

A Case Analysis of Disfluency in English-to-Chinese Simultaneous Interpreting Based on the Effort Model

YU Danni^{[a],*}

^[a]School of Foreign Languages, Zhijiang College of Zhejiang University of Technology, Shaoxing, Zhejiang, China.

* Corresponding author.

Received 12 June 2024; accepted 26 July 2024
Published online 26 August 2024

Abstract

Disfluency is quite common for student interpreters in simultaneous interpreting (SI), which is regarded as a critical parameter for interpreting quality. Therefore, it is of great significance to study disfluency in SI. Based on Gile's Effort Model, the thesis aims to shed light on causes to disfluency in SI by conducting a case study of the interpreting project on 2020 International Commercial and Legal Cooperation Forum.

With reference to the classification system of interpreting disfluency proposed by Dai Zhaohui, the author summarizes three major manifestations of disfluency in SI: pauses, repetitions and self-repairs. Under the guidance of Gile's Effort Model, interpreting performance deteriorates if processing capacity requirements exceed its supply. The author analyzes the problem triggers of disfluency from the perspective of cognitive overload, which involves Listening and Analysis Effort, Memory Effort, Production Effort and Coordination Effort. To be specific, it is attributed to the fact that the required processing capacity in Listening and Analysis Effort, Memory Effort, Production Effort and Coordination Effort exceeds the available one. Finally, the author proposes technical and non-technical coping strategies with an aim to improve delivery fluency in SI.

Key words: Simultaneous interpreting; Disfluency; Gile's Effort Model

Yu, D. N. (2024). A Case Analysis of Disfluency in English-to-Chinese Simultaneous Interpreting Based on the Effort Model. *Studies in Literature and Language*, 29(1), 47-58. Available from: <http://www.cscanada.net/index.php/sll/article/view/13481> DOI: <http://dx.doi.org/10.3968/13481>

1. INTRODUCTION

Interpreting is a communicative activity and interpreters play an interactive role in this process. In particular, simultaneous interpreting has played a more and more important role in achieving communication between different cultures.

Delivery fluency is an important parameter for interpreting quality assessment. However, for students majoring in interpreting, disfluency is a common phenomenon in interpreting practice. This may be due to multiple factors including overloaded tasks, strong accents and poor psychological quality. The author has found that disfluency phenomenon such as pauses, language fillers often occurs in interpreting output. To improve fluency delivery cannot be ignored, otherwise, interpreters may convey unclear messages to audience and undermine interpreters' credibility.

Although there are a great number of studies on disfluency in delivery, most of them focus on linguistic perspective. Moreover, many analyze disfluency in a general way. The author studies simultaneous interpreting (SI) from cognitive perspective based on Gile's Effort Model.

Gile's Effort Model serves as a theoretical framework to analyze problem triggers of disfluency and find out possible solutions. Cognitive psychology in interpreting research is introduced in the Effort Model, which divides the whole interpreting into three phases, in which Effort is consumed. By analyzing disfluency in each phase from the cognitive perspective, it provides a scientific account for the processing mechanism in interpreting.

The thesis tries to shed light on causes to disfluency in SI. The author hopes that this study can help interpreting learners better cope with disfluency in order to improve their interpreting quality. Based in the analysis of disfluency in English - Chinese simultaneous interpreting tasks, this thesis attempts to answer the following research questions:

- i. What are the contributing factors of disfluency in simultaneous interpreting?
- ii. What feasible coping strategies can be adopted to avoid unnecessary disfluency in simultaneous interpreting?

2. LITERATURE REVIEW

After reading relevant research and papers at home and abroad, the author draws on and summarizes research results concerning disfluency found by various scholars, including the definition and classification of disfluency.

2.1 Definition of Disfluency

The word “disfluency” is the antonym of “fluency”. Linguistic fluency is firstly defined by Leeson, who regards it as an ability a speaker has to produce multiple sentences that meet the phonetic, grammatical and semantic requirements of the language, when he/she has access to a limited corpus of a kind of language (Leeson, 1975). Based on the definition of fluency by Leeson, Fillmore (1979) argues that a fluent native speaker can keep the speech flow going, articulate on various occasions appropriately without being speechless or stammering, package the meaning into discourse coherently.

Disfluency is the opposite of the fluent speech. Shriberg (1994) regards that disfluency occurs when a contiguous stretch of linguistic material is deleted to interrupt the fluent flow of the speech. Fox Tree (2003) defines disfluency as a phenomenon that causes a break and interrupts the smooth flow of the speech. It is demonstrated by Yang (2004) that disfluency refers to a kind of speech production that is different from fluent speech in rhythm, sequence and timing. According to *Routledge Dictionary of Language and Linguistics* (Bussmann, 2006, p. 345), disfluency refers to a situation-specific speech production disorder that interrupts the fluent flow of language, due to a lack of motor coordination in the muscles involved in articulation, phonation, or respiration.

Generally speaking, although a great number of previous studies have been conducted on the definition of disfluency, scholars’ views vary, implicating a lack of generally agreed definition.

2.2 Classification of Disfluency

Based on the definition of disfluency, scholars have also categorized disfluency in different types. Maclay and Osgood (1959) categorize hesitation of speech into four types: (1) false starts including retraced false starts and non-retraced false starts; (2) repeats of syllables, words and expressions; (3) filled pauses; (4) unfilled pauses. Shriberg (1996) divides disfluency into six categories, namely, repetition, filled pause, insertion, substitution, speech error and deletion. Cecot (2001) summarizes

previous research and proposed that disfluency includes filled pauses, audible breaths, vowel and consonant lengthening, utterance interruptions, parenthetical sentences and glottical clicks. Dai Zhaohui (2011) introduces three types of disfluencies in consecutive interpreting, namely, pauses including filled pauses and silent pauses, repeats and self-repairs. Dai’s classification is agreed by many domestic scholars.

Meanwhile, a number of scholars have conducted research in one specific type of disfluency. With regard to pauses, it is illustrated that self-repair means the utterance that is discarded before completion and that may or may not be followed a reformulation (Foster & Skehan, 1996, p. 310). Levelt (1983) categorizes repairs into five types: (1) difference-repairs; (2) covert-repairs; (3) appropriateness-repairs; (4) error-repairs; (5) rest-repairs.

