linguistic perspectives of legal renditions (Hale, 2004; Mason, 2008). Previous studies have examined interpreters' treatment of discourse markers (Hale, 2004) and interpreters' additions/omissions of politeness markers (Mason, 2008). These studies have contributed to the field's better understanding how interpreters' handling of these features could potentially impact a court case. Yet, more aspects still await investigation.

Chang (2016) has pointed out that court interpreting involves a certain level of formulaic expressions that should be acquired by all court interpreters. Among the formulaic expressions, the most predictable aspects include the reading of a defendant's basic rights in court and certain phrases frequently incorporated in a summarized indictment. However, these two components differ greatly in their levels of complexity. The former is often presented in short sentences, while the latter is often presented as a long paragraph. Take a summarized indictment in Mandarin Chinese for example. It often

involves a long statement which covers multiple layers of information, making it very difficult for any beginning interpreter to gain the confidence required in handling such a piece of legal information.

The difficulty in processing a large, nonstop chunk of information found in a summarized indictment can be further analyzed from the angles of linguistic input (processing) and output (production). The first difficulty stems from how the information is presented. In a Chinese paragraph where a summarized indictment is conveyed, the 5W1H elements are often strung together with no specific breaking of the elements. In other words, an interpreter is very likely to hear a long string of facts in a statement with the structure of "*who* did *what* to *whom* at a certain time on a certain day (*when*) in a certain place (*where*) with certain consequences (*what*)." Even though the above broken-down parts of *wh*-statements may not seem very complicated when written out in such a manner, when the pieces of information presented in Mandarin Chinese are plugged into the listed *wh*-elements, the statement could immediately compound into three times the length of the above string (see examples given in Appendix A). When such a meaning-condensed text is read to a court interpreter-trainee, the multiple layers of richness embedded in a summarized indictment in Chinese often pose itself as a huge linguistic challenge. Worse yet, such a lengthy piece of legal message typically has to be processed in a very short time span.

Even when a court interpreter-trainee is able to process such a long piece of information, he/she still encounters the second difficulty. This challenge is pertinent to how such a message involving multiple

layers of *wh*-elements should be arranged and presented for the involved parties to understand in a court proceeding. Interpreting from Chinese into English, a court interpreter cannot simply produce the English rendition with the same syntactic structure of Chinese, the SL. In other words, most court interpreters working with the language combination of Chinese and English find themselves incapable of directly presenting the information expressed in Chinese (SL) into English (TL) due to both grammatical and syntactic differences between the two languages.

To address such difficulty commonly encountered by inexperienced court-certified interpreters, finding a suitable and feasible strategy to assist current and future court interpreters to process this fundamental yet key piece of legal information is of great importance. In the subsequent section, 5W1H, as an event extraction (EE) strategy, is introduced, especially its contribution to studies in different fields.

5W1H, as Information/Event Extraction Strategy

For most people, it is not hard to understand 5W1H (who, what, when, where, why, and how) as a basic concept for information gathering or text mining. Yet, such an approach of gathering information or mining texts has earned a solid place in the field of journalism. In journalism, the 5W1H approach is used as a strategy for the extraction of semantic elements in events (Chakma & Das, 2018; Sharma, Kumar, Bhadana, & Gupta, 2013; Wang, 2012; Zheng, Jin, Zhao, & Yue, 2014). A news story is considered complete only when it fully answers a checklist of these six questions. In other

words, the factual answers to these *wh*-questions must be elaborate enough for people to understand the whole story (Carmagnola, 2008). Furthermore, for such a goal to be realized, the information receivers need to know how to extract and describe an event and be able to build an event knowledge base at the semantic level. The employment of the 5W1H strategy seeks to extract the semantic information in a long message by distilling the details to answer all the listed *wh*-questions.

Though this approach may sound simple, its applications have become more complicated in the last decade. Today, the world has become globalized, and information is exchanged rapidly on the internet. With this development, people are exposed to an excessive amount of international as well as domestic news and events. In the attempt to relieve news information overload, several teams of researchers (Chakma & Das, 2018; Sharma et al., 2013; Wang, 2012; Wang & Zhao, 2012; Zheng et al., 2014) have applied the 5W1H approach for event extraction to process texts posted on Twitter, in news, and on microblogs, with an emphasis of information filtering. These researchers' findings have shown that this strategy offers the advantages of filtering through a large amount of information, gaining specific details, and reducing the likelihood of leaving out key information.

To understand how the 5W1H strategy is employed in the designing of different information-processing models, the concept of information extraction should be explained first. Information extraction, IE, is known as the automatic extraction of structured information from some unstructured sources, and the extracted information covers

entities, relationships between entities, and attributes which describe entities (Sarawagi, 2007). Another way to explain IE is that it is "the name given to any process that selectively structures and combines data" (Cowie & Wilkes, 2000, p. 249) found, explicitly stated, or implicitly expressed in one or more texts.

Furthermore, event extraction is a high-level IE task that tries to formulate an event as "who did what to whom, when and where" (Wang, 2012, p. 197). Event extraction is regarded as "a common application of text mining" and involves "deducing specific knowledge concerning incidents referred to in texts" (Hogenboom, Frasincar, Kaymak, & de Jong, 2011, p. 48). In practice, event extraction automatically identifies events in free texts to obtain detailed information, including time, location, participants, and their roles in the events. In the Seventh Message Understanding Conference, event extraction is defined as a template-filling task for a domain-dependent scenario (Chinchor & Marsh, 1998; Wang, Zhao, Zou, Wang, & Zheng, 2010), and it aims at "identifying event triggers of a certain event type in the text (Event Detection) and finding out related argument with different roles (Argument Identification)" (Ding & Li, 2018, p. 189). Take for instance the statement "Sara was injured in a car accident." In this statement, the event detection system is expected to detect an event *Injure* with the trigger word "injured," and the argument identification system is expected to identify "Sara" and "car accident" as event arguments with the roles of "person" and "place" separately.

In fields like information processing or text analysis, systems incorporating the 5W1H strategy are often adopted. In a study

conducted by Ikeda, Okumura, and Muraki (1998), the researchers employed the 5W1H approach to classify and navigate through Japanese language texts. Aimed at finding a more effective and efficient way to retrieve information for creating office documents, this team of researchers emphasized that the 5W1H information extracted from text data offered users a platform with three functions: episodic retrieval, multi-dimensional classification, and overall classification (Ikeda et al., 1998, p. 571). Figure 1 illustrates the classification and navigation enabled by the 5W1H strategy.
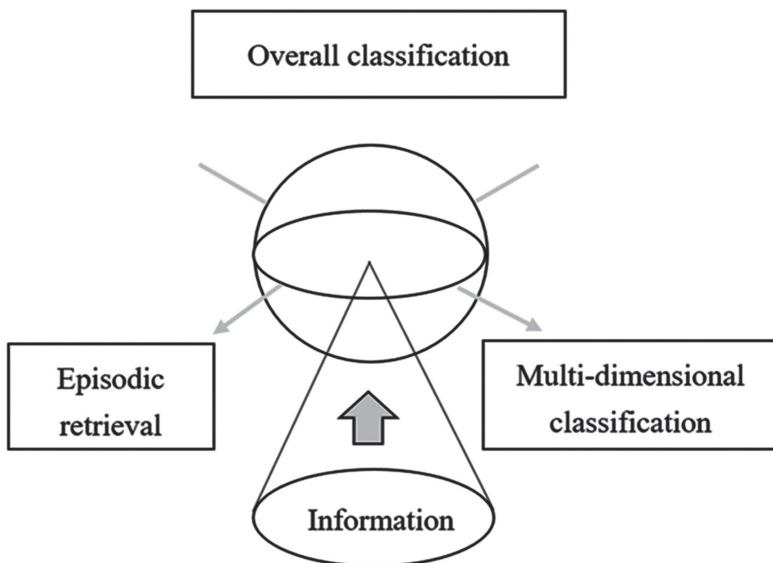

