WESTERN SYDNEY
UNIVERSITY

The Effect of Study Abroad Experience and Working Memory on Chinese-English Consecutive Interpreting Performance

Ruiyuan Wang

A thesis presented to the Western Sydney University in fulfillment of the requirements for the degree of Doctor of Philosophy

July 2020

To my late grandmother

Acknowledgements

My PhD journey has been challenging yet very rewarding. I have encountered many difficulties, but encouragement and support from people around me have helped me overcome all those difficulties and complete my thesis. I owe them my heartfelt gratitude and lifelong friendships.

Firstly, thanks to the joint PhD program between Western Sydney University (WSU) and Xi'an Jiaotong University that gave me an opportunity to pursue my PhD studies at WSU. My sincere thanks go to Professor Peter Hutchings, Associate Professor Wenxin Nie, Associate Professor Ying Li, Professor Xiangjing Chen and Associate Professor Fei Wang, for their valuable support and encouragement.

I wouldn't have been able to complete my journey without my supervisors Professor Jing Han, Associate Professor Bruno Di Biase and Associate Professor Mark Antoniou. As my principal supervisor, Professor Han has given me invaluable guidance, timely feedback and insightful comments. I am particularly grateful to her for being such a caring, supportive and encouraging mentor. She has always made herself available to me and supported and helped me to pull through challenging times and keep my research on the right track. Her insights and encouragement boosted my confidence whenever I was in self-doubt. Professor Han has not only guided me through my academic studies, she has also kindly provided me with opportunities to reach my potential including being an MC for an international conference at WSU.

I have been very fortunate to have had Associate Professor Bruno Di Biase as my co-supervisor. His knowledge and expertise in second language acquisition and bilingualism have been invaluable to the development of my research. His critical comments and constructive suggestions have been an immense help throughout my thesis writing and I have learnt so much from his feedback and guidance. I cannot thank him enough for his generosity with his time and help. I am particularly grateful for his meticulousness, encouragement and patience.

I am deeply indebted to Associate Professor Mark Antoniou as my co-supervisor for his instrumental guide in my research design, methodology, experiments and data analysis. He patiently taught me how to operate the experimental devices and walked me through the experiment process step by step. I have benefited enormously from his expertise and knowledge in working memory and psycholinguistics and from his invaluable feedback. His readiness to help, his resourcefulness, his encouragement and clear direction have helped me enormously during my research and thesis writing.

My heartfelt gratitude also goes to Dr Kenny Wang for his many practical suggestions and assistance during my thesis writing and data collection. I would also like to thank Alisa Tian for her help and support during my data collection. I am also grateful to those students who participated in my experiments.

During my years at WSU, my fellow PhD student Yanmei Liu and I have spent many hours together in discussing our researches and how to move forward. We have encouraged and supported each other, which is particularly important during the last few months of the lockdown due to Covid-19. I am grateful for and appreciative of our friendship.

I would also like to thank Juqiang Chen, Dr Ying Cao and Dr Yuan Ma for their help and advice on academic studies and their encouragement. My longstanding friends Xiaoling, Xinru, Yudi, Queenie and Bingqing have been my consistent source of love, support and inspiration.

Last but not least, I must thank my parents, who have always believed in me and encouraged me to follow my dreams. They have taught me to love knowledge, books and learning. I couldn’t have pursued my PhD studies without their financial support and their unconditional love.

I want to dedicate this thesis to my beloved grandmother who sadly didn’t live to see the completion of my PhD journey.

Statement of Authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

[redacted]

Ruiyuan WANG

July 2020

Table of Contents

List of Tables

List of Figures

Abbreviations

SAE	study abroad experience
SLA	second language acquisition
L1	first language
L2	second language
WM	working memory
NSA	non-study abroad group (participants without SAE)
SA	study abroad group (participants with SAE)
CI	consecutive interpreting
SI	simultaneous interpreting
IELTS	International English Language Testing System
NAATI	National Accreditation Authority for Translators and Interpreters

Abstract

This thesis investigates how study abroad experience (SAE) and working memory (WM) influence interpreting performance. Using a second language (L2) is cognitively demanding because it involves activation of a new language and the inhibition of the first language (L1). This is a general issue with all bilinguals, who have to suppress or control whichever language is currently not in use. As a special group of bilinguals, interpreters are expected to efficiently switch between the two languages by analysing input sound signals, extracting meaning, transforming, storing and retrieving the message in the input language, and then retrieving the lexicon in the target language that will be appropriate for expressing that message, (re)formulating it and finally conveying it in the target language. Moreover, some or all of these operations take place in parallel, and this multi-tasking heavily taxes interpreters' WM. The quality of interpreting performance is known to correlate with several variables, such as language proficiency, duration of training, and interpreting experience. One factor that has received little research attention is the effect of overseas experience: Does studying in a target-language environment benefit interpreting performance? Language learners, including interpreting students, are often advised to study abroad, but the benefits of this experience, especially for interpreters, is not well understood.

Taking an interdisciplinary approach, the present thesis examines the relationship between SAE, WM and interpreting performance. The main research questions examine whether students with SAE outperform those without such an experience in consecutive interpreting (CI), and how WM may be involved. To answer these questions, 25 Chinese (L1)-English (L2) interpreting and translation students were recruited in Australia and 25 were recruited in China. They were asked to complete CI tasks (in both directions), an online vocabulary knowledge test, and a self-report questionnaire evaluating their own language experience and proficiency. Two psycholinguistic experiments were also administered with the aim of assessing, more objectively, word translation efficiency and WM resource availability.

The results show that students with SAE surpassed their non-SAE counterparts in word translation efficiency, L2 fluency and L2 grammatical accuracy. A similar trend was observed in study abroad participants' overall CI performance from L2 to L1. It is worth noting that the tendency was independent of participants' WM. Concerning WM, the results indicate that it was strongly correlated with interpreters' bidirectional CI performance. That is, a larger WM could help achieve a better CI output in both language directions. Taken together, these findings suggest that two factors turn out to significantly influence CI performance, namely, prolonged and effective overseas study, and larger available WM resources. This research illustrates the importance of SAE and WM in interpreting, and sheds light on the relationships between language context, cognitive resources and interpreting performance. A better understanding of these relationships may have implications for future interpreting training and practice.

Chapter 1 Introduction

This thesis investigates the effects of study abroad experience (SAE) and working memory (WM) on interpreting performance. Specifically, it examines how Chinese (L1)-English (L2) interpreting students in masters degree courses vary in their consecutive interpreting (CI) performances in different language learning contexts, and how WM is involved. Unlike ordinary conversations in which both speaker and hearer use the same language, interpreting crosses language barriers and enables interlingual communication. Thus, it involves the management of at least two languages. Given interpreting's crucial dependence on bilingual performance, it is inevitably intertwined with bilingualism, second language acquisition, bilingual processing and language switching (Schwieter & Ferreira, 2017). From a cognitive perspective, interpreting, like other types of language production and comprehension, is constrained by WM resources, and thus the nature and function of this cognitive resource is often explored along with interpreting performance.

Spoken language interpreting can be divided into a simultaneous mode and a consecutive mode. In simultaneous interpreting, the interpreter's target language output takes place almost at the same time as the source text delivery. On the other hand, CI is the interpreting process which occurs after the speaker has completed one or more segments in the source language (Russell, 2005). Both modes are highly complex linguistic activities which comprise sub-components such as perception, comprehension, memorising and production. Currently, most master of translation and interpreting (MTI) programmes worldwide set CI as a compulsory mode which lasts the full length of a degree course (Jin, 2017). In addition, a good mastery of CI is usually considered (alongside sight translation) to be a foundation for learning simultaneous interpreting (Jin, 2017; Russell, 2005).

While researchers acknowledge the pedagogical significance of CI, a good number of the empirical studies exploring cognitive aspects of interpreting actually focus on the simultaneous mode, leaving CI relatively underexplored (Pöchhacker, 2016, p. 108). This thesis hopes to make a contribution to bridging this gap by concentrating on CI performance from the perspectives of language processing and WM resource availability.

Using a second language (L2) is effortful, even for the most fluent bilinguals (De Groot, 2000), as it involves inhibiting the dominant L1 and processing the L2 in a grammatically consistent manner. Both processes are governed by WM, a cognitive resource which is responsible for temporarily storing and manipulating the target information as well as concurrently inhibiting distractors (Baddeley, 1998, 2003, 2006, 2010). Research consistently demonstrates that, in the bilingual mind, linguistic elements (e.g., sounds, forms and concepts) in both languages are co-activated in a non-language-specific manner and they compete for selection during both speech comprehension and production (Jarvis & Pavlenko, 2008; Odlin, 2005, 2012). As a result, bilingual processing involves cross-linguistic interference between L1 and L2, and the selection of the target language at a given time. For instance, L2 use is achieved by successfully inhibiting the unintended L1 (Green, 1998). Moreover, L2 processing is a controlled operation during which learners have to draw on corresponding linguistic rules and assemble them in their WM (Levelt, 1989, p. 28). The burden imposed on WM is particularly salient when bilinguals are trying to produce a grammatically correct L2 sentence involving, for instance, grammatical agreement, which requires speakers to formulate the rest of the sentence while keeping in mind what they have already uttered, so that they can make the upcoming parts grammatically consistent with what they have already produced (De Abreu & Gathercole, 2012; Miyake & Friedman, 1998).

Interpreting is even more WM-demanding than ordinary bilingual processing. During the task, interpreters are expected to switch back and forth between languages quickly, comprehend and produce accurately in a well-formed manner without obvious crosslinguistic intrusions from either language (De Groot, 2000). Therefore, efficient lexical access is considered essential for interpreting performance (Christoffels et al., 2003; De Groot, 2000). When lexical access becomes too effortful, sound signals accumulate in interpreters' WM until they are able to process the signals into meaningful segments or take notes. In such situations, WM availability can be saturated rapidly. If this happens, either the incoming information cannot be attended to, or it is attended to at the expense of previously heard segments (Gile, 2009, p. 225).

Additionally, if too much of the interpreter's limited WM is consumed in word-level processing (in recognising words or retrieving the corresponding translation equivalents), then little remains for higher-level processing such as grammatical processing or cross-sentence integration of information. As a consequence, interpreting performance is impaired.

The WM tension can effectively be reconciled by automatic processing, which is effortless. In a task, the more the sub-processes are automatised, the more WM resources can be freed and redirected to other WM-taxing purposes (e.g., Antoniou et al., 2015; Christoffels et al., 2003; Orena et al., 2015; Segalowitz, 2003). Thus, if one can tackle the lexico-morphological processing of an L2 with greater efficiency – that is, with more automatised components, more resources can be left over for semantic, pragmatic and sociolinguistic levels of language use. Therefore, for complex linguistic activities such as interpreting, efficient and automatic linguistic processing has positive influence.

However, according to Interaction Theory, in second language acquisition (Carroll, 1999; Long, 1981, 1996) linguistic processing efficiency does not just happen: It is the result of effective and rich input, interaction and output. To improve processing efficiency, L2 learners consistently resort to methods such as taking target-language classes, watching target-language movies and undertaking repetitive practice. More radically, many others choose to go to study in a target language-speaking country for language immersion.

Study abroad experience, within the field of second language acquisition, is characterised as a type of L2 learning setting which differs from both purely natural exposure and classroom instruction. Learning through natural exposure refers to the untutored, spontaneous acquisition of a target language in the country where that language is spoken. In this kind of setting, learners are usually expected to analyse and generalise the linguistic rules by themselves, based on the input they receive (DeKeyser, 1997). Foreign language instruction, on the other hand, is typically conducted in classrooms within a country where the language to be learned is different from the dominant language commonly used in that country (e.g., Spanish taught in German classrooms), and the classroom typically serves as the sole source of exposure to the target language (Muñoz, 2008). This type of context is considered to be generally inflexible and test-oriented, where instructors explain linguistic rules to learners explicitly (DeKeyser, 1997). In contrast, SAE involves, characteristically, a mixture of language training in formal classroom settings and life experience in a country where the target language is the same as the dominant environmental language. Learners usually go abroad to a country where their L2 is spoken after a period of formal study in their country of origin. This is done because, given the constant availability of natural L2 in and out of the classroom environment, SAE has traditionally and theoretically been assumed to provide an optimal setting in which to acquire and/or further enhance one's L2.

Chinese young people, in particular, appear to be driving the trend of studying abroad (Bhandari, 2017; Button & Szego, 2020). In 2016, approximately 145,000 Chinese overseas students were pursuing English language courses for related academic purposes in English-speaking countries such as the UK, the USA, Australia and Canada, which, according to Pavlacic (2018) implies that almost every tenth English language learner around the world was from China.

In terms of interpreting, SAE is considered necessary, and may even be compulsory for training programmes in European countries. It is frequently a requirement that interpreting students spend time completely immersed in the country where the dominant language is their weaker language (Gile, 2005, 2009, p. 221; Napier, 2015). For instance, Heriot-Watt University in Scotland requires its third-year interpreting students to spend one year in the country of their weaker language (Napier, 2015). Also, it is mandatory for students majoring in Advanced Interpreting and Translation at the Université de Paris to have SAE for twelve consecutive months in a country where the dominant environmental language is the student's L2 (Toudic, 1997).

However, whether this experience provides greater linguistic benefits than domestic classroom instruction is a matter of ongoing debate. Moreover, the research in this regard is mainly conducted within the area of second language acquisition, where the empirical evidence suggests that students with SAE exhibit larger L2 gains than their classroom counterparts in specific areas such as vocabulary, oral fluency, accuracy, grammar, written performance and overall development. Also, studies have shown that an L2-dominant environment can effectively attenuate L1 interference when processing L2 (e.g., Baus et al., 2013; Linck et al., 2009; Sagarra & Labrozzi, 2018). Many other researchers question the benefits of SAE and offer evidence that such an experience provides little change in learners' L2 abilities, or that domestic classroom instruction leads to equal improvement. They argue that even though many L2 learners with overseas experience appear to speak with greater ease and confidence in language use, they are still error-prone in areas such as maintaining grammatical consistency (Arnett, 2013; Freed, 1995; Marqués-Pascual, 2011). However, these studies do not, by and large, look at potential improvements in interpreting students or in CI in particular.

Against this backdrop, this thesis endeavours to explore the effects that SAE, in conjunction with WM, may have on CI performance. Given that an increasing number of language learners, including interpreting students, are encouraged to cross borders to pursue SAE, it is worth investigating whether studying in an L2-dominant environment does indeed provide greater benefits for CI performance than the domestic classroom context. It would also be interesting to know how WM is involved in the way interpreting tasks are processed.

Employing an interdisciplinary approach, the present study compares the linguistic and CI performance of students with and without SAE. The correlation between their WM and their performance of these tasks is also analysed. Few researchers have combined student interpreters from different language learning contexts in a single study to examine the effectiveness of SAE and WM on interpreting. This study endeavours to gain further understanding of these relationships and contribute to filling this gap.

To that end, two groups of participants were recruited in the study: 25 Chinese students in Australia (the SA group) with 24 months' SAE on average, and 25 Chinese students in China (the NSA group) who had never been to any English-speaking country at the time of data collection. All fifty participants were Chinese full-time master's students majoring in Chinese (L1)-English (L2) interpreting and translation. To maintain objectivity, efforts were made to ensure that the SA participants were recruited randomly from three different universities in Sydney, Australia. Likewise, the NSA participants were recruited randomly from three different universities in Xi'an, China. Apart from that, the two groups of participants were matched in age and the duration of their interpreting training. Their L2 proficiency and L2 age of acquisition were also taken into account.

In order to evaluate their linguistic processing performance and WM resource availability, each participant was asked to complete a computer-based test on word knowledge, a word translation efficiency test, and a reading span test. In addition, participants' CI interpretation in the L1 to L2 direction was also used to give researcher an idea of their L2 fluency and grammatical accuracy. Participants' overall CI performance in both language directions was assessed using two CI tasks. In these tasks, the two input texts were comparable in terms of text type (monologic presentation), register (formal), context (speech), readability, length and delivery speed.

Three research questions guided the investigation. The first concerns the influence of SAE on interpreting students' linguistic performance (i.e., lexical and grammatical processing and fluency). If SAE contributes greater linguistic benefits than non-SAE instruction, then the SA group should outperform the NSA group in the linguistic tasks. The second question asks whether students with SAE will demonstrate superior CI performance than their domestic counterparts, realised as higher overall CI scores. If SAE is beneficial to interpreting, then participants in the SA group should obtain higher CI scores than participants in the NSA group. The third research question is about the significance of WM in these tasks. If WM is critically involved, participants with greater WM availability should surpass those with smaller WM in linguistic and CI performance.

This dissertation consists of six chapters. Following this introduction, the chapters are organised as follows:

Chapter 2 reviews the literature on the relationships between SAE, linguistic processing, WM and interpreting. Previous empirical studies related to these areas are also reviewed in this chapter, followed by an outline of the interdisciplinary framework used in the present study. Gaps in previous research are identified, located at the intersection between the three aspects considered here, namely: SAE, WM and CI performance.

Chapter 3 first presents the research questions and corresponding hypotheses. The chapter then describes the methods used to carry out this investigation of the interaction between SAE, WM resources, linguistic performance and CI performance. This description includes the overall design of the study, and a detailed description of the experimental tasks used to measure participants' linguistic and CI performance, as well as their WM resource availability. The justifications for the sample size and recruitment criteria are also stated, followed by explanations of task grading and the methods of data analysis adopted.

Chapter 4 presents the research data collected from the tasks completed by the participants and examines how these data address the research questions and their corresponding hypotheses. Quantitative measurements are used for data analysis.

A detailed analysis and interpretation of the results is presented in Chapter 5. The presentation is guided by the research questions and the relationship of the results to previous research findings. Findings are discussed regarding the effect of SAE on interpreting students' linguistic and CI performance, and how participants' WM interacts with their linguistic and CI performance in different language environments.

Finally, Chapter 6 summarises the major findings, focusing on the theoretical and pedagogical implications of the study, indicating its limitations, and offering suggestions for further research.

Chapter 2 Literature Review

This chapter firstly reviews the literature in the fields of study abroad experience (SAE) and second language acquisition (SLA). The focus is on the two components of linguistic performance: bilingual lexical processing and grammatical processing. Section 2.2 is devoted to issues, concepts and empirical studies concerning working memory (WM) and L2 processing. Section 2.3 introduces models and empirical studies on interpreting performance and WM. Related research regarding linguistic processing and interpreting performance is also presented. Section 2.4 describes the interdisciplinary frameworks related to the three main research areas of the current study: SLA, WM and interpreting. The research gap addressed in the study is identified in Section 2.5. Finally, this chapter concludes with a summary in Section 2.6.

2.1 Study Abroad Experience, Second Language Acquisition and Language Processing

2.1.1 Study Abroad Experience and Second Language Acquisition

Numerous linguistic studies into SAE support the view that substantial immersion experience is necessary to achieve a high level of L2 performance (D. E. Davidson, 2007). Indeed, mounting evidence suggests that being exposed to a language in its natural environment contributes to the development of proficiency in that language, including oral fluency (Juan-Garau, 2014, 2018; Llanes & Muñoz, 2009, 2013; Mora & Valls-Ferrer, 2012), accuracy (Mora & Valls-Ferrer, 2012; Pérez-Vidal & Juan-Garau, 2011), grammar development (Howard, 2005; Isabelli & Nishida, 2005), written performance (Godfrey et al., 2014; Grey et al., 2015; Pérez-Vidal & Juan-Garau, 2011), and overall development (J. B. Carroll, 1967). However, this assumption has also been challenged. A number of researchers question the benefits of SAE (Arnett, 2013; Collentine, 2004; Barbara Freed et al., 2003; Godfrey et al., 2014; Isabelli-García, 2010; Marqués-Pascual, 2011; Mora & Valls-Ferrer, 2012) citing evidence that study abroad appears to make little change to the L2 abilities of students, or that students with and without SAE perform equivalently.

Study abroad research, to a large degree, is built around the Interaction Hypothesis, also known as Interaction Theory (Carroll, 1999; Long, 1981, 1983, 1996) which describes the processes involved when learners encounter input (Krashen, 1985; Long, 1996), are involved in interactions (Gass, 1997, p. 104; Gass & Mackey, 2007), and receive feedback and produce output (Swain, 1985, 1995, 2005). Interaction Theory also attempts to elaborate on why interaction and learning can be linked by using cognitive concepts derived from psychology, such as noticing, WM, and attention (Izumi & Bigelow, 2000; Kormos, 1999; VanPatten, 1996, p. 24).

Input refers to the language that a learner is exposed to (Gass & Mackey, 2007), which is regarded as the primary source for language learning (Long, 1996; B. Schwartz, 1993). But input alone is insufficient for L2 development (Gor & Long, 2009; Harley & Hart, 1997; Harley & Swain, 1984); interaction is also required (S. Zhang, 2009).

Interaction refers to the conversations that learners participate in (Gass & Mackey, 2007). It provides learners with opportunities to draw attention to unknown parts of the language, and to make connections between form and meaning (Gass, 1997, p. 105; Pica, 1994). Through interaction with native speakers or more competent interlocutors, learners are able to gain access to multiple exemplars of target-like and comprehensible input, which may facilitate the development of SLA. For instance, in a study by Mackey (1999), students who were allowed to interact with native speakers on a task-based activity improved more in their development of English question formation than those who were only allowed to observe the interactions. Moreover, the interactors also performed better than others who received scripted pre-modified input on the same tasks, and they maintained their advantage on delayed posttests (see also Gass, 1997, p. 104; Gass & Mackey, 2007; Gor & Long, 2009; Mackey, 2007; Pica, 1994).

Last but not least, output in Interaction Theory refers to language use which is assumed to prompt learners to move from comprehension characterised as meaning-oriented, open-ended, non-deterministic, and strategic processing to the complete grammatical processing required for accurate production (Swain, 1995). Output, therefore, seems to play a potentially significant role in syntactic and morphological development (Gass & Glew, 2018). It is also believed to consolidate the linguistic knowledge that learners have not yet fully internalised, and thus boost processing automaticity (Swain, 1985, 2005). This is because automatic processing is attributable to consistent and successful mapping or practice (Loschky & Bley-Vroman, 1993). In the opinion of many researchers, automaticity is beneficial for linguistic processing. Firstly, automatic processing consumes fewer cognitive resources than the controlled processing which occurs when conscious effort and attention are needed to perform a task. The more automatic performance becomes, the more cognitive resources can be used for other purposes. For example, if a bilingual can manage the phonology and syntax of an L2 automatically, then more cognitive resources can be freed to process semantic, pragmatic, and sociolinguistic levels of communication (Segalowitz, 2003). Secondly, when a procedure becomes automatic it will process information with swiftness and accuracy. Thirdly, automaticity is strongly associated with fluency (Hulstijn, 1997; Skehan, 1998).

SAE, which is characterised by its potential for providing massive exposure to authentic language input and unlimited opportunities for target-language interaction, negotiation, and output (Fernández-García & Martínez-Arbelaiz, 2014; Rothman & Iverson, 2007; Zaytseva et al., 2018), is commonly thought to be beneficial for L2 acquisition and processing.

Before reviewing the studies regarding the effect of SAE on L2 development, it is useful to touch on some broad theoretical points to provide a background for their interpretation. For instance, one should distinguish between linguistic competence and performance. According to Chomsky (1965, p. 3), linguistic competence is the speaker-hearer's knowledge of the language, whereas performance refers to the real deployment of language knowledge in concrete situations. According to this view, having acquired linguistic knowledge in L2 is not equivalent to being able to make use of that knowledge stably and fluently in a given situation. Performance is susceptible to linguistically irrelevant factors such as cognitive loads (Ma, 2017), and thus may vary even for the same task administered on different occasions (Wright, 2018). This understanding of the disparity between competence and performance can be found in Hawkins and Liszka (2003), in which L2 learners performed well in their morphological knowledge tests but their performance was significantly unstable when it came to putting that particular knowledge into practice, for instance when retelling a movie or recounting an experience spontaneously. A similar discrepancy was reported earlier by Pienemann (1984) in the course of his 'teachability' experiments. Thus competence (i.e., linguistic knowledge) can be seen as providing the linguistic input for comprehension and production. Hence competence, which provides one of the components for the formulation or the interpretation of a message, is different from the production or comprehension process (i.e., performance). In the present subsection, we focus on the effect of SAE on learners' linguistic (lexical and grammatical) competence development, and the performance development will be discussed in the next two sessions.

An early major study looking on language development and the influence of study abroad was conducted by Carroll (1967) who tested 2,782 college seniors majoring in French, German, Italian, or Russian. The findings suggest that the amount of time spent abroad positively correlates with the level of language skills attained, as reflected in participants' test scores. In Carroll's study, participants who spent a year abroad had a mean score of nearly ten points higher than those who had never been abroad.

Since then, research on the effects of SAE on the acquisition of the target language has extended to more specific aspects of linguistic proficiency, such as lexical and grammatical development. Methods used to explore lexical and grammatical competence consist of lexical knowledge tests, grammaticality judgment tests, multiple choice tests, cloze tests, and so on. In the literature, two approaches to research design have usually been adopted: One is to investigate learners' linguistic development before and after SAE (e.g., Fitzpatrick, 2012; Grey et al., 2015). And the other is to compare a study abroad group, or group with immersion experience, with a control group of students who have never studied abroad (e.g., Barik & Swain, 1976; Harley & Jean, 1999; Lo & Murphy, 2010). Alternatively, an SA group may sometimes be compared to a group of native speakers (e.g., Jimenez-Jimenez, 2010; Pérez-Vidal & Juan-Garau, 2011; Song, 2015).