Pauses often occur in disfluent speech. Based on a specific time interval, the definition of pauses varies among scholars. For example, Goldman-Eisler (1968) regards pauses as intervals longer than 250 milliseconds (0.25 seconds). Pauses mean intermissions in sentences or between sentences longer than 0.3 seconds (Raupach, 1980). When it comes to the classification of pauses, it includes silent pauses lasting at least 0.25 seconds, filled pauses such as “um” and “uh”, and mixed pauses with silent pauses and filled pauses in a continuous manner (Mead, 2005). Fu Rongbo (2013) divides pauses into grammatical pauses and non-grammatical pauses. Many domestic researchers agree that pauses can be classified into filled pauses and silent pauses (Yang, 2004; Liao, 2009; Tan & Feng, 2013).

Repetition refers to the phenomenon that words, phrases or clauses are repeated with modifications to morphology, syntax or words order unchanged (Foster & Skehan, 1999). Dai (2011) puts forward four types of repetition, namely, repetition of syllables, repetition of one word, repetition of two words and repetition of more words.

2.3 Effort Model for interpreting

Effort Model was formulated by Daniel Gile in the 1970s. Introducing the basic concept “effort”, Gile develops the model from the two perspectives: “Interpreting requires some sort of mental energy that is only available in limited supply; and interpreting takes up almost all of this mental energy, and sometimes requires more than is available, at which times performance deteriorates.” (Gile, 1995, p. 161) In other words, since interpreters mental energy (or processing capacity) is limited, it is difficult for them to handle too many tasks at the same time. Interpreting is a multiple-task activity, where interpreters may deplete all of their mental energy.

Interpreting is an activity that involves multiple mental operations. Cognitive psychologists suggest that mental operations include automatic operations and non-automatic operations. Automatic operations mean that

when interpreters receive familiar information, their minds will respond automatically, which can save mental energy. However, when information of source texts is unfamiliar to interpreters, they will make a non-automatic response, which consumes some of their mental energy. Gile points out that interpreting mostly involves non-automatic operations. The concepts of mental energy and non-automatic operations serve as basis for establishing Gile's Effort Model.

Interpreting is a non-automatic operation which involves mental energy. The word "effort" is applied to describe the conceptual components in interpreting in order to emphasize the non-automatic operation (Gile, 2011, p. 148). The Effort Model consists of Listening and Analysis Effort, Memory Effort, Production Effort and Coordination Effort. Each effort is illustrated in the following.

Listening and analysis effort refers to "all comprehension-oriented operations, from the analysis of the sound waves carrying the source-language speech which reach the interpreter's ears, through the identification of words, to the final decisions about the 'meaning' of the utterance." (Gile, 1995, p. 162). The process begins from subconscious analysis of the sound signal carrying the source language information, to identifying the source language information. Interpreters should at least get the logical sense of each sentence which is invisible and underlying (Gile, 2011, p. 149).

Listening and analysis is the first step of interpreting. Interpreters must understand the message conveyed in source texts. Therefore, interpreters should allocate mental energy to this effort to ensure a smooth interpreting performance.

The memory effort refers to short-term memory operations in simultaneous interpreting. It is defined as the storage of information. When a sequence of source speech reaches interpreters, they firstly decode the segment to acquire the message. The sense of the source-speech unit must be stored in the short-term memory for later use. There is interval between the moment interpreters hear sounds of source language and the moment they produce target language (Gile, 2011, pp. 154-155). Short-term memory is needed when interpreters face unclear logic of the source speech, high information density, complicated sentence structure. Interpreters tend to wait for a moment in order to deliver a clear and smooth target speech to the audience.

The production effort is related to the output of interpreting. It is defined as the set of operations extending from conceptualizing the message to be delivered to planning of the target speech and the performance of speech reformulation (Gile, 2011, pp. 151-152).

It is pointed out that speech production is quite difficult for interpreters. Instead of speaking on their own freely, interpreters have to follow the source

language delivered by speakers, which means they cannot circumvent possible production difficulties (Gile, 1995, p. 165). Another difficulty is that habitual combinations of words are different from language to language (Gile, 2011, p. 152). Interpreters should be deprived of their own delivery habits to reformulate a target speech based on the semantic meaning of the source speech (Gile, 2011, p. 153). Therefore, interpreters should not be restricted by the grammar and syntactic structures of the source language. Instead, they are supposed to reformulate target language based on meaning.

The coordination effort refers to the effort which coordinates the mentioned four efforts. (Eysenck & Keane, 1990). Interpreters need to coordinate their mental energy to each part properly. If they fail to allocate mental energy properly to each part, they may not understand the source speech or recall the key message heard before.

2.4 Effort Model for Simultaneous Interpreting

Gile originally uses his Effort Model to analyze simultaneous interpreting (SI), assuming three basic efforts, namely, Listening and Analysis Effort (L), Production Effort (P) and Memory Effort (M) (Gile, 1995:162). Coordinated effort is added in subsequent refinements of the model. Gile proposes the following formula to demonstrate the relationship between the four efforts in SI Effort Model:

$$SI=L+P+M+C$$

It should be noted that the plus sign in the formula refers to "composition", rather than adding the four efforts together in the mathematical sense.

In simultaneous interpreting, these four efforts are not simply allocated sequentially. For instance, the source speech is divided into section A, section B, section C. When section A is produced, section B has been understood and stored in short-term memory, and section C is being analyzed through listening comprehension. Therefore, there is at least partial competition among three efforts, and processing capacity needs to be appropriately allocated, which will inevitably lead to the increase of processing capacity requirements. Thus, the formula of total processing capacity requirement is proposed by Gile as follows (Gile, 2011, pp. 158-159):

$$TR=LR+MR+PR+CR$$

TR: total processing capacity requirements

LR: processing capacity requirements for L

MR: processing capacity requirements for M

PR: processing capacity requirements for P

CR: processing capacity requirements for C

Likewise, the plus sign, in a broad sense, should be regarded more than addition.