Fig. 1 5W1H classification and navigation (Ikeda et al., 1998, p. 572)

First, from the episodic perspective, users collect information on the related events and arrange the details in a temporal order to create an episode. Second, most office workers are familiar with the

comparative viewpoint, the angle of multi-dimensional classification. In this sense, a retrieval act is performed with a different, or changed, retrieval viewpoint. Last, when users are faced with a large amount of classification data, they have to choose appropriate keywords to conduct repeated retrievals and classification for condensing, so the results can be narrowed down for easier understanding. In that project, the research findings showed that the above three functions were effective for documentation work at offices, and the precision of extraction generated by this IE approach was approximately 82%.

Apart from being used for designing information processing models, 5W1H is also used by professionals in other fields. For example, in sociology, to reach the goal of interdisciplinary content-sharing, Shimazu, Arisawa, and Saito (2006) developed a system, focusing its function on searching for specified information on Web documents. Through a module that converted tag-labels into 5W1H items, these researchers confirmed the module on this system was very useful for content-sharing across different disciplines. In addition, in the field of counseling, Han, Lee, Lee, and Lee (2013) incorporated the 5W1H strategy into their counseling techniques to interact with users, recognize what users say, predict context, and follow users' feelings. Using this approach, these researchers extracted 5W1H information and four basic emotions (happy, afraid, sad, and angry) from their Korean users for implementing a counseling dialog system. These researchers have advised to apply such an approach further to speakers of other languages. Furthermore, Han, Kim, and Lee (2015) introduced a text dialog system that provided counseling dialog based on the semantic

content of user utterances. The researchers generated micro-counseling system responses from the extracted emotion-, problem-, and reason-oriented semantic content from their users' utterances. In that study, the extraction of semantic content enabled their system to generate appropriate counseling responses for a variety of user utterances and showed the system could also function as a virtual counselor. These past studies have shown the extended application of the 5W1H strategy in information extraction.

Learning from the experience of those in other fields, this research aimed at investigating whether the employment of the 5W1H information extraction strategy could help student interpreters filter and parse the message and extract key information presented in a summarized indictment. This study posed the following two research questions:

1. How does the 5W1H strategy affect students in extracting key information in a summarized indictment?
2. To what extent does the 5W1H strategy help students improve their rendition performance at the sentence level?

Study Design and Procedures

This section covers four parts: settings, participants, data collection tools, and data analysis methods. The first part, settings, provides background information of how this study was set up and describes the training format, covering case selection criteria, training materials, and training content. The second part covers the information

of the participants in this study, especially their previous training in translation and interpretation. The third part explains the tools used for data collection in this study. The last part delineates how the collected data were analyzed for explaining the qualitative and quantitative results of this study.

Settings

This study was conducted as part of the training in Court Interpreting, a course formally offered to English-majors at a public university in northern Taiwan. The course covered a total of 36 hours in which the students were trained to assist legal cases in the capacity of future court interpreters. A collection of 13 cases was selected as training materials, and they were chosen based on previous research results (Chang, 2016) for the most-frequently handled case types by experienced court interpreters. All case scenarios were collected and adapted from real court cases by the course instructor who also worked as a high-court interpreter. The two-hour class time was allotted to case simulation (task-based approach training) and rendition review. The students met two hours weekly to interpret the assigned court cases.

Apart from the first four-week instruction on legal concepts and basic vocabulary/expressions, the case-simulation training of Court Interpreting lasted 13 weeks. In these 13 weeks, the in-class interpreting practice was carried out in two formats: individual and collective interpreting modes. The former covered four weeks in which the students individually completed the entire interpreting tasks. The latter covered nine sessions in which the entire class collectively

interpreted the selected cases. The instruction and training of Court Interpreting took place in a language-training classroom with the capacity of recording each student's in-class performance. For the two formats described above, the recordings were carried out in class. Moreover, all recorded tapes were downloaded from the system and uploaded to the Digital Learning Center (DLC) immediately after each class session for the students to transcribe their interpreting renditions. The current study is part of a large-scaled Court Interpreting research project, and only the data related to 5W1H strategy training were employed here.

The training in this course hoped to instill a sense of "realness" to the class practice. Consequently, for the individually-completed sessions, the focus was on cultivating each student's ability to handle an entire case on his/her own. As for the collectively-handled cases, the instructional format was similar to a "relay race" in which the members of a team took turns completing a given task or performing a certain action. When a judge, a prosecutor, a lawyer or an involved party made a statement, a student would take on the interpreter's role and provide a rendition (in the consecutive interpreting format). At any given turn, if a student failed to understand the given information, he/she would have to ask for a repetition or clarification as "an interpreter" for the proceeding to continue. In this "relay race" mode, the instructor would not interrupt to provide any assistance even when the situations of information omission/deletion or erroneous interpreting occurred. All the interpreting performance details were video-taped for the rendition review in the second hour of course instruction. Each week,

the class time was divided into one hour of case simulation and one hour of rendition review. In the first hour, the students took on the role of interpreters helping with a court case. In the second hour, the instructor went over taped interpreting renditions and commented on the strengths and weaknesses of delivered renditions. Although the 5W1H strategy was explained and practiced in the first four weeks of the semester as part of the basic instruction, further practices and enhanced instruction were furnished each week. During a rendition review (or performance review), the video-taped interpreting renditions delivered by the students were commented on by the instructor with the exception of the renditions for the summarized indictments. For each case, when the review came to the rendition of the summarized indictment, a mini-lesson was always provided as a tool of skill enhancement. The instructor first put the 5W1H components (who, what, when, where, why, and how) on the board, played the read-out part of the summarized indictment, and showed the entire class under which heading the message should go. This demonstration enabled the entire class to see how a summarized indictment was "dissected." Then the instructor played the content one more time and wrote brief notes on the board to call on different students to process the extracted 5W1H information. Once a summarized indictment was collectively interpreted by several students, the instructor would choose one student to deliver the entire content before she herself modeled how the message should be delivered. In this process, the students would see how an indictment was broken into pieces, was pieced back gradually, and was delivered all together. It was through this repetitive

process that the students' training of 5W1H strategy employment was solidified.

