

RESEARCH ARTICLE

The Influence of Pre-task Preparation on Cognitive Load in English to Chinese Consecutive Interpreting Process

Received: 18 April 2024; Revised: 09 November 2024; Published: 31 December 2024

Shuangshuang Yang

Macquarie University, Australia

Email: shuangshuang.yang@hdr.mq.edu.au

ORCID: <https://orcid.org/0000-0002-5522-8630>

Lei Mu

Guangdong University of Foreign Studies, Guangzhou, China

Email: mulei2002@139.com

ORCID:

Abstract:

Existing studies have revealed that sufficient pre-task preparation enhances interpreting task performance by improving accuracy and reducing information omissions. However, its influence on interpreters' cognitive load remains underexplored. Given that disfluency features in target speech can serve as key indicators of interpreters' cognitive load during the process, this study explores how disfluency patterns differ between interpreter groups employing different pre-task preparation approaches for an English to Chinese consecutive interpreting. We investigated the occurrence of filled and silent pauses, repairs, repetitions, and false starts in the target speech among three groups of interpreting trainees. One group was instructed to employ extra-linguistic knowledge-directed preparation strategies, another utilized linguistic knowledge-directed preparation strategies, while the third was prohibited from task preparation. The results indicate that the choice of preparation strategy can affect preparation efficacy, and sufficient preparation before the task can alleviate cognitive load as evidenced by reduced frequency of disfluency features. These findings offer invaluable insights into interpreting training, practice and research, providing evidence-based recommendations on effective pre-task preparation strategies and demonstrating how targeted preparation can alleviate cognitive load.

Keywords: extra-linguistic knowledge-directed preparation strategies, linguistic knowledge-directed preparation strategies, disfluency features, pauses, repairs, repetitions, false starts

1. Introduction

Consecutive interpreting is a mode of interpreting where interpreters convey information in the target language after the speaker pauses upon completing one or more ideas in the source language (Russell, 2005). This mode possesses cognitive challenges due to the interpreters' limited cognitive resources, potentially insufficient for the task, resulting in deteriorated task performance, such as increased errors and information omissions (Gile, 2009). To improve task performance, researchers, practitioners, and trainers are exploring methods to alleviate interpreters' cognitive load, the cognitive pressure imposed by tasks (Gile & Lei, 2021).



One potential solution is pre-task preparation, which is considered a crucial factor in determining the cognitive efforts required for completing an interpreting task (Chen, 2017), could be a possible solution. Previous studies on pre-task preparation have mainly focused on two aspects: its impact on interpreting task performance and the preparatory strategies employed by professional interpreters. Studies comparing task performance with and without ample time or technological aids have consistently shown that adequate pre-task preparation significantly enhances performance by improving accuracy (Pérez, 2018; Xu, 2018), especially regarding specialized terminology, and by reducing ear-voice span and omission rates (Galaz et al., 2015; Galaz, 2011; Lamberger-Felber, 2001). However, disfluency, defined as interruptions in speech plan execution (Postma et al., 1990) and regarded as an important indicator of cognitive overload, has received limited attention. Drawing from studies investigating pre-task preparation procedures and contents (Chang et al., 2018; Fantinuoli, 2017), two approaches can be utilized for preparing interpreting tasks: extra-linguistic knowledge-directed and linguistic knowledge-directed preparation strategies. Therefore, the objective of the current contribution is to investigate the impact of these two pre-task preparation approaches on cognitive load, as indicated by disfluency features, during the Chinese-to-English consecutive interpreting process.

2. Literature Review

Interpreting activity has always been considered cognitively demanding due to its dynamic and irreversible nature, which involves several interactive and independent sub-activities, such as analytical listening, comprehension, and reformulation (Gile, 2009; Setton & Dawrant, 2016). As these sub-tasks often overlap, interpreters are assumed to work close to saturation (Gile, 2009), leading to cognitive overload. Consequently, measuring cognitive load during the interpreting process can be utilized to gauge the dynamic nature of interpreting.

2.1 Measuring cognitive load in the interpreting process

2.1.1 Cognitive load measurement methods

As stated by Seeber (2015), the concept of cognitive load, initially introduced to Interpreting Studies by Gile (1985), was associated with observed information loss among professional interpreters, and later defined as “*the cognitive pressure that a process imposes by virtue of environmental and task-specific factors*” (Gile & Lei, 2021).

Four types of methods for measuring cognitive load during the interpreting process were summarized (Seeber, 2013): analytical method, subjective method, performance method, and psychological method. The analytical method involves utilizing the existing models and theories, such as Gile’s Effort Model (Gile, 2009), to gather subjective and analytical data to observe interpreting task characteristics, which can occur purely at a theoretical level and avoid empirical testing. The subjective method entails collecting self-reported data through introspection or rating scales, enabling performers to express their perceived feeling of effort or exertion. For instance, the NASA Task Load Index (Hart & Staveland, 1988) has been identified as one of the most widely used scales for measuring cognitive load. Performance methods use task performance features, such as disfluency features like pauses



(Mead, 2000, 2002, 2005) and fillers ‘uh (emm)’ (Moratto & Yang, 2023), to measure cognitive load in consecutive interpreting. Finally, the psycho-physiological method employs psychophysiological techniques, such as electroencephalography (Koshkin et al., 2018) and functional near-infrared spectroscopy (fNIRS) (Yan et al., 2024), to measure cognitive load, offering more objective results as they directly reflect performers’ physiological arousal.