Gile points out that each effort has specific processing capacity requirement. In order for interpretation to proceed smoothly, the following conditions need to be met:

- (1) $TR < TA$
 - (2) $LR < LA$
 - (3) $MR < MA$
 - (4) $PR < PA$
- $CR < CA$

TA: total available processing capacity
LA: processing capacity available for L
MA: processing capacity available for M
PA: processing capacity available for P
CA: processing capacity available for C

The explanation of the above several conditions are as follows: First, in the process of interpretation, the mental energy provided by the interpreter is not sequentially used for each task. Instead, it is allocated to different tasks at the same time; Second, the amount of mental energy the interpreter can provide is limited, so the total amount of processing capacity required for simultaneous interpretation must be less than the total amount of processing capacity available. Accordingly, each part of simultaneous interpretation, namely listening and analysis, short-term memory storage, target language production and coordination, must require less processing capacity than the interpreter's brain can provide. Third, the processing capacity of each task in the process of simultaneous interpretation of the interpreter must be sufficient to ensure the smooth interpretation, otherwise it will be easy to appear a certain energy saturation; Fourth, if mental energy is improperly allocated, the energy used to deal with one task will occupy the processing energy that should be allocated to the other tasks, resulting in energy imbalance and further triggering off the incomplete output of target information. For example, due to the large amount of information in source speech, the translator is overloaded with short-term memory. The Memory Effort increases, with Listening and Analysis Effort and Production Effort correspondingly reducing, which may cause the loss of information input and disfluency of target speech. Therefore, in order to guarantee the interpreting quality and the smooth process of interpreting, on one hand, the total capacity required should be less than the total capacity available. On the other hand, in each part, the capacity required for should be less than the capacity available. Otherwise, interpreters will be confronted with cognitive saturation, with their interpreting performance deteriorating.

2.5 Feasibility of the Effort Model for Analysis of Disfluency in SI

Simultaneous interpreting involves multiple cognitive processes including receiving source language, analyzing source information, decoding information and producing target language. These cognitive processes are intangible, which cannot be explained clearly from the linguistic perspective alone.

Disfluency represents a phenomenon that the interpreter fails to make fluent delivery, which is resulted

from deeply-rooted problems arising from the process of interpreting. Disfluency in target speech actually reflects the cognitive difficulties the interpreter encounters in the interpreting process. Gile's Effort Model serves as a theoretical framework that analyzes difficulties and problems in interpreting from the perspective of cognitive science.

According to the Effort Model, the whole process is divided into three phases, each of which entails a certain amount of processing capacity to deal with the task. If the interpreter's mental energy is insufficient, he cannot handle multiple tasks at the same time. Thus, the delivery related problems such as disfluency will occur. Therefore, the Effort Model can be applied to analyze disfluency and help the interpreter put forward corresponding coping strategies from the cognitive perspective.

3. METHODOLOGY

The project design includes the preparation before interpreting, simultaneous interpreting task in the form of recording, data collection and analysis, target text analysis. The author identifies disfluency phenomenon in the interpreting performance bases on the criteria of disfluency in SI. Then the interpreting performance of the fluency aspect is quantified. It should be noted that the target text analysis concentrates on linguistic errors caused by insufficient mental energy or improper energy distribution, excluding other factors such as personal behavior or psychological nervousness.

Two difficulties arise when the author conducts the analysis. One difficulty is to determine the criteria of disfluency in SI. Scholars have not proposed a unified classification of disfluency in interpreting output studies. However, based on relevant studies on disfluency in interpreting, the author finds that many scholars have agreed and adopted the classification proposed by Dai (2011) from the perspective of empirical study of interpretation, namely, pauses, repetition and self-repairs. In addition, these three types of disfluency phenomenon frequently arise in the author's interpreting performance, through analyzing the transcript of the target text.

The other difficulty is associated with the minimum threshold of silent pauses. Frankly speaking, there is no consensus on the specific time interval of silent pauses, but after reviewing a great number of past studies on spoken English fluency, the author finds that most scholars measures the time interval within sentence with over 0.3 seconds. Therefore, in this study, silent pauses refers to the intermissions lasting more than 0.3 seconds.

After determining the criteria of disfluency in SI, the interpretations are transcribed in a faithful manner. The silent pauses in the target speech production is measured by the software of *Adobe Audition 3.0*, which can be clearly visualized on screen. The fluctuating waves

indicate the interpreter is delivering, while the flat parts suggest silent pauses between sentences. It is noteworthy that the measuring tool may cause deviation, for both necessary pauses between sentences and unnecessary intervals are recorded. Therefore, the author needs to identify the latter and define as silent pauses manually. As for filled pauses, the author pays attention to utterances represented by non-lexical words and phrases like “um”, “uh” etc. Repetitions in this thesis mean that syllables, words, phrases and clauses are repeated, but morphology, syntax or word order remain unchanged. Self-repairs refer to the utterances that correct syntactic, lexical and phonetic errors.

Table 1
Disfluency in SI examined in the thesis

Disfluency phenomenon in simultaneous interpreting		
Pauses	Silent pauses	Intervals lasting over 0.3 seconds
	Filled pauses	Vocalized hesitation like “uh”
Repetitions		Syllable repetitions
		Word repetitions
		Sentence repetitions
Self-repairs		Syntactic error repairs
		Lexical error repairs
		Phonetic error repairs

As the text analysis is designed to analyze difficulties and problems based on the Effort Model, pauses, repetition and self-repairs are analyzed in terms of Listening and Analysis Effort, Memory Effort, Production Effort and Coordination Effort. In this way, the author can figure out whether the disfluent output is caused by insufficient processing capacity for a particular task, or the misallocation of the four Efforts to these tasks. Finally, in view of the analysis of specific problems, feasible improvement strategies are put forward.

4. DATA COLLECTION

Pauses are found to occur most frequently in the author’s interpreting task, accounting for the majority of the disfluency. As the interpreter needs time and mental energy to listening comprehension, information storage, speech production and coordination, it is understandable that pauses occur when the author fails to appropriately allocate mental energy. Repetitions are another disfluency phenomenon that occurs when the author fails to come up with the whole sentence in mind but she has to keep the speaking flow continuous due to time constraint. Self-repairs are also a major kind of disfluency phenomenon in the author’s interpreting, resulted from her intention to work out a better rendition to improve the original production and correct the wrong interpretation. Table 2 clearly describes the codifications of disfluency in the transcription, so as to facilitate analysis of problem triggers of disfluency.