Participants

The participants in this study covered 23 English-majors from a public university in northern Taiwan. These students took Court Interpreting in their senior year. In the department where this course is taught, students have two training specializations to choose from, with translation/interpretation (T&I) being one of them. Among the many T&I-related courses offered, Court Interpreting is the only course demanding students to switch back and forth constantly between Mandarin Chinese and English.[1] All the students who took Court Interpreting previously had taken Introductory Translation (72 hours in two semesters, one semester for each direction), Sight Translation (36 hours for both directions), Consecutive Interpretation (36 hours for both directions) and Advanced Interpretation Seminar (36 hours for both directions). In other words, all students were equipped with at least a total coursework of 180 hours.[2] Moreover, Court Interpreting is regarded as the culmination of T&I training at this department, for the students who take this course have to tackle the serious nature of court interpreting and cultivate the competence of working under high

---

[1] In almost all the translating/interpreting courses offered at this department, the training direction always starts with English into Chinese, followed by the training from Chinese into English. The rationale is that T&I learners should be trained to translate/interpret into their native language (in this case, Mandarin Chinese) first.

[2] For a few students who have missed one course listed above due to his/her exchange experience overseas, a test score equivalent to TOEIC 800 or higher can be submitted for meeting the competence requirement of taking Court Interpreting.

pressure and making constant switches between Mandarin Chinese and English.

Since the main goal of this study was to examine the effectiveness of the 5W1H strategy training in students' ability to extract information from a highly-condensed piece of information delivered in Chinese, it was necessary for the researcher to know how familiar these students were with the application of the 5W1H strategy. A few steps were taken to gather information on this issue. The initial step was a thorough examination of all required courses mandated by this department. Among the required courses, only in "Advanced Reading," a course offered in the freshman year, is 5W1H incorporated as a reading strategy for students to capture the main ideas delivered in any piece of reading. However, further inquiries with the two involved course instructors revealed that 5W1H was only referred to as a "handy" strategy for information grasp. The students were encouraged to filter the information they read through the 5W1H lens, but no further instruction was furnished. Moreover, in the students' junior year when they took Consecutive Interpretation, the course was taught by two instructors and neither put any emphasis on 5W1H training.[3] Therefore, it was concluded that, for this group of 23 students, 5W1H was a term that they had previously heard of but not an approach they had actively adopted for interpreting training purposes.

---

[3] Between the two instructors, one focused more on lectures and ST training for political speeches, while the other put more emphasis on rapid rendition formulation from one language to another.

Data Collection Tools

A simulated court case typically starts with either a judge or a prosecutor verifying the personal information of the case parties (the plaintiff, the defendant, the lawyers, or others pertinent to the case). Moreover, such procedural steps extend to the reading of the defendant's rights and the verification of the defendant's understanding of his/her rights, followed by a summarized indictment and the defendant's plea, all mandated by the Code of Criminal Procedure in the Republic of China. As all the linguistic expressions used up to this point are viewed as procedural and, to a certain degree, formulaic,[4] the training for Court Interpreting usually begins with the phrases and expressions pertinent to these three parts. In this study, the aim was to examine the effectiveness of the 5W1H information extraction strategy in training the student interpreters to handle summarized indictments. Three tools were used for data collection: a pre-test, 12 types of criminal cases (the last one as the post-test), and an interview.

The layout of the 18-week semester includes four weeks of basic instruction (including legal expressions/phrases and procedural knowledge), 13 weeks of case simulation, and the end-of-semester portfolio assessment (see Appendix B). The major instructional focus was placed on case simulation via the task-based learning (TBL)

---

[4] The language expressions are thought of as procedural and mostly "formulaic" because they are part of every criminal case, and such a process typically requires fixed expressions and phrases. The only details that may vary are related to the nature of different cases, for example, a traffic violation case and a robbery case (H. L. Tu, personal communication, January 18, 2020).

approach; the students were expected to acquire interpreting skills and legal concepts through case simulation. After four weeks of basic instruction, the students started their first case interpretation, the pre-test. In the pre-test (also the first simulated case), the participants interpreted a summarized indictment without receiving much instruction of the 5W1H strategy, and their renditions were recorded. Then, in the subsequent 12 weeks, the students processed a criminal case every week, each with a summarized indictment. The purpose of collecting these renditions was to chart the students' improvement over a span of 13 weeks. Among these 13 simulated cases, the last one was treated as the post-test. For data collection, the renditions of the summarized indictments for these 13 cases were transcribed. Last, one-on-one interviews were conducted for gaining insight from the participants, especially how their handling of summarized indictments was influenced by the employment of the 5W1H information extraction strategy.

During these 13 case-simulation weeks, no matter whether a case was individually-handled or collectively-interpreted, the part from the announcement of the beginning of a case to the complete reading of the summarized indictment was audio-taped through the recording equipment in the language classroom. After finishing each class session, the instructor uploaded all the audio-tapes to the DLC, making them available for the students' downloads and preparations of the transcripts. After that, all the student-transcribed renditions were double-checked by the researcher's teaching assistant. The recording and file-uploading processes were teacher-controlled to ensure the

transcripts reflected the students' in-class interpreting results.

Data Analysis Methods

This study covers both quantitative and qualitative analyses. For the quantitative analysis, attention was given to two aspects. First, through a contrastive analysis between the renditions collected from the pre-test and post-test, the students' performances were contrasted in two aspects: information extraction and grammar accuracy. For a better understanding of what the raters focused on in their examination of the renditions, these aspects are further explained. Before the analysis process is further explained, the raters' backgrounds are first provided. In this study, two raters were responsible for preparing the 5W1H information breakdown for all selected summarized indictments, scrutinizing the students' interpreting renditions, calculating the students' success rates for information extraction, and evaluating the accuracy of the renditions. The lead rater is the researcher/instructor who is trained in T&I and has a Ph.D. degree in linguistics, and the second rater is an English instructor with training in linguistics and TESOL. With a common background in linguistic training and further delineated preparation, the two raters conducted the analysis described below.

The first emphasis is placed on information extraction, which directly impacts the degree of information completeness in one's interpreting of a summarized indictment. Furthermore, this feature is directly linked to the effectiveness of a student's employment of the 5W1H strategy. As the 5W1H strategy focuses on the elements

of who, what, when, where, why, and how, the assumption is if a student can successfully parse the chunky message presented in a summarized indictment, he/she should be able to process, digest, and present the extracted information accordingly. In order to examine whether the participating students could extract the information they heard to achieve the effectiveness in parsing the heard message, all 13 indictments were read and extracted by the two raters separately. For all the chosen cases, the 5W1H elements are mostly present in the selected summarized indictments. Taking the following summarized indictment in Mandarin Chinese for example, the extracted 5W1H elements are marked in parenthesis:

> 被告，你 (who) 在民國 106 年 12 月 27 日 (when) 於新北市樹林區學府路與大學路交叉口、7-11 門口前 (where)，盜取他人摩托車一輛 (what)，車牌號碼為 BL-1234 (what)，經由摩托車持有人報案後 (why)，警方於柑園路一段 25 號附近尋獲 (where)。經警方調閱 7-11 門口監視器 (how)，確認摩托車為你所竊 (who)，檢察官起訴你犯竊盜罪 (why, the reason／charge)，你有何辯解 (what)？

> Defendant, you (who) were involved in an incident taking place in front of the 7-11 at the intersection of Xuefu Road and Daxue Road in Shu-lin District, New Taipei City (where). The incident happened on December 27, 2017 (when). You stole a motorcycle with the plate number BL-1234 (what). The owner of the motorcycle reported the case (why). The police found

the missing motorcycle near No. 25, Section 1, Gangyuan Road (where). The police checked the CCTV installed on the door of that 7-11 (how) and confirmed that the motorcycle was stolen by you (who). You have been charged with theft (why, the reason/charge). What is your plea (what)?