2.1.2 Measuring interpreters’ cognitive load through disfluency features

Measuring the performance of interpreters as a reflection of cognitive load has a long history in the field of interpreting studies (Seeber, 2013). As explained in the theoretical framework outlined by Chen (2017), the relationship between cognitive load and interpreting performance suggests that cognitive load impacts interpreting performance only when it starts to exceed the interpreter’s available capacity. At this threshold, a decline in performance becomes noticeable. However, when interpreters possess sufficient capacity to manage the effort demands, performance remains stable. Conversely, when cognitive load is exceptionally high, performance may deteriorate to such an extent that differentiating between varying levels of load as performance may deteriorate to a critical level. In this latter case, performance degradation may reach a critical level, making it challenging to observe distinct changes in performance. One of the manifestations of performance decline in the process of interpreting is disfluency or disfluency features. Common types of speech disfluencies include hesitations, pauses, fillers (such as um and ah), corrections, false starts, repetitions, interjections, stuttering, and slips of tongue (Garnham, 2013). In the interpreting process, four types of disfluencies are frequently observed, namely false starts, repetitions, repairs, and pauses (Song & Cheung, 2019). False starts occur when the initial syllable is incorrectly pronounced and replaced with a correct one. Repetitions involve repeatedly producing the same information in speech. The definition of repairs and pauses is more complicated:

Repairs, or “self-corrections”, are defined as one of the discourse features used to measure spoken languages (Foster et al., 2000), involving speakers correcting errors made during or immediately after speech production. Therefore, the occurrences of repairs interrupt the flow of speech and prompt a new reformulation. Levelt (1983) observed that repairs often coincide with pauses. According to his framework, repairs that occur during the pre-articulatory phase are classified as covert repairs, which are akin to filled pauses. Conversely, repairs occurring during the post-articulatory phase are classified as overt repairs, which are further subdivided into three types by Song and Cheung (2019): A-repairs, E-repairs, and D-repairs. A-repairs, short for appropriateness-repairs, serve to clarify potential ambiguity in context and correct inappropriate use of terminology, as well as inconsistencies with previously used terms and expressions (Levelt, 1983). E-repairs involve correcting errors made by speakers and can be further divided into corrections of errors at the lexical, syntactic, and phonetic levels (Levelt, 1983). D-repairs denote corrections of errors in the conceptualizer (Kormos, 1999). Therefore, speakers substitute the produced message with a new and different message when such an error occurs.

Pauses serve as indicators of uncertainty and hesitation, often arising when speakers encounter challenges in speech production or experience cognitive overload. Initial research on pauses dates to the 1960s, with more detailed studies emerging in the 1980s (Cecot, 2001). Pauses are generally



categorized into two types: silent pauses and filled pauses. Silent pauses, also known as unfilled pauses, are periods of silence between vocalizations lasting at least 200-270 milliseconds (Hargrove & McGarr, 1994). Setton (1999) identified three functions of silent pauses: 1) directing exclusive attention to input information; 2) special formulation or information retrieval; and 3) controlled ideal breaks at constituent or discourse boundaries. In the current study, which aims to examine silent pauses resulting from cognitive overload, the third category is excluded during data collection and analysis. Filled pauses involve interruptions in speech flow by non-lexical sounds and meaningless sounds, such as "ah," "er," or "uh," or the elongation of syllables and words (Cenoz, 1998). In their study comparing disfluencies between relay and non-relay simultaneous interpreting, Song and Cheung (2019) considered filled pauses a significant disfluency type, as they may reflect brief shifts in attention during planning or retrieval, thus contributing to cognitive overload.

Pauses including silent pauses and filled pauses, repetitions, restructuring, and false starts were measured to examine the differences between source and target speech (Tissi, 2000). The study revealed that target speech contained significantly fewer but longer silent pauses compared to source speech, with the frequency of filled pauses varying considerably among individuals. Cecot (2001) replicated these findings by analyzing the source and target speech of eleven professional interpreters. Mead (2000) explored the impact of interpreting directionality on cognitive load using a sample of 15 interpreting students. This study found that filled pauses increased when interpreting into the B language, as opposed to interpreting from the B to A language. as opposed to interpreting from the B to A language. Bóna and Bakti (2020) examined the differences of cognitive load revealed through the frequency of pauses, fillers, repetitions, broken words, prolonged sounds, and revisions across four types of speech tasks: consecutive interpreting, sight translation, spontaneous speech, and extemporaneous speech. Their findings indicated that these four types of speech production can be represented on a continuum of cognitive load, with sight translation at the lower end and spontaneous speech at the higher end. Extemporaneous speech exhibited a higher cognitive load than spontaneous speech but a lower cognitive load than consecutive interpreting.

In addition to experimental settings, studies in natural settings have employed corpus-based approaches. Plevoets and Defrancq (2016) investigated the triggers of disfluencies for both spontaneous speakers and interpreters, by probing the differences between source speech and target speech disfluencies for interpreters. Their findings indicated that interpreters produce more 'uh(m)s' than non-interpreters, supporting the notion that interpreting is more cognitively demanding than spontaneous speech. Triggers of cognitive overload in interpreting output included delivery rate, lexical density, and sentence length, although the proportion of numerals showed a decreasing trend. When examining the influence of cognitive overload triggers in the source speech, a higher proportion of numerals led to increased 'uh(m)' occurrences. Plevoets and Defrancq (2018) identified three limitations in their earlier study (Plevoets & Defrancq, 2016) —unit of analysis, data analysis model, and the measurement of syntactic complexity. These issues were addressed in their subsequent study (Plevoets & Defrancq, 2018), which revealed that lexical density, particularly in the source speech, heightened interpreters' cognitive load, while formulaicity had a mitigating effect. Additionally, Moratto and Yang (2023) explored the impact of interpreting directionality on cognitive load as indicated by 'uh(m)' frequency. They investigated four cognitive overload triggers: lexical density, proportion of numbers, sentence length, and delivery rate. Their findings indicated that lexical density



in both source and target speeches significantly influenced cognitive load when interpreting from language A to language B. However, this effect was not observed when the directionality was reversed. Furthermore, the impact of delivery rate on cognitive load displayed an inverse trend, with significant effects found only when interpreters worked from language B to language A. In contrast, the proportion of numbers and sentence length both demonstrated a significant impact on cognitive load.

2.2 Pre-task preparation as a factor influencing cognitive load

Cognitive load was perceived as a complex construct that reflected the interplay of task, environmental, and interpreter-related characteristics (Chen, 2017). Task and environmental characteristics determine the cognitive effort needed to accomplish a given task under specific circumstances, whereas interpreter characteristics pertain to the cognitive effort exerted and experienced by the interpreter (see Table 1). While preparation is identified as one of the task characteristics influencing the cognitive effort required for the task, there is limited research examining its potential impact on cognitive load.