Table 2
Disfluency codifications in the transcription

	Codification	Examples
Silent pause	x(...)y	I (...)want
Filled pause	x(uh)y	I (uh) like
Repetition	x	我
Self-repair	x	我

Then, the author collects and double-checks all disfluency phenomenon in the interpreting transcript, which is listed in Table 3. It is obvious that pauses, especially silent pauses occur far more frequently than repetitions and self-repairs.

Table 3
Descriptive statistics of disfluency in SI

	Frequency	Speech 1	Speech 2	Speech 3	Total frequency
Silent pause	47	88	47	182	
Filled pause	14	19	17	50	
Repetition	8	13	4	25	
Self-repair	5	15	4	24	
Total frequency	74	135	72	281	

5. FINDINGS AND DISCUSSION

The investigation into the interpreting recordings reveals that disfluency phenomenon occurs more often in some places than others. The author figures out that disfluency often occurs in the following circumstances:

- i. Disfluency is inclined to occur in long and complicated sentences.
- ii. Disfluency tends to occur in sentences with high information density and professional terminology and knowledge.
- iii. Disfluency may occur when the interpreter fails to convert Chinese-featured linguistic structure to English-featured one.
- iv. Disfluency may also occur when slips of tongue are corrected.

The research is trying to analyze the causes of disfluency in simultaneous interpreting listed based on Gile’s Effort Model. The author finds that disfluency frequently arises when she is saturated with cognition in the interpreting. Based on the Effort Model, the four efforts engaged in interpreting require a certain amount of mental energy in order to complete the interpreting task successfully. Therefore, the total available processing capacity should exceed the total amount of processing capacity requirements of the four efforts. Cognitive saturation appears when the processing capacity requirements are larger than the available mental energy, or when the available processing capacity for each effort is insufficient to handle each stage of interpreting. The author tries to figure out problem triggers of disfluency in

each stage, from the perspective of cognitive saturation. As shown in Table 4, all contributing factors identifies in the research are classified into four categories.

Table 4
Classification of contributing factors of disfluency

Cognitive saturation in SI	Contributing factors of disfluency in SI
Listening and comprehension of source speech	High information density Failure of recognizing terminology Different syntactic structure
Storage of short-term memory	High information density Unfamiliar professional knowledge Complicating syntactic structure of target language
Production of target speech	Inflexible conversion Correcting slips of tongue
Coordination	Incoordinate mental energies

5.1 Listening and comprehension of source speech

Listening and comprehension of the source speech is the first and foremost stage of interpreting, which is associated with the Listening and Analysis Effort in the Effort Model. Gill believes that this stage refers to the whole process from the interpreter identifying the words to the sound waves carrying the source language in the interpreter's ear through subconscious analysis, to the interpreter finally understanding the meaning of the speaker. To put it simply, the interpreter should not only hear the source language, but also analyze the meaning carried by the source language (Gile, 2009, p.160). Therefore, "Deverbalization" proposed by Seleskovitch can be adopted in listening comprehension, a strategy that requires the interpreter to get rid of the bondage of language form so as to grasp the sense of language. By analyzing the author's disfluency phenomenon in interpreting, the author concludes that the failure to understand the source speech is resulted from high information density, absence of recognizing terminology, interference of different sentence structure between Chinese and English.

5.1.1 High information density

As it is mentioned above, speakers delivered their speeches very fluently by reading the script which has been written down before, so they give more information in per unit of time compared with impromptu speeches, posing pressure on the author who needs more available processing capacity within the limited time. Suppose that the author has limited processing capacity for processing a certain amount of information in listening comprehension stage, if more information is produced flocking into the author's brain, the author will be confronted with an overload that requires the author to increase the amount of processing capacity exceeding available capacity to handle information. Therefore, the interpreting quality

deteriorates and disfluency occurs in the author's interpreting reproduction.

Example 1:

SL (source language): Investors from China have been active in the Bangladesh market for many years now and they have invested in many key sectors including energy, infrastructure, telecommunication, stock exchange."

TL (target language): 中国投资者(...)多年来一直活跃于孟加拉国, 在很多关键领域进行投资, 包括(uh)能源, (uh)电信, (uh)股票等等。

Analysis: Generally speaking, the main idea of the sentence is reproduced, with some pauses in delivery. The problem lies in the author's limited processing capacity for listening and analysis, due to the high information density. A considerable amount of information containing a string of nouns pours into the author's brain, making the author struggled in listening comprehension during which more processing capacity is in need in listening comprehension phase. The author finds it difficult to process all noun phrases in brain, although she hears every word. Thus, the word *infrastructure* is missed and three filled pauses arise before "能源" "电信" "股票" are spoken out.

5.1.2 Absence of recognizing terminology

The three professional speeches are associated with legal field including alternative dispute resolution, legal institutions and investment law. The author is exposed to the general topic and background information before the interpreting, but it is inevitable that the author encounters a certain number of unfamiliar words and phrases. It pays mental extra energy to listen and comprehend terminology, which is always of no help because of the lack of relevant background knowledge. In this case, the processing capacity for Listening and Analysis Effort is less than its required processing capacity. To explain with the Effort Model, the inequality $LR \leq LA$ cannot be met. In addition, the more mental energy the author pays to listening and comprehending terminology, the less mental energy is allocated to speech production, leading to the occurrence of disfluency in speech delivery.

By analyzing problem triggers of disfluency in the speech production, the author concludes two major types of disfluency caused by the absence of recognizing terminology. Firstly, the author encounters professional words that are included in the glossary prepared beforehand, but fails to identify them due to inadequate preparation. Secondly, unfamiliar phrases containing totally unknown concepts make the author at a loss, setting obstacles to delivering the message clearly and fluently.

Example 2

SL: It is commonly heard in business that time equals money, in other words, a solution that is less time consuming or quick and does not cost too much is an efficient solution. As investors enter into more and more contracts, they run the risk of facing potential disputes and

the best way to prevent that is to include an Alternative Dispute Resolution or ADR clause in all contracts.

TL: 在商业领域有这样一句话, 时间就是金钱, 也就是说, 如果一个方法它(...)消耗的时间少, 那么(...)就是一个高效的方法。投资者可能面临着(...)潜在的争端, 所以ADR, 也就是替代性争端解决, 也就是至关重要的。

Analysis: In addition to a few silent pauses occurring *Example 2*, repetition is another deficiency in delivery that cannot be ignored. To be frank, the author gets familiar in the terminology *Alternative Dispute Resolution*, but still fails to receive the information loaded in the proper nouns and thinks out the equivalent instantly, which challenges the author's cognitive load in listening and analysis. In accordance with the Effort Model, the author's processing capacity available for Listening and Analysis fails to meet the requirements, leading to the cognitive saturation. As a result, repetition such as “也就是” occurs in the target language delivery.