Clearly, several elements can be identified and extracted in the long message above. Moreover, as each case is different in nature, some elements, like the element of "what," may appear more than one time. For evaluating the students' effectiveness in information extraction, the two raters first made attempts to interpret three indictments to see how they would apply the 5W1H strategy. After that, with their separate results, the raters discussed the disagreed points and aligned their results, placing a special emphasis on tallying the total counts of the *wh*-elements in these cases. Then the two raters processed the remaining 10 indictments. Using the Krippendorff's alpha test (Hayes & Krippendorff, 2007) to estimate the inter-rater reliability, this study obtained a high level of inter-rater reliability of .93 for the two raters' alignment in information extraction. With the agreed information extraction results, the raters went through a total of 299 copies of renditions (13 copies from 23 participants). The results were further analyzed in two ways. For one, the pre-test and post-test results were contrasted to show students' performance differences before and after the 5W1H strategy training. On this part, attention was given to the students' information extraction performances specifically (see Table 1). For the other, the students' performances during the 11-week practices

(excluding the pre-test and post-test) were tallied and compiled into Table 2 to provide a more complete picture about how the students improved their information extraction skills over time. In addition, Figure 2 plots the students' overall performance improvement in this semester.

The second emphasis is grammar accuracy. This feature can be further divided into two aspects: "tense use" and "sentence flow." Students' errors in interpreting renditions can be examined through many angles; however, this study limits the error analysis to these two aspects. In the aspect of tense use, any court case can be described as a "happened event," meaning that the event described in the corresponding indictment is a past event that has been brought to the court for investigation and ruling. Under such a circumstance, grammar, especially the use of tense, plays an important role in reflecting what has already happened. Since it is a common practice for a prosecutor to present, or just read out, a summarized indictment in a fast and non-stop manner, many student interpreters may not be able to catch the event sequences or express them using correct grammar. This study presupposes when a student interpreter is able to use the 5W1H strategy to filter the heard information correctly, the extracted information should be able to help him/her to picture or organize the happened event.

Furthermore, in the aspect of sentence flow, this criterion is chosen to reflect the difficulty in interpreting a Chinese message into English. In this study, the two languages, Mandarin Chinese and English, are

distinctly different in their syntactic structures. For instance, a time adverbial is often inserted in the middle of a sentence, as shown in "我今天下午三點回家" (I this afternoon 3 o'clock went home). In this given example, if it is directly translated into English, the syntactic structure in English will not be evaluated as correct, for the adverbial phrase is in the wrong place. Worse yet, a summarized indictment is usually presented as a short paragraph that is often read as one extremely long sentence. Such a long sentence blends in all 5Ws and 1H in a random sequence. For student interpreters who do not know how to filter information into specific categories, the direct rendition of such a message is likely to be incomprehensible and cause further problems in the communication among different parties in a court case.

For this part of the analysis, the two raters first separately interpreted the parsed information and went through 15 indictment renditions (five copies from three different cases) produced by the students. Attention was given to tense use and sentence flow. For instance, when a student used a wrong verb tense or had a sentence with an incorrect placement of an adverbial phrase, it was counted as one error. In their evaluation of the first lot of 15 indictment renditions, both raters read through the renditions at least three times. Their inter-rater reliability was established at the level of .89. Then the raters, adopting the same approach, completed the examination and evaluation of the remaining 284 copies of indictment renditions. The results were compiled into Table 3 and Table 4.

Findings and Discussion

This section is organized according to the two research questions. Table 1 shows the performance differences gathered from the pre-test and the post-test, with an emphasis given to students' information extraction. Moreover, the total counts of 5W1H information extraction for the two tests are provided to put the students' performances into perspective. With the rater-extracted total counts, the students' performances were converted into percentages for a clearer and more straightforward comparison between their performances in the two tests. Then a column with the students' semester improvement records was provided. In addition, to gain insight into the students' overall learning progresses, their performances for the 11 weeks in between the two tests were compiled in Table 2. Similar to how Table 1 was organized, the students' weekly performances in their employment of the 5W1H strategy were further scrutinized in Table 2. Every student's learning progress in this regard was calculated by dividing his/her total of information extraction in each summarized indictment into the rater-identified total of information extraction.

Students' Improvement From Pre-test to Post-test

One major aim of this study was to examine if the 5W1H strategy could help the students improve their performances in extracting key information from the orally presented summarized indictments. In this section, the students' performances in the pre-test/post-test were contrasted to evaluate the effectiveness of the 5W1H strategy

training. Moreover, the interview results were provided to account for the students' evaluation of their learning and the fluctuations in their performances.

In Table 1, the students were coded from A to W for privacy concerns. In the pre-test, the total data pool covered 23 students, and their renditions were compared to the total counts of key information identified by the two raters.[5] In the post-test, the students' renditions were evaluated in the same manner. When the two sets of results were contrasted, the students' percentages of improvement were calculated and provided in the last column of Table 1. The results shown in the column of "Improvement Percentage" indicate when the student interpreters became more keenly aware of the target information, they became more prepared in handling this challenging task of information extraction. In Table 1, the students' improved performances ranged from 19% to 38%. Among the 23 students, the training effectiveness is especially evident for the students who initially were not able to extract the target information successfully. In Table 1, 12 students (A, D, G, H, I, K, L, N, O, S, V and W) were only able to extract 55% (or less) of the 5W1H-related information in the pre-test. Nevertheless, at the end of the training, they demonstrated improved performances, ranging from 20% to 38%. Compared to this group of students who started out with a relatively lower level of information extraction skills, the

---

[5] When identifying key information in the summarized indictments from the pre-test and the post-test, the two raters identified 11 and 12 *wh*-elements separately. When the two raters examined the transcripts of the students' renditions, attention was paid to the extraction of key information, covering both completeness and accuracy. If a student, for example, only captured the year and month but missed the date of the incident, it was counted as one error.

Table 1

Students' Performances in the Pre-test and the Post-test (Identifying 5W1H Information)

Student Code	Pre-test 5W1H total count: 11		Post-test 5W1H total count: 12		Improvement Percentage (%)
	S-I-*wh*	%	S-I-*wh*	%	
A	6	55	10	83	28
B	8	73	11	92	19
C	7	64	10	83	19
D	6	55	10	83	28
E	7	64	11	92	28
F	7	64	10	83	19
G	6	55	10	83	28
H	5	45	10	83	38
I	6	55	10	83	28
J	8	73	12	100	27
K	6	55	10	83	28
L	6	55	11	92	37
M	7	64	11	92	28
N	5	45	9	75	30
O	6	55	10	83	28
P	7	64	11	92	28
Q	7	64	11	92	28
R	8	73	12	100	27
S	6	55	10	83	28
T	8	73	12	100	27
U	7	64	11	92	28
V	6	55	9	75	20
W	6	55	10	83	28

Note. S-I-*wh* represents the *wh*-elements identified by participating students.

other 11 students' improvement range was 19% to 28%. While such a range of improvement might seem insignificant, the performances

have to be interpreted through another lens. Take three students (J, R, and T) for example. They were able to extract all needed information in their post-test of the assigned summarized indictment. From the angle of 5W1H strategy training, such performance results were and should be considered "highly satisfactory." Another seven students (B, E, L, M, P, Q, and U) also were able to extract 11 out of 12 pieces of 5W1H information presented in the assigned task. Overall, the findings presented in Table 1 have attested to the usefulness of the 5W1H strategy in training students to tackle summarized indictments.