Table 1. Task, Environmental and Interpreter Characteristics Related to Cognitive Load in the Interpreting Process (Chen, 2017)

Task characteristics	Environmental characteristics	Interpreter characteristics
Interpreting mode	Physical environment or conditions	Cognitive abilities
Language pair	Visibility of speaker or audience	Motivation
Interpreting direction	Equipment	Experience
Speech features		State of arousal or Activation level
Speaker features		
Expected response		
Time on task		
Preparation		
Task criticality		
Task novelty		

Most of the existing studies focused on its effects on accuracy and information omission in the target speech, with little attention given to the disfluency features. These studies aimed to prove the hypothesis that interpreters demonstrate enhanced performance when adequately prepared for interpreting tasks (the efficacy of pre-task preparation). Researchers have typically conducted comparisons between groups to assess this. For example, Anderson (1994) examined the interpreting performances of two cohorts of professional interpreters: one received a manuscript or summary of the source speech immediately before the task, while the other did not receive any materials. In contrast to expectations, the former cohort did not surpass the latter in terms of intelligibility and accuracy. This outcome may be attributed to the limited preparation time, as interpreters only accessed the manuscript just before the task. Consequently, they had to both listen to the source speech and read the manuscript during interpretation. Processing input from two channels simultaneously might impose greater cognitive demand compared to single-channel input (Gerver, 1975), resulting in no discernible discrepancy between the two groups. Lamberger-Felber's (2017) investigation reinforced this inference



by analyzing the performances of a cohort of professional interpreters across three interpreting tasks under varying conditions: receiving the source text with adequate preparation time, receiving the source text without preparation time, and not receiving the source text at all. The study revealed a decrease in overall erroneous renditions and omissions of proper names and numbers when interpreters were provided with sufficient time to prepare using the source text. Nonetheless, there were no discernible distinctions among the three conditions regarding semantic deviations. Considering that the studies discussed above exclusively focused on professional interpreters, there arises a question about the applicability of their findings to interpreting students. This question was addressed by studies comparing the performance of interpreting students across two tasks: one allowing preparation and the other not. Galaz (2011) and Díaz-Galaz et al. (2015) discovered that preparation significantly impacted the performance of interpreting students, reducing the ear-voice span, increasing the accuracy rate, and decreasing the omission rate when students were allowed to prepare for the task.

In recent years, as technology has become increasingly prevalent across all industries, its influence and application in the field of interpreting have also gained traction. Researchers have been particularly interested in exploring whether new preparation methods, which involve the use of technological tools such as corpus management tools, could offer greater benefits to interpreting task performance compared to traditional methods mainly depending on manual preparation using pen and paper (Pérez-Pérez, 2018; Xu, 2018; Liu, 2011). The findings of these studies have been consistent, demonstrating that corpus management tools and online learning resources like Webquest have enhanced interpreters' task performance by improving the accuracy of specific terminologies and expressions.

2.3 Pre-task preparation strategies

Despite the established benefits of pre-task preparation in interpreting tasks, both interpreting practitioners and researchers often take it for granted. According to a survey on the workload of professional interpreters conducted by AIIC (2002), 60 percent of the respondents reported a high frequency of “*not receiving background material*”, while 40 percent indicated a high frequency of “*not having enough time for to prepare*”. Moreover, another survey on glossary usage among interpreters found that none of the responding interpreters created a glossary for every meeting they worked at (Jiang, 2013). Specifically, 62.5 percent of them reported generating glossaries only for technical or unfamiliar meetings, while only 31.3 percent of them did so for most meetings (Jiang, 2013).

Furthermore, there is a notable lack of research into effective methods for interpreting preparation. Existing studies often provide insufficient guidance on the procedures and strategies for preparing for interpreting tasks. For example, an exploratory survey conducted in Taiwan (Chang et al., 2018) investigated how professional interpreters prepare for domain-specific tasks through interviews. The findings indicated that interpreters strategically prepared for unfamiliar topics by utilizing conference documents, and compiling glossaries to organize domain-specific concepts and terminology, thereby facilitating comprehension and reformulation. They also incorporated language usage of speakers and participants into the preparation process, viewing interpreting as a long-term activity where knowledge from one task could be applied to prepare for future tasks on similar topics. Thus, effective interpreting preparation involves efforts before, during, and after conferences to acquire task-related knowledge. In addition to understanding what professional interpreters do to prepare for the tasks, attention was also



given to the technological tools that should be employed. Moser-Mercer (1992) compared the workflow of interpreters and translators, suggesting although electronic tools were beneficial to translators, they often did not meet the needs of interpreters. This led to advice on designing terminological management software or a databank specifically for interpreting purposes. Subsequently, Fantinuoli (2012) introduced a computer-assisted interpreting tool, namely InterpretBank, which can enhance glossary searching efficiency, even for novice interpreters, and improve students' workflow.

Although the number of studies detailing strategies for preparing interpreting tasks is limited, there is consensus that pre-task preparation involves acquiring domain-specific knowledge (Chang et al., 2018; Díaz-Galaz et al., 2015; Fantinuoli, 2017; Galaz, 2011). This encompasses two types of knowledge: 1) extra-linguistic knowledge, also known as topic-specific knowledge, concerning the expertise in a particular topic; 2) linguistic knowledge, encompassing the correct usage of terms, jargon, and professional expression (Fantinuoli, 2017). Based on this categorization, pre-task preparation can be conducted through two approaches, namely the extra-linguistic knowledge-directed, and linguistic knowledge-directed preparation strategies. The former approach prioritizes comprehending concepts so as to construct a knowledge framework of a particular domain. This can be achieved using materials provided by conference organizers and speakers, information searched and learned through online methods, as well as other information obtained from other forms. Conversely, the latter primarily concerns the memorization of terms and expressions listed in the glossary.

The above-discussed review revealed that investigations on the impact of pre-task preparation strategies on cognitive load revealed by the disfluency features are scarce. To fill in this gap, the current paper aims at examining the impact of two different approaches of pre-task preparation strategies on the cognitive load in Chinese to English consecutive interpreting process, through the perspective of disfluency features. In order to achieve this aim, the following research questions should be answered:

- 1) Do different pre-task preparation methods result in varying preparation outcomes?
- 2) How do various preparation methods influence interpreters' cognitive load?