Studies on lexical knowledge development as a result of SAE have investigated participants with different language proficiency levels and exposure lengths. Some studies have yielded positive evidence that SAE is beneficial for improving L2 lexical knowledge, including lexical size (e.g., Barik & Swain, 1976; Ife et al., 2000; Lo & Murphy, 2010) and depth ('how well' a word is known by the learner) (e.g., Collentine, 2004; Jimenez-Jimenez, 2010; Lo & Murphy, 2010). Results vary regarding whether or not SAE facilitates the widening of learners' lexical repertoires in the target language (e.g., Dewey, 2008; Lara, 2014; Segalowitz et al., 2004). Ife et al. (2000), for instance, investigated the impact of SAE on the lexical development of intermediate versus advanced L2 learners. Participants' length of immersion in the L2-speaking country ranged from four to eight months. The results suggested that participants in both intermediate and advanced groups achieved a remarkably similar patterns of improvements in their lexical size during their overseas periods, even though the gains of the advanced group were always slightly higher than those of the intermediate one. The length of time spent abroad was also found to be positively associated with lexical gains. Similarly, Milton and Meara (1995) observed improvement among SA learners. However, in their study, learners with smaller initial vocabularies tended to make more progress than those with larger starting vocabularies.

According to Jiménez-Jiménez (2010), the soundness of Ife et al. and Milton and Meara's studies was compromised by the lack of a control group. There was no evidence available to support the claim that the lexical gains of participants during their SAE would not have emerged if they had continued their learning experience in a classroom context. Therefore, Jiménez-Jiménez (2010) recruited a control group and compared the lexical size and depth of SA and NSA learners. The findings suggest that classroom instruction did not contribute to participants' development in lexical size nor depth. SAE, on the other hand, contributed to both levels. Collentine (2004) also found a facilitative role for SAE in participants' lexical depth improvement: the SA group in Collentine's study were able to generate more semantically dense words than the domestic classroom group. These findings are seen as evidence that an extended stay (over six months) in the L2 community is necessary to trigger lexical size development, and foster improvement in lexical depth.

Generally, researchers attribute the larger lexical size of SA groups to the fact that they had more opportunities to encounter new words. SA participants are expected to have significant amounts and wide varieties of L2 exposure. The distinct role played by L2 in the curriculum during SAE also contributes to this result. The L2 is used as the medium of instruction, meaning some or all of the academic subjects studied by the students are taught in L2 (Tedick, 1998), so that students can be 'immersed' in the L2 input and then learn the language incidentally (Genesee, 1985), as the focus is on academic content, not the language itself (Lo & Murphy, 2010). In other words, it is assumed that even though students in both learning contexts encounter L2 vocabulary in language lessons, SA learners also come across novel words in other academic subjects and in the whole out-of-school environment.

In contrast, other studies question the effect of SAE by providing evidence that it does not benefit learners' lexical development (e.g., Briggs, 2015; Fitzpatrick, 2012), or that the effect of SAE can also be achieved through intensive classroom instruction (e.g., Dewey, 2008; Serrano et al., 2011). Serrano, Llanes and Tragant (2011) reported that intensive domestic classroom instruction can provide similar improvements to SAE in lexical development. They examined L2 lexical complexity, defined as different words being employed (Dewi, 2017), and found the SA and NSA groups to be equivalent despite the fact that the NSA received twice as many hours of L2 instruction per week. Briggs (2015) investigated the most frequent types of out-of-class interactions during SAE reported by 241 participants, and examined how these interactions related to vocabulary gains. Briggs found that interactions did not contribute to learners' vocabulary improvement, which led her to conclude that out-of-class interaction may provide limited opportunities for learners to acquire novel words.

There have also been inconsistent findings regarding SAE and grammatical knowledge development. Grey et al. (2015) administered grammaticality judgment tasks to advanced L2 learners before and after their five-week SAE. Significant improvements were observed, but were limited to judgments of word order and number agreement, with no changes in the accuracy of gender agreement judgments. Grey et al. explained that number agreement is less cognitively demanding than gender agreement during language acquisition and processing. This explanation has been supported by numerous studies (e.g., Antón-Méndez et al., 2002; Dowens et al., 2010; Rossi et al., 2014; Sagarra & Herschensohn, 2010). Also, the SA length in Grey et al.'s research only lasted five weeks which was relatively short. Thus, the lack of SA benefits in this regard is unsurprising (Sagarra & Labrozzi, 2018). Similar results were later reported by Faretta-Stutenberg and Morgan-Short (2018) when the SAE lasted for 12 to 15 weeks. However, Isabelli-García (2010) found that both SA and NSA intermediate learners, after four months of learning in different language contexts, improved in their gender agreement acquisition. Moreover, in Isabelli-García's study, NSA students demonstrated an even higher accuracy rate in the tests than SA participants.

A speculative explanation for these inconsistent results is that they may stem from the nature of grammatical knowledge. Grammatical knowledge is a type of declarative knowledge, and the acquisition of such knowledge relies largely on the conscious building of novel rules and drawing analogy between them (Ellis, 2011, p. 44). Converging evidence also demonstrates that for late L2 learners, classroom instruction, characterised by explicit explanation and modified input, is more helpful for learning new linguistic structures than language immersion (Doughty, 1991; Hilton, 2011; Kormos, 2006, p. 167; Marqués-Pascual, 2011; Sanz & Leow, 2011; Ullman, 2001). Pure immersion places high demands on cognitive resources (e.g., WM) when acquiring or generalising linguistic rules from the environment without explicit instructions (see Section 2.2). Moreover, if learners cannot understand the language that is being addressed to them, they cannot employ that input to acquire the L2 structure (Gass & Mackey, 2007). Therefore, in terms of the relationship between SAE and language competence, some researchers argue that SAE does not materially affect late L2 learners' underlying linguistic knowledge but acts on how accurately they are able to apply the knowledge they have already acquired (Wright, 2013). This argument suggests that immersion does not increase grammatical knowledge (Rothman & Iverson, 2007), but may lead to enhancement in language automaticity (Segalowitz, 2003) and in oral fluency (e.g., Collentine, 2004; B. Freed et al., 2004; Llanes & Muñoz, 2009; Marqués-Pascual, 2011; Mora & Valls-Ferrer, 2012; O'Brien et al., 2006; Pérez-Vidal & Juan-Garau, 2011; Segalowitz et al., 2004; Segalowitz & Freed, 2004; Wright, 2013, 2018). This appears to be true regardless of the proficiency level of the learner (e.g., Llanes & Muñoz, 2009; Mora & Valls-Ferrer, 2012), and regardless of their immersion length (e.g., Llanes & Muñoz, 2009; Montero et al., 2017; Segalowitz & Freed, 2004; Serrano et al., 2011).

In sum, the above literature mainly focuses on L2 lexical and grammatical knowledge development in SAE, but offers no explanation of why a learner's L2 performance is variable and inconsistent in different situations, regardless of their linguistic knowledge. An examination from a psycholinguistic processing perspective will shed light on this issue. Furthermore, language processing can also illuminate the complex ways in which SAE and cognitive resources impact language performance.

2.1.2 Lexical Processing

Psycholinguistics investigates and describes the psychological process underlying the language use. It tells how speakers/listeners employ their linguistic knowledge during the process of comprehension and production (E. M. Fernández & Cairns, 2011, p. 9; Ratner & Gleason, 2004). There is broad agreement in this field that monolingual production is the result of a highly automatised system similar to that constructed by Levelt (1989, p. 9), shown in Figure 2.1. In Levelt's model, speakers first construct a pre-verbal message in the *Conceptualiser.* This pre-verbal message is fed forward to the *Formulator* from which a search is instigated into the mental lexicon to retrieve the necessary linguistic elements (*lemmas*, in Levelt's model) to express those concepts. Lemmas are bundles of declarative knowledge about a word's meaning and grammar (Levelt, 1989, p. 236). Once the lemmas enter the *Formulator*, they are assembled into syntactic structures according to the syntactic information carried by the lemmas, and are subsequently organised for the articulation of the message itself through the *Articulator* (Levelt, 1989, p. 12; Vigliocco & Hartsuiker, 2002). During these processes, linguistic knowledge is not the only input to be operated on, since the *Formulator* also receives non-linguistic input from the *Conceptualiser* (through the pre-verbal message). This input includes the communicative intention of the speaker, the pragmatic situation (where, when, why and with whom is the speaker is speaking) and the meaning intended by the speaker. The combined linguistic and non-linguistic inputs constitute the total input to the processor, which composes the message for production (or for interpretation of somebody else's message, in comprehension). A core component of this complex process concerns lexical access, as grammatical encoding; that is, the construction of phrases, clauses, and whole sentences, depend largely on lexical processing (Bock & Levelt, 1994; Levelt, 2001).

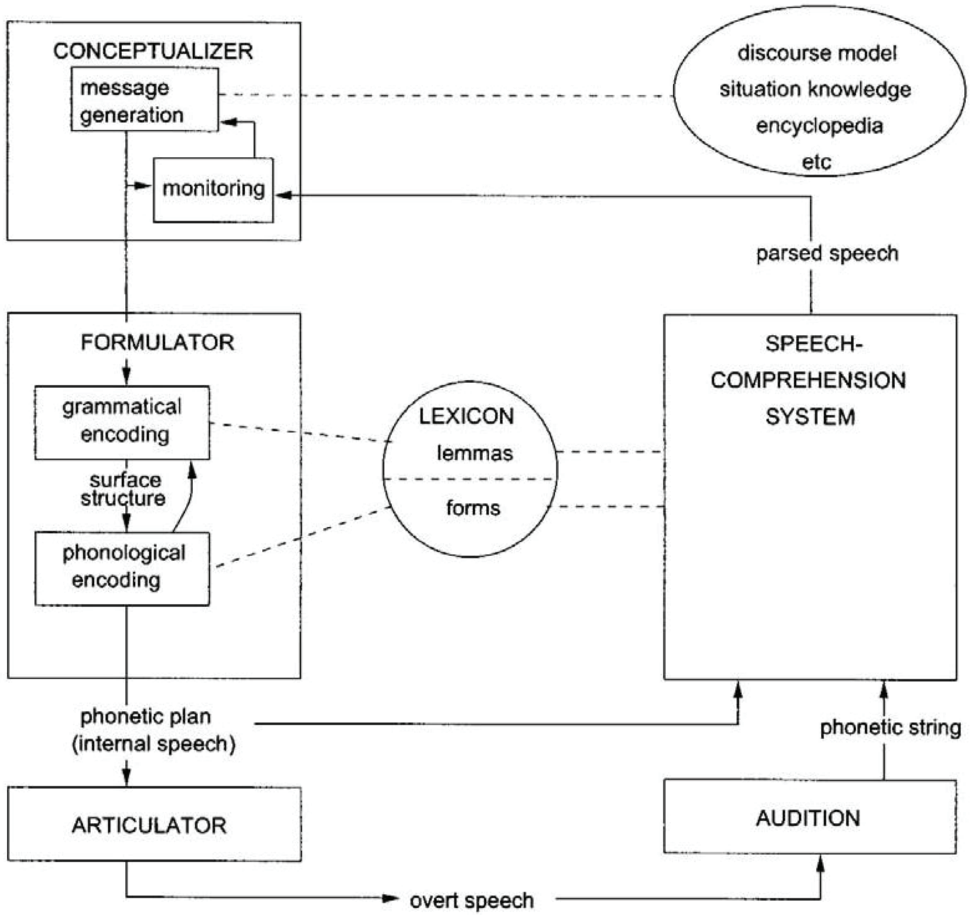


Figure 2.1. A blueprint for the speaker: Levelt's (1989, p. 9) model of speech production.

Within this broad framework, Levelt, Roelofs, and Meyer (1999) propose a lexical access model, specifically to explain the process of lexical access in language production (Figure 2.2). They argue that language production proceeds through the stages of conceptual preparation, lexical selection, morphological and phonological encoding, phonetic encoding, and then to the articulation. In parallel with these processes, self-monitoring of the speaker's internal and overt speech occurs.

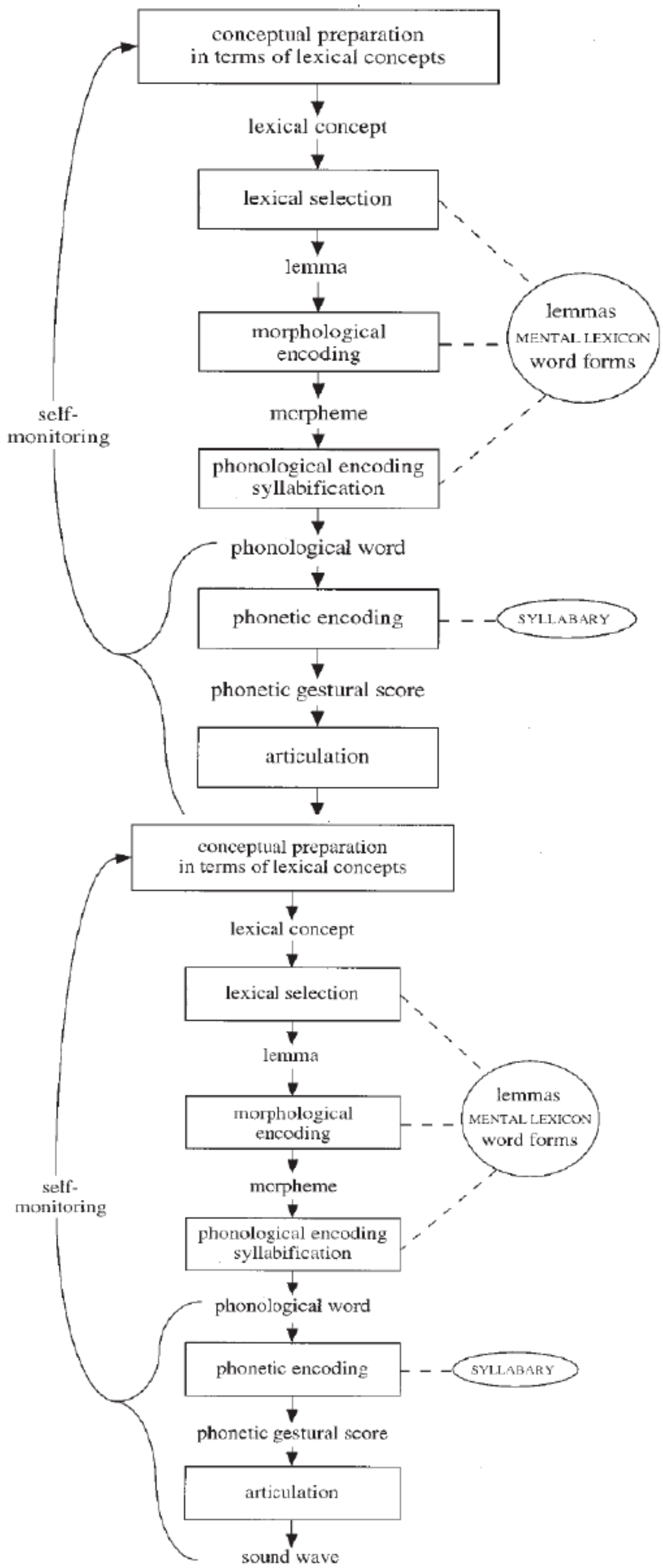


Figure 2.2. A theory of lexical access in speech production (Levelt et al., 1999).

According to this model, lexical selection refers to a word or a lemma being retrieved from the mental lexicon in order to express the intended lexical concept.

Levelt et al. claim that at this stage, lemma selection is frequency-dependent, in that it favours the most frequently activated lemma. Once a lemma is selected, its syntactic properties are simultaneously retrieved for further processing (Levelt, 2001; Levelt et al., 1999). In English, for example, verb lemmas have properties for number, person, tense, and mood, and it is obligatory for further morphological encoding to value these syntactic properties. Levelt (1989, p. 182) uses the verb *eat* as an example. The syntactic properties of *eat* include the category of the entry (verb), the syntactic arguments it can take (the verb is transitive) and so forth. The fulfilment of these syntactic conditions is achieved through the activation of certain items during grammatical encoding, including the morphological specification of the item. For *eat*, the third person present-tense is *eats*; and its past-tense inflexion is *ate*. The detailed mechanisms of grammatical encoding are discussed in Section 2.1.4. After the construction of a morphological target frame, the incrementally retrieved morphophonological codes are inserted. The composed syllables are then fed into the final stage of lexical access, that is, phonetic encoding.

These processes are different in mature adult L1 and L2 processing. On the one hand, L1 processing operates with a very high degree of automaticity and multiple forms of processing typically occur in parallel (Levelt, 1989, p. 28). Our lack of introspection about how we retrieve grammatical structures and generate the appropriate articulatory gestures to produce target words attests to this automaticity (Gathercole & Baddeley, 1993, p. 98). On the other hand, L2 processing is slower, due to its lower degree of automaticity. Less automatic processing demands more WM (cognitive) resources (see Section 2.2), which results in what is mostly perceived as nonnative-like processing (Di Biase & Kawaguchi, 2002; Gathercole & Baddeley, 1993, p. 98; Leeman, 2003; Leonard & Shea, 2017; Levelt, 1989, p. 21; McDonald, 2006; Segalowitz, 2003).

It is often assumed that in bilingual language processing, there is crosslinguistic interference between L1 and L2, both in production and comprehension (e.g., Chang, 2012; De Groot & Starreveld, 2015; Giezen & Emmorey, 2016; Jarvis & Pavlenko, 2008; Odlin, 2003, 2012; Poarch & Van Hell, 2012; Starreveld et al., 2014); the L1 may influence the L2 (Poulisse, 1999; Poulisse & Bongaerts, 1994) and the L2 may impact on the L1 as well (Brown & Gullberg, 2008; Cook, 2003; Grosjean, 2013, p.

120; Magiste, 1986). This has led some researchers to conclude that a bilingual speaker is not the sum of two monolinguals but a unique and specific speaker-hearer (Grosjean, 1992, 1997).

There is a general consensus that the bilingual mental lexicon is activated in a non-language-specific manner (e.g., Brysbaert et al., 1999; Costa et al., 2000, 2003; De Groot et al., 2000; Dijkstra et al., 1999; Dijkstra & Van Heuven, 2002, 2012; Grosjean, 2013, p. 34; Hopp, 2018; Jared & Kroll, 2001; Linck et al., 2008; Schwieter & Ferreira, 2017), which implies that an input language can non-selectively co-activate representations from both the target and non-target languages. For instance, Van Heuven, Dijkstra and Grainger (1998) found that even when only one language was required for lexical decisions, the performance of fluent Dutch-English bilinguals was influenced by the presence of orthographic neighbours in both languages. Van Heuven et al. saw the cross-language effects of lexical form as evidence that access to the lexicon is non-selective for bilinguals, and that their lexicon in both languages may be integrated at the early stages of L2 acquisition. Furthermore, in a more recent eye-tracking study, Wu, Cristino, Leek and Thierry (2013) found that non-selective language activation exists not only among bilinguals whose L1 and L2 share orthographic and phonological features, but also among bilinguals whose two languages have radically different lexical graphemes and phonological features, (e.g. Chinese and English), as long as the relevant words in the languages share overlapping semantic concepts (e.g., Moon & Jiang, 2012; Thierry & Wu, 2007; Wu et al., 2013; T. Zhang et al., 2011).

A consequence of such non-selective language activation is that lexical alternatives in the both languages of the bilingual may become available and compete for selection (e.g., Abutalebi et al., 2008; Kroll et al., 2008; Misra et al., 2012). How are lemmas retrieved in the minds of bilinguals? There are a number of proposals regarding the nature of the bilingual lexico-semantic system (e.g., Vigliocco & Vinson, 2007; Votaw, 1992). The conventional assumption is that words in each of the bilingual's languages share a common semantic code, and the semantic alternatives of both languages are available and competing for lexical selection, but they are different in the degree of activation (Kroll & Sunderman, 2003). Researchers suggest that a word's degree of concreteness impacts on the strength of semantic links between L1 and L2 (e.g., Tokowicz & Kroll, 2007; Van Hell & De Groot, 2008). Also, the number of word translation equivalents plays a role (e.g., Kroll & Tokowicz, 2005; Tokowicz & Kroll, 2007). Some studies have found that bilinguals take longer to translate words with more than one dominant translation equivalent.

The fundamental question in the study of bilinguals' linguistic processing is: How do they resolve this competition and control two language systems to serve their purpose in relation to the intended language? For example, when presented with a particular lexical stimulus, how does a bilingual translate it into the other language instead of simply reading it in the same language?

According to the view of language non-specific activation, lexical selection in the intended language is generally believed to be achieved by means of the inhibition of non-target language words (Costa, 2005; Costa & Santesteban, 2004; De Groot & Starreveld, 2015). Some researchers, such as Green (1998) and Thomas and Allport (2000) believe that this process is accomplished with the help of an external cognitive mechanism which falls outside the mental lexicon. This mechanism takes charge of regulating whichever the language to use at a given time, and prevents the non-target one from being selected as outlined in the Inhibitory Control model (Figure 2.3) (Green, 1998).

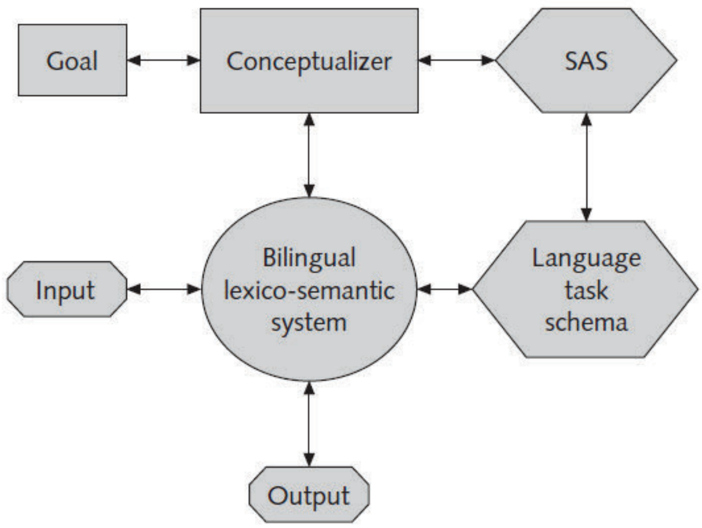


Figure 2.3. Inhibitory Control model.

The *Conceptualiser* in Green's model has the same function as Levelt's (1989), in that it is responsible for building the meaning of a message based on the speaker's communicative goal. This mechanism is assumed to be independent of language; that is, bilinguals can express the same idea with different words, both within (i.e., by using synonyms) and across languages (i.e., by using translation equivalents). The Inhibitory Control model further assumes that the mental lexicons of both languages are situated in the *Bilingual lexico-semantic system*, and each lemma has a language tag indicating which language it belongs to. In line with Levelt et al.'s (1999) monolingual lexical access proposal, each lexical item is associated with its corresponding lemma, which has syntactic properties for further higher-level construction attached. Green's Inhibitory Control model thus posits that the *language task schema* modulates which task the bilingual needs to perform at a given time (Most relevant to the present thesis are tasks assessing word recognition and word translation). A word recognition task involves one language only, which requires participants to decide whether a string of letters is a real word in a designated language or not. Experimental word translation task, on the other hand, involves two languages. In the task, a word is shown in one language followed by another word in the other language, and then participants are asked to decide whether the second word is the correct translation of the first one or not. Process of deciding which task to perform is controlled by the *Supervisory Attentional System* (*SAS*), a cognitive mechanism which plays a role resembling that of the *central executive* in the WM framework (Baddeley, 2006) (discussed in Section 2.2).

Green's model then offers an explanation for direction-dependent translational asymmetry, that is, for why translation is slower and error-prone from L1 to L2 than in the opposite direction (Green, 1998; Kroll & Stewart, 1994). For late bilinguals, generally their dominant L1 is more active than their weaker L2. Thus, the imbalance between the two languages results in the differential inhibitory efforts demanded for the two translational directions (Christoffels et al., 2007; Costa & Santesteban, 2004; Meuter & Allport, 1999; Peeters et al., 2014). Translation from L1 to L2 involves inhibiting the dominant L1, which is effortful, whereas the L2 to L1 direction entails suppressing the weaker L2, and thus is easier to accomplish. This position is shared by several researchers (e.g., Bultena et al., 2015; Kroll et al., 2002; Kroll & Sunderman, 2003; Peeters et al., 2014).

An essential feature of Green's (1998) model is that the inhibited lemmas can be re-activated. According to the model, once the dominant L1 is inhibited, it will take more cognitive resources to be re-activated. Numerous studies confirm the supposition that switches from the weaker L2 back to the dominant L1 take longer and are more error-prone than switches in the opposite direction (e.g., Costa, 2005; Costa & Santesteban, 2004; Kroll & Dijkstra, 2002; Kroll & Sunderman, 2003; Kroll & Tokowicz, 2005; Meuter & Allport, 1999; Olson, 2017; Peeters et al., 2014; Tokowicz, 2014, p. 15). For instance, Peeters et al. (2014) observed robust asymmetrical switching costs in a picture naming task when participants shifted into their L1. Likewise, in an often-cited number naming study, Meuter and Allport (1999) found that language-switching costs were consistently larger when participants switched back into their dominant L1 from their L2 than in the opposite direction. Meuter and Allport suggested that it was the relative strength of the two languages that determined the switching costs, and bilinguals who were more balanced would exhibit less asymmetry in their switching costs.

Some researchers support this assumption (e.g., Costa, 2005; Costa & Santesteban, 2004). For instance, Costa and Santesteban (2004) administered a naming task to two groups of participants with varying L2 proficiency. An asymmetry was only found among the less-proficient participants, replicating Meuter and Allport (1999), and no asymmetry was observed for the more-proficient bilinguals. However, some challenge this interpretation. In a picture naming task, Christoffels, Firk and Schiller (2007) did not observe a switching cost asymmetry in intermediate bilinguals. But what was special about these bilinguals was that they frequently switched between their two languages in daily life. Christoffels et al. therefore suggested that daily language switching habits, in addition to language proficiency, may be an important factor influencing inhibitory control and switching costs.

Some researchers have also explored the effect of language contexts on inhibitory control. They examined whether, in the presence of cues to the L2, the relative activation of competitors in the dominant L1 can be attenuated, thereby reducing the inhibitory control requirements for that language.