Example 3

SL: Party autonomy is a priority, which gives importance to the parties needs. The most common forms of ADR are mediation and arbitration which as we all know are also practiced routinely in many provinces in China. ADR is part of international best practice and has border enforceability under the New York Convention on the Recognition and Enforcement of Foreign Arbitral Awards and the new Singapore Convention on Mediation.

TL: 它自主(...)非常重要。最常见的形式是仲裁和调解, 在中国的很多省份也已经应用这种形式了。在纽约公约(...)在纽约公约和新加坡公, 调解公约的框架之下, 这种形式已经为世界公认。

Analysis: This example is characterized by legal terminology related to arbitration and mediation. The great challenge in interpreting this paragraph lies in understanding the concept of *party autonomy* as well as coming up with the equivalents for *the New York Convention on the Recognition and Enforcement of Foreign Arbitral Awards* and *the new Singapore Convention on Mediation*. It is obvious that silent pauses, repetitions and self-repairs occur mainly in interpreting the above two difficult points, resulted from the author's exhausted cognitive load allocated between listening and production. The author faces great challenge in recognizing these unfamiliar terminologies. In other words, the processing capacity requirements for Listening and Analysis Effort exceeds the processing capacity available for identifying the terminologies. In addition, as the author is unable to interpret *the New York Convention on the Recognition and Enforcement of Foreign Arbitral Awards*, a reaction chain is triggered with continuous input of another unfamiliar terminology, increasing pressure on Listening and Analysis Effort. Thus, the author gives the wrong rendition at first and then corrects it.

5.1.3 Different sentence structure

Another mistake the author makes in listening and analysis stage is associated with the weak grasp of English sentence structure. Because of the difference in sentence structure between English and Chinese, the author sometimes cannot understand the message loaded in the complex sentence. Even though the author knows what every word means, she still finds it difficult to grasp the message conveyed by speakers since what she have heard is a patchwork of individual words. Under this circumstance, the author has to consume more mental energy to analyze English syntactic structure. However, the available processing capacity for Listening and Analysis Effort is limited, failing to meet the increased requirement.

Example 4

SL: A four-hour flight from Istanbul can reach 57 countries in Europe, the Middle East, North Africa, and Central Asia, a large market with a population of 1.5 billion. Istanbul New Airport is one of the largest airports in the world, with a capacity of 200 million passengers per year.

TL: 从伊斯坦布尔坐飞机(...)可以(...)到57个国家和地区, 包括欧洲(...)等国家, 伊斯坦布尔机场每年可容纳(...)2亿(...)乘客。

Analysis: A great number of silent pauses occur in this example, partly resulted from the fact that the author fails to figure out what *a large market with a population of 1.5 billion* refers to, although she understands the meaning of every word. After reflecting on the source text, the author recognizes that this complex sentence contains the apposition *a large market with a population of 1.5 billion* which refers to Europe, the Middle East, North Africa, and Central Asia. It is easy to identify the apposition by reading the source text, but challenging to have an clear command of the logic during interpretation. Thus, the author allocates plenty of effort into trying to piece together words heard while less effort is distributed to listening and producing the following message, with the consequence that the author omits some countries' names and pauses for a while.

5.2 Storage of short-term memory

The immediacy of interpretation puts forward high requirements on the short-term memory of interpreters, while the working memory of short-term memory performs multiple functions of storing information temporarily, processing and coordinating information. Storage of short-term memory is the core of information processing. Short-term memory is an important part in simultaneous interpreting and the storage of short-term memory underpins the speech production of target language, which greatly affects the quality of target language. Although the mental energy allocated to short-term memory in simultaneous interpreting is less than that in consecutive interpreting, it is still challenging to store

information segments in short-term memory. The author fails to reproduce target language fluently because she cannot memorize all pieces of information in such a short time resulted from the short-term memory overload. Two causes are concluded after the author analyzes the source text and the recording of target language: high information density and unfamiliar professional knowledge.

5.2.1 High information density

High information density may lead to memory overload of the interpreter. Gile (1995) points out that a cognitive saturation can arise through an increase in processing capacity requirements in Short-term Memory Effort, which forces the interpreter to store enormous information segments in short-term memory for some time until he/she is able to reproduce it in target language. In other words, the available processing capacity for short-term memory is insufficient to meet the required processing capacity if the source speech contains a large amount of key information. Storing massive information segments effectively before reformulating in the target language is essential and can ensure the quality of interpreting production. However, it is quite difficult for the author to keep all information in mind and reproduce it fluently.

Example 5

SL: Our cooperation extends to Singapore, Malaysia, India, Vietnam, Philippines, Indonesia, Thailand, Sri Lanka among others. We would like to highlight here that we have a cooperation agreement with the Permanent Court of Arbitration which is the first permanent intergovernmental organization to provide a forum for the resolution of international disputes through arbitration and other peaceful means.

TL: 我们的合作也拓展到了新加坡, 印度尼西亚, (uh)泰国, 印度尼西亚, 斯里兰卡等等。我们(...)也和常设, 常设仲裁法院(...)达成了合作协议, 来(uh)提供和平的(...)争端解决方法。

Analysis: To be honest, the short and simple sentence is easy to translated in the written form. However, the author encounters great difficulty in reproducing smoothly and fluently as eight countries are mentioned in the short simple sentence. Due to the limited time, the author is unable to be reminded of Chinese equivalences for every country's name correctly and instantly. It requires more processing capacity for short-term memory which exceeds the author's available processing capacity. As a result, disfluency occurs ranging from filled pauses to repetitions, and only three countries' names are interpreted. The other terms in English are compressed as “等等” in Chinese.

Example 6

SL: The development of science and technology, transportation and communication has made the geographical distance between the two countries closer and closer. This is also part of the reason why we regard China as a close and reliable partner and continue to promote economic integration and economic cooperation in Asia and the rest of the world.