3. Research Design

3.1 Participants and procedures

Thirty interpreting trainees (female N=26, male N=4, average age of 23.9) enrolled in an MTI (Master of Translation and Interpreting) program at a university in China are recruited and compensated for participating in the experiment. These trainees are selected because they successfully completed and passed the final examination of the Consecutive Interpreting Course, indicating they possessed the necessary skills for consecutive interpreting. The participants are divided into three groups: 1) Group 1, utilizing the extra-linguistic knowledge-directed preparation strategies; 2) Group 2, employing the linguistic knowledge-directed strategies; and 3) Group 3, serving as control groups, prohibited from preparing for the interpreting task.

Each participant is tasked with a 5-minute consecutive interpreting assignment from English to Chinese. The experimental groups received preparation materials and instructions on preparation



strategies three days before the task. All participants' interpretations were audio-recorded and transcribed for subsequent data analysis. After the preparation and just before the interpreting task, participants in the experimental groups are required to complete an online questionnaire aiming at examining the results of interpreting preparation. The quantitative data from the questionnaire was analyzed using variance analysis to address the first research question. Additionally, the frequency of disfluency features in the transcriptions of participants' interpretations was assessed and analyzed using variance analysis to answer the second research question.

3.2 Consecutive interpreting task material

A consecutive interpreting test was adapted from a 30-minute keynote speech on the topic of decarbonization in the shipping industry. Since interpreting preparation has been proven to be more efficient and helpful for technical conferences, the authors of the current paper opted for a technological topic to maximize the potential impact of preparation on disfluency features. Economic, business, environmental, and political themes, commonly covered in interpreting courses and likely familiar to interpreting trainees, were therefore excluded from the resource pool. Consequently, the shipping industry, a relatively new and technical subject for the trainees, with an international conference held on this topic being recent at the time of the experiment, was chosen.

The original speech was divided into six segments, each approximately five minutes long. These segments were then evaluated by a group of nine trainees (who were not participants in the experiment) to determine the most suitable segment for testing based on difficulty level. These trainees were asked to listen to each segment as an interpreting task material, after which they completed a simple questionnaire to categorize each segment as easy, appropriate, or difficult through a wholistic perspective. The segment that received the highest consensus as being "appropriate" was selected for the study. The second segment of the speech, which six of the students categorized as appropriate, was selected and transcribed into written text to exclude disfluencies of the original speech, as these disfluencies could potentially affect the fluency of the target language production. The written text was carefully reviewed, revised, and recorded by a native speaker to ensure grammatical accuracy, appropriate speech rate, and clear sound quality. The final version, lasting approximately 4.86 minutes, was divided into five segments, each lasting one minute, for the consecutive interpreting task.

3.3 Pre-task preparation instructions

Two types of pre-task preparation strategies were identified in the existing literature, namely the extra-linguistic knowledge-directed preparation strategies and the linguistic knowledge-directed preparation strategies. For each type, the authors have developed detailed instructions outlining step-by-step operational procedures.

3.3.1 Extra-linguistic knowledge-directed preparation strategies

The focus of extra-linguistic knowledge-directed interpreting preparation strategies is the contextual information and the knowledge structure of a specific domain within which linguistic knowledge is embedded. Participants preparing for a consecutive interpreting task using this approach are required to thoroughly seek relevant information. This includes understanding the conference topic, acquiring



domain knowledge related to the subject, and achieving a broad and in-depth understanding of the topic and conference proceedings, including operational dynamics within the domain.

Utilizing this approach, participants receive the conference topic and name three days in advance. They are given one day to gather information on the topic. Subsequently, the speaker's PowerPoint (PPT) slides are provided two days before the task. Participants are instructed to refine their preparation based on the content of the slides, with a primary focus on the information presented within them.

Before completing the interpreting task, participants are required to submit three types of preparation products to the researchers. The first is a knowledge map of the entire topic, encompassing information on the conference, frequently discussed sub-topics, relevant theories, their applications, and details about the speaker. This map is drawn based on the preparation undertaken on the first day. The second product is another knowledge map, focusing specifically on the content of the PowerPoint slides. Participants must not only anticipate the structure of the speech based on the PPT slides but also integrate this information with the broader domain knowledge map to identify relationships between the speech and the whole domain. Finally, participants are required to provide a glossary, which is considered a by-product of the knowledge acquisition process. This glossary may be monolingual, bilingual, or even multilingual.

3.3.2 Linguistic knowledge-directed preparation strategies

This preparation method focuses on acquiring and memorizing equivalent expressions of terminologies, in-house jargon, and key concepts. Consequently, participants are not required to comprehend knowledge to the level of an expert in the domain but rather at a linguistic level.

Participants using this approach receive the speaker's PowerPoint slides directly three days in advance. Their focuses are directed toward identifying unfamiliar and unknown words, expressions, and terminologies within the slides. They are tasked with discovering equivalent expressions for these linguistic components and memorizing them before the task. However, participants are permitted to search for relevant knowledge of these linguistic components if necessary.

After preparation, participants are required to submit only one product to the researcher: a list of glossaries. This list may comprise words, phrases, terms, and even sentences, all presented in both the source and target language.

3.4 Questionnaire for measuring interpreting preparation results

The questionnaire comprises three sections. The first section collects personal information from each participant, including their experiment number, age, gender, and group number. The second section consists of eight single-choice questions designed to assess linguistic knowledge, focusing on terminologies and expressions extracted from the material used in the interpreting task. The final section includes two single-choice questions, and three multiple-choice questions, aimed at evaluating extra-linguistic knowledge based on the interpreting task material.

The questionnaire is presented in Chinese to accommodate the fact that all participants are Chinese students, thus minimizing potential language-related misunderstandings. Items within the questionnaire are organized into distinct clusters: items 1 through 4 gather personal information, items



5 through 12 assess participants' familiarity with linguistic knowledge related to the interpreting task, and the last five items evaluate participants' domain-specific knowledge. The questionnaire was created using the online software "WEN JUAN XING" and distributed directly to participants via WeChat.