In investigating the influence of language conditions on language learning results, Kroll et al. (1998) measured monolinguals' L1-L2 word translation performance after they learnt novel L2 words in two opposing language conditions. In the first condition, participants associated new L2 words with their L1 translation equivalents or with pictures. In the other condition, words were paired with upside-down pictures to inhibit participants' L1-picture associations, thereby facilitating L2-picture associations. Participants were faster when translating the L1 word into the L2 equivalent that they had learnt with upside-down pictures, confirming that L2 retrieval benefited when the L1 was inhibited during L2 processing (Kroll & Sunderman, 2003). This also suggests that SAE, which is characterised by an abundant activation of L2 stimuli and constant attenuation of dominant L1 (e.g., Baus et al., 2013; Linck et al., 2009; Sagarra & Labrozzi, 2018), may be favorable to L2 processing. Similar results have been observed in study abroad research.

Linck et al. (2009) compared two groups of L2 learners matched for L2 proficiency and WM span. However, one group had a three-month SAE, whereas the other had none. The results showed that the L2 context helped learners achieve a deeper L2 processing. In addition, participants in the L2 context were less affected by interference from L1 neighbour distractors. These findings indicated that the L2 environment effectively inhibited learners' L1, and thereby facilitated their processing of words in the target L2. By contrast, participants in the dominant L1 environment had to exert greater efforts to overcome the external L1 cues in inhibiting the L1 when processing in their L2.

SA learners were also reported to be more efficient in L2 lexical processing than their NSA peers. Antoniou, Wong and Wang (2015), for instance, recruited two groups of Mandarin participants. One group had been living in the United States for 6.5 years on average, and the other group of participants were living in their home country China. Participants with SAE were faster and more accurate in spotting target L2 words from a sequence of spoken words than their NSA counterparts, indicating that the efficiency of a non-native speech processing can be improved by intensified and prolonged exposure to that language within an immersion context. Similar results were also found in a longitudinal study: Grey et al. (2015) observed a positive correlation between participants' SAE and their word recognition performance, as evidenced by increased accuracy and decreased reaction times after a short SAE of merely five weeks.

In contrast, Segalowitz, Freed, Collentine, Lafford, Lazar and Díaz-Campos (2004) observed no significant effect from language contexts on participants' L2 lexical processing performance after a 13-week SAE. More complex results were reported by Sunderman and Kroll (2009). In their study, participants' SAE was 3.8 months on average. Their results showed that SAE alone was sufficient for enhancing participants' lexical comprehension but not for promoting their lexical production performance. Similar mixed results were reported by Tokowicz, Michael and Kroll (2004), who found SAE alone did not improve lexical processing accuracy, even though the participants in this study had eight months of SAE on average. In explaining the complex effect of SAE on learners' lexical processing performance, the researchers in both studies assumed that other factors such as participants' WM resources played an influential role in their processing performance. They suggested that the learning context may help promote language processing but is not sufficient to result in a significant improvement. Interaction of WM and bilingual processing, including inhibitory control, is dealt with in greater detail in Section 2.2.

Another interesting finding reported by Tokowicz et al. (2004) was that learners in different language contexts made different types of errors. They found that the SA group in their study tended to give answers during translation tasks even though they sometimes did not know the precise translation equivalents. When facing the same situation, participants without SAE were inclined to be more conservative and gave non-responses. This result replicated that of DeKeyser (1991) who suggested SAE may encourage individuals to use an alternative word or phrase with a similar meaning when they do not know the correct word. Tokowicz et al. (2004) interpreted their result as evidence that SAE may increase participants' desire to communicate despite inaccuracy; thus, SA participants may set a lower threshold for translation selection and allow less precise translation to be produced.

In sum, the process of lexical access is complex in monolingual production, and is even more complex in the bilingual domain, because the language has to be selected. Bilingual production involves the activation of both languages and the inhibitory control of lexical items in the unrequired language. Since bilinguals predominantly have unbalanced proficiency in their two languages, there are asymmetric switching costs when bilinguals undertake linguistic tasks that involve language-switching.

Factors such as relative linguistic strengths, L2 proficiency and their daily language switching habits are all considered to be important in influencing the degree of asymmetry. Moreover, researchers have found that SAE has a significant effect on bilinguals' language modulation and L2 processing. Lexical processing is critical in speech comprehension and production on a syntactic level, as its syntactic properties are also triggered for later grammatical encoding when the lemma is selected. Section 2.1.3, focuses on bilingual grammatical processing and its correlation with SAE.

2.1.3 Grammatical Processing

Grammatical encoding and lexical processing are closely linked in both comprehension and production (Borovsky et al., 2013; Elman, 2009; Fedorenko et al., 2012; Konopka & Meyer, 2014). Upon retrieval of a lemma, its meaning and grammatical properties become available, and the appropriate syntactic environment is also established for the word (Levelt, 1989, p. 236; Levelt et al., 1999).

Even though the categorisation may vary, linguists have reached a general consensus on the processes involved in grammatical encoding. They include lemma selection, function assignment, constituent assembly and inflexional morphology (e.g., Bock & Levelt, 1994; V. Ferreira & Slevc, 2007; Levelt et al., 1999; R. C. Martin & Slevc, 2014). However, these processes are not strictly serial; that is, the processing of different parts of an utterance may occur concurrently at different stages (R. C. Martin & Slevc, 2014). Among these processes, function assignment refers to assigning syntactic relations or grammatical functions to each lemma (e.g., subject-nominative, object-dative); the constituent assembly is assumed to create a control hierarchy for phrasal constituents which manages the word order and dependencies in sentences during production; and the inflexional morphology entails the generation of fine-grained details (Bock & Levelt, 1994; Levelt et al., 1999).

Inflexional morphemes are used to indicate morphological categories such as the number of the subject or object for verbs, and this is a rule in many languages, including English (Hartsuiker & Barkhuysen, 2006). However, no such inflexional categories exist for Chinese (C. Li & Thompson, 2008). Therefore, there is no standardised translation of such categories as number or person from English into Chinese. For instance, the English word 'vehicle' in singular form can be translated as Chinese ‘chēliàng--车辆’, but ‘chēliàng’ also corresponds to the English word ‘vehicles’ in plural form which is marked by the suffix *–s*. At the morphosyntactic level, English also marks the relation between subject and verb (plural or singular) in the present tense with a morphological inflextion by adding or not adding the morpheme *-s* (Fromkin, 2012, p. 75; C. Li & Thompson, 2008), as can be appreciated in example (1a-b).

(1) a. Mary speaks French.

b. They speak French.

In the first sentence the *–s* at the end of the verb ‘speak’ is an agreement marker. It does not add lexical meaning but signifies that the subject of the verb ‘Mary’ is being referred to in the third person, and is singular and that the verb is in the present tense.

This is different from Chinese where there is no ‘agreement’ on the verb to indicate what is the subject and what is the object (C. Li & Thompson, 2008), as can be appreciated in (2a-b) which represents the Chinese version of (1a-b).

(2) a. Mǎlì shuō fǎyǔ.（玛丽说法语。）

Mary speak* French.

b. Tāmen shuō fǎyǔ.（他们说法语。）

They speak French.

Chinese presents the subject-verb relationship ordinarily (subject comes before the verb in default structures), or by means of prosody, or function words, instead of through morphological inflexions (C. Li & Thompson, 2008; Wiedenhof, 2015, p. 118). In this Chinese example, the grammatical person and number of the subject ‘Mary’ does not affect the verbal morphology: there is only one morphological form in Chinese verbs. In other words, Chinese native speakers do not have the habit of attending to features of the subject when they are selecting a verbal form for a sentence.

If Chinese speakers transfer their L1 processing habits to processing English, which is always the case for late L2 learners, then this crosslinguistic difference is inclined to result in errors like subject-verb number disagreement, an error made by many Chinese native speakers when processing English (e.g., Chen et al., 2007; H. Liu et al., 1992). Also, such a transfer of an inappropriate habit between languages could be a barrier to acquiring full native-likeness in the L2 (Clahsen & Felser, 2006; Frenck-Mestre & Pynte, 1997; Juffs, 1998).

Developing a new linguistic habit takes continual practice and repetition; thus, SAE, characterised by varied and plentiful L2 input and practice opportunities, has attracted researchers' attention. Numerous studies have explored the effects of SAE on grammatical accuracy in L2 learners and have generated positive findings (e.g., Howard, 2005; Isabelli & Nishida, 2005; Isabelli-García, 2010; Linck et al., 2009; Mora & Valls-Ferrer, 2012; Pérez-Vidal & Juan-Garau, 2011; Sagarra & Labrozzi, 2018; Song, 2015).

In empirical studies, grammatical processing assessments are often associated with measuring the learner's sensitivity to grammatical violations during reading, or calculating their grammatical errors in speech production (e.g., Ellis & Sagarra, 2010; Isabelli & Nishida, 2005; Juan-Garau, 2014; Labrozzi, 2009; Linck et al., 2009; Sagarra & Labrozzi, 2018; Song, 2015). For instance, using eye-tracking, Song (2015) compared the plural-inflexion sensitivity of a group of native English speakers to a group of Korean advanced English learners who were living in an English-speaking country. Korean, like Chinese, does not mark number information morphologically on the noun and/or its dependants (Liter et al., 2017). In that study, those English learners with SAE demonstrated native-like grammatical sensitivity by slowing down their reading pace like native English speakers did when they were presented with inflexional violation sentences. Similar improvements in grammatical sensitivity during SAE were reported by Sagarra and LaBrozzi (2018) and LaBrozzi (2009). Both studies made a comparison between participants with and without SAE. Eye-tracking revealed different processing patterns between the two groups, indicating that NSA learners continued to use lexical cues to assign temporal reference, whereas SAE learners began to turn to morphological cues, like native speakers, when an adverb-verb incongruence occurred in a sentence. The researchers concluded that the SA group was experiencing a transitional stage in progress towards 'native-likeness'. These combined results suggest that SAE is able to refine the grammatical processing of L2 learners (e.g., Bardovi-Harlig, 1992; Bordag & Pechmann, 2007; Duperron, 2006; Ellis & Sagarra, 2010; Gor & Long, 2009; Guntermann, 1995; Isabelli & Nishida, 2005; Parodi et al., 2004; Yager, 1998).

There is evidence that language comprehension is easier to be improved than production (M. Muñoz & Marquardt, 2003; Paradis, 2004), as comprehension consists of negative activation of phonological and/or lexical cues; whereas, in the case of production, initiatively accurate recall is required (Atkins & Baddeley, 1998). Indeed, this argument is supported by findings which indicate that in L2 speech production, grammatical processing is much more inconsistent than it is in L2 speech comprehension. Some studies have revealed salient SAE benefits compared to classroom instructions (Pliatsikas, 2010). Howard (2005), for example, reported that participants with SAE were capable of marking past tense with the correct forms to a greater extent in a speech-eliciting test than those without language immersion. Likewise, Isabelli and Nishida (2005) and Pérez-Vidal and Juan-Garau (2011) also observed superior grammatical performance and grammatical complexity among SA learners in the use of the subjunctive mood in Spanish. Moreover, SAE benefits may emerge relatively quickly, even in the case of short-term SAE of three to four weeks (Llanes & Muñoz, 2009).

In contrast to the above results, some other researchers suggest that SAE is not superior to classroom instruction in terms of grammatical processing (e.g., Arnett, 2013; DeKeyser, 2010; Isabelli-García, 2010; Marqués-Pascual, 2011; Rothman & Iverson, 2007; Segalowitz et al., 2004). Rothman and Iverson (2007) conducted a longitudinal study of a group of L2 learners before and after their SAE. The findings of the study suggested that two-thirds of participants had acquired the target grammatical structures before their SAE, which indicated that although exposure to naturalistic input is invaluable, this type of exposure is not necessary or exceptionally beneficial for attaining a command of the target grammatical structure. Rothman and Iverson argued that classroom instruction was grammatically oriented, and contained the necessary grammatical features for L2 learners to acquire the target grammatical structure. However, what was lacking in Rothman and Iverson's study was a classroom comparison group; thus, it is unknown whether, after acquisition, SAE is more effective than classroom instructions in promoting L2 grammatical processing.

A convincing body of evidence suggests that a participant's state of grammatical readiness at the time of commencing their study abroad is critical for their grammatical gains during that study abroad period. Empirical studies show that SAE does not substantially improve learners' underlying grammatical knowledge, but is beneficial for proceduralising the already learnt grammatical knowledge (e.g., Arnett, 2013; Collentine, 2004; Marqués-Pascual, 2011; Wright, 2013). This is because procedural knowledge is highly dependent on the frequency of stimulus occurrence (Hilton, 2011, p. 146), and SAE is assumed to provide rich L2 stimulation. Therefore, in exploring the effect of SAE on grammatical accuracy, it is necessary to evaluate whether participants have acquired the target grammatical structure or not before measuring their accuracy. Rarely have previous studies done so, and perhaps this is an alternative explanation for their inconsistent results. Moreover, DeKeyser (2010) reports that learners who have already acquired the linguistic knowledge before exposure are more likely to be happy to engage in communicative interactions with native speakers than those who have not acquired the target knowledge, and it is this engagement that makes it possible for them to put what they have already learnt into practice, and learn from the interactions and feedback. But before the proceduralisation of the declarative knowledge (Section 2.1.5), the role of WM is critical (Section 2.2). This is because learners have to draw on their WM whenever they consciously construct an utterance based on these grammatical rules (Section 2.2.4).

Generally, the literature reviewed above suggests that studies examining the effect of SAE on grammatical processing has produced varying results. In some instances, SAE has been found to benefit grammatical processing, whereas in others no difference has been reported between learning contexts, and sometimes domestic classroom environments have been found to produce better results than SAE. The benefits of SAE on L2 learners' grammatical processing are often theoretically assumed rather than empirically established, which provides a window into the complex and multidimensional nature of grammatical processing. Some researchers suggest that SAE can help learners to be better at processing grammatical structures that have already been acquired. Therefore, besides SAE, factors such as learners' grammatical readiness may also be influential. Other researchers have emphasised the important role of learners' WM resources in processing grammatical structures. Thus, factors such as

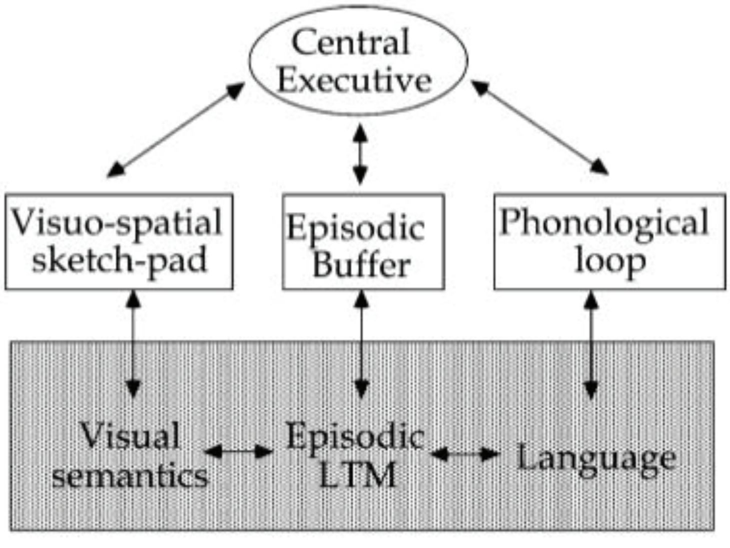
WM should also be considered before any strong conclusions concerning the relationship between grammatical processing and language environment can be drawn.

2.2 Working Memory, Study Abroad Experience and Second Language Acquisition

2.2.1 Working Memory

Working memory (WM) was first formalised by Miller, Galanter and Pribram (1962) to differentiate it from short-term memory. The latter of which is primarily concerned with storage. WM, on the other hand, involves the temporary storage and manipulation of information at the same time (e.g., Baddeley, 1998, 2003, 2006; De Abreu, 2011; De Abreu & Gathercole, 2012), and it is crucial for complex tasks (Cowan, 2014).

The most influential model of WM is the multi-component model, proposed by Baddeley and Hitch (1974). As the name suggests, WM is thought to be composed of several inter-related components (Figure 2.4).



Shaded area: 'crystallised cognitive systems capable of accumulating long-term knowledge'

Figure 2.4. Development of the WM multi-component model (Baddeley, 2003).

The *central executive* is an attentional controller (Baddeley, 1996, 2010) in charge of selective attention, inhibition of irrelevant or distracting stimuli, allocating attention, and coordination of multiple tasks (Baddeley, 1996, 1998, 2006, 2010). The *visuo-*

spatial sketchpad stores and manipulates visual and spatial information (Baddeley, 2009). The *phonological loop* stores and processes verbal information. Thus, it is crucial for language acquisition and processing (Baddeley, 2009; Baddeley et al., 1998; Miyake & Friedman, 1998), and is of great importance to the present study. Functions subserved by the phonological loop include performing sub-vocal rehearsal to offset the effects of memory decay, and recoding written input into a phonological form that can be retained in the phonological store (Norris et al., 2018). The *episodic buffer* is a limited storage system that can hold information in a multidimensional code, such as visual, verbal and semantic codes from perception and long-term memory. These codes can be combined, and they can interact with long-term memory (LTM) within the episodic buffer (Baddeley, 2009).

Previous literature has demonstrated that WM is one of the most critical components of cognitive achievement, and serves as an excellent predictor of performance in many complex cognitive tasks, including reasoning (Ardila, 2003; Kyllonen & Christal, 1990), spatial ability (Miyake et al., 2001), general fluid intelligence (Engle et al., 1999) and language processing (Ardila, 2003; Daneman & Merikle, 1996; Miyake & Friedman, 1998). Tasks such as shifting between information, updating stimuli and planning also rely critically on WM (Miyake & Shah, 1999). Even though WM plays such a significant role in many daily tasks, its availability for processing and storage functions is limited (Baddeley, 2009; Barrouillet et al., 2007; Gathercole & Baddeley, 1993, p. 223; Gile, 2009; Hudjetz & Oberauer, 2007; Miller, 1994; F. Wang, 2017). Given that both processing and storage of relevant information are assumed to draw on the same resource supply, a shortage of available resources could lead to a negative impact on task performance, namely a trade-off effect (Baddeley & Hitch, 1974; Barrouillet et al., 2007; Gathercole & Baddeley, 1993, p. 223; Marcel A Just & Carpenter, 1993; Mead, 2005). Performance deteriorates when cognitive load increases, which leads to a gradual loss of necessary information or a processing slowdown (Anderson et al., 1996; Barrouillet et al., 2007; Carpenter et al., 1995; Case et al., 1982; Conway & Engle, 1994; Cowan, 2008; Daneman & Carpenter, 1980; Marcel Adam Just & Carpenter, 1992; Labrozzi, 2009; M. Liu, 2001; Robinson, 2003). Put simply, when the processing demands exceed the cognitive resources, the processing is both slow and error prone.

With respect to language, resources may be allocated to specific linguistic aspects such as interlanguage in L1, and phonology, syntax, morphology etc. in L2 (Kormos, 2006; Robinson, 2003). Further, individual differences in WM resource availability or the efficiency with which cognitive processes are executed are assumed to influence how trade-offs occur during their processing and the ultimate result of the language tasks (Gathercole & Baddeley, 1993, p. 222; M. Liu, 2001).

2.2.2 Working Memory Tests

There are three broad categories of WM tasks, and each relies on different stimuli such as verbal, numerical or visual-spatial reasoning or some combination of two of the above to evaluate WM resource availability.

Recall N-back tasks require participants to store and make judgments about the identity of a stimulus presented *n* trials previously. The task involves encoding and the temporary storage of each stimulus *n* of the stimulus sequence in WM and a continuous updating of incoming stimuli.

Binding tasks test participants' ability to form and maintain associations between the features of stimuli in the WM.

Complex span tasks require participants to store and manipulate information simultaneously. From this point of view, a span task is a measure of the availability of WM resources that can be flexibly deployed to information processing and/or storage. During the presentation of to-be-remembered stimuli, participants are also required to complete a relatively simple secondary task. The point of the secondary task is to engage their attention and, therefore, disrupt active refresh of the to-be-remembered stimuli, and this process constitutes a trade-off between the processing and storage of WM resources (Daneman & Carpenter, 1980). Performance on complex span tasks indicates the amount of WM resources available after the processing requirements of the tasks are met (Padilla et al., 2005), and this understanding is compatible to the WM model proposed by Baddeley and Hitch (1974).

Importantly, these three categories of WM tasks (recall N-back, binding and complex span) produce scores that are highly correlated, resulting in good convergent validity (Michael & Gollan, 2005; Tokowicz, 2014, p. 64; Turner & Engle, 1989;

Wilhelm et al., 2013). This suggests that individual differences in WM resources are independent of the task. Of the three categories of WM, the complex span task is the best-known and most frequently used category for measuring WM (Wilhelm et al., 2013). The complex span task has also been widely used in linguistic studies to measure learners' verbal WM.

Many complex span tasks have been developed. These include reading span (Christoffels et al., 2006; Daneman & Carpenter, 1980; Signorelli & Obler, 2012), listening span (Daneman & Carpenter, 1980; Kopke & Nespoulous, 2006; M. Liu et al., 2004), operation span (Turner & Engle, 1989), and counting span tasks (Case et al., 1982). All of these tasks are similar in their basic structure but are different in the type of stimuli that are required to recall (digits, letters, words, spatial locations). These tasks also vary in the type of secondary task that is involved to engage participants' attention and disrupt their memory refresh (e.g., counting the number of objects in an array, solving simple math problems, judging whether a figure is symmetrical or not, or reading sentences aloud). In most cases, the number of stimuli presented for recall increases as the task progresses.

Of the aforementioned complex span tasks, the reading span task is believed to be the most appropriate type of WM test for L2 processing research (Juffs & Harrington, 2011), and has been used in many studies of L2 processing. Subjects read aloud successive sentences in a set and make decisions about the semantic plausibility of each sentence while simultaneously remembering the final word of each sentence. For example, in Harrington and Sawyer (1992), a two-sentence set (3a-b) required subjects to read aloud the following two sentences:

(3) a. He played baseball all day at the park and got a sore <u>arm</u>. (makes sense? Yes)

b. The clerk in the department store put the presents in a <u>toilet</u>. (makes sense? No)

At the end of the set, the subject must recall each of the final words in the sentences within that set (e.g., *arm, toilet*). The number of sentences in a set—and thus the number of sentence-final words to be remembered—gradually increases, placing greater cognitive demands on subjects. The reason subjects are asked to read aloud is based on the operational nature of WM. According to the dominant WM model proposed by Baddeley (2006, 2010), the *phonological loop* has a limited phonological store, and individuals tend to extend the retention period of verbal information in the phonological loop by rehearsing the information (either aloud or silently), and this process is called articulatory rehearsal (Baddeley, 1998, 2006). Asking subjects to read aloud disrupts the articulatory rehearsal process.

The number of correctly recalled words is deemed to represent the subject's WM resource availability, either in terms of the maximum set size of all or a portion of correctly recalled sentence-final words (Daneman & Carpenter, 1980; Marcel Adam Just & Carpenter, 1992), or in terms of the absolute number of final words recalled (Turner & Engle, 1989). Numerous studies have found a high degree of correlation between reading span and overall reading comprehension, and between reading span and specific reading skills (Daneman & Carpenter, 1980; Daneman & Green, 1986; Daneman & Merikle, 1996; Friedman & Miyake, 2004; Harrington & Sawyer, 1992; Turner & Engle, 1989).

2.2.3 Working Memory, Study Abroad Experience and Second Language Acquisition

The success with which individuals acquire an L2 varies greatly, and individual differences in cognitive abilities have been studied as a potential source of this variability (Tokowicz, 2014, p. 63). The assumption is that because L2 acquisition and processing encompass an integration of a new set of linguistic representations and rules into an already-existing L1 system, and because learners are expected to overcome the crosslinguistic interference and establish a new linguistic system, there must be some cognitive mechanism at work during this challenging and cognitively demanding process (e.g., Brooks et al., 2006; Harrington & Sawyer, 1992; Michael & Gollan, 2005; Miyake & Friedman, 1998; Trude & Tokowicz, 2011).

Robust WM effects (executive and storing functions) have been found across a range of complex linguistic tasks (Linck et al., 2014). WM has consistently been reported to be positively associated with a variety of L2 phenomena (Park et al., 2020), including L2 knowledge acquisition (e.g., Linck et al., 2014; Linck & Weiss, 2011; K. Martin & Ellis, 2012; Miyake & Friedman, 1998; Sagarra, 2017), linguistic skills (e.g., Ardila, 2003; Kyllonen & Christal, 1990; Miyake & Friedman, 1998; Trude &

Tokowicz, 2011), and L2 processing (comprehension and production) (e.g., Collentine & Freed, 2004; Miyake & Friedman, 1998). Researchers also suggest that large WM resource availability is usually accompanied by rapid L2 development (e.g., Juffs & Harrington, 2011; Linck et al., 2014; Linck & Weiss, 2011; Miyake & Friedman, 1998). Take vocabulary learning as an example. According to the Inhibitory Control model (Green, 1998; see Section 2.1.3), successful inhibition of language interference facilitates identification and acquisition of novel words (Bartolotti et al., 2011), and this function is subserved by the *Supervisory Attentional System*, which equates to the *central executive* in Baddeley's WM framework. Moreover, when L2 learners encounter unfamiliar words, the *central executive* in the WM system allows them to infer the meaning of the words from their context by recalling previous information and making new associations (Daneman & Carpenter, 1983; Daneman & Green, 1986; Gathercole & Baddeley, 1993, p. 73).

Few studies have explored the relationship between WM and SLA in different language environment, and those that have have generally focused on grammatical acquisition (e.g., Brooks et al., 2006; French & O'Brien, 2008; Linck & Weiss, 2011; K. Martin & Ellis, 2012; O'Brien et al., 2006; Sanz et al., 2016; Sunderman & Kroll, 2009; J. Williams & Lovatt, 2003). Studies reveal that WM is salient when learners are required to generalise the grammatical rules in an L2 immersion context without any explicit instruction (e.g., Brooks et al., 2006; K. Martin & Ellis, 2012; Sanz et al., 2016; Sunderman & Kroll, 2009); however, when an explicit explanation is provided, the advantage of high WM availability disappears (Sanz et al., 2016). Such a correlation is evident in laboratory settings where the intensity and quality of input L2 can be manipulated by researchers (K. Martin & Ellis, 2012; Sanz et al., 2016).