TL: 我们的重点也就是与中国开展建设性对话。科技的发展(uh)使得两国的距离越来越短。这也是为什么我们把中国看作是最可靠的合作伙伴(...)由此进一步(...)促进亚洲经济的(...)发展。

Analysis: In addition to *example 5* in which the problem trigger lies in a string of English terms in parallel structure, such as *science and technology, transportation, communication*, this example is typical because of the lengthy English expression *economic integration and economic cooperation in Asia and the rest of the world*. It poses pressure to the author's cognitive overload in short-term memory as these notional words pour into the brain continuously, decreasing the mental energy that should be allocated to reproduction. Although the author hears words and expressions like *integration, cooperation and the rest of the world*, she still cannot come up with corresponding equivalents immediately under great pressure. A cognitive saturation arises as the author need to store such a large amount of information in short-term memory before she can reformulate the complete message. Thus, the processing capacity requirements for Short-term Memory increase, exceeding the author's available processing capacity. Under this circumstance, silent pauses occur deteriorating the fluency of the author's interpreting production.

5.2.2 Unfamiliar professional knowledge

Due to the absence of professional knowledge, although the author can clearly hear and understand each word, she cannot find the corresponding understanding or doubt the correctness of her own understanding when these words are combined into the author's ear. In addition, the author cannot activate the corresponding long-term memory, and it is difficult to store information in short-term memory. At the same time, it will occupy too much mental energy and affect the implementation of other translation programs, resulting in the fluency target language production.

Example 7

SL: Chinese investment in Turkey can enjoy tariff exemption, corporate income tax deduction, land allocation and social security support. Although these are the main incentives, we also provide tailored investment incentives. In other words, we will base on the nature of each investment and the requirements of each investor.

TL: 中国投资者可以享受关税豁免和(...)社会保障等支持。除了这些, 我们也会为他们量身定制一些投资经历, 激励。我们会依据每个投资项目(...)的性质来决定。

Analysis: Four professional terms are listed in this sentence, including *tariff exemption, corporate income tax deduction, land allocation and security support*. Each of these four juxtapositions is not difficult to translate in written form, but when they are stacked together, the author nearly runs out of available processing capacity for short-term memory as the author feels struggled to store these unfamiliar expressions. The author is unable to think out equivalents for all terms, with only “关税豁免”

and “社会保障” produced. It is obvious that silent pauses occur as the author does not know how to interpret *corporate income tax deduction* and *land allocation* under the time pressure.

5.3 Production of the target speech

After listening and storing information in short-term memory, the interpreter needs to convey the message spoken by the native speaker in an organized manner, with the content faithful to the source language speech. Speech production is the final stage in simultaneous interpreting, in which the interpreter tries the best to make the audience understand the message of the source language speech to the maximum extent.

Levelt (1989) has proposed a widely accepted speech production model, which consists of three main components: Conceptualizer responsible for producing and monitoring preverbal messages, Formulator which encodes preverbal messages as internal speech, and Articulator outputting overt speech. Moreover, a monitor device is involved in the Conceptualizer to monitor the speech process and speech result, in order to prevent some output errors from being articulated.

The speech production model involving three stages of speech production can be applied to analyze mental energy allocation in Production Effort. As for native speakers who are equipped with high-level language automaticity, the three components are operated at the same time. Nevertheless, as Efforts required by Conceptualizer, Formulator and Articulator compete with each other for the interpreter's limited processing capacity, the interpreter does not acquire the high-level automaticity, resulting in disfluency in speech production.

In this analysis, the author finds that she sometimes consumes excess processing capacity in production, particularly in formulation and articulation process. With a focus on the author's language drawback, it is concluded that the occurrence of disfluency can be attributed to complicating target language syntactic structure, inflexible conversion as well as correcting slips of tongue.

5.3.1 Complicating target language syntactic structure

According to Levelt (1989), Formulator encodes and processes preverbal messages that are generated by Conceptualizer previously and produces the internal speech. This process involves making lexical, syntactical and phonological choices. Formulator in the Production Effort, processing capacity requirements for it are related with retrieving the optimal expressions. Searching for appropriate syntactic structure is included in the process of internal speech formulation. Since there exists difference in syntactic structure between Chinese and English, it sets obstacles for the author to reorganize target language structure. Therefore, it takes more mental energy for the author to select proper syntactic structures from the long-term memory.

In this analysis, the author finds that disfluency arises in some cases when the processing capacity which is available to Formulator does not meet the requirements. Poor language proficiency drives the author to complicate syntactic structure that should be simplified. Not only does it make the target language delivery redundant, but it also reduces the processing capacity that is designed to be allocated in other parts by devoting a large amount of processing capacity to speech production.

Example 8

SL: Our founding Chairman, Mr. Mahbubur Rahman is the only businessman from Bangladesh to be listed recently in the 150 “high impact” leaders under the “Creating Emerging Market Project (CEMP) at the Harvard Business School (HBS).

TL: 我们的创始人(...)是唯一一个(...)来自孟加拉(...)国籍的人, 他在, 他是150名最具影响力的领导(...)人中唯一一位孟加拉(...)国人。

Analysis: The occurrence of silent pauses in the example undermines the author's interpreting performance. This long sentence is complicated in English syntactic structure, so it takes more time and mental energy for the author to retrieve the corresponding Chinese syntactic structure. The author can easily understand the general message conveyed in the source language speech, nevertheless, she fails to organize the syntactic structure in target language under pressure. The difficulty lies in the selection of Chinese syntactic structure. The simple expression *businessman from Bangladesh* is interpreted into “来自孟加拉国籍的人” in a complicated way. The author intends to cut the long sentence into several segments to make the message more understandable to the Chinese audience. The choice of segmentation leads to the difficulty in selecting a different structure for the second half of the sentence, which is obviously shown by a series of silent pauses.

5.3.2 Inflexible conversion

As is mentioned above, Formulator in Production Effort involves grammatical and phonological encoding which calls for a certain amount of processing capacity. In addition to complicating syntactic structure of target language, the failure of meeting the processing capacity requirement of Formulator in Production Effort can also be manifested in inflexible conversion. To be specific, due to the author's low availability of working language, the limited processing capacity designed to produce target language speech fails to meet the requirement, so it takes more mental energy to make lexical and grammatical choices from the long-term memory.