When the students' semester learning results on their employment of the 5W1H strategy are examined more closely, the records of information extraction (in %) have revealed that almost all the participating student interpreters demonstrated some fluctuations in their information extraction of the handled indictments as shown in Table 2. A conclusion which can be tentatively drawn from the implementation of the 5W1H skill training is that an overall improvement can be observed in nearly all the participants. Although the participants did display some fluctuations in their performances from week to week, when their performances in Week 2, Week 7, and Week 12 were singled out for a comparison,[6] all the student interpreters showed improvement during this period of time (see Figure 2). Taking Week 12 for instance, only four students' (C, I, O, and R) performances

---

[6] Because of the fluctuations in the students' performances from one week to another and the total of 23 students, a line graph with the learning results (in %) of the entire class is not an effective means. However, when the students' performances in Week 2, Week 7, and Week 12 were selected, the overall improvement in their performances can be observed and examined more easily.

dipped slightly. Yet, when these four students' learning results were examined over the entire training period, their learning results still demonstrated an upward trend.

Table 2

Students' Information Extraction Performances in 11 Weeks

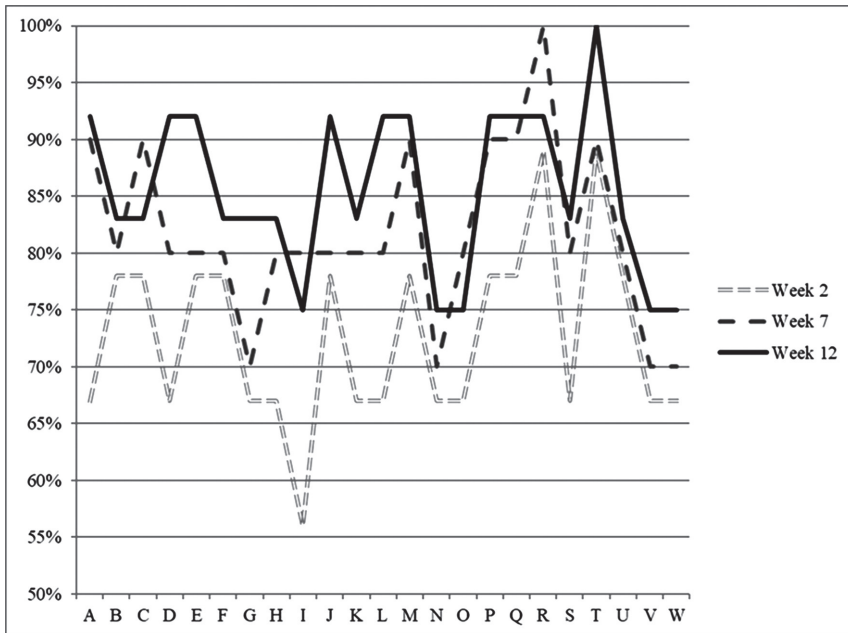| Student Code	5W1H Information Extraction Counts (%)										
	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12
A	67	73	80	77	82	90	80	82	85	91	92
B	78	73	80	85	82	80	80	82	85	82	83
C	78	82	80	77	91	90	90	82	92	91	83
D	67	73	70	69	73	80	90	82	92	82	92
E	78	82	80	69	73	80	80	82	85	91	92
F	78	82	80	77	82	80	80	82	85	82	83
G	67	64	70	69	73	70	80	73	77	82	83
H	67	64	80	77	82	80	90	82	77	82	83
I	56	64	70	69	73	80	80	82	77	82	75
J	78	73	80	85	82	80	90	82	85	91	92
K	67	64	70	69	82	80	80	82	77	82	83
L	67	73	70	77	82	80	90	91	77	82	92
M	78	82	80	77	82	90	90	91	92	91	92
N	67	73	70	69	73	70	80	73	77	82	75
O	67	64	70	77	82	80	90	82	85	82	75
P	78	82	80	85	91	90	90	91	85	91	92
Q	78	73	80	77	82	90	100	91	92	100	92
R	89	91	90	85	91	100	100	91	92	100	92
S	67	64	70	69	73	80	80	82	77	82	83
T	89	82	80	85	91	90	90	91	85	91	100
U	78	73	70	77	82	80	90	82	77	82	83
V	67	64	70	69	73	70	80	82	77	82	75
W	67	64	60	69	73	70	70	73	77	73	75

Fig. 2 A comparison of the students' performances in information extraction in Week 2, Week 7, and Week 12

In the one-on-one interview sessions, the students were asked three questions (see Appendix C) for their evaluation of the 5W1H strategy training, their observation of self-improvement, and their accounts on the fluctuations identified in their performances. Looking back at their performances during the 13-week case simulation training, 18 out of 23 students rated the 5W1H strategy training as "Very Useful," while another five students gave the rating of "Useful." When asked to elaborate their answers, the students provided further explanations of their perceived helpfulness. Two students (A and L) provided the following details:

The 5W1H strategy is extremely helpful because I could quickly categorize the information that I heard from an orally presented summarized indictment. I remember my first time trying to interpret a summarized indictment. I felt overwhelmed by such a lengthy piece of information. I almost froze there and didn't know what to do or what to jot down. Now, with the habit of using this strategy, I know I am much less anxious when listening to a summarized indictment. (A)

When the teacher first taught us the concept of the 5W1H strategy, I thought the strategy was too simple to be effective. After all, we all learned the elements of who, what, when, where, why and how. It was not until I actually tried to put what I heard into these categories of elements that I began to realize its effectiveness. With the pre-sorted categories, I feel more at ease and more ready to handle a summarized indictment. As I become calmer in expecting to hear certain information involved in a case, I don't feel as panicked. I think this change in the level of readiness has helped me the most. (L)

Other students shared similar attitudes through their use of keywords like "less worried," "a peace of mind," "less burdened," "more prepared," and "less afraid." In particular, two students (V and W) expressed that though they were still not able to catch all the key information towards the end of the semester, they still felt "much more

assured" about their handling of summarized indictments. One of them (V) said, "I like the feeling of making progresses, especially when the progresses are measurable. I know I have the ability to catch more and more information now."

When asked why their performances fluctuated in different weeks, the students' answers pointed to a few common factors. For instance, some students (C, E, O, and R) recalled that their performances suffered when their concentration levels were influenced by "a coming exam," "poor sleep quality the night before," "being late for class that day," and "other distractions." For these students, when their minds were on some other matters, their concentration levels were compromised, directly impacting their performances on the specific indictment interpreting. In addition, some students (I, N, O, V and W) referred to "the nature of different court cases" as the reason for their performance fluctuations. During the entire period of 5W1H strategy implementation, the students received the vocabulary list each week for the specified court case and were told to memorize the vocabulary items. However, some students indicated when they heard a corresponding summarized indictment, if they hesitated in selecting the suitable vocabulary items or expressions for interpreting, such a short pause would cause them to fail to catch all the key information.