2.2.4 Working Memory, Study Abroad Experience and Linguistic Processing

A growing number of empirical studies have explored the executive and storing functions of WM in linguistic processing when learners are applying their L2 knowledge into concrete situations (e.g., Abutalebi et al., 2008; Hernandez & Meschyan, 2006; Michael & Gollan, 2005; Tokowicz et al., 2004).

The bilingual processing literature reviewed in Section 2.1 revealed that lexical candidates from both languages are activated in parallel and they compete for selection (e.g., Brysbaert et al., 1999; Costa et al., 2000, 2003; Dijkstra et al., 1999; Dijkstra & Van Heuven, 2002, 2012; Grosjean, 2013, p. 34; Hopp, 2018; Jared & Kroll, 2001; Linck et al., 2008). To reconcile this competition, bilinguals have to inhibit the activation of the unintended language (Green, 1998). It has been suggested that WM plays an essential role in language inhibition (Michael & Gollan, 2005). If we accept this premise, then L2 learners with higher WM resource availability may have an advantage in lexical and higher-order language processing (Gathercole & Baddeley, 1993, p. 222; Tokowicz, 2014).

This account of WM influencing lexical processing is supported by an increasing number of empirical studies (e.g., Kroll et al., 2002; Linck et al., 2008; Michael et al., 2002, 2003; Sunderman & Kroll, 2009; Trude & Tokowicz, 2011). Linck et al. (2008) report that bilinguals who had larger WM resource availability were better able to suppress the activation of lexical competitors from the non-target language during L2 picture naming. Compatibly, Michael, Tokowicz and Kroll (2003) also observed that participants with higher WM were faster and more accurate when translating words in both directions than their lower WM counterparts.

The relationship between WM and lexical processing becomes more complicated when we consider changes in language environment. Sunderman and Kroll (2009) examined the role of WM and SAE in L2 learners' lexical processing. Participants' performance on word comprehension were measured via a translation recognition task, and production was evaluated via a picture-naming task. Sunderman and Kroll proposed three theoretical hypotheses for how WM and SAE could relate to L2 processing. First, the internal resources hypothesis states that there is a general effect of WM in that the greater the internal resources of the learner, the faster and more accurate their processing will be. This hypothesis focuses on the WM of the learner rather than the language environment. Second, the external cue hypothesis suggests that the learner's language environment (e.g., classroom vs. study abroad) will predict performance, such that participants with SAE will exhibit a processing advantage. This hypothesis, therefore, focuses on the language environment rather than the learner's WM. Third, the interaction hypothesis suggests that both factors matter, and that they interact to influence processing. In Sunderman and Kroll's (2009), the results from the translation recognition task supported the internal resource and external cue hypotheses, because both WM and SAE made independent contributions to participants' L2 lexical comprehension. The results from the picture naming task supported the interaction hypothesis, which suggests that WM resources or SAE alone are not sufficient to affect L2 lexical production and the two factors function in unison.

Lexical processing serves as a core component of higher-order processing such as language comprehension and production (Bock & Levelt, 1994; Levelt, 2001), and if lexical processing becomes easier, WM resources can be freed for higher-order structures such as syntactic and morphological processing (O'Brien et al., 2006). For instance, to correctly comprehend a 'garden-path' sentence (A typical example is 'The horse raced past the barn fell'), the listener needs to be able to successfully retrieve the beginning of the sentence from their WM for information rehearsal after they have listened to the whole sentence. It is also assumed that in production, bilinguals have to produce new words while referring to some of the previously uttered words to ascertain their grammatical relationship in the sentence in order to make the upcoming parts of the sentence grammatically consistent (e.g., regular past tense, third-person agreement) (Ellis, 2011, p. 44; Hartsuiker & Barkhuysen, 2006; McDonald, 2006). All these tasks are done by maintaining verbal information in WM (Obler, 2012).

In subject-verb number agreement, for instance, the verb of the sentence is often separated from its subject by several intervening words or phrases, which requires speakers to retrieve or maintain the number of the subject while formulating the rest of the sentence (Badecker & Kuminiak, 2007; Hartsuiker & Barkhuysen, 2006; R. C. Martin & Slevc, 2014). In normal situation, this process may be effortless for L1 grammatical processing, which is generally not available to conscious access and is largely implicit (e.g., Fodor, 2008; Pliatsikas, 2010; Ullman, 2001). However, L2 processing is WM-demanding, and learners have to constantly draw on their declarative knowledge of L2 morphosyntactic rules (e.g., Hilton, 2011, p. 146; Kormos, 2006, p. 167; Sanz & Leow, 2011; Ullman, 2001). Therefore, WM relates to the ability to keep relevant information active and enables leaners to efficiently form a grammatically correct utterance (e.g., Badecker & Kuminiak, 2007; De Abreu & Gathercole, 2012; Hartsuiker & Barkhuysen, 2006; Jin, 2010; Kellogg, 2004). Additionally, Moreno, Bialystok, Wodniecka, and Alain (2010) reported that bilinguals depend on their executive function of WM when detecting grammatical violations during language processing.

There is a small, but growing body of work showing that WM is involved in L2 agreement processing (e.g., Hartsuiker & Barkhuysen, 2006; Hopp, 2010; Keating, 2009; Sagarra & Herschensohn, 2010; Sagarra & Labrozzi, 2018). Most of the work on agreement processing and WM has focused on participants' sensitivity to agreement violations (R. C. Martin & Slevc, 2014). In an L2 classroom setting research, Sagarra and Herschensohn (2010) investigated the role of WM in the processing of gender and number agreement by asking beginning and intermediate adult L2 learners of Spanish and Spanish monolinguals to complete a self-paced reading (performance) and a grammaticality judgment task (competence). These tasks measured participants' sensitivity to Spanish noun–adjective gender/number disagreement. All participants were found to be capable of making highly accurate grammatical judgements; however, only the intermediate L2 learners and Spanish monolinguals were sensitive to gender and number violations in their self-paced reading tasks. Additionally, intermediate learners with higher WM demonstrated higher accuracy in L2 comprehension questions than lower WM participants. The researchers interpreted the findings as evidence that adult L2 learners are capable of achieving native-like grammatical processing patterns as long as they possess sufficient language knowledge and WM availability.

However, not all studies support the existence of a positive link between WM and language processing (e.g., Chun & Payne, 2004; Park et al., 2020). For instance, Park et al. (2020) only observed marginal correlations between WM and elicited imitation performance. Elicited imitation requires participants to repeat a set of sentences as accurately as possible after listening. Park et al. (2020) argued that the reason for detecting marginal WM effects in their study was due to participants' high language proficiency. They claimed that high L2 proficiency learners tended to rely less on WM during processing. However, their explanation is inconsistent with evidence that WM plays an increasingly large role in L2 processing as L2 proficiency increases (e.g., Linck & Weiss, 2011; O'Brien et al., 2006). Furthermore, the lack of correlation may be due to the cognitive load imposed by the task (Gathercole, 2007). Simple repetition of recent language inputs does not require the full involvement of WM, and in order to engage WM, the task should implicate the processing and storage availability of WM (Gathercole, 2007).

Numerous studies demonstrate that even native speakers' grammatical accuracy varies as cognitive load increases (e.g., Hartsuiker & Barkhuysen, 2006; Hopp, 2010; Keating, 2009; Tanner et al., 2012). Hartsuiker and Barkhuysen (2006) explored the role of WM in keeping subject-verb number agreement in a spoken sentence completion task among a group of L1 speakers. Hartsuiker and Barkhuysen proposed two hypotheses. The resource-constrained hypothesis states that participants' production of subject-verb number agreement will be impaired when they are required to simultaneously attend to a secondary task (memorising three words); Alternatively, the resource-free hypothesis predicts that a secondary task will make no difference to participants' agreement accuracy. The researchers also predicted that low-WM participants would be more susceptible to being hindered by the secondary task, and would make more errors than their high-WM counterparts. The results supported the resource-constrained hypothesis that both WM availability and WM load affected participants' subject-verb number agreement performance in their native language. Also, cognitive load impinged on low-WM speakers only, because their inadequate WM availability hindered them when attempting to store additional words and process sentences at the same time. In other words, language processing difficulties (under relatively heavy processing burdens) seem not to be L2 learner-specific problems; they appear to be related to the cognitive demands of the task and the availability of cognitive resources.

The relationship between WM and L2 grammatical processing is relatively well documented in classroom and laboratory settings (Juffs & Harrington, 2011; Linck et al., 2014), but studies exploring the effect of WM and L2 grammatical processing in study abroad settings have been rare, and have yielded inconsistent results (e.g., LaBrozzi, 2012; O'Brien et al., 2006; Segalowitz & Freed, 2004; Sunderman & Kroll, 2009; Tokowicz et al., 2004). On the one hand, there is evidence that SAE benefits grammatical processing and WM has no effect (Grey et al., 2015), whereas other work has found that WM contributes to grammatical processing but SAE adds no significant benefit (Wright, 2013), and yet other studies suggest that SAE and WM interact (Marqués-Pascual, 2011; Sagarra & Labrozzi, 2018; Segalowitz & Freed, 2004). These studies also suggest that immersion intensity (Sunderman & Kroll, 2009; Wright, 2013) and immersion age (Cheung, 1996; De Abreu & Gathercole, 2012; Kormos & Sáfár, 2008; Verhagen & Leseman, 2016) should be taken into account.

Grey et al. (2015) found that participants exhibited improvement in their L2 morphosyntactic processing after a five-week intensive SAE. However, correlation analysis revealed no significant relationships between these improvements and their WM. Grey et al. attributed this lack of correlation between WM and grammatical gains to the strict entrance requirements of the study abroad program in their study, speculating that the strict entrance examination had limited variability amongst students, resulting in medium to strong effects in L2 grammatical development, but potentially levelling out variation in cognitive capacity.

In contrast, Sagarra and LaBrozzi (2018) found that a higher-WM group was significantly more sensitive to morphosyntactic violations than a lower-WM one; however, participants with higher-WM and SAE demonstrated better sensitivity than those high WM participants who had no SAE. When grouping participants according to their SAE, the SA group demonstrated higher sensitivity to grammatical violations than their NSA peers, even though the two groups were comparable in their WM.

How might SAE influence WM? Some researchers suggest that owing to the constant regulation of two language systems as a result of SAE, bilinguals in this language environment are exposed to extensive practice of executive functions of language control on daily basis. The regular, habitual use of the bilingual control mechanism to reconciles the L1 and L2 competition should have cognitive benefits (Bartolotti et al., 2011; Bialystok et al., 2009; Linck et al., 2008; Ransdell et al., 2006; Xie, 2018; Xie & Dong, 2017). However, studies that have examined whether SAE improves WM resource availability have generally yielded negative results (e.g., Linck et al., 2008; Segalowitz & Freed, 2004; Sunderman & Kroll, 2009; Tokowicz et al., 2004).

Linck et al. (2008) conducted two experiments on L2 learners who varied in their L2 proficiency, and neither experiment found any correlation between participants' SAE (three months) and their WM availability. They suggested that the lack of an SAE effect may stem from the relatively short exposure length. Being immersed in an L2 context may initially induce costs to WM, but then WM may recover with the time when the bilingual becomes more experienced in controlling the two languages within the L2 environment. The authors suggested that future research should explore if an extended L2 immersion experience could actively improve WM. However, studies with extended SAE did not demonstrate significant correlation either. Sunderman and Kroll (2009) reported that participants without SAE surprisingly had a marginally higher WM than participants with overseas experience for an average length of 3.8 months. Similarly, Tokowicz et al. (2004) found no correlation between length of SAE and WM in participants with an average study abroad length of eight months.

Even though studies suggest that language environment does not differentially affect WM resources, some studies do appear to indicate that language environment does affect cognitive processing efficiency. For example, participants with SAE showed significantly faster performance in mental flexibility than a control group (e.g., Xie, 2018; Xie & Dong, 2017). Other studies have also found that SAE influences learners' usage of the underlying L2 knowledge, according to both behavioural assessments (Wright, 2013) and brain image data (Morgan-Short, 2007). Thus, an alternative possibility regarding the relationship between SAE and WM could be that rich and effective input, interaction and output in SAE contribute to native-like linguistic processing automaticity (e.g., Juan-Garau, 2014, 2018; Labrozzi, 2009; LaBrozzi, 2012; Mora & Valls-Ferrer, 2012; Sagarra & Labrozzi, 2018; Segalowitz, 2003; Serrano et al., 2011). With more elements of linguistic processing being automatised, more WM resources can be freed to engage in other elements of linguistic processing or other cognitively demanding tasks (Antoniou et al., 2015; Christoffels et al., 2003; O'Brien et al., 2006). This assumption can help explain why, in previous studies, SA groups have demonstrated better linguistic performance than NSA groups even though the two groups had comparable WM availability and language proficiency (e.g., LaBrozzi, 2012; Linck et al., 2009; Sagarra & Labrozzi, 2018), but more evidence is needed to test this assumption.

If this assumption holds true, SAE and WM are expected to have salient effects on language tasks that are demanding both linguistically and cognitively, such as interpreting. Being a special form of language processing, interpreting involves concurrently listening, analysing, and memorising the input information in one language, and orally transmitting that information in another language (Russell, 2005; Russell & Takeda, 2015). The relationships between linguistic processing, WM and interpreting will be illustrated in detail in the following section.

In summary, previous literature suggests that WM and SAE are both important for L2 learning and processing. However, it is not understood how WM affects the L2 in study abroad contexts. Some studies suggest that WM mediates the effects of language context on language development, whereas others argue that language context cancels out the effect of WM on language development. Learning context and WM may have an interactive impact on learner's performance. Section 2.3 introduces a type of complex linguistic task which pushes the limits of WM—interpreting. Section 2.3 also reviews previous literature regarding the relationship between interpreting and WM.

2.3 Effects of Working Memory and Linguistic Processing on Interpreting Performance

2.3.1 Interpreting as a Special Form of Language Use

Interpreting is claimed to be a special form of language use, which aims to facilitate communication by overcoming barriers of different languages (Pöchhacker, 2004, p. 10). Interpreting involves, in a broad sense, comprehension of the source language, information storage, processing, and production of the target language (Christoffels & De Groot, 2005; Russell, 2005). During the production phase, interpreting has the appearance of spontaneous speech (Kirchhoff, 1976), which consists of speech planning, grammatical encoding and articulation (Bygate, 2002; Levelt, 1989, p. 11), but it also has several significant features which make it unique (Gile, 2009, p. 239; Jin, 2010). The most salient and robust feature of interpreting is its language transformation. Different from ordinary conversation in which both speakers are using the same language, interpreting involves at least two languages (Schwieter & Ferreira, 2017). Thus, interpreters must quickly and accurately reformulate information from the input language into the output one, and this code switching and reformulation place high demands on interpreters' linguistic processing as well as their cognitive resources (Christoffels & De Groot, 2004, 2005; Gile, 2009). Christoffels and Groot (2004) found that paraphrasing and interpreting were more cognitively demanding than shadowing (repeating sentences), as both paraphrasing and interpreting tasks involved information reformulation and grammatical structures changing. Christoffels and Groot concluded that it was not the language switch per se that added substantial difficulty to the interpreting task, but the combination of reformulating the input together with the language switch that taxed the interpreters' cognitive resources most. The second feature that distinguish interpreting from spontaneous speech production manifests in the intention of speaking (Gile, 2009, p. 163; Jin, 2010). In spontaneous speech, speakers construct a preverbal message based on what they want to communicate (Levelt, 1989, p. 9), and then the message activates lexical concepts (Levelt et al., 1999). During the formulation phase, speakers convey the conceptual content of their intended message by retrieving lexical items and assembling them in accordance with grammatical and phonological rules, and eventually, speakers execute the pre-articulatory plan prepared in the formulation phase into speech (Levelt, 1989, p. 11). The advantage of speaking one's own thoughts is that the speaker can always rearrange or modify or even omit some of the information to bypass the difficulties in lexical or grammatical retrieving (Costa, 2005; Gile, 2009, p. 163). But interpreters are expected to render an accurate and faithful interpretation of information given by an external source (the speaker); in other words, interpreters' speech is not based on their own intention of communication, but is impacted by how well they are able to retrieve the information stored in their memory and notes, based on their comprehension of the input. Hence, interpreters rely much more on their comprehension of the input than speakers do during spontaneous speech production. Additionally, the to-be-interpreted information stored in memory decays with time, which leads to poorer recall (Barrouillet et al., 2007). This constitutes the third feature of interpreting, high temporal pressure (Christoffels, 2004, p. 1; Gumul & Lyda, 2007; Henrard & Van Daele, 2017; X. Li, 2013; Schwieter & Ferreira, 2017).

These three features contribute to the fourth one, which is particularly seen in interpreting but not in other SLA tasks: high cognitive pressure (e.g., Autin & Croizet, 2012; De Bot, 2000; M. Liu & Chiu, 2011; Tzou, 2008). This cognitive pressure can sometimes be mediated by the degree of familiarity and manageability of the input information (Jin, 2010). When information is familiar to interpreters or is very logical, then long-term memory may lend support to comprehension by relieving the load placed on cognitive resources when storing and retrieving the information for later use. However, when the input information is unfamiliar to interpreters or beyond their ability to process (e.g., too dense or too long), interpreters are less capable of integrating the incoming information with their pre-existing knowledge. As a result, they are more subject to memory decay and retrieval failure. On the other hand, speakers performing other linguistic tasks such as spontaneous speech production, repetition or story-telling are largely language-dependent (e.g., Marian & Neisser, 2000; Park et al., 2020). In other words, the constraints placed on memory are often much more substantial in interpreting tasks than speech comprehension and production in other linguistic settings (Autin & Croizet, 2012; Gile, 2009, p. 221; Yu & Van Heuven, 2017).

The high cognitive pressure associated with interpreting also stems from the requirement to multi-task (the simultaneity of interpreting will be illustrated in Gile's models in Section 2.3.2). Both simultaneous and consecutive interpreting is a multi-tasking discourse interchange where language perception, comprehension, language transformation are carried out concurrently (Russell, 2005). Multi-tasking is highly cognitively demanding, because all concurrent tasks compete for limited cognitive resources and interfere with each other, which easily results in impaired performance such as information loss (Barrouillet et al., 2007).

In sum, the linguistic demands on language transformation and the different intentions for speaking, together with temporal and cognitive pressures, contribute to the uniqueness which distinguishes interpreting from other linguistic processing. It also raises theoretically important questions to be addressed in this thesis concerning the relationship between WM, interpreting performance and language context.

2.3.2 Theoretical models of Interpreting and Working Memory

Since the 1980s, the cognitive processes underlying interpreting have attracted increasing research attention and interest (Schwieter & Ferreira, 2017; Zheng & Xiang, 2017). To provide an explanatory account of the particularities involved in this cognitive-demanding language use, an interdisciplinary approach has been developed in interpreting studies by means of borrowing research paradigms, theoretical frameworks and data collection and analysis methods from closely related fields, such as psycholinguistics and cognitive science (A. Ferreira et al., 2015; A. Ferreira & Schwieter, 2017; M. Liu, 2008; Schwieter & Ferreira, 2017; Zheng & Xiang, 2017).

In the long tradition of research into the cognitive processes involved in interpreting, WM is among the factors that have gained prominent theoretical and empirical impetus (A. Ferreira et al., 2015; Jin, 2017; Mizuno, 2005; Timarová, 2008; Timarová et al., 2015). As reviewed in the previous sections, the sub-components of the WM system, the *central executive* and the *phonological loop* are actively involved during language processing (Baddeley, 1998, 2003, 2006; Gathercole & Baddeley, 1993). Interpreters are expected to produce 'pure' target language after receiving the source language (Christoffels & De Groot, 2005), even though the nature of the task leads to co-activation of both languages. Therefore, controlling languages is crucial to a satisfactory interpreting performance. According to Green's (1998) Inhibitory Control model (Section 2.1.3), the resolution of language competition is achieved by inhibiting the unintended language, thus enabling interpreters to focus on the target one during the task. This process is assumed to be governed by *central executive* in the WM framework (Baddeley, 2006).

The *phonological loop* specialises in maintaining verbally coded information and is therefore considered to be another essential sub-system for interpreting (Christoffels & De Groot, 2005; Injoque-Ricle et al., 2015). It is assumed that the phonological store holds source language messages in a phonological form, and these memories decay over time. The sub-vocal rehearsal process is responsible for refreshing the decaying representations maintained in interpreters' phonological store.

Despite different perspectives and conceptualisations of the specific tasks WM performs during the interpreting process, basically, all major cognitive processing models of interpreting have incorporated one or both of the aforementioned functions of WM system (e.g., Darò & Fabbro, 1994; Gerver, 1976; Gile, 2009; Moser-Mercer, 1978; F. Wang, 2017). Some of the cognitive processing models consider that WM and its role in interpreting are limited to storage functions (e.g., Gerver, 1976; Moser-Mercer, 1978), while others argue that it performs both storage and executive functions (e.g., Darò & Fabbro, 1994; Gile, 2009; Moser-Mercer, 1978; F. Wang, 2017). Even though researchers differ in the way they believe WM functions during interpreting, it is undeniably considered to be important in the interpreting process. The aim of this section is to explore the role of WM in interpreting and contrast theoretical models with empirical results in interpreting research. Instead of an exhaustive overview of all previous cognitive processing models of interpreting, this section is very selective and just reviews some of the major models that have incorporated WM, including three simultaneous interpreting (SI) models and two consecutive interpreting (CI) models.

One of the earliest cognitive models of SI was formulated by Gerver in 1976 (Figure 2.5). This model is a sequential model of mental processing during interpreting (Timarová, 2008). The major contribution of the model is that it depicts the order of interpreting processes and explain why certain situations occur during interpreting.

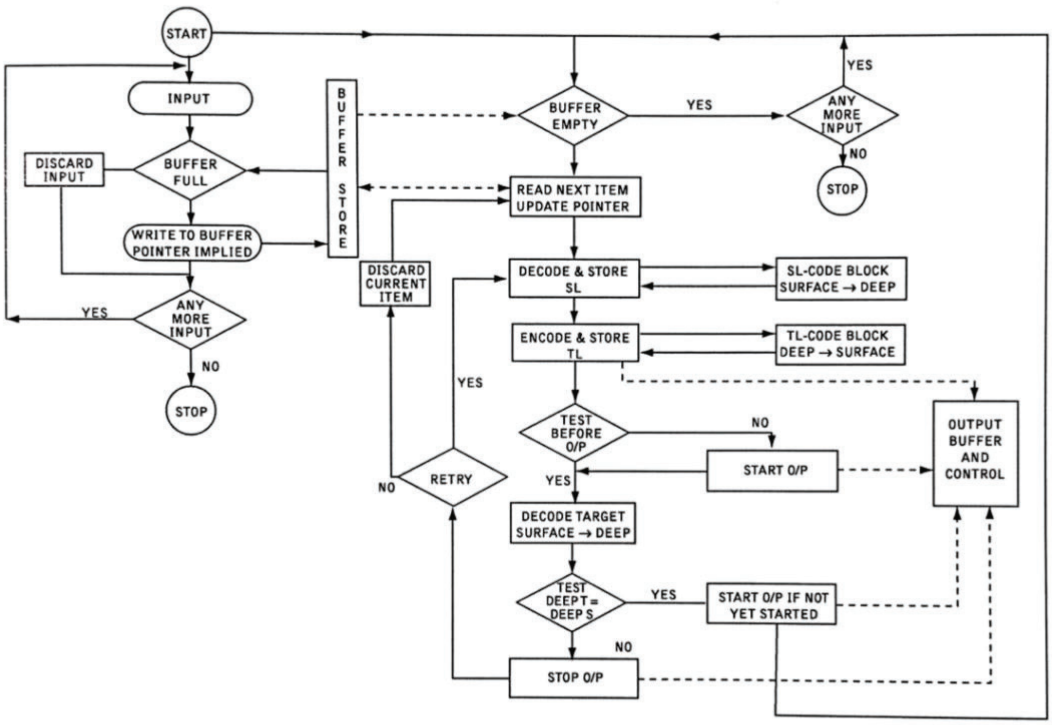


Figure 2.5. Gerver's model of SI (adapted from Moser-Mercer, 1978).

This model incorporates separate buffer (temporary) storage, for the input and output languages. Buffer storage in Gerver's model serves a similar role to the phonological loop sketched by Baddeley and Hitch (1974) in their WM model, which is necessary for temporarily maintaining the results of the interpreter's intermediate analysis of the information. By including a discard mechanism, the model illustrates the situation when buffer storage is full as a result of too much source language input, and the situation of failing to encode target language structures after another attempt. The Gerver model has been very influential in laying stress on the role of memory systems in SI (De Bot, 2000), even though it is not explicit concerning the language transfer process (Jin, 2010; Moser-Mercer, 1978; Pöchhacker, 2004, p. 101).

Another widely cited simultaneous model that incorporates memory structures is Moser-Mercer’s (1978) model of the SI process (Figure 2.6), which is based on Massaro’s (1975) comprehension theory. Unlike Gerver, Moser-Mercer is more explicit as to specific process segments. WM in this model is responsible for storing and recoding information, and is assumed to be critical to every process in the central column of the model (Moser-Mercer, 1978). By including the recoding function, Moser-Mercer proposes a very modern concept of WM, which seems to include executive functions (Timarová, 2008). According to Moser-Mercer, throughout the interpreting process, WM is believed to interact constantly and closely with long-term memory, and this helps interpreters to construct pre-linguistic semantic structures, activate target language elements and prepare output during interpreting. Moser-Mercer’s model addresses the important question of how syntactic and semantic information is organised and accessed (Mackintosh, 1985).