3.5 Data analysis tool: ELAN

The recordings of the interpreting renditions are transcribed and synchronized using ELAN 5.0.0, a widely used tool for annotating video and audio resources. ELAN 5.0.0 converts acoustic signals into oscillograms, enabling researchers to visualize the sounds and annotate markers with a precision of up to 1 millisecond. An example of the annotation in ELAN is illustrated in Figure 1.

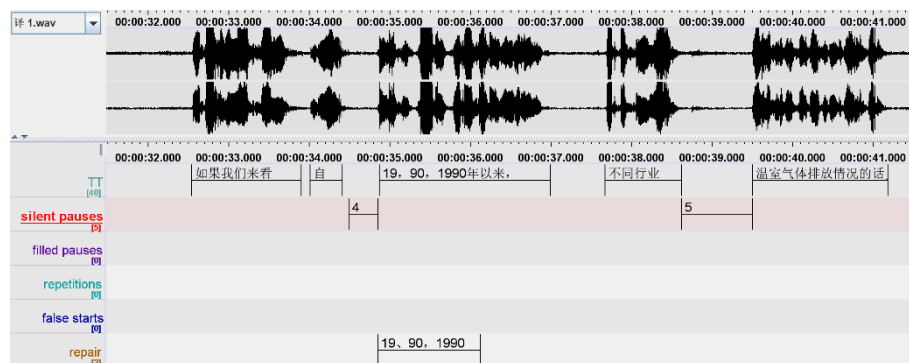


Figure 1 Annotation Example in ELAN

Notes: Tier 1(TT) contains the transcription of the interpreting renditions; Tiers 2, 3, 4, 5, and 6 annotate disfluencies features within the renditions including pauses (silent and filled pauses), repetitions, false starts, and repairs. Specifically, a silent pause denotes a period during which no acoustic signal is present, except for those occurring after a complete sentence or phrase.

4. Results

4.1 Intergroup differences of interpreting preparation results

Although 30 participants were involved in the experiment, only 20 were permitted to prepare for the interpreting task. Among them, the first 10 participants followed the instructions for extra-linguistic knowledge-directed preparation, while the subsequent 10 participants adhered to the instructions for linguistic knowledge-directed preparation. Consequently, only these 20 participants completed the preparation questionnaire.

The descriptive statistics of the interpreting preparation results, including average, minimum, and maximum scores, and standard deviations, are presented in Table 2 and Table 3. The analysis reveals that the average total score of Group 1 slightly exceeds that of Group 2. However, Group 2 scored higher on the second part of the questionnaire, which assesses linguistic knowledge, such as expressions and terminologies in the shipping industry. Conversely, Group 2 scored lower on the third part, which evaluates the extra-linguistic component including thematic information, key concepts and their definitions discussed in the speech, and the knowledge structure of the shipping industry. This suggests that although participants in group 2 attained a lower total score, they performed better on



items assessing linguistic components, likely due to emphasis on searching for, understanding, and memorizing the terminologies, jargon, and specific expressions commonly used in the shipping industry when they prepare the task. However, participants in group 1, who were instructed to prioritize extra-linguistic knowledge preparation with linguistic preparation as the by-product, exhibited better overall performance and particularly excelled in the last part of the questionnaire.

Table 1. Descriptive Statistics about Interpreting Preparation Result of Group1

	Min.	Max.	M	SD.
Total Score	50.00	75.00	63.50	7.47
Part2 Score	25.00	40.00	30.00	4.08
Part3 Score	20.00	40.00	33.50	6.26

Table 2. Descriptive Statistics about Interpreting Preparation Result of Group2

	Min.	Max.	M	SD.
Total Score	55.00	75.00	60.50	6.43
Part2 Score	35.00	40.00	38.00	2.58
Part3 Score	20.00	35.00	22.50	4.86

To determine whether there were statistical differences between the scores of Group 1 and Group 2, all the data were initially analyzed using SPSS 26.0 to assess their normal distribution. The Normality Test results, as shown in Table 4, indicated that only the total score of the interpreting preparation questionnaire followed a normal distribution ($p > .05$). In contrast, the scores of Part 2 and Part 3 within the questionnaire were not normally distributed ($p < .05$). Consequently, the Independent Samples T-Test was employed to analyze the total score, while the Mann-Whitney U-test was utilized for analyzing the scores of Part 2 and Part 3.

Following the Normality Test, Levene's test was performed to assess the homogeneity of variances for the total pre-task preparation scores. The results yielded a p-value of 0.57, indicating no statistically significant difference in variance between Group 1 and Group 2. Subsequently, an Independent Samples T-Test was conducted to evaluate differences in total scores between the groups. The analysis revealed no significant difference ($p = .349 > .05$ as shown in Table 5 Variation Analysis Results). In contrast, the Mann-Whitney U-test was utilized to investigate group differences in scores for Parts 2 and 3. The findings indicated significant group differences, with p-values of .005 and .003, respectively, demonstrating notable discrepancies between Group 1 and Group 2 in these two parts.

Table 3. Normality Test Result of Interpreting Preparation Results

	Min.	Max.	M	SD.	Shapiro-Wilk (Sig.)
Total Score	50.00	75.00	62.00	6.96	.191
Part2 Score	25.00	40.00	34.00	5.28	.005
Part3 Score	20.00	40.00	28.00	7.85	.003



Table 4. Variation Analysis about Interpreting Preparation

Group	N	Mean	SD.	P of Total Score	P of Part1 Score	P of Part2 Score
1	10	63.50	6.96	.349	.000	.002
2	10	60.50	5.28			

4.2 Intergroup differences of the distribution of disfluency features

Tables 6, 7, and 8 present the average, minimum and maximum frequencies of all types of disfluencies features in participants' interpretations, along with their standard deviations. Descriptive analysis indicated that the average frequency of total disfluencies features in interpreting renditions of Group 1 was lower than that of Group 2 and 3, with Group 2 exhibiting a lower average frequency than Group 3. For individual disfluency features, the average frequency was consistently lowest in Group 1, followed by Group 2, and the highest in Group 3. The only exception is false starts, where the average frequency remains the same across all three groups (M=.50).