Through analyzing the source text and target text, the author figures out that some occurrences of disfluency are caused by the challenge that the author cannot retrieve lexical and grammatical expressions flexibly because of limited processing capacity in Formulator.

Example 9

SL: Since 2016, Philippines-China Relations have entered a new phase of close and friendly status, which provides a conducive platform for enhancing business collaboration and partnerships.

TL: 2016年以来, 我们的双边关系进入了, 进入了友好阶段, 为(...)加强合作关系提供了(...)好的平台。

Example 10

SL: In the later stages of the epidemic, our economic indicators showed good development prospects.

TL: 我们的数据显示, (uh)经济发展前景良好。

Analysis: The author is puzzled when analyzing the cause of disfluency in *example 9* and *example 10* as both sentences are short, with simple words and clear message. By reviewing the recording of target language delivery, the author identifies the problem which lies in slow and inflexible lexical conversion from source language to target language. The author does hear the expressions such as *close*, *conducive* and knows their meanings, but it is challenging for the author to immediately think out appropriate equivalents for these expressions, which exhausts the total amount of mental energy left for production. The limited processing capacity in Formulator makes the speech flow interrupted by the repetition of “进入了” and the silent pause before “好的”.

Likewise, the occurrence of filled pause in *example 10* manifests that the author’s capability of grammatical conversion need to be improved. The expression in source text *good development prospects* is in modifier-center structure which consists of a modifier *good* and a noun phrase *development prospects*. The author intends to convert the grammatical structure into subject-predicate structure “发展前景良好”. However, due to the lack of second language automaticity and inflexible conversion, the author consumes more processing capacity for production than expected, thus results in filled pause.

5.3.3 Correcting slips of tongue

Based on Levelt’s speech production model, Articulator is the third stage of production which executes the previous encoding and produce external speech. When the available processing capacity for Articulator is not enough to meet its requirement, slips of tongue may occur unconsciously.

Example 11

SL: Despite the setbacks and challenges brought about by the pandemic, the Philippines remains in the middle of the pack, placing next to Vietnam and Indonesia in terms of prospects for recovery. We are working for V-shaped recovery to bounce back to pre-COVID levels of economic growth.

TL: 尽管疫情带来了挑战, 菲律宾仍然(...)仍然在中间(...), 我们努力实现V型复苏, 来fei恢复到疫情前的(...)水平。

Analysis: The first sentence in *example 11* is a long complicated sentence, which contains *Vietnam* and *Indonesia*, making the author confronted with cognitive

overload. There is a slip of tongue in the second sentence. The processing capacity distributed to Articulator in Production Effort is insufficient to meet its requirement. Thus, an error in speech production arises with a wrong initial consonant *f* pronounced instead of *h*. Moreover, the processing capacity available to Conceptualizer is insufficient, which is held accountable for the slip of tongue due to the absence of internal monitoring that is involved in Conceptualizer.

5.4 Coordination

The most typical feature of simultaneous interpreting is simultaneity of listening and speaking, which has been taken into consideration in an attempt to figure out reasons for disfluency in SI (Tissi, 2000, pp. 105). The SI interpreter requires more coordination effort than the CI interpreter, so it is essential that the interpreter should manage his/her mental energy appropriately to conduct multitasks in SI. Based on Gile’s Effort Model, the Coordination Effort is proposed to describe the interpreter’s management to divide attention in order to carry out multi-tasks in SI. The author finds that she cannot coordinate her mental energy into each part on some occasions. In particular, when the author encounters difficulties in listening and comprehension as well as organizing target language, little mental energy is allocated to speech production, resulting in disfluency in delivery.

Example 12

SL: We can provide assistance and consultancy services in addition to case facilitation so that you do not have to worry about what clause to incorporate and how. We have a readymade clause that we can share with you on request.

TL: 我们(...)也提供咨询服务, 所以你们也不用担心(uh)不用担心(...)条款的问题。我们(...).

Analysis: Various types of disfluency occur in this example including three silent pauses, a filled pause and a repetition. The contributory factor is the inappropriate attention allocation among Listening and Analysis Effort as well as Production Effort. As a complex sentence, it contains an object clause *what clause to incorporate and how*, which is not difficult to understand. However, the author hesitates for a while because she fails to think out the corresponding equivalent for this object clause immediately. Due to the different syntactical structure between English and Chinese, it is challenging for the author to come up with a correct rendition flexibly. Reorganizing the syntactical structure of target language exhausts the author’s mental energy. Therefore, the mental energy which is intended to be allocated to Production Effort is partially transferred to Listening and Analysis Effort. The insufficient processing capacity left for Production Effort does not meet the requirements. In addition, because of the unbalanced allocation of mental energy, the author pays too much attention to the previous

sentence and thus misses the following sentence *We have a readymade clause that we can share with you on request*, resulting in a long pause.

Example 13

SL: In 2019, with the increasingly close relations between the two countries, the number of Chinese tourists to Turkey hit a record high, exceeding 565,000. We want to increase the momentum of economic development. Our goal is to bring the exchange of visits between the two countries' ministers and heads of state back to the highest level in 2018 and 2019.

TL: 2, 2019年, 中国向土耳其的(...)中国游客到访土耳其达历史新高。同时我们也想使两国元首和(uh)部长的互访数回到2018年和2019年的最高水平。

Analysis: This example is featured with big numbers. It is obvious that silent pauses, repetitions as well as self-repairs occur in the author's target language production. Frankly speaking, the numbers are not uttered in high density. However, as the author is insensitive to English figures, particularly big figures, she needs excess processing capacity for listening and analyzing figures. It is difficult for the author to render them into target language fluently and correctly. Too much concentration on figures results in the fact that the available processing capacity for production is exhausted, without monitoring the target language delivery which is interrupted with silent pauses, repetitions and self-repairs.

6. CONCLUSION AND SUGGESTIONS

Based on the Effort Model, the interpreter's mental energy is limited and simultaneous interpreting that requires multitasking capability, consumes all the mental energy, as a result of which, the interpreter is often caught in cognitive saturation. Thus, interpreters may encounter great challenges and have poor performance in simultaneous interpreting. In this case, non-fluent interpretation of source language is easily produced. After the analysis of the source language text and the recording of the author's interpretation, she summarizes three types of disfluency that are manifested in the target language speech: pauses including silent ones and filled ones, repetitions and self-repairs. They are triggered by the author's cognitive saturation. To be specific, the author is confronted with cognitive overload in listening and comprehension, storing information segments in short-term memory, speech production and coordinating mental energy.