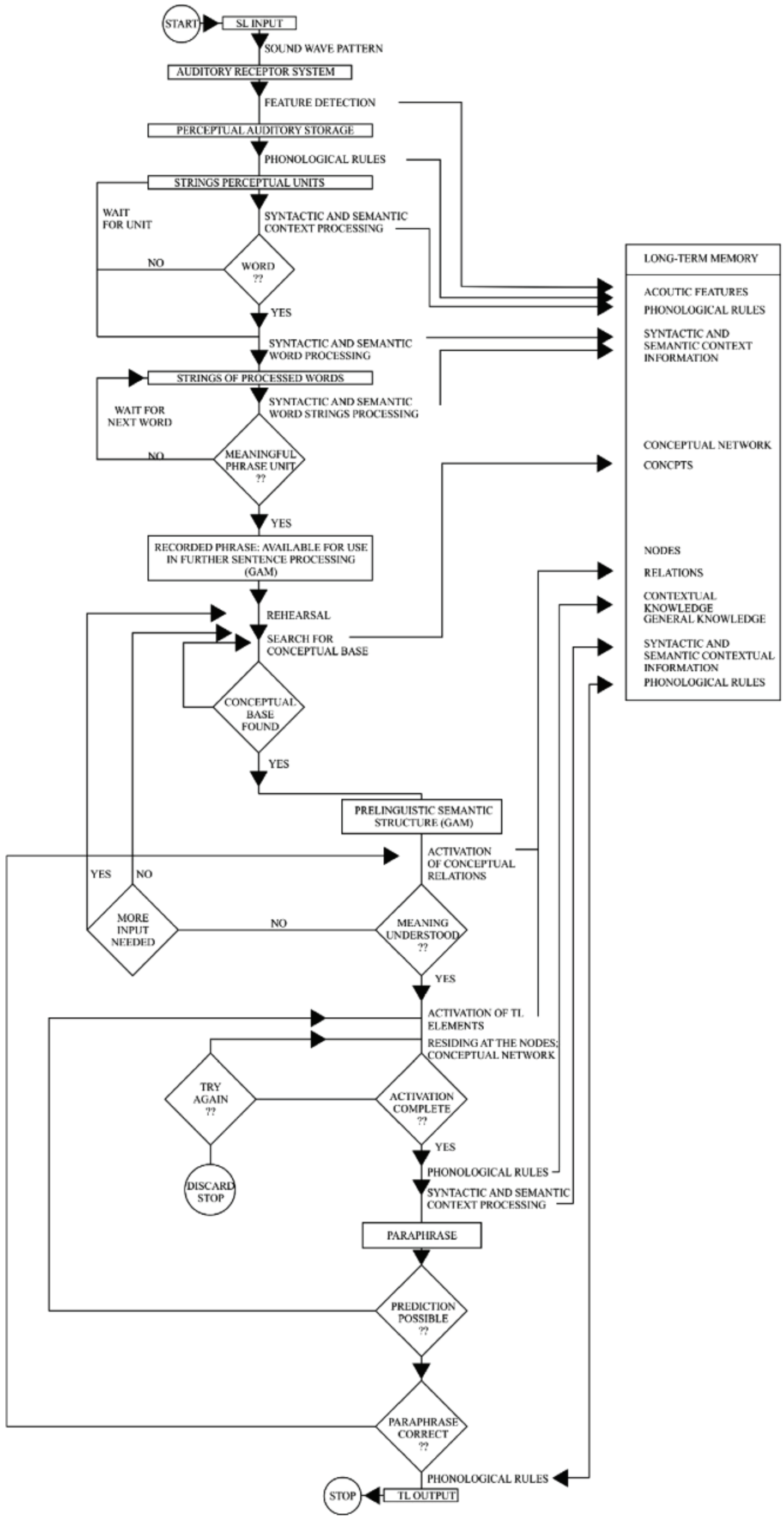


Figure 2.6. Moser-Mercer's Model of SI (Moser-Mercer, 1978).

The third SI model was proposed by Darò and Fabbro (1994). This model, like Gerver's, centres on memory and leaves other processes and structural components out of the account (Figure 2.7). The model includes two memory systems: WM and long-term memory. Each is further categorised into sub-systems. The WM system is again based on the model put forward by Baddeley and Hitch (1974), and both executive and storing functions are included. The *central executive*, with limited processing resources, is responsible for allocating inputs to the *phonological loop*, and the information that awaits processing is also passively stored there. Darò and Fabbro (1994) claim that listening to input with concurrent production of the target language interferes with the process of sub-vocal rehearsing of the message retained within the *phonological store*, and thus leads to reduced recall.

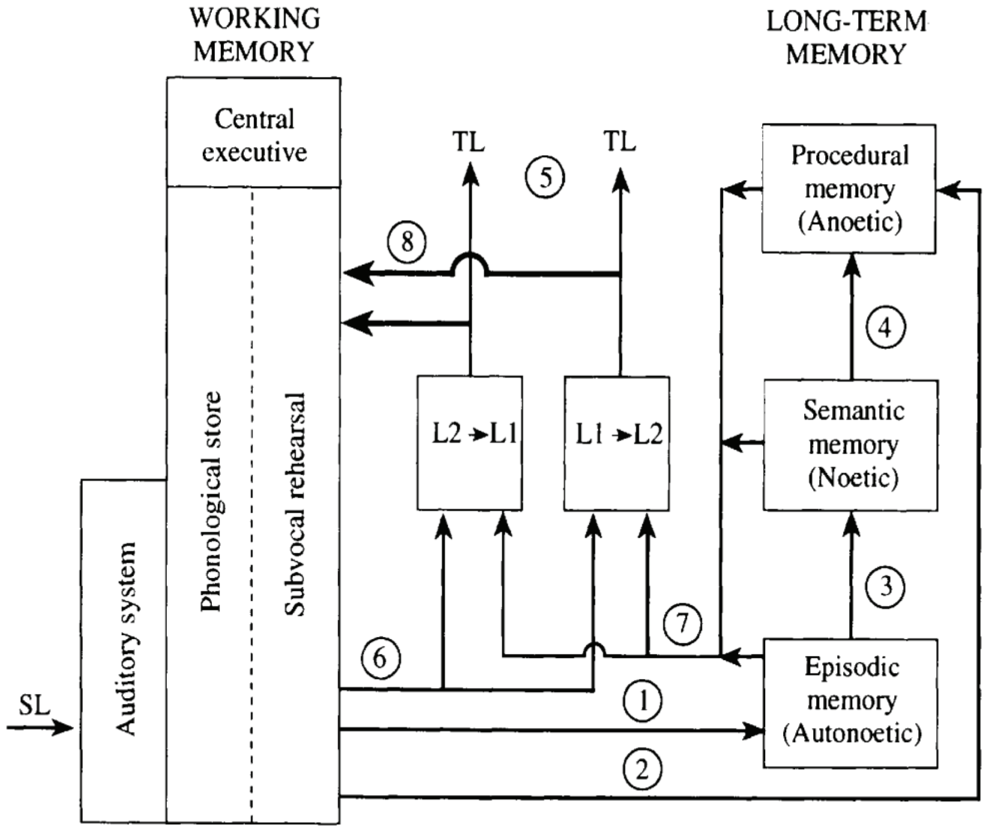


Figure 2.7. A general model of SI (Darò & Fabbro, 1994).

The vital role of WM in CI can also be appreciated from the incorporation of WM system in the following two CI models.

Influenced by the WM multi-component model (Baddeley & Hitch, 1974), and Levelt's production model (Levelt, 1989), Wang (2017, p. 182) put forward a holistic CI model (Figure 2.8) which illustrates the information processing routine from source language input to target language output, and in two directions (Chinese-English and English-Chinese). Wang's model consists of two major systems that combine and interact: the WM system and the long-term memory system. According to the model, WM plays a significant role in facilitating source language decoding (left column), processing and message generation (conceptualiser), as well as target language lexical mapping, syntactic encoding, and the final target-language output (right column). Specifically, in this model, the *central executive* takes charge of source language comprehension, conceptualisation, and target-language output preparation; the *phonological loop* temporarily stores the intermediate analysis of syntactic and lexical information; and the *episodic buffer* temporarily stores the linguistic representations constructed during source language input. This model includes a *notetaking* component which represents the facilitative role of notetaking in alleviating the memory burden imposed by the linear nature of the source language input. Wang's model features a decision point termed '*comprehensible?*' that separates it from other interpreting models. This decision point is followed by another decision point termed as '*prediction possible?*' in the production phase of the flow chart. If the responses are 'no' at both points, information loss caused by cognitive overloading occurs. The major contribution of this model is that it provides a discrete view of the various stages of CI and links the stages. It also provides possible explanations for interpreters' information loss. However, information loss is only presented in the production phase in the flow chart, and it does not explicate what happens when interpreters are overloaded with the input source language. According to the Speaker Model (Levelt, 1989, p. 9), interpreters (as a type of interlocutor) monitor not only for meaning but also for grammatical accuracy (Laver, 1973; Levelt, 1989, p. 14), and they constantly compare the meaning of what was said or internally prepared to what was intended, and also detect self-generated form failures that are well represented by self-correction (Levelt, 1989, p. 13; Levelt et al., 1999) if they have sufficient underlying WM resources (Kormos, 1999). Therefore, what is lacking in this CI model is a monitor mechanism in the production phase in representing interpreters' self-monitoring. However, the model develops a more comprehensive framework for a better understanding of the cognitive processes underlying CI.

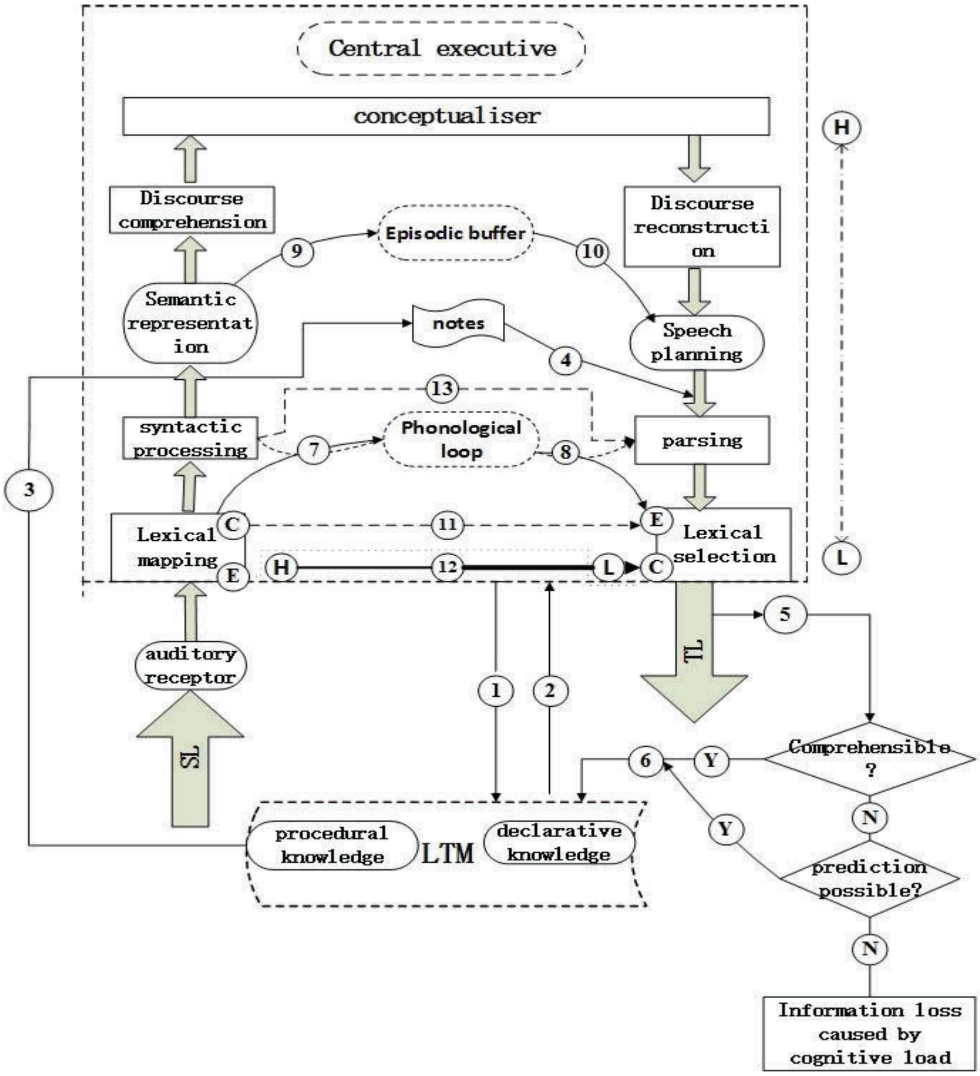


Figure 2.8. A general model of information processing during CI (F. Wang, 2017, p. 182).

The last model to be reviewed in this chapter is the Effort Model of CI (Figure 2.9), developed by Gile (1995, p. 179) who based the model on close scrutiny of interpreting errors. Gile found that interpreting mistakes occurred not only in speeches with fast speed, high information density or difficult technical terms, but also in speech segments which were clear, slow-paced and non-technical (Gile, 2009, p. 157). Moreover, erroneous interpretation happened to both interpreting students and experienced interpreters with excellent working language proficiency. This rules out inadequate target-language production and insufficient understanding of the source speech as explanations to interpreting errors (Gile, 2009, p. 157). Also, when asked to interpret the same text twice, the interpreters tended to make different mistakes in both interpretation tests, which indicates that the cause of the errors also goes beyond a lack of (extra)linguistic knowledge of the source text (Gile, 2009, p. 158).

On the basis of these observations, Gile claims that interpreting difficulties may stem from insufficient WM (cognitive) resource allocation. He puts forward the 'Tightrope Hypothesis' which states that most of the time during interpreting, interpreters are suffering from chronic cognitive tension and are working close to the maximum of their WM (Gile, 2009, p. 183). To explain his hypothesis, and to elaborate on how the errors occur, Gile (1995, p. 179) develops a set of formulas to represent the relationship between each component involved during interpreting. The Effort Model is straightforward and easily understood, and thus has been adopted as a conceptual framework by many interpreting instructors and researchers, including the author of the present study.

Gile refers to the components in the model as 'Efforts' in order to highlight their effortful nature, as they involve attentional action which draws on WM resources (Gile, 2009, p. 160). According to Gile, CI consists of a comprehension phase and a production phase.

Comprehension phase:

$$CI = L + N + M + C$$

$$TR = LR + NR + MR + CR$$

$$TA \geq TR$$

Production phase:

$$CI = Rem + Read + P + C$$

Figure 2.9. Effort Model of CI (Gile, 2009, p. 174).

The comprehension phase in Gile's CI model involves all comprehension-oriented components, including L (Listening and Analysis), N (Notetaking), M (Memory) and C (Coordination). M and C refer to the storing function and executive function of WM, respectively. Each of the four components in this phase shares a portion of WM resources. The total WM (cognitive) resource requirements (TR) during this phase is equal to the sum of the resource requirements for each component, which varies according to the input speech rate and the interpreter's segmentation of the input. To guarantee a smooth interpretation, total resources available (TA) for each Effort must be equal to or larger than its requirements for the task at hand (TR). The production phase in the CI model includes Rem (Recall the information from memory or/and notes or/ and the source speech), Read (Read or decipher the notes taken in the listening phase) and P (Production). Gile assumes that the cognitive load of CI falls mainly in the comprehension phase rather than the production phase (Gile, 2009, p. 176). The comprehension phase, according to him, is paced by the speaker, and so all components are competing for limited resources under the pressure of time; but in the production, the interpreter is free to allocate WM to each effort at his/her own pace, which reduces the cognitive pressure.

The CI Effort Model highlights the simultaneity of the interpreting process, during which two or more components are running concurrently, and the finite WM can be allocated flexibly to these components. This conception builds on, and is highly compatible with, models relating to allocation and management of mental resources developed by cognitive psychologists (Anderson, 2015, p. 129; Baddeley, 2000; Broadbent, 1958, p. 228). Therefore, how WM is allocated during interpreting seems to significantly impact the performance.

Admittedly, the Gile's Effort Model has its limitations in relation to specificity and sophistication, but the significance of this conceptual framework should be recognised due to its potential to bridge interpreting research and scientific paradigms (Jin, 2010).

In general, despite differences in theoretical approaches in the abovementioned interpreting cognitive processing models, the primary similarity between them is that WM plays a vital role in interpreting, and this assertion is supported by empirical research.

2.3.3 Empirical Studies of Working Memory in Interpreting Performance

The role of WM in interpreting has not only been speculated in theoretical models, but also been examined in many empirical studies (Dong & Cai, 2015), where concepts and tools from cognitive psychology are employed to explain and measure WM resource availability (Gile, 2009).

Several researchers have reviewed the current empirical studies on WM and interpreting in an effort to identify their relationship (e.g., Dong & Cai, 2015; Mellinger & Hanson, 2019; Timarová, 2008). Generally, these researchers classified studies in this regard into three main lines: the first line principally investigates the relationship between WM and measures of interpreting quality (e.g., Cai et al., 2015; Christoffels et al., 2003; Injoque-Ricle et al., 2015; Liu, 2001; Macnamara & Conway, 2016; Tzou et al., 2012; Van Dijk et al., 2012). The second line of research involves studies that compare the performance of professional interpreters to various comparison groups including novice interpreters and non-interpreters (e.g., Christoffels et al., 2003; Köpke & Nespoulous, 2006; Köpke & Signorelli, 2012). The third line of studies explores the relationship between WM and other interpreting-related sub-skills such as lexical processing, L2 proficiency and interpreters' psychological competence (e.g., Cai et al., 2015; Christoffels et al., 2003; Dong et al., 2013; Kállay & Visu-Petra, 2014).

The overarching approach to explore the relationship between WM and interpreting performance is by analysing the correlation between them. Two professional interpreters as individual raters are invited, and both raters were asked to give scores on participants' holistic interpreting performances separately. The inter-rater reliability is calculated afterwards. Even though not every individual study finds a strong correlation, the overall combination of studies provides convincing evidence that participants' interpreting performance is positively correlated with their WM regardless of their interpreting training history and expertise (e.g., Christoffels et al., 2003, 2006; Injoque-Ricle et al., 2015; Mellinger & Hanson, 2019; Tzou et al., 2012). For instance, Tzou et al. (2012) reported a significant positive correlation between participants' WM and their SI performances for interpreting students as well as bilinguals without any interpreting training. This result echoes that of Christoffels et al. (2003), who also observed an equally positive relationship between SI performance and WM in a group of bilinguals who had no prior interpreting training. Similar correlations have been identified among professional and novice interpreters (e.g., Injoque-Ricle et al., 2015; Timarová et al., 2014; Zhang, 2009). Injoque-Ricle et al. (2015) categorised participants by their WM into two groups. The high WM group consists of participants with a WM span equal or higher than percentile 75, and the low WM group involves those whose WM span equal or lower than percentile 25. Correlation analyses showed positive significant correlations between SI and the WM and between SI and the number of days worked per month. Additionally, the high WM group outperformed the low WM group on all tasks involved.

As well as focusing on the relationship between WM and spoken language interpreting performance, researchers have also examined the relationship between WM and sign-language interpreting performance, and these researchers have obtained similar findings. In a longitudinal sign language interpreting study, Macnamara and Conway (2016) collected participants' data at four time points, and their results again demonstrated the importance of WM in interpreting performance. They found that participants' initial WM resource availability strongly predicted their final SI performance.

Of all empirical studies exploring the cognitive aspects of interpreting, a disproportional majority focuses on the simultaneous mode, leaving CI relatively underexplored (Pöchhacker, 2016, p. 108; Zheng & Xiang, 2017). While fewer studies have been conducted in CI than SI, findings also suggest that participants' CI performance is positively correlated with their WM. For instance, in a recent study Cai et al. (2015) found a significant correlation between WM and interpreting students' CI performances. This result was later echoed by Wang (2017) who reported significant correlation between WM and CI performance among interpreting learners. In a longitudinal study, Dong, Liu and Cai (2018) recruited two groups of Chinese learners of English. One group received CI training for 16 weeks, and participants in the control group received general English training for the same period of time. Participants' WM was measured at both pre-test and post-test, and their CI performance was assessed only at post-test. The results indicated that participants' WM significantly correlated with their CI performance. Given the finding, Dong et al. suggested that with longer period of training, or with higher L2 proficiency, the relationship between CI performance and WM could be stronger.

If it is true that large WM is strongly correlated with the quality of interpreting performance, researchers then speculate that professional interpreters should exhibit some advantage in their WM resource availability compared to non-interpreters, reflecting either an inherent aptitude or the effects of extensive practice (Darò & Fabbro, 1994; Wen & Dong, 2019). Therefore, the second line of research involves studies that compared the WM of professional interpreters with interpreting students or non-interpreting bilinguals in order to gain more insight into how important WM is to interpreting performance. This type of approach is referred as expert–novice comparison (Liu, 2008; Moser-Mercer et al., 2000). However, results have been inconsistent.

As expected, numerous studies have provided rather strong evidence that professional interpreters exhibit superior WM resource availability than bilingual students or L2 teachers (e.g., Bajo et al., 2000; Christoffels et al., 2006; Padilla et al., 2005; Stavrakaki et al., 2012), bilingual or multilingual lay subjects (e.g., Becker et al., 2016; Dong & Xie, 2014; Kállay & Visu-Petra, 2014; Signorelli et al., 2012), monolinguals (e.g., Henrard & Van Daele, 2017), and translators (e.g., Henrard & Van Daele, 2017).

An interpreter WM advantage was observed by Christoffels et al. (2006), who compared professional interpreters with unbalanced bilingual students as well as language teachers who were of similar educational background and professional experience. An interpreter advantage in WM was also reported by Signorelli et al. (2012). Signorelli, Haarmann, and Obler (2012) compared professional interpreters and non-interpreters' WM measured by reading span task, and they found that interpreters demonstrated larger reading spans than the non-interpreters. These results were taken as evidence that professional interpreters are better at storing and manipulating information in their WM. Additionally, some studies have further suggested that the WM advantage observed among professional interpreters compared to other bilingual groups does not stem from language proficiency, as participants in other two non-interpreter groups did not differ from each other in WM or language processing (Bajo et al., 2000; Stavrakaki et al., 2012). Therefore, WM availability should be related to the interpreting expertise.

Conversely, the opposite pattern has been reported by other studies (e.g., Chincotta & Underwood, 1998; Köpke & Nespoulous, 2006; Liu, 2001, p. 20; Liu et al., 2004; Timarová et al., 2015). For example, Chincotta and Underwood (1998) predicted that due to an WM advantage derived from the occupation, professional interpreters would demonstrate better cognitive control and be less affected by having to perform concurrent tasks than interpreting students or non-interpreter bilinguals. However, their results rejected this prediction, and they found that the performance of experienced interpreters was similar to that of other groups under concurrent interference. Likewise, Köpke and Nespoulous (2006) compared the WM and short-term memories of professional interpreters with those of interpreting students, and found that the best cognitive performance was always produced by the interpreting students rather than the professional interpreters.

However, exploring the relationship between WM and interpreting using the expert-novice paradigm has received some criticism. Dong and Cai (2015) pointed to participants' age range as a potential source of mixed results. They argued that many studies did not take the variance in groups' age ranges into account. Novice or student interpreters are generally younger than professional interpreters. Research on individual differences in WM suggest that WM is closely related to age: WM peaks at the age of 24 (Alloway & Alloway, 2013) and then begins to decline; Therefore, the lack of WM advantage among professional interpreters when compared with younger novice interpreters and untrained bilinguals may be caused by their older age. Thus, Signorelli et al. (2012) suggested that when investigating WM, participants' ages need to be taken into consideration. Another factor that may account for the mixed results is that most of these empirical studies compared the WM of professional interpreters with the WM of other groups without measuring either group's overall interpreting performance. In other words, professional interpreters are often assumed to perform better at interpreting than novice interpreters but without any measurement (Liu, 2001; Wang, 2016).

Liu et al. (2004) recruited three groups of participants who were varied in their training length and professional expertise: professional interpreters, student interpreters with longer period of training and beginning student interpreters. The results showed that three groups did not differ significantly in their WM, but professional interpreters exhibited significant better SI performance than the other two groups. The authors attributed the difference in SI performance, at least in part, to the development of specific interpreting expertise rather than to WM availability.

Moreover, some studies also suggest that interpreting training (e.g., Bajo et al., 2000; Macnamara & Conway, 2016; Tzou et al., 2012; Wen & Dong, 2019) and expertise (e.g., Signorelli et al., 2012; Yudes et al., 2011; Zhang, 2008) can also effectively influence WM availability. Therefore, the length or amount of training and level of interpreter expertise should both be considered before investigating the presence/absence of an interpreter advantage (Wen & Dong, 2019).

To gain a deeper understanding of the role of WM in interpreting, researchers also examined the possible relationships between WM and other interpreting-related sub-skills such as lexical processing, L2 proficiency, interpreters' psychological state, as well as how these sub-skills contribute to interpreting performance (e.g., Cai et al., 2015; Christoffels et al., 2003; Dong et al., 2013; Kállay & Visu-Petra, 2014).

Christoffels, De Groot and Waldorp (2003) focused on the roles of WM and lexical processing for bilinguals without interpreting training. Results suggested that these two factors formed independent sub-skills of SI and were both significantly related to SI performance. Unlike the untrained bilinguals in Christoffels et al. (2003), participants in Cai et al. (2015) were all beginner interpreting students. In this study, WM, lexical processing and L2 proficiency were tapped and examined. Even though WM positively correlated with CI performance, researchers found that participants' L2 proficiency seemed to play a more important role than WM.

Instead of merely probing language skills and WM, Dong et al. (2013) also examined participants' anxiety during interpreting. Participants in their study were also beginner student interpreters. Based on the analysis of correlations between the data of interpreting performance and the other test, researchers emphasised the significant roles of participants' WM and anxiety. They concluded that even though language competence effectively influenced interpreting performance, the language competence mostly functioned through the mediation of interpreters' psychological competence such as WM and anxiety. And finally Dong et al. suggested that the function of interpreting training is to teach learners to coordinate these relevant capabilities properly during interpreting.