Table 5. Descriptive Statistics about Disfluencies Phenomena of Group1

Disfluency Type		Min.	Max.	M	SD.
Pauses	Silent Pauses	2.00	14.00	7.90	4.07
	Filled Pauses	.00	12.00	2.00	3.68
Repairs		1.00	9.00	3.40	2.55
Repetitions		.00	1.00	.30	.48
False Starts		.00	3.00	.50	1.08
Total		3.00	30.00	9.90	6.49

Table 6. Descriptive Statistics about Disfluencies Phenomena of Group2

Disfluency Type		Min.	Max.	M	SD.
Pauses	Silent Pauses	4.00	35.00	24.50	9.14
	Filled Pauses	.00	23.00	9.90	8.53
Repairs		.00	32.00	5.90	9.42
Repetitions		.00	12.00	2.10	3.70
False Starts		.00	1.00	.50	.53
Total		4.00	97.00	42.90	23.6

Table 7. Descriptive Statistics about Disfluencies Phenomena of Group3

Disfluency Type		Min.	Max.	M	SD.
Pauses	Silent Pauses	6.00	43.00	28.90	9.30
	Filled Pauses	.00	28.00	16.20	8.42
Repairs		.00	15.00	8.80	4.76



Repetitions	.00	13.00	6.70	4.32
False Starts	.00	2.00	.50	.71
Total	8.00	92.00	61.10	22.03

To assess the significance of between-group differences in disfluency features are significant, variation analyses were conducted. Before these analyses, normality tests were performed to determine the most appropriate analytical methods. The results (refer to Table 9 Normality Test Results on the Frequencies of Disfluency Phenomena) indicated that, except for the frequencies of total disfluency phenomena ($p=1.21 > .05$), the frequencies of each individual disfluency type were normally distributed. Therefore, intergroup differences in the frequencies of each disfluency type were examined using Independent Samples T-Tests, while the frequencies of total disfluency phenomena were analyzed using the Mann-Whitney U-test.

Table 8. Normality Test Results on the Frequencies of Disfluencies Phenomena

Disfluency Type		Min.	Max.	M	SD.	Shapiro-Wilk
Pauses	Silent Pauses	2.00	43.00	20.43	11.94	.017
	Filled Pauses	.00	28.00	9.37	9.15	.002
Repairs		.00	32.00	6.03	6.45	.000
Repetitions		.00	13.00	3.03	4.20	.000
False Starts		.00	3.00	.50	.78	.000
Total		3.00	97.00	39.37	27.07	.121

Based on the results of various analyses on the frequencies of disfluency phenomena (as shown in Table 10 Variation Analysis Results on Disfluency Phenomena), all three groups exhibit significant differences from one another in the frequencies of total disfluency phenomena ($p=.002$, $.001$, and $.019 < .05$). Regarding the frequencies of each disfluency type, a significant difference between Group 1 and Group 2 was observed only in the frequencies of silent pauses ($p=.00 < .05$) and filled pauses ($p=.015 < .05$). Additionally, a significant intergroup difference is found between Group 2 and Group 3 in terms of the frequencies of repetitions ($p=.02 < .05$). However, Group 1 and Group 3 differed significantly from each other in most types of disfluency frequencies, except for false starts ($p=1.00 > .05$).

Table 9. Variation Analyses Results on Disfluencies Phenomena

Disfluency Type			Group2	Group3
Pauses	Silent Pauses	Group1	.000	.000
		Group2		.300
		Group1	.015	.000



	Filled Pauses	Group2		.114
Repairs		Group1	.429	.005
		Group2		.396
Repetitions		Group1	.114	.00
		Group2		.020
False Starts		Group1	1.00	1.00
		Group2		1.00
Total		Group1	.002	.001
		Group2		.019

5. Discussions

5.1 Interpreting preparation outcomes

This study aims to investigate the impact of two interpreting preparation methods on the cognitive load, as measured by disfluency features. The hypothesis was that these two preparation methods would yield different preparation outcomes due to their distinct procedural focuses. While some literature advocates for prioritizing advance preparation focusing on extra-linguistic information, such as domain-specific knowledge, others highlight the importance of linguistic preparation, particularly the terminological component (Gile, 2009). However, no direct comparison between the effects of linguistic and extra-linguistic preparation components had been made.

To address the first question of this study, such a comparison was undertaken, revealing that interpreting trainees who conducted interpreting preparation guided by the extra-linguistic knowledge-directed method achieved higher scores in tests assessing preparation effects. This can be attributed to the fact that when interpreters utilize the linguistic knowledge-directed preparation method, their attention is primarily devoted to searching for and memorizing terms and expressions, thereby potentially overlooking the extra-linguistic knowledge embedded within. This finding aligns with the suggestion that terminological preparation for interpreting tasks should involve knowledge system building (Will, 2007).

Furthermore, besides enhancing performance in specific interpreting tasks, the domain knowledge-directed interpreting preparation strategies also prove effective in improving interpreting competence by fostering the development of a systematic and comprehensive world knowledge system. World knowledge, including domain-specific knowledge, is recognized as a crucial component of interpreting competence (Wang, 2007), which interpreting trainees must acquire and cultivate during their studies. Adopting the extra-linguistic knowledge-directed preparation strategies thus represents a means to construct and broaden interpreters' knowledge systems.

Contrary to our expectations, interpreting students and professional interpreters who used a linguistic knowledge-directed preparation approach achieved higher scores on items assessing only the linguistic components. This result is consistent with existing findings conducted by Anderson (1994), Galaz



(2011), and Galaz et al. (2015), in which preparation materials were given to the participants just before the interpreting task. Although the researchers of these studies claimed that the preparation time for the interpreting task was sufficient, participants could only prepare the interpreting task based on the provided materials, focusing on linguistic elements. The findings of these studies consistently demonstrated a higher overall accuracy rate of the interpretations (Galaz, 2011; Galaz et al., 2015) and of the numbers and in the target speech names (Anderson, 1994) for participants who conducted preparation in this manner to those who did not.

5.2 The influence of pre-task preparation on disfluency features

Based on the data analysis, interpreting trainees who prepared for the interpreting task using the extra-linguistic knowledge-directed preparation strategies exhibited significantly fewer occurrences of disfluency features compared to those who prepared using linguistic knowledge-directed strategies and those who were not permitted to prepare for the task. The findings suggest that thorough preparation before an interpreting task reduces cognitive load, leading to improved performance in interpreting tasks.