Among these 23 students, some clearly demonstrated much more improvement over the span of 13 weeks. When asked to reflect on their improvement as well as their experience in putting the 5W1H strategy to work, four students (H, N, R, and T) provided the assessment of their own changes and growth:

Personally, I have always found multi-tasking a great challenge. In this semester's training, we had to quickly categorize the information we heard in a summarized indictment into 5W1H. For me, I am just not fast enough. It is really difficult to digest the information, categorize the details, and take notes at the same time. I think the strategy is useful but I will need more time to cultivate it into a habit. (H)

I like putting the heard information into the 5W1H categories but taking notes is hard because, in previous interpreting courses, we were not trained to take notes. That's why I often had a hard time figuring out my handwriting. Another problem is my poor memory. If I couldn't recall what my notes were about, I would have missing parts in my renditions. (N)

I feel quite satisfied with my learning progress this semester. When I tried to apply the 5W1H strategy to the handling of a summarized indictment, I felt excited because I knew I was able to capture at least more than 85% of the content. Then all I needed to do was to produce what I had grasped. I felt assured. (R)

At the beginning, I disliked the fact that we had to process a short paragraph-like piece of information. I still remember, in the first few cases, I dreaded this task because I knew it was coming right after the reading of the defendant's rights. The 5W1H strategy

training really helped in lowering that anxiety. The best part is I could see my own improvement in breaking down a heavily-loaded piece of information. (T)

These interview responses from the students have aided the understanding of Court Interpreting training in two aspects. First, to enhance the training results, the instruction and practice on note-taking should be added. Second, mini-lessons and practices on strengthening the students' working memory are necessary for building the students' confidence in information capture. That is, both elements can further enhance the production quality of interpreting renditions. To sum up, in the aspect of extracting key information (or preserving the completeness of heard information), both the students' improvement over the span of 13 weeks (see Table 1, Table 2 and Figure 2) and their reflections collected in the interviews have shown and confirmed the effectiveness of the 5W1H strategy training.

Students' Improvement in Grammar Accuracy

Apart from examining if the students could effectively use the 5W1H strategy to extract key information from a heard indictment, this study set another goal of evaluating whether the students' interpreting renditions revealed improvement in grammar accuracy. In this aspect, the raters' attention was directed to two issues: "verb tense" and "sentence flow." When analyzing these two issues, the raters focused on how the students performed in the pre-test and post-test. In Table 3, the types of errors related to the incorrect use of verb

tense are sorted and tallied. In both the pre-test and post-test, the two summarized indictments were filtered into 11 sentences and 12 sentences separately. One special note is that not every student had the same sentence parsing results. While most students were found to have parsed the two indictments in a very similar way, some chose to combine information when they deemed the information as related. For instance, in the post-test, more than half of the students (a total of 14) lumped the information of the incident (a car accident) and the location (the intersection of Hsin-yi Road Section 5 and Sungren Road) into one long sentence. Consequently, the totals of sentences gathered from the pre-test and post-test were 197 and 232 separately. When the two raters evaluated the collected sentences for grammar accuracy, attention was given to verb-related errors (see Table 3) and problems affecting sentence flow (see Table 4).

Table 3

Verb-Related Errors in Students' Interpreting Renditions

Verb-Related Errors		Total Sentence Count	
		Pre-test (11)	Post-test (12)
Incorrect handling of past tense	in present tense	52	30
	with the incomplete verb form (double verbs)	47	29
	in past perfect tense	29	14
	in present progressive	25	14
	with incorrect spelling (irregular verbs)	20	8
Incorrect use of passive voice (in active voice)		25	13

In Table 3, the errors were mainly related to the students' use of verb tense. As explained previously, a summarized indictment describes a "happened" case. Naturally, the use of verb tense becomes an important feature in a piece of interpreting rendition. In this group of student interpreters, the verb-related errors basically can be observed in five aspects. In the pre-test, 52 errors were identified when the students misused the present tense in the place of past tense. The second frequently occurred error is related to the use of double verbs. Such an error is brought by the influence of Mandarin Chinese, their L1, in which two verbs can be strung together as the sentence " 我想回家 ". Whereas such a sentence is expressed as "I want to go home" in English, it is literally translated as "I want go home" in Chinese. Perhaps influenced by this language feature, 47 errors were found with this problem. Moreover, 29 errors were related to the incorrect use of the past perfect tense for the past tense; 25 errors were associated with the incorrect use of the present progressive tense. In almost all "happened" cases, the present progressive tense is rarely used. Last, 20 errors were linked to the students' incorrect use of the past tense for irregular verbs. While, in the interviews, the students acknowledged that time pressure was the culprit for such a mistake, the occurrence of this error actually highlights their weakness in the use of irregular verbs. It is true that learners' challenges and difficulties with verb tenses are related to their overall learning of English; however, for this study, since these issues were highlighted in the in-class performance review every week, the students' improvement in these areas should be able to reflect their awareness in the extracted information. In other

words, as the interpreting-trainees listened to a summarized indictment, used the 5W1H strategy to filter and extract information, and paid attention to the sequenced details, their raised awareness in how to process a summarized indictment could be reflected in their language accuracy.

Table 4

Errors Affecting Sentence Flow

Errors Affecting Sentence Flow	Total Sentence Count	
	Pre-test (11)	Post-test (12)
Incorrect placement of adverbial phrases (location & time)	44	29
Incorrect stringing of words	31	19
Incorrect use of appositives	21	11

In the aspect of errors that affect the sentence flow in a piece of interpreting rendition, the raters found three major problematic issues: the incorrect placement of adverbial phrases (including time and location adverbials), the incorrect stringing of words, and the incorrect use of appositives. Among the sentences collected from the pre-test, the counts for these three types of errors were 44, 31, and 21 respectively. Again, in the placement of an adverbial phrase, the students' renditions into English were heavily influenced by what they heard in Mandarin Chinese. For example, in the sentence "告訴人手肘、手臂與手掌有多處擦傷以及瘀青" (the accuser suffered multiple scratches and bruises on his elbows, arms, and palms),

because of the influence from the SL, nine students' renditions were seriously affected by Mandarin Chinese in their sentence structures. Their renditions shared the similarity of placing the adverbial phrase of "on his elbows, arms, and palms" immediately after the subject, the accuser. Such a sentence does not flow well in English and may impact the understanding of the information receiver. Furthermore, a typical error of stringing words incorrectly can be found in the expression of a date. In Mandarin Chinese, the expression of a date is in the sequence of "year, month, hour, and minute." However, in English, depending on the use of British English or American English, these details can be strung in different orders. Yet, neither takes on the sequence of the Chinese date expression. Consequently, for the students who rushed into a statement in which a date was included, their renditions were negatively impacted.

Last, the use of appositives is less commonly employed in Mandarin Chinese but can become quite handy for an English sentence in which an appositive furnishes some additional information to the noun it modifies. When this rule is applied to the interpreting of a summarized indictment, it is commonly seen in the case where an accuser and his/her name are often strung together. In such a case, the use of an appositive can be employed to address such a sequence in a message. The students in this study were often influenced by Mandarin Chinese when a message included the structure of "被告(即○○○)." In the rendition transcripts, most frequently, the students rendered the statement "告訴人 (即○○○) 發生車禍" as "the accuser was ○○○ and happened a car accident" rather than processing the sentence as

"the accuser, ○○○, was involved in a car accident." In the post-test, errors related to the use of appositives decreased greatly thanks to both the use of the 5W1H strategy and the increased awareness of such a sentence structure.