Empirical studies also explored the relationship between WM burden and the interpreting directionality, which can easily be reflected in interpreters' direction-dependence in their interpreting performance (Chabasse & Dingfelder Stone, 2015). WM burden imposed on interpreters is highly contingent upon factors such as topic, density of input and rate of delivery (Han, 2015a; Han & Riazi, 2017; M. Liu, 2001). But fundamentally it depends on directionality (Chabasse & Dingfelder Stone, 2015), which refers to whether interpreting is conducted into or from one's 'native language' or 'language of habitual use' (Pavlović, 2007). The debate about which direction of interpreting is more cognitively demanding is still ongoing (Pöchhacker, 2016, p. 20). Unlike word translation asymmetry, which is assumed to be largely influenced by the inhibition efforts of language switches, translation at a higher level is more complicated (De Bot, 2000). Some researchers claim that interpreting into one's weaker language (L1-L2) is more cognitively demanding than the opposite direction. They argue that retrieving corresponding equivalents in the L2 places a heavier burden on speech production and presentation than interpreting in the opposite direction (L2-L1). Proponents state that when interpreting from the weaker language (L2-L1), the target language is the interpreter's mother tongue, and thus the notetaking becomes more feasible, which facilitates target language production in the production phase of CI (e.g., Chabasse & Dingfelder Stone, 2015; Donovan, 2005, 2003; Seleskovitch, 1999). Conversely, some other researchers argue that interpreting into one's weaker language (L1-L2) is easier than the other direction, as interpreting from one's native language can reduce the cognitive resources needed for listening comprehension and analysis of the input, and thus allow interpreters to generate a more complete and accurate output (e.g., Barik, 1975; Williams, 1995). In particular, this direction should be advantageous for CI where the cognitive load is mainly in the comprehension phase rather than the production phase (Gile, 2009, p. 176). Also, it is suggested that the L1-L2 direction of CI is easier for notetaking, as the latter is facilitated if interpreters' L1 serves as the source language (Dam, 2004).

Even though few empirical studies have examined the question of which direction is more demanding, results can be found to support both sides of debate. For instance, interpreters were found to make fewer omissions and meaning errors when interpreting from L1 to L2, and their interpretations were conceptually more adequate than the opposite direction, although interpreters made more language errors when interpreting from L1 to L2 (Barik, 1975). An interesting result was reported by Zhang (2009) in an SI study. At first participants demonstrated a superior performance from L2 to L1 direction when delivery speed was slow (140 words/min); however, all participants performed better from L1 to L2 when the delivery rate of the source texts increased to 166 words/min, indicating that input language speed is an influential factor in interpreters' performance.

In terms of CI, if one assumes that Gile is correct in suggesting that cognitive load falls mainly in the comprehension phase rather than the production phase, and if the comprehension phase places a light burden on the interpreter, such as when the input is in his/her native language, then the interpreter should be less subject to cognitive resource saturation and demonstrate better performance (Chabasse & Dingfelder Stone, 2015; Russell & Takeda, 2015). However, complicated results were reported in this regard. In a bidirectional CI study, Fu (2012) used fluency as an indicator of CI performance and found that participants tended to make more expression-related pauses in their L1-L2 than in L2-L1 direction. But the number of their logical-related pauses was significantly larger in L2-L1 direction. This result echoes that of Mead (2002), which also revealed that the source speech in L2 was more demanding in terms of logical analysis. Mead concluded that both directions were cognitive-demanding because interpreting performance was largely dependent on non-automatic processes in both L1 and L2. Also, comprehension and production were both effortful operations for interpreters.

To summarise, even though there are still mixed and often controversial findings that need to be clarified, empirical results of previous studies generally support the importance of WM in interpreting models by revealing a positive correlation between WM and interpreting performance. However, this correlation is constantly influenced by factors such as participants' age and interpreting expertise. Conflicting results regarding the interpreting directions and their corresponding cognitive demands have also been reported by previous researchers.

2.3.4 Linguistic Processing and Interpreting Performance

2.3.4.1 Lexical Accessing and Interpreting Performance

Experiments on multi-tasking reveal that when some of an individual's cognitive resources are engaged by a secondary task, their performance is slowed down or impaired (Marcel Adam Just et al., 2008; Newman et al., 2007). This is also the case in interpreting where some or all of the processes such as listening, comprehension, and production are taking place at the same time. As the input unfolds, WM, which is responsible for information processing and storing, is used to analyse sound signals and accesses the corresponding words so as to extract meaning out of them (Gile, 2009, p.

224). This process is not automatic but WM-consuming. If the interpreter has trouble in turning the sound signals into meaningful words or segments, they are forced to keep an accumulation of unprocessed sound signals in their WM. The problem is that WM is a finite resource, indicating that its maximum availability can be exceeded. When the WM is saturated, either the incoming information can no longer be attended to, or interpreters attend to the new information at the expense of previously received segments (Gathercole & Baddeley, 1993, p. 225), resulting in errors or information loss.

Due to the effortful nature of interpreting, some researchers suggest that efficient linguistic processing such as high lexical availability may be an important factor in alleviating interpreters' cognitive burden (Gile, 2009, p. 222; Mead, 2002). Despite the obvious, fundamental differences between word translation and the actual tasks of interpreting which involves comprehension and production of complete discourse (Gile, 1998), lexical accessing is associated with higher-order processing. When only small amounts of cognitive resources are devoted to lexical tasks, more resources will be available for other task components, and those processing results will be more efficiently available for further processing (Antoniou et al., 2015; Christoffels et al., 2003; Christoffels & De Groot, 2005; O'Brien et al., 2006).

Similar to SLA studies, in the interpreting field, lexical processing performance is often assessed by setting tasks such as word translation or word translation recognition (e.g., Bajo et al., 2000; R. Cai et al., 2015; Christoffels et al., 2003; Santilli et al., 2019). And there is broad agreement that satisfactory interpreting performance cannot occur without efficient lexical processing (Bajo et al., 2000; Christoffels et al., 2003; Fabbro & Darò, 1995; Santilli et al., 2019).

In interpreting studies that have tested participants' lexical accessing performance, professional interpreters mostly demonstrate more efficient lexical processing than non-interpreting bilinguals in the form of faster and more accurate word translation responses (Bajo et al., 2000; Santilli et al., 2019). Moreover, some studies have revealed significant correlation between participants' lexical accessing and interpreting performance (e.g., Bajo et al., 2000; Christoffels et al., 2003; Fabbro & Darò, 1995; but see Cai et al., 2015). In an SI study, Christoffels et al. (2003) measured participants' response times in retrieving word translation equivalents upon reading source language words and found that both directions of word retrieval response times significantly correlated with participants' SI performance. However, such a correlation was not found by Cai et al. (2015), who tested participants' CI performances.

Given the importance of lexical processing in interpreting, researchers have also examined factors that may impact on interpreters' interlingual links in the mental lexicon. Chmiel (2016) investigated the role of interpreters' habitual language use in their lexical processing. She assessed two groups of professional interpreters who had different interpreting directions in their practice. One group worked unidirectionally (L2–L1 direction), and the other group worked bidirectionally (in the L2–L1 and L1–L2 directions). The two groups' lexical retrieval efficiency was calculated. Chmiel predicted that the unidirectional interpreters would exhibit a higher degree of asymmetry in word translation directionality, as they consistently worked only in one direction. However, the results showed that bidirectional interpreters were more unbalanced in their lexical retrieval performance than their unidirectional counterparts. Chmiel suggested this result might stem from the two groups' different language exposure. The unidirectional participants had lived in a L2-speaking country for at least five years prior to the experiment. That is, participants in this group utilised both languages in their professional environments outside the interpreting booth on a daily basis. In contrast, the bidirectional participants always lived in their home country and used their L1 predominantly in both professional and non-professional settings.

Chmiel's study provides us with a perspective on the influence of language environment on interpreters' performance. Apart from language proficiency and daily language use, language exposure may also be an influential factor impacting interpreters' (lexical or sentence) translation performance. However, what is missing in Chmiel's study is that interpreters' age and cognitive resources were not taken into account, and these two factors are considered important in affecting linguistic performance.

2.3.4.2 Fluency and Interpreting Performance

Production disfluency phenomena such as pauses, hesitation and lack of accuracy are triggered by many factors, and one of them is interpreters' lexical accessing problems (Gile, 2009, p. 223). An important characteristic of efficient lexical accessing is fluency (DeKeyser, 2001; Fuchs et al., 2001), which refers to the speed and the ease with which a speaker can generate words and sentences (Moser-Mercer et al., 2000).

Fluency is also considered one of the most important factors contributing to the overall quality of interpreting both from the perspective of the end users' expectations and from an academic and professional perspective (E. I. Fernández, 2013; Han, 2015b; Han et al., 2020; Kurz, 2001; Rennert, 2010; Wu et al., 2013; Yu & Van Heuven, 2017).

The first empirical study aiming to elucidate quality criteria for interpreting was carried out by Bühler (1986) on a sample of professional interpreters. Criteria listed by Bühler included interpreters' accent, voice, fluency, cohesion, sense consistency with the original message, completeness of interpretation, grammatical accuracy, terminology, interpreting style, preparation of conference materials, endurance, poise, appearance, reliability, teamwork ability and positive feedback from delegates. Following a similar methodology, Kurz (1993) used the first eight criteria in Bühler's study to canvass the end users' of interpreting services for their thoughts on which criteria were the most important. Responses were on an ordinal scale from 1 to 4 (irrelevant, less important, important, highly important). Fluency was deemed to be an very important criterion, and it received and average score of 3.1, following content-related criteria such as correct terminology and completeness of interpretation. In a more recent global survey study, Chiaro and Nocella (2004) recruited 286 professional interpreters across five continents, and they found that after the content-related criteria (consistency with the original, completeness of information and logical cohesion), fluency of delivery, correct terminology and correct grammatical usage were the three second-most important factors. The importance of fluency in interpreting utterance as a criterion was later confirmed by Pöchhacker and Zwischenberger (2010), who surveyed 704 interpreters worldwide from The International Association of Conference Interpreters, and the findings indicated that fluency was perceived as very important by 71% of participants, only following content-related criteria (sense consistency with the original and logical cohesion), the form-related criterion and correct terminology.

Even though fluency is important in interpreting performance, it is an ambiguous concept without widely accepted evaluating criteria (Han, 2015b). Existing studies on interpreting mostly adopt (para)linguistic parameters in a given interpretation sample to assess fluency (Fu, 2012; Han, 2015b; Macías, 2006; B. Wang & Li, 2015; Xu, 2010). However, researchers suggested that, for speech tasks imposing different degrees of loads on speakers' cognitive resources, there should be corresponding differences in fluency rating (Grosjean, 1980; Yu & Van Heuven, 2017). For example, it is reported that pauses are less frequent but longer in interpreting output than in the original spontaneous speech (B. Wang & Li, 2015).

Across all interpreting studies, articulatory rate and unnatural pauses are two critical parameters of interpreting fluency (Han, 2015b; Rennert, 2010). Articulatory rate is the number of syllables, including disfluencies, divided by total duration of speech. Some studies calculate the duration with pauses (e.g., Han, 2015b), whereas some do so without (e.g., Yu & Van Heuven, 2017). Unnatural pauses are indicators for effortful speech production (Wang & Li, 2015), where speakers need more accessing time (Bygate, 2002). Unnatural pauses include silent and filled pauses; however, there is no clear-cut minimum threshold in interpreting studies as to how long a silent pause constitutes an unnatural one. Across the literature, minimum durations have included 0.18 seconds (e.g., Duez, 1982), 0.25 seconds (e.g., Duez, 1982; Goldman-Eisler, 1972; Mead, 2005; Tissi, 2000), 0.3 seconds (Tannenbaum et al., 1967), 0.5 seconds (Han, 2015b), and 0.56 seconds and 1.4 seconds (J. Wang, 2016). The concept of filled pauses is relatively consistent, and commonly refers to meaningless 'ums, ahs, errs' etc. (e.g., Han, 2015b; Rennert, 2010; J. Wang, 2016). Besides unnatural pauses and articulatory rate, phonation ratio (Han, 2015b) and mean length of a run (Han, 2015b; Yu & Van Heuven, 2017) have also been used as parameters of interpreting fluency.

Fluency is associated with efficient and automatic linguistic processing (Hulstijn, 1997; Skehan, 1998). To improve linguistic processing efficiency and automaticity, studies suggest that consistent practice is essential (Section 2.1.3); however, rote repetition is not an effective form of practice (Bygate, 2002). One possible way of ensuring both quantity and quality of L2 practice (input and output) is through SAE (Section 2.1), which is characterised by massive exposure to authentic L2 input and unlimited opportunities for L2 interaction and output. And empirical studies indicate that this kind of exposure provides effective inhibition of dominant L1 and thus is beneficial for the target L2 processing (Section 2.1). Even though SAE has been investigated in the field of SLA and has also been strongly recommended for interpreting students (e.g., Napier, 2015; Gile, 2009, p. 221; Gile, 2005), the effect of SAE on linguistic processing underlying interpreting is often theoretically assumed and rarely being empirically explored. The same is true of the relationship between SAE and holistic interpreting performance.

The reason we cannot assume that interpreting students will benefit from SAE in the same way as other bilinguals is that interpreting students are a special subset of bilinguals. For interpreters, two languages are constantly activated even when they are in a monolingual environment (García, 2019, p. 199; Grosjean, 1997; Schwieter & Ferreira, 2017), and they have to deliberately and consciously switch between the two languages on daily basis. This unique attribute of interpreting students may make the effect of SAE complex. Another issue that is worth exploring is the role of WM in interpreting performance under study abroad conditions. Will the involvement of SAE mediate the role of WM in interpreting, as reported in some of the SLA studies? Additionally, given that interpreting learners need to constantly switch between two languages through inhibiting the unintended one, will there be any cognitive benefit for interpreting students as a result of SAE? This thesis attempts to tentatively investigate these questions and bring us one step closer to a more comprehensive understanding of the contextual effect on interpreting performance and language processing.

2.4 Theoretical Framework

On the basis of the literature reviewed on SAE, WM, and interpreting, this thesis presents an interdisciplinary framework for understanding the complicated relationship between learning context, interpreting performance and learners' cognitive resources. This framework synthesises concepts from SLA and WM theories to advance knowledge by providing a better understanding of the impact of SLA and WM on interpreting performance, and by filling knowledge gaps in research highlighted in Section 2.5.

Interdisciplinary approaches are suitable for the present study because they allow the researcher to develop a deep and multi-dimensional understanding of the underlying phenomenon of interpreting. Also, such approaches are in line with the research trends of interpreting studies which are informed by approaches from cognitive psychology and linguistics (A. Ferreira et al., 2015; Han, 2018; Zheng & Xiang, 2017). The theory underpinning SAE and L2 processing is Interaction Theory (Carroll, 1999; Long, 1981, 1983, 1996) (Section 2.1). According to this theory, language input (Krashen, 1985; Long, 1996), interaction (Gass & Mackey, 2007), and output (Swain, 1985, 1995) are critical for language acquisition and processing. Interaction can provide learners with opportunities to create or strengthen the link between language form and meaning (Gass, 1997, p. 104; Pica, 1994), through which learners are able to gain access to multiple exemplars of target-like input. meanwhile, output helps promote learners' linguistic processing automaticity (Swain, 1985, 1998, 2005). Interaction Theory also emphasises the role of WM in assisting bilinguals to learn from their interactions. Therefore, SAE, characterised by its rich exposure to authentic target language use (Fernández-García & Martínez-Arbelaiz, 2014; Zaytseva et al., 2018), is commonly assumed to facilitate L2 acquisition and the efficiency of L2 processing.

The theory for cognitive resources in the present study is based on the WM multi-component model proposed by Baddeley and Hitch (1974, 2000) (Section 2.2). According to this model, WM is responsible for the temporary storage and processing of target information as well as the concurrent inhibition of distractors (Baddeley, 1998, 2003, 2006, 2010). Thus, WM is considered to be critical for complex tasks such as interpreting. Even though WM is essential, it is a limited availability system; thus, efficient deployment of such resources for linguistic processing (Gathercole & Baddeley, 1993, p. 99) directly affects interpreting performance.

The multi-component WM model fits well with the view of the interpreting Effort Model proposed by Gile (1995, 2009) (Section 2.3). The gist of Gile's model is that the total processing capacity required for an interpreting task is the combined processing capacity of all sub-processes. Moreover, to ensure smooth interpreting performance, an interpreter's total processing capacity for each sub-process should not exceed their available WM resources. This is because WM is limited, and when a task exceeds the available resources, performance will deteriorate. To utilise WM efficiently, it is suggested that interpreters should enhance their linguistic processing automaticity, as interpreting is, essentially, a form of linguistic processing. As elements of linguistic processing become automatised, more WM resources can be freed to be reallocated to other elements or other cognitively demanding tasks (Antoniou et al., 2015; Christoffels et al., 2003; O'Brien et al., 2006).

Based on the above framework, it is predicted that for two groups of interpreting students of comparable WM resources, those with SAE will be more likely to demonstrate better linguistic processing abilities than interpreting students without

SAE. Furthermore, study abroad students are also predicted to perform better in their interpreting. Likewise, students with greater WM availability should have an advantage in their linguistic and interpreting performance relative to low-WM peers, since high-WM learners are more likely to have sufficient availability to meet task demands.

2.5 Research Gap

Studies investigating the influence of language context and learners' cognitive resources generally adopt L2 processing tasks under normal conditions, such as spontaneous speech or reading comprehension. These studies mainly reveal that SAE is generally beneficial for bilinguals' L1 inhibition and L2 processing (e.g., Baus et al., 2013; Linck et al., 2009; Sagarra & Labrozzi, 2018). However, less attention has been paid to the effect of L2 context and WM resources on interpreting students' linguistic and interpreting performance. Interpreting distinguishes itself from other linguistic tasks by its multi-tasking nature, which imposes heavy pressure on interpreters' working knowledge of languages and their WM (Gile, 2009, p. 222). Additionally, interpreters are a special subset of bilinguals because they must keep their two working languages activated even when in a monolingual environment (García, 2019, p. 199; Grosjean, 1997). Therefore, due to the intricacies and complexities of interpreting, it is essential to examine linguistic and cognitive variables (and their interactions) in order to develop a holistic understanding of the interpreting process in the study abroad context.

To address this need, the current study advances our knowledge by comparing linguistic and CI performance, and to compare the cognitive processing abilities of two groups of interpreting students with and without SAE. The aim of this thesis is to explore the relationships between SAE, linguistic processing, CI performance and WM in interpreting students living in different language environments.

Based on the literature, the theoretical framework and research gap identified above, the research questions and the corresponding hypotheses, as well as the research methods adopted for the present study, will be presented in Chapter 3.

2.6 Summary

This chapter has reviewed theories and studies concerning three critical areas. Firstly, Section 2.1 addresses the relationship between SAE and SLA, with a particular focus on the two components of linguistic performance, lexical and grammatical processing, and the influence of SAE on these two components. In relation to lexical processing, this section describes three models in order to demonstrate how language is processed. Bilingual linguistic processing takes place through inhibiting the 'unintended' language and only keeping the intended one activated. During this process, inhibition costs occur, which are reflected as direction-dependent translational asymmetry in lexical processing tasks. Additionally, lexical selection is closely linked to aspects of grammatical processing such as morphological encoding. The efficiency and automaticity of both components of linguistic processing are assumed to be influenced by the language environment. SAE is an L1 inhibitory environment which also provides rich and natural L2 input, interaction and output opportunities. Thus, such an experience is considered beneficial for L2 processing. This view has been partially confirmed by many empirical studies reviewed above. Inconsistent results mainly appeared in grammatical development, suggesting an interactive effect from learners' WM resources in this linguistic domain.

Secondly, Section 2.2 reviews the WM multi-component model proposed by Baddeley and Hitch (1974). Three broad categories of WM tasks were also reviewed in this section, followed by empirical studies concerning the relationship between WM and SLA as well as the potential linguistic benefits of WM and SAE. To date, the results of studies examining the relationship between WM, linguistic processing and SA contexts have been mixed. Some studies suggest that the SA environment cancels out the effect of WM, whereas others indicate an interactive effect from both immersion context and WM.

Section 2.3 reviews a unique linguistic task: consecutive interpreting, which has the appearance of spontaneous speech but is unique in terms of high demands on working languages and WM, and many interpreting cognitive-processing models are essentially centred around the WM system. Empirical studies regarding the influence of WM on interpreting performance have enjoyed a growing interest among interpreting researchers in recent decades. The majority of these studies support the view that WM is linked to interpreting performance. A similar positive relationship has also been identified between linguistic processing efficiency and overall interpreting performance. However, empirical studies about interpreting performance take place mainly in at-home or laboratory contexts, with few studies exploring the effect of WM on interpreting learners' CI performance in a study abroad context.

The theoretical framework that underpins the present study is described in Section 2.4, followed by the research gap being identified in Section 2.5. Based on the research gap and the literature review, the research questions are presented in Chapter 3. The corresponding research hypotheses and the research methods adopted for this study will also be introduced in the next chapter.

Chapter 3 Research Questions and Methodology

Based on the literature review in Chapter 2, three research questions are proposed here that will guide the research in this thesis in an effort to fill some of the gaps in the literature. This chapter also includes a description of the methodology adopted to answer the research questions. The chapter consists of five sections (1) research questions and hypotheses, (2) research design, (3) ethics approval, (4) method of data analysis, and (5) summary. First, the research questions are put forward in Section 3.1. The questions concern the impact of study abroad experience (SAE) on working memory (WM), lexical retrieval, and consecutive interpreting (CI) performance. For each research question, we present several hypotheses to be tested in the present study. Second, an interdisciplinary design is proposed in Section 3.2 in order to explore the research questions and the corresponding hypotheses. We outline the experimental methods used, which include two psycholinguistic experiments (reading span and word translation recognition tasks), interpreting tasks in both directions (i.e., CI from L1 to L2 and the opposite direction), a self-report questionnaire measuring language experience and proficiency (LEAP-Q), and an online vocabulary knowledge test (LexTALE). A detailed description of stimuli and materials is also provided, followed by descriptions of the experimental procedures, details about participants and a description of task grading. Section 3.3 discusses ethical considerations, and Section 3.4 briefly illustrates the methods used for data analysis. Finally, the chapter ends with a summary in Section 3.5.

3.1 Research Questions and Hypotheses

Based on the literature review in Chapter 2, this thesis seeks to address three research questions:

(Q1) Compared with interpreting students without SAE, do interpreting students with SAE demonstrate better linguistic performance (i.e. better lexical and grammatical processing and fluency)?

(Q2) Compared with interpreting students without SAE, do interpreting students with SAE exhibit better CI performance?

(Q3) Does WM affect performance on linguistic and CI tasks?

The above research questions will be addressed by the following hypotheses.

For Research Question (Q1):

Hypothesis 1:

Students with SAE will outperform students without SAE:

SA students will respond faster, be more accurate, and more balanced (i.e., exhibit less asymmetry between the two directions of translation) in the word translation recognition task; they will have higher subject-verb number agreement rates and be more fluent in their L2 production in L1-L2 CI output.

This hypothesis is based on the Interaction Theory and the Inhibitory Control model (discussed in Section 2.1). SAE provides participants with abundant opportunities to interact with native speakers in the target language, and according to the Interaction Theory (S. E. Carroll, 1999), frequent target-language interaction contributes to success in many areas of linguistic learning and processing, such as listening comprehension, fluency, and consolidation and internalisation (automatisation) of existing linguistic knowledge. According to the Inhibitory Control model (Green, 1998), the processing of the intended language is achieved by suppressing the co-activated but unwanted language. Inhibiting the dominant L1 is more effortful than inhibiting the weaker L2, but SAE can effectively attenuate learners' dominant L1 and promote L2 processing (e.g., Baus et al., 2013; Linck et al., 2009; Sagarra & Labrozzi, 2018). Therefore, we hypothesised that SAE will automatise parts of interpreting students' lexico-grammatical performance through natural practice with other speakers, which makes their processing of L2 more efficient than participants who have not been to a country where the target language is dominant.

For Research Question (Q2):

Hypothesis 2:

Interpreting students with SAE will demonstrate better CI performance than their NSA counterparts, as measured by the assessments of two expert raters.

This hypothesis is based on the interpreting Effort Model (Gile, 1995, 2009). According to the model, the total cognitive resource required for an interpreting task is the sum of the cognitive resources required for each of the sub-processes. An enhancement in linguistic processing automaticity is believed to facilitate quality interpreting performance. As elements of linguistic processing become automatised, more cognitive resources can be directed to other elements or other cognitively demanding tasks (Antoniou et al., 2015; Christoffels et al., 2003; O'Brien et al., 2006). If the previous hypothesis is correct, that SAE benefits linguistic performance, then it is predicted that CI will also be influenced by participants' language environments. In other words, students with SAE will demonstrate better CI performance than their counterparts in the L1 context.

For Research Question (Q3):

Hypothesis 3:

WM resource availability will significantly correlate with linguistic and CI performance.

Hypothesis 4:

There will be a tendency for high WM participants with SAE to outperform the rest of the participants in linguistic and CI tasks.

Hypothesis 5:

WM resource availability in the SA group will be larger than that in the NSA group.

These hypotheses are based on the WM multi-component model proposed by Baddeley and Hitch (1974), which suggests that WM is a finite cognitive resource which is responsible for inhibition, activation, information storage and manipulation. These functions are believed to be important for both linguistic and CI performance. Furthermore, WM is also assumed to be a flexible resource which can be allocated to specific linguistic tasks so as to achieve an optimal processing result (Gathercole & Baddeley, 1993, p. 99; Kormos, 2006; Robinson, 2003). Therefore, if SAE promotes the efficiency of linguistic processing as predicted above, WM resources required for those linguistic tasks can be saved for other purposes which is assumed to contribute to

CI performance. However, regarding Hypothesis 4, the design of this research does not permit a definitive response, but it allows us, as a first step, to speculate about how this complex relationship works, which will provide a possible direction for future research.

3.2 Research Design

As mentioned earlier, the purpose of this study is to explore the effect of SAE and WM on interpreting students' linguistic and CI performance.

In order to fulfil the purpose of the study, two groups of Chinese-English (L1-L2) interpreting students (n=25/group) matched in age, interpreting training history, highest level of education and language proficiency were recruited. The students in one group had more than 24 months' SAE on average, and the students in the other group had no SAE.