Originally investigated in second language acquisition as indicators of the second language production process, disfluency features have increasingly intrigued researchers in the field of interpreting studies, who use them as indicators of cognitive overload during interpreting tasks, producing fruitful findings. For instance, Pöchhacker (1995) observed more slips and shifts in interpreting production compared to monolingual speech, indicating that the interpreting task is more challenging and demanding. Similar results were obtained when disfluency features of four types of speech production were compared, identifying sight translation and consecutive interpreting as more cognitively demanding than extemporaneous and spontaneous speech (Bóna & Bakti, 2020). Song and Cheung (2019) found different patterns of disfluencies when comparing the output of relay and non-relay simultaneous interpreting. In addition to speech types, comparisons of disfluency features between source speech and target speech of an interpreting task were also conducted, revealing similar results (Tissi 2000) study. Although these studies investigated the triggers of cognitive overload in the interpreting process, what can be done to ease the cognitive demand awaits further investigation. Therefore, the current study expanded the findings of existing studies by providing empirical evidence on the effectiveness of pre-task preparation in reducing cognitive load. Regarding specific types of disfluency features, the present study found significant differences in pauses only between Group 1 and Group 2. No significant differences were observed between Group 2 and Group 3. Pauses, both silent and filled, are common responses to lexical, morphological, and planning difficulties, with filled pauses being more indicative of planning difficulties (Cenoz, 1998; Setton, 1999). With better preparation, interpreters may experience fewer difficulties in linguistic and planning aspects, thereby reducing cognitive load. The lack of intergroup differences in pauses could be attributed to the preparation method employed by Group 2, which focused on searching and memorizing terms and expressions. Since speech production planning heavily relies on comprehension of the speech content and effective comprehension during interpreting requires not only language proficiency but also extra-linguistic knowledge and deliberate analysis of the entire knowledge system (Gile, 2009), Group 2 students may have experienced stalled logic analysis due to insufficient knowledge.



The only significant difference in interpreting production between Group 2 and Group 3 was in repetitions, which are also considered a strategy to manage time pressure in the interpreting process (Cecot, 2001). Despite lacking extra-linguistic knowledge, Group 2 students possessed ample linguistic knowledge, which may have saved them energy when encountering terminologies and expressions from a specific community, thereby reducing time pressure.

In addition to pauses and repetitions, differences in the frequencies of repairs distinguished the output of Group 1 and Group 3. Repairs are self-correction behaviors when errors are detected in speech production. Petite (2005) investigated different categories of repair mechanisms in simultaneous interpreting and found that interpreters consistently invested processing resources in producing repairs, even when it was not cost-effective. Given that Group 1 students underwent the most preparation in this experiment while Group 3 students did not prepare at all, Group 1 students were more contextualized in their task, better at monitoring the accuracy and completeness of their productions and made fewer errors in their renditions. However, Group 3 students, lacking knowledge of their interpreting task, may have experienced more uncertainty, leading to a higher frequency of repairs during interpretation.

6. Conclusion

This study examined how the two distinct pre-task preparation strategies influence the frequency of disfluency features in English to Chinese consecutive interpreting. Given that this study is an initial one to examine the effectiveness of linguistic knowledge-directed and extra-linguistic knowledge-directed preparation strategies in reducing cognitive load reflected by disfluency features and the sample size is relatively small, the findings of the current study are preliminary. It demonstrates that differences in the focus of pre-task preparation strategies can lead to variations in preparation efficacy, consequently impacting the experienced cognitive load as reflected by disfluency features.

However, several limitations in this study warrant acknowledgment. First, the sample size and diversity were limited. The participants were recruited solely from one university, totaling 30 individuals with 10 participants in each group, chosen for convenience. The sample size was relatively small and lacked diversity, potentially affecting the external validity of the study. Future research should aim to address this limitation by recruiting a larger and more diverse sample. Second, the limitation pertains to the tool used for interpreting preparation. Given that interpreting preparation can be a time-consuming and resource-intensive endeavor, leveraging modern technology could enhance preparation efficiency. Previous literature has highlighted the importance of modern tools in interpreting preparation and proposed various technological solutions. Therefore, future research on this topic could incorporate the utilization of these modern tools into the research design to optimize the preparation process. Despite being a preliminary study, this paper offers significant insights into interpreting training, practice and research by providing systematic and detailed frameworks for effective pre-task preparation strategies and demonstrating how these strategies can help alleviate interpreters' cognitive load.



Declarations and Acknowledgment:

This article is supported by the 2022 National Social Science Foundation of China (No. 22AYY006). The authors declare that there is no conflict of interest regarding the publication of this article.

References

- Anderson, L. (1994). Simultaneous interpretation: Contextual and translation aspects. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 101–120). John Benjamins B.V.
- Bakti, M. (2009). Speech Disfluencies in Simultaneous Interpretation. *Selected Papers of the CETRA Research Seminar in Translation Studies 2008*.
- Bóna, J., & Bakti, M. (2020). The effect of cognitive load on temporal and disfluency patterns of speech: Evidence from consecutive interpreting and sight translation. *Target. International Journal of Translation Studies*, 32(3), 482–506.
- Cecot, M. (2001). Pauses in simultaneous interpretation: A contrastive analysis of professional interpreters' performances. Eric Archive, University of the Basque Country.
- Cenoz, J. (1998). Pauses and Communication Strategies in Second Language Speech. Eric Archive, University of the Basque Country.
- Chang, C., Wu, M. M., & Kuo, T. G. (2018). Conference interpreting and knowledge acquisition: How professional interpreters tackle unfamiliar topics. *Interpreting*, 20(2), 204–231.
- Chen, S. (2017). The construct of cognitive load in interpreting and its measurement. *Perspectives*, 25(4), 640–657.
- Díaz-Galaz, S., Padilla, P., & Bajo, M. T. (2015). The role of advance preparation in simultaneous interpreting: A comparison of professional interpreters and interpreting students. *Interpreting*, 17(1), 1–25.
- Fantinuoli, C. (2012). InterpretBank: Design and implementation of a terminology and knowledge management software for conference interpreters [D., University of Mainz].
- Fantinuoli, C. (2017). Computer-assisted preparation in conference interpreting. *Translation & Interpreting*, 9(2), 24–37.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354–375.
- Galaz, S. (2011). The effect of previous preparation in simultaneous interpreting: Preliminary results. *Across Languages and Cultures*, 12(2), 173–191.
- Garnham, A. (2013). *Psycholinguistics (PLE: Psycholinguistics): Central Topics*. Psychology Press.
- Gerver, D. (1975). A Psychological Approach to Simultaneous Interpretation. *Meta: Journal des traducteurs*, 20(2), 119.
- Gile, D. (1985). Le modèle d'efforts et l'équilibre d'interprétation en interprétation simultanée. *Meta: journal des traducteurs / Meta: Translators' Journal*, 30(1), 44–48.
- Gile, D. (2009). *Basic concepts and models for interpreter and translator training* (Rev. ed). John Benjamins Pub. Co.
- Gile, D., & Lei, V. (2021). Translation, effort and cognition. In F. Alves & A. L. Jakobsen (Eds.), *The Routledge Handbook of Translation and Cognition* (pp. 263–278). Routledge.



- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland.
- Jiang, H. (2013). The interpreter's glossary in simultaneous interpreting: A survey. *Interpreting*, 15(1), 74–93.
- Kormos, J. (1999). Monitoring and Self-Repair in L2. *Language Learning*, 49(2), 303–342.
- Koshkin, R., Shtyrov, Y., Myachykov, A., & Ossadtchi, A. (2018). Testing the efforts model of simultaneous interpreting: An ERP study. *PLOS ONE*, 13(10), e0206129.
- Lamberger-Felber, H. (2001). Text-oriented research into interpreting—Examples from a case-study. *HERMES - Journal of Language and Communication in Business*, 26, Article 26.
- Levitt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41–104.
- Mead, P. (2000). Control of pauses by trainee interpreters in their A and B languages. *The Interpreters' Newsletter* 10, 89–102.
- Mead, P. (2002). Exploring hesitation in consecutive interpreting: An empirical study. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century: Challenges and opportunities* (pp. 73–82).
- Mead, P. (2005). Methodological issues in the study of interpreters' fluency. *The Interpreters' Newsletter*, 13, 39–63.
- Moratto, R., & Yang, Z. (2023). Probing the cognitive load of consecutive interpreters: A corpus-based study. *Translation and Interpreting Studies*.
- Moser-Mercer, B. (1992). Banking on Terminology Conference Interpreters in the Electronic Age. *Meta: Journal Des Traducteurs / Meta: Translators' Journal*, 37(3), 507–522.
- Pérez, P. S. P. (2018). The use of a corpus management tool for the preparation of interpreting assignments: A case study. *Translation & Interpreting*, 10(1), 137–151.
- Petite, C. (2005). Evidence of repair mechanisms in simultaneous interpreting: A corpus-based analysis. *Interpreting*, 7(1), 27–49.
- Plevoets, K., & Defrancq, B. (2016). The effect of informational load on disfluencies in interpreting: A corpus-based regression analysis. *Translation and Interpreting Studies*, 11(2), 202–224.
- Plevoets, K., & Defrancq, B. (2018). The cognitive load of interpreters in the European Parliament: A corpus-based study of predictors for the disfluency uh(m). *Interpreting*, 20(1), 1–32.
- Pöchhacker, F. (1995). “Those Who Do...”: A Profile of Research(ers) in Interpreting. *Target*, 7(1), 47–64.
- Postma, A., Kolk, H., & Povel, D. J. (1990). On the relation among speech errors, disfluencies, and self-repairs. *Language and Speech*, 33 (1), 19–29.
- Russell, D. (2005). Consecutive and simultaneous interpreting. In J. T. (Ed.), *Topics in Signed Language Interpreting* (pp. 135–164). John Benjamins.
- Seeber, K. G. (2013). Cognitive load in simultaneous interpreting: Measures and methods. *Target. International Journal of Translation Studies*, 25(1), 18–32.
- Seeber, K. G. (2015). SIMULTANEOUS INTERPRETING. In H. Mikkelsen & R. Jourdenais (Eds.), *The Routledge Handbook of Interpreting* (pp. 79–95). Routledge.
- Setton, R. (1999). *Simultaneous Interpretation: A cognitive-pragmatic analysis* (Vol. 28). John Benjamins Publishing Company.
- Setton, R., & Dawrant, A. (2016). *Conference Interpreting – A Complete Course* (Vol. 120). John Benjamins Publishing Company.
- Song, S., & Cheung, A. K. F. (2019). Disfluency in relay and non-relay simultaneous interpreting: An initial exploration. *FORUM. Revue Internationale d'interprétation et de Traduction / International Journal of Interpretation and Translation*, 17(1), 1–19.



- Tissi, B. (2000). Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. *The Interpreters' Newsletter*, 10(4), 103–127.
- Will, M. (2007). Terminology Work for Simultaneous Interpreters in LSP Conferences: Model and Method. Conference Proceedings.
- Xu, R. (2018). Corpus-based terminological preparation for simultaneous interpreting. *Interpreting*, 20(1), 33–62.
- Yan, H., Zhang, Y., Feng, Y., Li, Y., Zhang, Y., Lee, Y., Chen, M., Shi, Z., Liang, Y., Hei, Y., & Duan, X. (2024). Assessing mental demand in consecutive interpreting: Insights from an fNIRS study. *Acta Psychologica*, 243, 104-132.
- Wang, B.H. (2007). From interpreting competence to interpreter competence——A tentative model for objective assessment of interpreting. *Foreign Language World*, 3, 44–50.

About the Authors:

Shuangshuang Yang: a PhD student of the Translation and Interpreting Program in the Department of Linguistics, in the Faculty of Medicine, Health and Human Sciences, at Macquarie University. She is interested in the cognitive process of interpreting, especially simultaneous interpreting.

Lei Mu: PhD, is a Professor at Guangdong University of Foreign Studies. Additionally, she serves as both an expert member and committee member within the Translators Association of China. Her research focuses on translation and interpreting training, theories in translation and interpreting, and the history of translation studies.