In short, in the aspect of "grammar accuracy," because the students applied the 5W1H strategy for information extraction, their information extraction results further helped the production of more succinct and smooth renditions.

Conclusions and Limits

This study has generated a few interesting findings. First, as English learners, most students in Taiwan have been instructed to pay attention to the *wh*-elements in reading comprehension. That is why, for this group of students, many of them had heard of 5W1H but failed to see the connection between this strategy and the task of interpreting a summarized indictment. It is no exaggeration to say that applying the 5W1H concept as an event extraction strategy was completely novel to these participating students at the beginning of this study. Before they learned to apply the 5W1H information extraction strategy for the purpose of filtering the heard information (like that in a summarized indictment), the students' renditions in the pre-test were problematic in the aspects of information completeness, message parsing, and grammar accuracy. After 12 weeks of 5W1H training (excluding the pre-test), all students, to varying degrees, showed improvement in the aspect of information completeness. The interview responses also

revealed the students themselves acknowledged the effectiveness of the 5W1H strategy in information extraction. It is true that the learning and acquisition of any new language skill can never take place in a vacuum, meaning the above observed learning effectiveness of the 5W1H information extraction strategy could potentially be a result from the students' learning in other classes. However, during this specific semester, all the courses made available for these students in their senior year did not put any focus on the training of information extraction.[7] Therefore, as the influences from other courses in that semester are minimized, the semester-long training effectiveness of the 5W1H strategy is established.

Second, the analysis of the renditions produced by the student interpreters has offered informative results to show the types of errors frequently related to verb tense, placement of adverbial phrases, and the use of appositives. For translation and interpretation instructors, these findings can help them design mini-lessons to cultivate learners' awareness of these potential problems when handling a summarized indictment. Moreover, the findings can be used to raise learners' awareness of their errors and weaknesses for future improvement.

However, this study also faces some limitations, specifically in the number of participants and the length of the study. First, because of the interpreting nature in this training, the class size was small, only

---

[7] In the semester when this study took place, the courses offered to these seniors included "Shakespeare," "American Literature," "Appreciation and Studies of English Picture Storybooks for Children," "Humor Studies," "Advanced Writing for Business Situations," and "Advanced Studies in Translation." None of them, inclusive of "Advanced Studies in Translation," put an emphasis on the training of information extraction.

23 students. To gain a fuller understanding of the effectiveness and potential limits of this strategy, future studies should be carried out on a larger pool of participants or those learning to interpret court cases in other language combinations. Moreover, due to the time constraint, only 13 simulated cases were used in this training. To obtain more solid and concrete conclusions on the effectiveness of this training approach, future studies of similar kinds should be carried out over a longer time span. A more detailed and longer study may help court interpreting instructors and future trainees identify problematic issues affecting rendition accuracy.

References

Alvarez, R., & Vidal, M. C. (1996). *Translation, power, subversion.* Philadelphia, PA: Multilingual Matters.

Berk-Seligson, S. (1999). The impact of court interpreting on the coerciveness of leading questions. *The Journal of Forensic Linguistics, 6*, 30-56.

Carmagnola, F. (2008, January). *The five Ws in user model interoperability*. Paper presented at the 5th International Workshop on Ubiquitous User Modeling, Gran Canaria, Spain.

Chakma, K., & Das, A. (2018). A 5W1H based annotation scheme for semantic role labeling of English tweets. *Computacióny Sistemas, 22*(3), 747-755.

Chang, K. C. C. (2013). Current practices of court interpreting in Taiwan: Challenges and possible solutions. *Compilation and Translation Review, 6*(2), 127-164.

Chang, K. C. C. (2016). Needs analysis for the training of court interpreters. *Compilation and Translation Review, 9*(2), 93-136.

Chen, Y. L. (2018). *Court interpreting: Theory and practice*. Taipei, Taiwan: Wu-Nan Book.

Chen, Y. L., & Chen, T. (2013). Enhancing the quality of court interpretation – A functionalist approach. *Compilation and Translation Review, 6*(2), 99-126.

Chen, Y. L., & Liao, P. (2016). A revised model for the professionalization of court interpreting in Taiwan. *Compilation and Translation Review, 9*(2), 137-164.

Chen, Y. T. (2018). *Working with court interpreters in Taiwan – A survey of judges, prosecutors, and lawyers* (Unpublished master's thesis). National Taiwan Normal University, Taiwan.

Chinchor, N., & Marsh, E. (1998). *MUC-7 information extraction task definition (version 5.1)*. Paper presented in Seventh Message Understanding Conference (MUC-7). https://www.aclweb.org/anthology/M98-1027.pdf

Cowie, J., & Wilkes, Y. (2000). Information extraction. In R. Dale, H. Moisl, & H. Somers (Eds.), *Handbook of natural language processing* (pp. 249-269). New York, NY: Marcel Dekker.

Ding, R., & Li, Z. (2018). Event extraction with deep contextualized word representation and multi-attention layer. In G. Gan, B. Li, X. Li, & S. Wang (Eds.), *14th International Conference: Advanced Data Mining and Applications (ADMA 2018)* (pp. 189-201). Cham, Switzerland: Springer.

González, R. D., Vásquez, V. F., & Mikkelson, H. (1991). *Fundamentals of court interpretation: Theory, policy and practice*. Durham, NC: Carolina Academic Press.

Hale, S. (1999). Interpreters' treatment of discourse markers in courtroom questions. *The Journal of Forensic Linguistics, 6*(1), 57-82.

Hale, S. (2004). *The discourse of court interpreting*. Amsterdam, the Netherlands: John Benjamins.

Han, S., Kim, Y., & Lee, G. G. (2015). Micro-counseling dialog system based on semantic content. In G. Lee, H. Kim, M. Jeong, & J. H. Kim (Eds.), *Natural language dialog systems and intelligent*

assistants (pp. 63-72). Cham, Switzerland: Springer.

Han, S., Lee, K., Lee, D., & Lee, G. G. (2013). Counseling dialog system with 5W1H extraction. In M. Eskenazi, M. Strube, B. Di Eugenio, & J. D. Williams (Eds.), *Proceedings of the SIGDIAL 2013 Conference* (pp. 349-353). Stroudsburg, PA: Association for Computational Linguistics.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*, 77-89.

Hogenboom, F. P., Frasincar, F., Kaymak, U., & de Jong, F. M. G. (2011). An overview of event extraction from text. In M. Erp, V. W. R. Hage, V. L. Hollink, A. Jameson, & R. Troncy (Eds.), *Proceedings of Detection, Representation, and Exploitation of Events in the Semantic Web (DeRiVE 2011), workshop in conjunction with the 10th International Semantic Web Conference 2011 (ISWC 2011)* (pp. 48-57). Aachen, Germany: CEUR-WS. org.