Data regarding participants' English vocabulary knowledge, lexical processing (Chinese-English and English-Chinese) accuracy and response time, English grammatical accuracy and fluency in Chinese-English (L1-L2) CI production were collected as indicators of linguistics performance. Their overall CI scores in two language directions were used as indicators of CI performance. Apart from that, participants' WM and language experience were also collected for later analysis.

In the first task, participants' lexical processing accuracy and response times were measured through their accuracy and speed in judging Yes/No to 60 English and 60 Chinese word translation equivalents.

In the second task, how participants' WM was used in their L2 was measured by the English reading span task which was developed by Daneman and Carpenter (1980).

In the third task, participants completed an online lexical decision task by judging whether strings of letters on the screen were English words or not. This was to test their lexical knowledge and this task was not timed.

In the fourth task, participants were presented with auditory material via headphones and they were asked to undertake CI tasks. Their performances were recorded with a voice recorder for later analysis.

In the last task, participants filled out a Language Experience and Proficiency Questionnaire which asked them about their daily language usage and language learning history.

Each of these tasks is described in detail in Section 3.2.1.

3.2.1 Stimuli and Materials

3.2.1.1 Word Translation Recognition

In this study, participants' lexical processing performances were measured by word translation recognition tasks presented via E-Prime Professional 2.0 and Chronos. Participants' response times (ms) and accuracy (%) were recorded and used for further analysis. Word translation recognition tasks were administered in both language directions (English-Chinese and Chinese-English), and the stimuli consisted of 60 English and 60 Chinese auditorily presented words. On each trial, participants heard a word through headphones, and either its correct translational equivalent or an incorrect misleading word/non-word was presented onscreen using printed text. Participants were asked to indicate whether or not the word/non-word on the screen was the correct translation of the word presented via headphones by pressing buttons labelled Yes or No as quickly and accurately as possible. There was no response time limit.

Lexical accessing is word frequency-dependent (Levelt, 2001; Levelt et al., 1999), such that more commonly used words are responded to more quickly than words that we rarely encounter (Gile, 2009, p. 230; Matthei & Roeper, 1985). In other words, the frequency of occurrence of a word in a language influences the time it takes to retrieve that word in the mental lexicon. Previous studies also suggest that word length and definition accuracy also influence the process of lexical access (De Groot et al., 2002). Therefore, this study controlled word frequency, word length and definition accuracy across stimuli and target translation equivalents in both language directions.

The English and Chinese words used in this task were taken from the SUBTLEX-UK and SUBTLEX-CH word frequency databases respectively. As shown in Table 3.1, the mean frequency of auditory English stimuli in the English-Chinese lexical task was 3.6 per million in a corpus of 160 thousand words (Van Heuven et al., 2014), the mean frequency of their Chinese target translation equivalents in this direction was 2.4 per million in a 33.5-million-character database (Q. Cai & Brysbaert, 2010), and the mean frequency of their Chinese incorrect misleading words was 2.4 per million in the same database.

In the other direction, the mean frequency of auditory Chinese stimuli in Chinese-English word translation recognition task was 2.1 per million, the mean frequency of the English target translation equivalents was 3.5 per million, and the frequency of incorrect misleading English word/non-word was 3.6 per million. Therefore, the frequency of misleading words was similar to the frequency of target translation equivalents in both language directions. This similarity of frequency suggests that differences in accuracy or response time while performing the task are not due to differences in word frequency between the misleading words and target translation equivalents but a true reflection of participants' lexical processing.

Table 3.1. *Mean word frequency of stimuli.*

Word frequency	English-Chinese	Chinese-English
Audio presented words	3.6	2.3
Written misleading words	2.4	3.2
Target translation equivalents	2.4	3.5

The Chinese words used in both directions consisted of similar numbers of characters, and the lengths of the English words used in this task were also similar.

Table 3.2. shows the average word length in English and the average number of Chinese characters in this task. In English-Chinese direction, the English words that participants heard consisted of seven letters on average, and the number of Chinese characters that appeared on the screen that were meant to mislead participants was two on average, and the average number of characters in correct Chinese translations also consisted of two characters. In the opposite direction, the average word length of Chinese words that participants heard through the headphones was two Chinese characters, and the English words that were meant to mislead participants consisted of seven letters on average. The average word length of the correct English word translations was also seven letters. It is clear that both the misleading words and the

correct translation equivalents adopted in this task had similar word lengths (two characters in Chinese and seven letters in English); therefore, it is reasonable to assume that any inaccurate or slow judgments arose, not because of differences in word length between the misleading words and the target translation equivalents, but because of participants' lexical processing.

Table 3.2. *Mean word length of stimuli.*

Word length	English-Chinese	Chinese-English
Audio presented words	7	2
Written misleading words	2	7
Target translation equivalents	2	7

Further information on these stimuli can be found in Appendix A. Stimuli were presented in random order for each participant. In each trial, words were triggered as soon as participants made the decision and pressed the button.

3.2.1.2 Working Memory Resource Availability

The English reading span task (Daneman & Carpenter, 1980) was administered via E-Prime Professional 2.0 to measure participants' WM resource availability. The reason for using a reading span task to tap WM is that it is a complex storage-and-processing task (Daneman & Hannon, 2007), which involves simultaneous information processing and storage; hence, this task involves processing akin to what interpreters must do when they are interpreting (Signorelli & Obler, 2012). As introduced in Section 2.2.2 WM Tests, in this reading span task, sentences were presented one by one in the centre on a computer screen. Each sentence contained 11 to 13 words, and each sentence ended with a different word. As the test progressed, the number of sentences within each set increased from initially two to a maximum of five. There were three sets of sentences for each set size. Participants were instructed to read aloud each sentence in the set and make a decision on the semantic plausibility of that sentence. Half of the experimental sentences in this test were plausible, while the other half were implausible. In addition, at the end of each set, participants were asked to recall as many of the sentences' final words as possible. There were no restrictions on recall order or recall time. When participants were ready to proceed, they were instructed to press any key to start the next set. In total, there were a maximum of 42 sentences (requiring yes/no decision-making) plus 42 recall words in this task (e.g., Christoffels et al., 2003; M. Liu, 2001; Tzou, 2008). This task constitutes a concurrent load on WM because the words from previous sentences must be stored and maintained during comprehension of the current sentence (M. Liu, 2001), a process which resembles the process of interpreting.

The sentences of the English task were partly derived from Harrington and Sawyer (1992) (Appendix B). All of them were written by native English speakers and contained no technical terms and were simple in structure and content, as in (4a-b).

(4) a. Drinking a lot of liquor could prevent us from getting sunstroke. (Make sense? No)

b. All of us should be treated with patience, respect and kindness. (Make sense? Yes)

3.2.1.3 Consecutive Interpreting Performance

In the CI task, materials were presented auditorily to each participant via headphones. Participants were required to translate the source language message into the target language, and their spoken responses in the target language were recorded using a portable digital audio voice recorder.

CI passages were about general and topical current affairs and involved little specialised knowledge (see Appendix C). This was to ensure that participants would be familiar with the topics. In both directions, recordings contained four parts in total, with only the last three parts used for data analysis in this study and the first part considered as a familiarisation phase.

The experimental CI (Chinese-English) material-recording took 2 minutes and 15 seconds to listen to. It concerned the emergence of new energy cars. Three Chinese native speakers who were of similar educational level to our recruited participants rated the readability and level of comprehension difficulty of the Chinese material on a scale from 1 (very easy to understand) to 5 (very complex to understand). The raters judged the materials to be moderately easy/complex to understand, assigning a mean score of 3, with an inter-rater reliability of 0.9 on Kappa. Additionally, none of the participants recruited for the study reported having encountered any difficult terminology after finishing the experiment. A native speaker of Chinese (Mandarin) recorded the speech material in a sound-proof booth, and her delivery rate was on average 180 words per minute.

The experimental CI (English-Chinese) task-recording, likewise, was 2 minutes and 8 seconds long. The text was about a brief overview of artificial intelligence. Three English native speakers who were of similar educational level to our participants rated the readability and comprehension difficulty of the English materials on a scale from 1 to 5. The inter-rater reliability was 0.8 on Kappa, that material was assessed to be not too easy or too difficult to understand, with a mean score of 3. Once again, none of the participants reported having encountered difficult terminology after finishing the experiment. A native speaker of English produced the speech material in a sound-proof booth, and her delivery rate was on average 180 syllables per minute.

The two tasks were presented at a faster speed than considered as comfortable for SI (120 words/minute) (De Groot, 2000; Pöchhacker, 2004), but was kept below 200 words/minute, so as not to compromise comprehension (Griffiths, 1992).This is because, firstly, the present study is of CI, and the participants did not have to produce while listening. Secondly, the researcher intended to include some difficulties in the task to challenge the participants. As shown by previous research, delivery speed is an influential factor in interpreting performance (Gerver, 1969; Han & Riazi, 2017; M. Liu, 2001; W. Zhang, 2009).

Even though measurements of the difficulty of a source text is elusive and prone to subjective judgement (Hale & Campbell, 2002), and comparison of text difficulty across language modalities is even harder (J. Wang, 2016), every effort was made in the present study to make sure that the two source texts were comparable in terms of text type (monologic presentation), register (formal), context (speech), readability, length and delivery speed.

Participants were not given time to prepare for the interpreting task, and the background information about the interpreting materials was explained just before the task. Participants were administered CI tasks in counterbalanced order to control for fatigue which may have influenced performance at the group level. For example, the first participant received the Chinese-English CI before the English-Chinese CI, and the next participant received the English-Chinese CI before the Chinese-English CI, and so on.

3.2.1.4 Vocabulary Knowledge and Language Experience

Participants also completed LexTALE, a valid and standardised online test of English vocabulary knowledge widely used in linguistic studies (e.g., Christoffels et al., 2007; De Bruin et al., 2014; Keuleers et al., 2015; Ramscar et al., 2014). This task required participants to judge whether a string of letters that appeared on a screen was or was not an English word. The score is calculated immediately and shown on the screen upon completion of the task. This task has been shown to provide an accurate indication of general English proficiency (Lemhöfer & Broersma, 2012). Further details on LexTALE can be found in Appendix D.

Participants' language experience was self-reported by using the Language Experience and Proficiency Questionnaire (LEAP-Q). This questionnaire has been established as a valid, reliable and efficient tool for assessing the language profiles of participants in research settings (Kaushanskaya et al., 2019; Marian et al., 2007), and has been widely used in linguistic and psycholinguistic research (e.g., Bialystok et al., 2009; Maitreyee & Goswami, 2009). Of particular interest to the present thesis, LEAP-Q assesses individuals' language-related information including language learning history, language usage, language exposure, and proficiency. Participants first reported their gender, age, age of L2 acquisition, and years of L2 study. The self-report measure also captured detailed information about L2 learning and self-ratings of participants' proficiency in speaking, listening and reading in their L2 using a ten-point scale (1= very little knowledge, 10 = like a native speaker). Participants were also asked to estimate how many hours they spent each day interacting with friends, and watching, reading, or listening to media (e.g., TV, written texts, radio) and in self-instruction in the L2. Further information on the LEAP-Q can be found in Appendix E.

3.2.2 Procedures

Participants gave informed consent to participate in the research (see Appendix F). Each session was conducted individually within a quiet testing space at a university location convenient to each participant. Participants were informed that they were free to withdraw from the research study at any time and without penalty.

All verbal instructions were given in the participants' native language (Chinese).

Participants completed the tasks in the following order: word translation recognition task for measuring lexical processing, reading span task for WM resource availability, LexTALE online task of vocabulary knowledge, CI tasks (counterbalanced order), and finally the LEAP-Q for assessing language experience and proficiency.

The experiment was run on a Lenovo laptop. The word translation recognition and reading span tasks were presented via E-Prime Professional 2.0 software. In the word translation recognition task, a Chronos response box was used to record participants' response times and accuracy. All tasks were conducted in one session that lasted approximately 50 minutes as illustrated in Table 3.3.

Table 3.3. *Procedures of the tasks in the present study.*

Word translation recognition (two directions)	Reading span task (L2 only)	LexTALE (L2 only)	CI task (two directions)	LEAP-Q
10 mins	10 mins	5 mins	15 mins	10 mins

3.2.3 Participants

3.2.3.1 Sample Size Justification

We estimated the required sample size by performing a power analysis based on the data of Cai et al. (2015) with an effect size (dz) of 0.5916 in a CI task comparable to that used in the present study. In order to achieve 95% power with an α of .05 for a *t*-test (comparing the group members' learning), we concluded that we required a total sample size of 37. We recruited additional participants to ensure we had sufficient power to detect any statistical differences in interpreting performance.

The sample pool of the present study consisted of 50 unbalanced bilinguals, with Chinese as their L1 and English as their L2, and this size is deemed sufficient in order to yield meaningful results for analysis.

3.2.3.2 Recruitment Criteria

Flyers with recruitment criteria were posted on the university's bulletin boards. The brief introduction of the study, the duration of the experiment, and participants' recruitment criteria were all clearly stipulated. The researcher's email and mobile number were also listed on flyers. Prospective participants who were interested in the research contacted the researcher and expressed their willingness to participate in the research. In order to be recruited in the study, participants needed to be full-time master's students majoring in interpreting and translation studies. Half of the participants were enrolled in their master's degrees in Sydney, Australia, a predominantly English-speaking environment, for at least one year (SA group). The other half were studying in Xi'an, China, and had never lived in an English-speaking country (NSA group), because staying even a short period of time in a foreign country is sufficient for language improvement (Labrozzi, 2009; Llanes & Muñoz, 2009). Thus, all participants had been admitted into their respective universities with a required IELTS (International English Language Testing System) academic score of $\geq$ 6.5 or had passed their entrance examination for a master's degree. Recruitment criteria also stipulated that participants should be either at the end of their second year or the beginning of their third year of interpreting training. In order to maintain objectivity, the 25 SA participants were recruited randomly from three different universities in Sydney, and likewise the other 25 NSA participants were recruited randomly from three different universities in Xi'an. Given the recruitment had specific criteria, it wasn't possible to do strict random sampling but quasi-random sampling

To be included in this study, participants were required to be native Chinese speakers between 20 and 30 years of age. This is because age is known to influence WM and processing speed (Signorelli et al., 2011). For instance, Alloway and Alloway (2013) observed that WM peaks in the 20s and declines after that. Furthermore, all of the participants recruited in the present study had to begin learning English as a compulsory L2 course, starting at the age of 12 in middle school. This was to make sure that all of them began learning English post-puberty, once their L1 had been established (e.g., Crain & Lillo-Martin, 1999; Sagarra & Herschensohn, 2010).

3.2.4 Task Grading

3.2.4.1 Word Translation Recognition

Word translation recognition tasks were adopted to measure participants' lexical processing performance. Accuracy (%) and response times (ms) of word translation recognition tasks were calculated by E-Prime Professional 2.0, and only the response times of correct trials were included in the data analysis. The response times were recorded to the nearest millisecond. Participants were instructed to make their responses as quickly and accurately as possible.

3.2.4.2 Reading Span

The author conducted reading span to measure participants' WM resource availability. In calculating participants' reading spans, two criteria were used: (a) correctly judged sentences and (b) correctly recalled final words. This follows the scoring protocol of Kroll et al. (2002). The span score was calculated according to the total number of words participants could recall in trials for which the plausibility judgment was correct. If the participant did not judge the sentence correctly but recalled the final word, this word was not included in the total number of items recalled. In other words, participants' recalled final words were considered valid only when they also made correct judgments on the plausibility of sentences. The number of validly recalled final words was used as a measure of a participant's WM span.

3.2.4.3 Consecutive Interpreting Performance

Participants' CI performances in the L1-L2 direction were assessed in two tiers, with both objective and subjective measurements involved. The objective measurements included grammatical accuracy (subject-verb number agreement) rate and fluency of L2 production. The subjective measurement was based on their overall interpreting performance in both language directions.

3.2.4.3.1 Subject-Verb Number Agreement Scoring

Participants' English output for Chinese-English (L1-L2) CI was analysed to see whether interpreting students with SAE might be more accurate in marking subject-verb number agreement than those without SAE. Participants' verbatim transcriptions of their English production in L1-L2 CI were used for analysis.

This is an interpreting task which requires delivery of the source language message in the target language (English), instead of simply speaking English; thus the identification of correct or incorrect target grammatical form (subject-verb number agreement) could be ambiguous, as can be appreciated in (4a-c).

(4) a. ***中国****已经在新能源汽车这一领域实现了快速发展。*

b. China has achieved rapid development in electric vehicles. (third person singular form, correct)

c. Many countries have achieved rapid development in electric vehicles.

Sentence (4a) is from the Chinese source text. Sentence (4b) is a translation of Sentence (4a). Sentence (4c) is the translation from one of the participants. Although this participant translated 'China' as 'many countries', the plural form of verb 'have' nevertheless agrees with the plural form of subject 'many countries' in the sentence; therefore, no subject-verb disagreement was detected in this sentence. In other words, the researcher focused on participants' actual production in measuring the subject-verb number agreement.

In assessing participants' performance accuracy, their state of acquisition should first be determined. One well-known theory of SLA is Processability Theory (Pienemann, 1998) which asserts that learners can only acquire what they can process. Processability Theory makes a strong distinction between acquisition and accuracy in performance. Accuracy is not a measure of acquisition but a measure of performance. What defines acquisition is the emergence criterion (Di Biase & Kawaguchi, 2002) which indicates that a structure is acquired when the oral production of the learner shows lexical and formal variation with that structure. For instance, in the acquisition of English subject-verb number agreement in the third person singular marking with -

s, emergence would require that the -*s* appears in at least two different verbs (e.g., *she cook*s, *he eats*). This satisfies the requirement for appropriate 'lexical' variation. Formal variation is also required for the emergence criterion, which means that a verb appears in contexts requiring marking (suffix -*s*) and others that do not (e.g., *she cook*s; *they cook*). However, even when the criterion is satisfied, a learner's performance often varies between use and non-use of the structure they have just acquired. Or learners may even generalise the marking to inappropriate positions (e.g., *We often cooks*). But this variation does not invalidate their acquisition. After determining a participant's state of acquisition, their performance accuracy could be investigated.

Obligatory uses of the targeted verb form were identified for the participants of each group. All verbs were coded and the subject-verb number agreement marking by participants was counted. At the coding phase the transcripts were double-checked, and error identification was performed by the researcher. A native speaker of English in the researcher's university did an inter-rater reliability check on the error identification. The initial agreement rate was 90%, and a discussion of discrepancies resulted in 100% agreement.

To examine participants' subject-verb number agreement accuracy, a percentage of correct usage was calculated (Bitchener, 2008; Kao, 2019). The following equation shows how the percentage accuracy was calculated.

$$\frac{\text{number of correct usage in contexts}}{\text{number of obligatory contexts}} \times 100 = \text{percent accuacy}$$

Accuracy rate was calculated as a percentage of correct usage for the target form given the range of obligatory occasions arising in participants' L2 outputs. For instance, if a participant produced eight correct uses of the target forms of verbs from ten obligatory occasions, their accuracy rate would be 80%.

3.2.4.3.2 Fluency

For interpreting fluency, the present study employed the two critical parameters as observable and specific measures: articulatory rate and filled pauses. These two parameters have been shown to be effective in representing fluency by previous researchers both in SLA and interpreting studies (e.g., Han, 2015b; Juan-Garau, 2018;

Lin et al., 2018; Llanes & Muñoz, 2009; Marqués-Pascual, 2011; Pérez-Vidal & Juan-Garau, 2011; Rennert, 2010, p. 200; Yu & Van Heuven, 2017).

The articulatory rate was the total number of English (L2) syllables produced by participants in one minute, including pauses. The number of syllables was counted by the researcher based on the transcripts of each participant's output. At the counting phase the transcripts were double-checked. Any repetitions performed by the participants were counted only once. The same principle was applied in the case of false starts and rephrasing. That is, the only syllables that were counted were the ones in the final clause. The average rate was calculated by dividing the total number of syllables by the total time in minutes that elapsed from the beginning to end of each participant's speech. Using the syllable as a standard unit of measurement here is in line with Pöchhacker's (1993) recommendation that this can avoid the confusion caused by the variability in word length across different languages, as 'fluent' is a relative rather than absolute sense (Færch et al., 1984, p. 142; W. Zhang, 2009). It is, for example, logically possible for an individual to exhibit slower production than others if this person is a slow talker, even in his/her native language. Hence, in order to minimise the risk of confusing a slow speech rate with poor fluency in English (L2), we also measured participants' rate of producing Chinese (L1) syllables per minute in the L2-L1 CI task and correlated this with their L2 fluency. The L1 average rate was calculated in the same manner, by dividing the total number of Chinese characters by the total time in minutes that elapsed from the beginning to the end of each participant's speech.

Measurements of filled pauses per minute were based on participants' L2 outputs. Firstly, the researcher counted the number of filled pauses based on the transcripts of each participant's output. At the counting phase the transcripts were double-checked. Secondly, the number of filled pauses per minute was calculated by dividing the total number of filled pauses by the total time in minutes that elapsed from the beginning to the end of each participant's speech. A filled pause takes place when the speaker fills his/her speech with non-lexical fillers such as mmm, ah or um.

3.2.4.3.3 Overall Grading

To obtain an overall evaluation of CI performance, two experienced professional NAATI (National Accreditation Authority for Translators and Interpreters) assessors were invited to listen to participants' interpretation output in both language directions and they rated the participants' performances independently on a holistic basis. Both assessors were native Chinese speakers who had been practising, teaching and assessing interpreting in an English-speaking country for over 10 years. The assessors rated participants' performances on a scale from 1 (low) to 10 (high) according to their information accuracy, language quality, comprehension, expression, and delivery. Participants' accents were not counted in the ratings. Half points were used in the ratings.

3.3 Human Research Ethics Approval

The research reported within this thesis was conducted in accordance with the ethical guidelines issued by the Western Sydney University Human Research Ethics Committee (approval number H12405). The privacy and confidentiality of participants were respected throughout the research process. Recruitment flyers were posted on the university's bulletin boards. The nature of the study, the duration of the experiment, and the criteria for participation were clearly stipulated. The researcher's email and mobile number were also listed on flyers. Prospective participants who were interested in the research contacted the researcher and expressed their willingness to participate in the research. Participants were assured that participation or non-participation would not in any way impact on their studies at their universities. On the day of the experiment, participants were provided with copies of Participant Information Sheet (Appendix G) and Consent Form (Appendix F). All participants were contacted through email or mobile phone, and they signed consent forms before participating in the study. All participants were assigned an anonymised number to ensure their identity would not be revealed, and after the experiment they were offered $AUD25 or 100 RMB as a token of appreciation for their effort and time.

3.4 Data Analysis

The measures of CI performance, WM, and lexical processing were examined individually. To analyse the relationship between SAE and CI performance, the author conducted a 2 × (2) mixed factorial ANOVA with the between-subjects factor of group (NSA group and SA group) and the within-subjects factor of direction of interpreting (L2-L1 vs. L1-L2). Also, to analyse the relationship between SAE and lexical processing accuracy and response time, the author conducted a 2 × (2) mixed factorial

ANOVA with the between-subjects factor of group and the within-subjects factor of direction of lexical retrieval (L2-L1 vs. L1-L2). Pairwise comparisons were also conducted to explore significant interactions between the factors. Furthermore, the measures of CI performance, WM, and lexical processing were combined in a correlational analysis to explore the relationship between WM and CI performance, and the relationship between lexical processing and CI performance. Descriptive statistics for each combination of factors were computed. An α level of .05 was set for all analyses. In exploring the relationships between WM, SAE, subject-verb number agreement, and fluency, a correlational analysis was conducted for each variable. Statistical analyses will be discussed in detail in Chapter 4.

3.5 Summary

On the basis of the literature reviewed in the previous chapter, this chapter has presented three research questions on (1) the effect of SAE on linguistic performance, (2) the effect of SAE on CI performance, and (3) the correlation between WM and linguistic and CI performance in different language contexts.

To answer the research questions and their corresponding hypotheses, an inter-disciplinary design was developed, involving two psycholinguistic experiments, CI tasks with two directions, a self-report questionnaire survey and an online vocabulary test. This chapter describes all the methodological aspects of the study, including materials, procedures, sample selection, grouping, tasks grading and ethics considerations.

This chapter ends with a brief introduction to data analysis, illustrating how the abovementioned factors are analysed to answer the research questions.

Chapter 4 Results and Data Analysis

This chapter presents the research data gathered from the experiments outlined in Chapter 3 and examines how these data address the research questions and their corresponding hypotheses specified in Section 3.1. Fifty interpreting students with different language environment completed a series of experimental tasks, including a working memory (WM) test (reading span test), a lexical processing test (word translation recognition test) in both language directions, a consecutive interpreting (CI) test, also in both language directions, a vocabulary knowledge online test (LexTALE), and a language experience questionnaire (LEAP-Q). To control for variables other than those under investigation, participants were matched in terms of their age, age of onset of L2 learning, major of their university studies, and the length of interpreting learning at the time of data collection. This chapter has three sections: Section 4.1 examines participants' language and training backgrounds by presenting the data collected from LEAP-Q and the LexTALE online vocabulary test. Experimental results, including correlation and regression analyses, are presented in Section 4.2. Finally, the chapter ends with a summary in Section 4.3.

4.1 Participants' Language and Training Backgrounds

Data regarding participants' language and training backgrounds were collected from the LEAP-Q questionnaire and LexTALE online test. As mentioned in Chapter 3, all participants were native Chinese speakers who had been admitted into universities in either Australia (SA group, n = 25) or China (NSA group, n = 25), and all of them were full-time master's students majoring in Chinese-English Interpreting and Translation studies either at the end of their second year or the beginning of their third year at the time of data collection.

Group-level data for variables of interest that are known to modulate interpreting performance are reported in Table 4.1. These variables include participants' age, semesters of interpreting learning, age at commencement of English (L2) learning, vocabulary knowledge, self-rated L2 proficiency, hours per week spent interacting with English native speakers, watching TV, reading or listening to media (e.g., TV, written texts, radio) and self-instruction in the L2.