In response to avoid unnecessary disfluency, feasible coping strategies are put forward under the guidance of Gile's Effort Model, which are divided into technical strategies and non-technical strategies.

Technical strategies are related to simultaneous interpreting skills and tactics, such as conversion, omission, anticipation, segmentation, restructuring and so

on. These tactics are practical and critical components that help largely improve the interpreting quality. For instance, omission can be adopted on the occasion that part of source language is redundant. The redundant parts of the source language cannot be completely interpreted into target language in simultaneous interpreting. However, omitting certain redundant elements in source language can ease the interpreters' allocation burden of the mental energy on these redundant constituents, thus improve the fluency of interpreting production and guarantee the successful and smooth conveyance of information. In particular, the high information density or complexity requires the interpreter to simplify target language in order to convey the correct message of source language. Apart from omission, segmentation is also a useful tactic in SI practice, which can relieve the interpreters' cognitive load in listening and analysis as well as short-term memory. Segmentation is applied to cut the source language into several information unit, as long information unit will increase the pressure of listening and information storage. If the mental energy is allocated to listening and short-term memory in an unbalanced manner, the energy left for speech production will be insufficient, leading to a series of problems in interpreting delivery. On the contrary, if the information unit is too short, interpreters may fail to understand the meaning expressed in the source language or remember message, hindering the interpreting output. Therefore, appropriate segmentation can largely contribute to the improvement of interpreting quality.

In addition to technical strategies, non-technical strategies should also be emphasized with an aim to ensure a pleasant and high-quality delivery. Non-technical strategies are concluded as three aspects: improving language proficiency, make full preliminary preparation and conducting self-monitoring.

Language proficiency must be enhanced, which can decrease the processing capacity requirements in simultaneous interpreting. In other words, interpreters need to have an excellent mastery of working languages. Compared with the output of the first language, the output of the second language always occupies more processing capacity, because the second language availability is relatively lower. Therefore, the occurrences of disfluency arise, making the speech flow intermittent and unpleasant to listen to. Interpreters should improve their language proficiency in multiple ways, to name a few, strengthening listening exercises, mastering public speaking skills and so on.

Language proficiency cannot ensure the high-quality interpreting performance. It is necessary that interpreters should make full preliminary preparation and work on extralinguistic knowledge which is beneficial to listening and comprehension. That is to say, exposure to terminology and expansion of background information can enhance the familiarity with the source language

speech. In this case, interpreters are equipped with higher level of comprehension availability by accumulating background knowledge, which can save the available processing capacity and reduce the required processing capacity in listening and analysis, short-term memory and even speech production.

Last but not least, self-monitoring is a requisite coping strategy that can improve the quality of interpreting performance. Interpreters are required to allocate a certain amount of mental energy in listening and analysis during the process of simultaneous interpreting. It should be pointed out that part of the Listening and Analysis Effort should be allocated to self-monitoring of interpreters' target language production. It can be applied when interpreters reflect on what they have interpreted and make an effort to correct any errors and avoid disfluency in the target language output, using tactics like omission, supplement and simplification.

On one hand, interpreters should choose SI technical strategies appropriately in practice so as to establish the rule of selecting coping strategies. On the other hand, non-technical strategies cannot be ignored. In particular, language proficiency lays a solid foundation for interpreters' fluency improvement in delivery, which entails long-term self-training.

REFERENCES

- Bussmann, H. (2006). *Routledge Dictionary of Language and Linguistics*. London and New York: Routledge.
- Cecot, M. (2001). Pauses in simultaneous interpretation: A contrastive analysis of professional interpreters' performance. *The Interpreters' Newsletter*, 11, 63-85.
- Dai, Z. H. (2011). A Study on Disfluency in Chinese to English interpretations of Chinese EFL Learners. *Shanghai Journal of Translators*, (01), 138-143.
- Fillmore, C. J. (1979). On fluency. *Individual Differences in Language Ability & Language Behavior*. New York: Academic Press.
- Foster, P., & Skehan, P. (1996). The influence of planning on performance in target-based learning. *Studies in Second Language Acquisition*, 18(2), 299-423.
- Foster, P., & Skehan, P. (1999). The effect of source of planning and focus on planning on task-based performance. *Language Teaching Research*, 185-215.
- Fox Tree, J. E. (2003). Disfluencies in spoken language. In Nadel, L. (Ed.), *Encyclopedia of Cognitive Science*, 983-986. London: Nature Publishing Group.
- Fu, R. B. (2013). Directional Effects on Disfluencies in Consecutive Interpreting. *Modern English*, (02), 198-205.
- Gile, D. (1995). *Basic Concepts and Models for Interpreters and Translator Training*. Amsterdam and Philadelphia: John Benjamins.
- Gile, D. (2011). *Basic Concepts and Models for Interpreters and Translator Training* (Revised ed.). Shanghai: Shanghai Foreign Language Education Press.
- Goldman-Eisler, F. (1968). Speech production and the predictability of words in context. *Quarterly Journal of Experimental Psychology*, 10, 96-106.
- Lesson, R. (1975) *Fluency and Language Teaching*, p. 136. London: Longman Group Limited.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 33, 41-103.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15, 19-44.
- Mead, P. (2005). Methodological issues in the study of interpreters' fluency. *The Interpreters' Newsletter*, 12(1), 497-498.
- Miao, H. Y. (2009). Pauses in L2 Non-fluent Oral Production. *Journal of PLA University of Foreign Languages*, (04), 56-60.
- Raupach, M. (1980). Temporal variables in first and second language speech production. *Temporal variables in speech* (pp.263-270). New York: Mouton.
- Shrimberg, Elizabeth, Elle. (1994). *Preliminaries to a Theory of Speech Disfluencies*. Berkeley: University of California.
- Tan, Y. Z., & Feng, E. Y. (2013). A Study of Unnatural Pauses in Chinese-to-English Consecutive Interpreting. *Journal of Huaihua University*, (03), 89-91.
- Yang, J. (2004). A Review of Studies on Disfluencies in Speech Production. *Foreign Language Teaching and Research*, (04), 278-284.