Ikeda, T., Okumura, A., & Muraki, K. (1998). Information classification and navigation based on 5W1H of the target information. *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, 1*, 571-577. https://www.aclweb.org/anthology/P98-1093.pdf

Li, K. X., & Zhang, X. H. (2006). *Legal texts and legal translation*. Beijing, China: China Translation and Publishing Corporation.

Mason, M. (2008). *Courtroom interpreting*. New York, NY: University Press of America.

Morris, R. (1999). The gum syndrome: Predicaments in court interpreting. *The Journal of Forensic Linguistics, 6*(1), 7-29.

Sarawagi, S. (2007). Information extraction. *Foundations and Trends in Databases, 1*(3), 261-377.

Sharma, S., Kumar, R., Bhadana, P., & Gupta, S. (2013). News event extraction using 5W1H approach & its analysis. *International Journal of Scientific & Engineering Research, 4*(5), 2064-2067.

Shimazu, K., Arisawa, T., & Saito, I. (2006). Interdisciplinary contents management using 5W1H interface for metadata. In T. Nishida, Z. Shi, U. Visser, X. Wu, J. Liu, B. Wah, W. Cheung, & Y. M. Cheung (Eds.), *2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI2006 main conference proceedings) (WI'06)* (pp. 909-912). Los Angeles, CA: IEEE Computer Society.

Tu, H. L. (2019). *A study on the legal interpreter training program – Interpreters' perspectives* (Unpublished master's thesis). National Taiwan Normal University, Taiwan.

Wang, W. (2012). Chinese news event 5W1H semantic elements extraction for event ontology population. In M. Rabinovich & S. Staab (Eds.), *WWW'12-Proceedings of the 21st Annual Conference on World Wide Web Companion* (pp. 197-202). New York, NY: The Association for Computing Machinery.

Wang, W., & Zhao, D. (2012). Ontology-based event modeling for semantic understanding of Chinese news story. In M. Zhou, G. Zhao, D. Zhao, Q. Liu, & L. Zou (Eds.), *NLPCC 2012, CCIS 333* (pp. 58-68). Heidelberg, Germany: Springer.

Wang, W., Zhao, D., Zou, L., Wang, D., & Zheng, W. (2010).

Extracting 5W1H event semantic elements for Chinese online news. In L. Chen, C. Tang, J. Yang, & Y. Gao (Eds.), *Web-Age Information Management: 11th International Conference, WAIM 2010 proceedings* (pp. 644-655). Berlin, Germany: Springer-Verlag.

Yule, G. (1996). *Pragmatics*. Oxford, UK: Oxford University Press.

Zheng, L., Jin, P., Zhao, J., & Yue, L. (2014). A fine-grained approach for extracting events on Microblogs. In H. Decker, L. Lhotská, S. Link, M. Spies, & R. R. Wagner (Eds.), *Database and Expert Systems Applications, 25th International Conference, DEXA 2014, Munich, Germany, September 1-4, 2014. Proceedings, part I* (pp. 275-283). Cham, Switzerland: Springer.

Appendix A

Summarized Indictment Example 1

告訴人○○○ (who) 於中華民國 107 年 3 月 1 日早上 8 點 15 分 (when) 於信義區之信義路五段與松仁路交叉口 (where) 發生車禍 (what)，車禍原因為當時被告停車，將其車輛暫停路邊，在開車門時不慎撞到告訴人之摩托車 (why)，致使告訴人手肘、手臂與手掌有多處擦傷以及瘀青 (what)，檢察官起訴被告○○○ (who) 此車禍案件 (why, the reason/charge)，被告做何辯解 (how/what)？

The accuser is ○○○ (who). He suffered a car accident (what) at 8:15 am on March 1, 2018 (when). The location was the intersection of Hsin-yi Road Section 5 and Songren Road (where). The cause for this accident was that the defendant parked his car on the roadside and when he opened the door, he bumped into the accuser's motorcycle (why). Such a move caused the accuser to suffer multiple abrasions and bruises on his elbows, arms and palms (what). Defendant, you have been indicted for this car accident (why, the reason/charge). What is your plea (how/what)?

Summarized Indictment Example 2

被告○○○ (who) 於中華民國 107 年 12 月 18 日晚間 7 點 50 分 (when) 在其自家門口 (where)，地址為新北市土城區中央路一段 45 號 8 樓 (where)，因挪動摩托車 (why) 而與鄰居（即告訴人）(who) 發生口角，進而產生肢體衝突 (what)。經告訴人至醫院驗傷 (what)，醫師診斷為手臂、身體多處挫傷、瘀青 (what)，後被提告 (why)。檢察官起訴被告違反刑法之傷害罪 (why, the reason/charge)，被告做何答辯 (how/what)？

Defendant (who), this case happened at 7:50pm on December 18, 2019 (when). The location was in front of your residence (where). The address is Floor 8, Number 45, Section 1, Chungyang Road, Tu-cheng District, New Taipei City (where). You had a quarrel with your neighbor (what), also the accuser (who), for moving his motorcycle (why). The quarrel escalated into a physical confrontation (what). The accuser went to the hospital for an inspection of injury (what). The doctor's diagnosis showed that he suffered multiple contusions and bruises on his arms and body (what). Consequently, the accuser filed this case against you (why). You have been indicted with violation of the criminal law, and the charge is assault (why, the reason/charge). What is your plea (how/what)?

Appendix B

Course Syllabus of Court Interpreting

Weekly Schedule	Tentative Teaching Schedule	Teaching Methods and Activities
Week 1	Course Introduction/Legal Terminology & Expressions	■ Lecture
Week 2	Procedural/Formulaic Expressions	■ Lecture
Week 3	Legal Concepts	■ Lecture
Week 4	5W1H Strategy Introduction/Practice	■ Lecture ■ Strategy Practice
Week 5	Case 1: Theft (pre-test)	■ Case Simulation (CS) ■ RenditionReview(RR)
Week 6	Case 2: Assault/attack	■ CS & RR
Week 7	Case 3: Business Negligence	■ CS & RR
Week 8	Case 4: Traffic Rule Violation	■ CS & RR
Week 9	Case 5: Document Counterfeiting	■ CS & RR
Week 10	Case 6: Driving Under Influence	■ CS & RR
Week 11	Case 7: Drug Trafficking	■ CS & RR
Week 12	Case 8: Drug Possession	■ CS & RR
Week 13	Case 9: Violation of Sexual Autonomy	■ CS & RR
Week 14	Case 10: Fraud	■ CS & RR
Week 15	Case 11: Compulsory Indecency	■ CS & RR
Week 16	Case 12: Patent Violation	■ CS & RR
Week 17	Case 13: Traffic Accident (post-test)	■ CS & RR
Week 18	End-of-Semester Portfolio Assessment	■ Assessment

Appendix C

Interview Questions

1. Looking back at your learning this semester, how would you rate the training method, the 5W1H information extraction strategy, in the Likert scale of one to five (not useful at all, not useful, neutral, useful, and very useful)? Why?

2. Using the scale of 1 to 4 (poor, fair, great, excellent), how do you rate your improvement in the 13 weeks of case simulation training?

3. During this period of training, you interpreted one summarized indictment each week. They involved different case scenarios, but the key information shared some commonality. Can you recall why your performances fluctuated a bit in the xx week? [8]

---

[8] Based on the students' individual performances, the necessary information about their performance fluctuation was provided in this question.