To determine whether the two groups of participants were comparable in terms of L2 listening, speaking, reading and vocabulary knowledge, a series of independent samples *t*-tests were conducted. The results suggested statistical non-significance between the groups in terms of their self-rated L2 abilities for speaking, $t(24) = .97$, $p = .34$, listening, $t(24) = .00$, $p = 1.00$, reading, $t(24) = 1.00$, $p = .41$, and vocabulary knowledge, $t(24) = -.85$, $p = .33$. These findings confirm that at the time of data collection, the SA and NSA groups were matched in terms of their L2 speaking, listening, reading and vocabulary knowledge. The two groups were also matched in their age, length of interpreting training, and age of onset of L2 learning. Thus, any group statistical difference observed in the experimental tasks cannot be accounted for by any of the aforementioned variables and are likely to be due to differences in language contexts. It is noteworthy that when looking at the frequency of participants' L2 activities per week, no significant differences were identified between the two groups' reported hours spent watching TV, $t(24) = 1.28$, $p = .21$, listening to radio/music, $t(24) = .20$, $p = .85$, reading, $t(24) = .51$, $p = .62$, or self-instruction, $t(24) = .00$, $p = 1.00$. The only statistical significance observed was in participants' hours of interacting with native speakers in English, $t(24) = 2.90$, $p = .008$, with the SA group spending more time interacting with English-speaking friends, as would be expected.

Table 4.1. *Characteristics of participant groups.*

Characteristic	SA Group (Mean)	NSA Group (Mean)
Number of participants	25.0	25.0
Age (years)	25.0	24.0
University semester of interpreting learning	4.0	4.0
Age of onset (learning L2)	12.1	12.1
Vocabulary knowledge	63.1	65.5
Self-report proficiency in		
L2 Speaking	6.1	5.7
L2 Listening	6.3	6.3
L2 Reading	6.9	6.6

Self-reported time spent	SA Group (Mean)	NSA Group (Mean)
L2 Interacting with English speakers	37.5**	19.9**
Watching TV	35.0	27.7
Listening to radio/music	35.7	35.0
Reading	42.0	39.9
Self-instruction	36.4	36.4

*$p < .05$, ** $p < .01$, *** $p < .001$

4.2 Results of Experimental Tasks

Experimental results and analyses of participants' WM, word translation recognition as well as their CI performances are presented in this section.

4.2.1 Descriptive Analysis

Descriptive analysis of participants' experimental data including WM resource availability, word translation recognition and CI performance in both directions are presented here. Table 4.2 provides the descriptive statistics for experimental tasks for the two groups. Descriptive statistics for the WM test were calculated from the participants' scores on the reading span task (Daneman & Carpenter, 1980). The mean WM availability for the two groups was similar (SA group: $M = 24.96$, $SD = 5.90$; NSA group: $M = 24.44$, $SD = 5.46$).

Lexical processing performance was measured using a word translation recognition task. This task required participants to judge as accurately and quickly as possible whether a word that appeared on a laptop screen in one language was a correct translation equivalent of the word presented through their headset in another language. The word translation recognition accuracy from the Chinese to English (L1-L2) direction (SA group: $M = 65\%$, $SD = 0.08$; NSA group: $M = 72\%$, $SD = 0.08$), and English to Chinese (L2-L1) direction (SA group: $M = 70\%$, $SD = 0.10$; NSA group: $M = 77\%$, $SD = 0.10$) are presented in Table 4.2. The L1-L2 word translation recognition response times (SA group: $M = 948.65$, $SD = 193.95$; NSA group: $M = 1167.87$, $SD =$

224.44), and the L2-L1 word translation recognition response times (SA group: *M* = 929.94, *SD* = 214.25; NSA group: *M* = 1032.53, *SD* = 205.76) are also presented in Table 4.2. Only the response times of correct trials were included in the data analysis.

The holistic scores of participants' CI performance were provided by two professional NAATI assessors. Both of them listened to and marked all 50 participants' CI performances in both directions on a scale from 1 (low) to 10 (high) based on the accuracy, language quality, understanding, expression and delivery of their interpreting recordings. Participants' accents in their production were not considered. Decimal points were used whenever necessary. Final scores used in checking inter-rater reliability were based on the rounded-up scores from each assessor. Inter-rater reliability agreement for English-Chinese CI (L2-L1 CI) was .80 for the SA group and .88 for the NSA group. The inter-rater reliability in Chinese-English CI (L1-L2 CI) agreement was .88 for the SA group, and .84 for the NSA group. Given such satisfactory levels of inter-rater reliability, the means of the two assessors' scores were used for the final analyses. As presented in Table 4.2, the mean score of the SA was higher than the mean score of the NSA in the L1-L2 CI direction (SA group: *M* = 5.62, *SD* = 0.23; NSA group: *M* = 5.39, *SD* = 0.25). The situation was the same with the other direction (L2-L1 CI), where the mean score of the SA was again higher than that of the NSA (SA group: *M* = 5.22, *SD* = 1.75; NSA group: *M* = 4.14, *SD* = 1.56).

4.2.2 Testing Assumptions of Parametric Statistics

In addition to descriptive statistics such as the mean (*M*) and standard deviation (*SD*), skewness and kurtosis are also reported in Table 4.2. Skewness is a measure of the asymmetry of a distribution, and kurtosis is a measure of how peaked or flat the data are relative to a normal distribution. It is important that these be between -2 and +2 to ensure that the data are normally distributed. Normal distribution is a prerequisite for running parametric statistical analyses such as *t*-tests and ANOVA.

As shown in Table 4.2, the value for skewness and kurtosis of all variables are within the normal range of -2 and +2, which indicates that the means are normally distributed in terms of symmetry and peakedness in a bell-shape curve, and that parametric statistical analyses can be conducted to find out whether the mean differences are significant.

Table 4.2. *Descriptive analysis of study abroad (SA) and non-study abroad (NSA) groups' working memory (WM) spans, word translation recognition accuracy (%) and response times (ms) in both language directions, and consecutive interpreting (CI) scores in both language directions.*

	Group	*M*	*SD*	Kurtosis	Skewness
WM					
	SA	24.96	5.90	-0.63	-0.04
	NSA	24.44	5.46	-1.09	0.31
Word Translation Recognition					
L1-L2 (% correct)					
	SA	65%	0.08	0.06	0.69
	NSA	72%	0.08	-0.72	0.17
L2-L1 (% correct)					
	SA	70%	0.10	-1.21	-0.20
	NSA	77%	0.10	0.39	-0.40
L1-L2 (ms)					
	SA	948.65	193.95	0.59	-0.03
	NSA	1167.87	224.44	-0.30	-0.20
L2-L1 (ms)					
	SA	929.94	214.25	0.31	0.25
	NSA	1032.53	205.76	-0.27	-0.31
CI					
L1-L2					
	SA	5.62	0.23	-0.45	0.38
	NSA	5.39	0.25	-0.38	0.44
L2-L1					
	SA	5.22	1.75	-0.67	0.40
	NSA	4.14	1.56	-1.11	0.66

4.2.3 Correlation and Regression Analyses

To analyse the relationships between variables, correlation and regression analyses were conducted. The correlations between WM, word translation recognition tasks and CI performance are presented in Table 4.3. The alpha level was set at .05 for all analyses.

Table 4.3. *Correlation between the performance on WM, word translation recognition and CI performance for SA and NSA groups.*

Task	L1-L2 CI		L2-L1 CI	
	SA	NSA	SA	NSA
WM	.58 **	.44 *	.56**	.53**
Lexical Processing				
L2-L1 (% correct)	.78***	.49*	.70***	.45*
L1-L2 (% correct)	.58 **	.38~	.36 ~	.24
L2-L1 (ms)	-.21	-.01	-.30	-.28
L1-L2 (ms)	-.22	-.33	-.21	-.22
Interpreting				
L2-L1 CI	.76***	.82***		

$*p < .05$, $** p < .01$, $*** p < .001$, $\sim .05 < p < .1$.

On inspection, Table 4.3 shows that WM was an important variable which correlated significantly with both directions of CI performance for both SA and NSA groups (L1-L2 CI: SA: $r = .58$, $p = .002$ vs. NSA: $r = .44$, $p = .03$; L2-L1 CI: SA: $r = .56$, p = .003 vs. NSA: $r = .53$, $p = .006$). This result suggests that interpreting students who have greater WM availability are more likely to perform well in both directions of CI, independently of their language environment.

L2-L1 word translation recognition accuracy also correlated significantly with both directions of CI for both SA and NSA groups (L1-L2 CI: SA: $r = .78$, $p = .000$ vs. NSA:

r = .49, p = .012; L2-L1 CI: SA: r = .70, p = .000 vs. NSA: r = .45, p = .023). In contrast, L1-L2 word translation recognition accuracy only correlated significantly with L1-L2 CI measures in the SA group (r = .58, p = .002). No other significant correlations were observed between L1-L2 word translation recognition accuracy with either direction of CI measures in the NSA group. These results indicate that better L2-L1 word translation recognition accuracy was linked to better CI performance in both directions, independently of the learning context, but this link was stronger in the SA group. Moreover, the link between the L1-L2 word translation recognition accuracy and CI performance only appeared in the SA group.

No significant correlations were observed between L2-L1 word translation recognition response times and CI measures in either CI direction, and this held for both groups (L1-L2 CI: SA: r = -.21, p = .31 vs. NSA: r = -.01, p = .95; L2-L1 CI: SA: r = -.30, p = .13 vs. NSA: r = -.28, p = .17). Similar results were also found between L1-L2 word translation recognition response times and CI performance (L1-L2 CI: SA: r = -.22, p = .27 vs. NSA: r = -.33, p = .112; L2-L1 CI: SA: r = -.21, p > .31 vs. NSA r = -.22, p = .28). These results indicate that speed in finding target translation equivalents had no significant impact on interpreting students' CI performance.

In terms of the two groups' L1-L2 and L2-L1 CI performance, a strong positive correlation was found between the two interpreting directions (SA: r = .76, p = .000; NSA: r = .82, p = .000), which indicated that interpreting students with strong CI performance in one language direction were likely to perform strongly in the other language direction as well.

4.2.4 Study Abroad Experience and Lexical Processing Performance

To analyse the relationship between SAE and lexical processing performance, the author conducted a 2 × (2) mixed factorial ANOVA with the between-subjects factor of group (NSA vs. SA) and the within-subjects factor of direction of word translation recognition (L2-L1 vs. L1-L2). A significant main effect of direction, $F(1, 48) = 35.02$, $p < .001$, $\eta_p^2 = .422$, indicated that overall participants from both groups were more accurate at retrieving words from L2 to L1 than from L1 to L2 (74% vs. 68%, respectively), as shown in Figure 4.1.

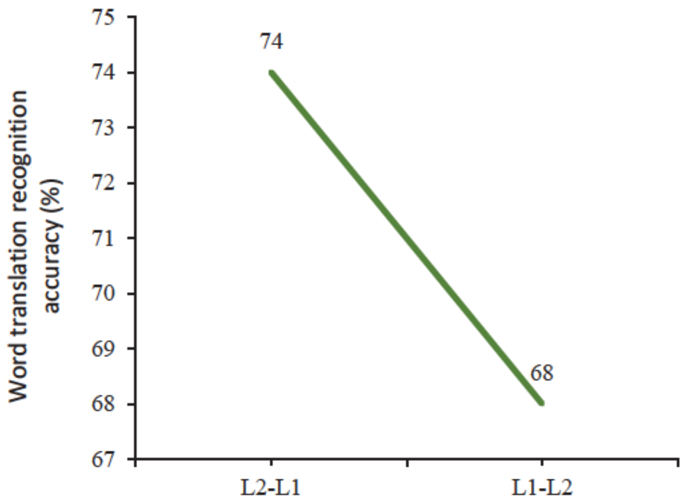


Figure 4.1. Overall word translation recognition accuracy (%) in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions across all participants regardless of language environment.

Table 4.4. provides the accuracy rates of SA and NSA groups in two translation recognition directions. A significant main effect of group, $F(1, 48) = 8.51, p = .005, \eta_p^2 = .151$, suggests that the NSA group outperformed the SA group in terms of accuracy in retrieving words in both directions (mean accuracy rate: 75% vs. 68%, respectively), as shown in Figure 4.2. The group × direction interaction was not significant, $F(1, 48) = .05, p = .83, \eta_p^2 = .001$.

Table 4.4. *Performance on word translation recognition accuracy (%) in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for SA and NSA groups.*

Group	L2-L1	L1-L2
SA	70%	65 %
NSA	77%	72 %

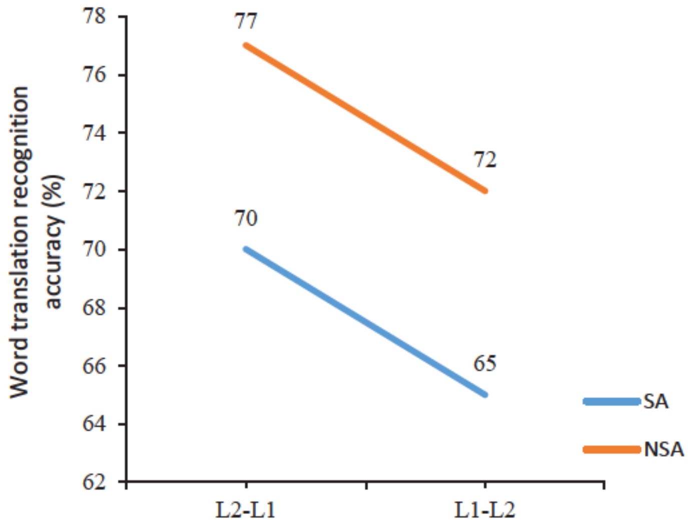


Figure 4.2. Word translation recognition accuracy (%) in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for SA and NSA groups.

To analyse the relationship between SAE and lexical processing response times, the author also conducted a 2 × (2) mixed factorial ANOVA with the between-subjects factor of group (NSA vs. SA) and the within-subjects factor of direction of word translation recognition (L2-L1 vs. L1-L2). A significant main effect of direction, $F(1, 48) = 8.96$, $p = .004$, $\eta_p^2 = .157$, confirmed that overall, participants were significantly faster at retrieving words from L2 to L1 than from L1 to L2 (981.2 vs. 1058.3, respectively), as depicted in Figure 4.3. This result is consistent with previous findings from word translation studies (e.g., R. Cai et al., 2015; Kroll et al., 2002).

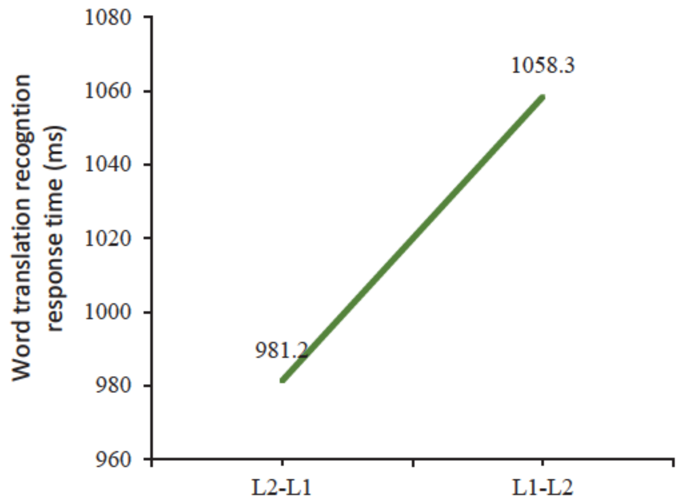


Figure 4.3. Word translation recognition response times in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for all participants (times are given in miliseconds).

There was also a significant main effect of group, $F(1, 48) = 9.05$, $p = .004$, $\eta_p^2 = .159$. Table 4.5. provides each group's word translation recognition response times in both directions. This main effect suggests that the SA group outperformed NSA group in terms of response times in correctly retrieving words in both directions (939.3 vs. 1100.2, respectively), as shown in Figure 4.4.

There was also a significant group × direction interaction, $F(1, 48) = 5.134$, $p = .028$, $\eta_p^2 = .097$. To explore this interaction, the author conducted pairwise comparisons with an adjusted alpha level of .025. The NSA group showed a significant difference in their word translation recognition response times for L1-L2 vs. L2-L1, $t(24) = 4.17$, $p < .001$; whereas no statistical difference was detected for the SA group, $t(24) = .47$, $p = .32$, indicating a more balanced performance across the two language directions.

Given that the SA group was faster than the NSA group but was less accurate, we considered it a speed-accuracy trade off.

Table 4.5. *Performance on word translation recognition response time in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for SA and NSA groups (times are given in milliseconds).*

Group	L2-L1	L1-L2
SA	929.9	948.7
NSA	1032.5	1167.9

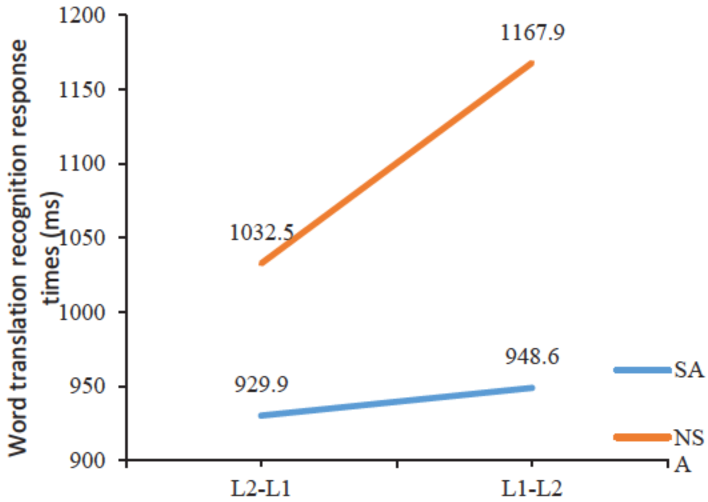


Figure 4.4. Word translation recognition response time in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for SA and NSA groups (times are given in miliseconds).

4.2.5 Study Abroad Experience and Grammatical Processing

In further analysing participants' linguistic performance, their verbal accuracy of L1-L2 CI production was coded and investigated for inflexional form (subject-verb number agreement). The subject-verb number agreement accuracy rate of each participant was calculated as a percentage of correctly inflected verbs given the range of obligatory occasions arising in participants' L2 output. For instance, if a participant produced eight correctly inflected verbs from ten obligatory occasions, this participant's accuracy rate would be 80%.

Before comparing the two groups' subject-verb number agreement accuracy rate, emergence criterion (Di Biase & Kawaguchi, 2002) was applied to check whether all participants had acquired this linguistic structure or not. According to the verbatim transcriptions of participants' English productions, both groups met the emergence criterion, indicating they had acquired this structure. After that, we compared the accuracy rate of the two groups.

As shown in Table 4.6, the mean accuracy rate of the SA group was higher than that of the NSA group. To explore the correlation between SAE and subject-verb number agreement, a *t*-test was conducted to determine if the two groups differed in their accuracy rates. The difference between the two groups was significant, $t(24) = 2.45$, $p = .02$, indicating that the SA group was more successfully attending to this grammatical structure while performing the cognitively demanding interpreting task.

Table 4.6. *Subject-verb number agreement (%) in Chinese to English (L1-L2) CI for SA and NSA groups.*

Groups	Subject-Verb Number Agreement
SA	71.3 %
NSA	57.6 %

4.2.6 Study Abroad Experience and Fluency

Another linguistic performance investigated in the present study was participants' L2 fluency in their L1-L2 CI production.

In order to quantify participants' fluency in their L2 production, two measures were computed: syllables per minute and filled pauses per minute (e.g., Han, 2015b; Lin et al., 2018; Rennert, 2010; Yu & Van Heuven, 2017).

As shown in Table 4.7, the mean syllables/min produced by the SA group was higher than that of the NSA group. The results of a *t*-test revealed that this difference between the two groups was not significant, $t(24) = 1.087$, $p = .288$.

Table 4.7. *Syllables/minute in Chinese to English (L1-L2) CI for SA and NSA group.*

Group	Syllables/ min
SA	140
NSA	130

So as not to confuse a slow speech rate with poor fluency in L2 production, we also measured participants' rates of producing L1 (Chinese) syllables per minute in the L2-L1 CI task and correlated this with their L2 fluency. Both groups' L1 production rate (syllables/min) was higher than their L2 production rate (syllables/min) (SA: 190/min vs. 140/min; NSA: 186/min vs. 130/min). Group correlations between the L1 and L2 production rates are presented in Table 4.8. A significant correlation was found for the SA group, which may suggest that the speech rate of SA participants was closer to their native language performance. No significant correlation was found in the NSA group.

Table 4.8. *Correlation between English (L2) (syllabus)/min and Chinese (L1) (characters)/min for SA and NSA groups.*

Group	Correlation
SA	0.53**
NSA	0.39

*$p < .05$, ** $p < .01$

Another parameter of fluency used in this study was participants' filled pauses/min. Participants in the NSA group produced 4.8 filled pauses/min on average, and the SA group produced 2.6 filled pauses/min on average.

To explore the relationship between SAE and filled pauses per minute, the author conducted a *t*-test. The results confirmed that the difference between the two groups in their filled pauses is marginally significant, $t(24) = -2.130$, $p = .05$, which suggests that the SA group made fewer pauses (a measure of fluency) than the NSA group in fluency in their L2 production for L1-L2 CI task.

4.2.7 Study Abroad Experience and Consecutive Interpreting Performance

To analyse the relationship between SAE and CI performance, the author conducted a 2 × (2) mixed factorial ANOVA with the between-subjects factor of group (NSA vs. SA) and the within-subjects factor of direction of interpreting (L2-L1 vs. L1-L2). Recall that CI performance was quantified on a scale ranging from 1 (low) to 10 (high) according to their information accuracy, language quality, comprehension, expression, and delivery. A significant main effect of direction, $F(1, 48) = 36.76$, $p < .001$, $\eta_p^2 = .434$, indicated that overall, participants were more successful at interpreting in the L1-L2 direction than the opposite direction (5.51 vs. 4.68, respectively). There was no significant main effect of group, $F(1, 48) = 0.46$, $p = .50$, $\eta_p^2 = .01$). However, there was a significant group × direction interaction, $F(1, 48) = 9.76$, $p = .003$, $\eta_p^2 = .169$. To explore the interaction, pairwise comparisons were conducted with an adjusted alpha level of .025. The NSA group showed a significant difference in their interpreting performance for L1-L2 vs. L2-L1, $t(24) = -7.23$, $p < .001$; whereas the SA group did not differ statistically in their two directions of CI, $t(24) = -1.4$, $p = .21$. As shown in Table 4.9 below, the SA group was more balanced and consistent in the two directions of their CI performance, as opposed to NSA group which exhibited inconsistency.

Table 4.9. *Consecutive interpreting performance in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for SA and NSA groups.*

Group	L2-L1	L1-L2
SA	5.22	5.62
NSA	4.14	5.39

4.2.7.1 Lexical Processing and CI Performance

4.2.7.1.1 L2-L1 Lexical Processing Accuracy and CI Performance

Section 4.2.3 shows that lexical processing accuracy rather than response times is correlated with CI performance in the present study. As shown in Table 4.10, significant strong correlations between L2-L1 word translation recognition accuracy (Word Acc) and CI performance in both directions were observed in the SA group. In the NSA group, less strong correlations but still significant correlations were observed.

Table 4.10. *Correlation between English-Chinese (L2-L1) word translation recognition accuracy and two directions of CI performance for SA and NSA groups (with WM).*

L2-L1 Word Acc	CI	
	L2-L1	L1-L2
SA	.70**	.78**
NSA	.45*	.49*

*$p < .05$, ** $p < .01$, *** $p < .001$, ~$.05 < p < .1$.

To determine if the correlations were primarily due to lexical processing performance or the contribution of WM, we re-did all correlation analyses but with WM partialled out. When the contribution of WM is statistically controlled (partialled out), as is shown in Table 4.11, significant correlations between L2-L1 lexical processing accuracy and CI performance in the SA group persisted, but the correlations were fragile in the NSA group.

In sum, the correlation between L2-L1 word translation recognition accuracy and CI performance was statistically robust. Even though the correlation crept down and the *p* value crept up above .05 in the NSA group in their L2-L1 CI, the pattern did not change. However, the data suggests that the correlation in the NSA group was less statistically reliable than the correlation in the SA group.

Table 4.11. *Correlation between* English-Chinese *(L2-L1) word translation recognition accuracy and two directions of CI performance for SA and NSA groups (WM partialled out).*

L2-L1 Word Acc	CI	
	L2-L1	L1-L2
SA	.58**	.69**
NSA	.38~	.44*

*$p < .05$, ** $p < .01$, *** $p < .001$, ~$.05 < p < .1$.

4.2.7.1.2 L1-L2 Lexical Processing Accuracy and CI performance

Table 4.12 presents the correlation between L1-L2 word translation recognition accuracy (Word Acc) and two directions of CI performance for both groups. When analysing correlations between L1-L2 word translation recognition accuracy and CI performance in both directions, we only found a significant correlation in the SA group and only in the L1-L2 direction of CI performance. For the NSA group, an insignificant correlation was found in both directions of CI performance. In general, the link between CI performance was stronger with L2-L1 word translation recognition accuracy than with L1-L2 word translation recognition accuracy.

Table 4.12. *Correlation between Chinese-English (L1-L2) word translation recognition accuracy and two directions of CI performance for SA and NSA groups (with WM).*

L1-L2 Word Acc	CI	
	L2-L1	L1-L2
SA	.35~	.58**
NSA	.24	.38~

*$p < .05$, ** $p < .01$, *** $p < .001$, ~$.05 < p < .1$.

When WM was partialled out, as shown in Table 4.13, this weak correlation was further weakened, and the significance was even lower. Correlation between L1-L2 word translation recognition accuracy (Word Acc) and L2-L1 CI was non-significant in both groups, but the L1-L2 CI still significantly correlated with L1-L2 word translation recognition accuracy in the SA group, but the correlation was weaker than when WM was not partialled out (.49 vs. .58, respectively). No significant correlation was observed in the NSA group in either CI direction.

Table 4.13. *Correlation between Chinese-English (L1-L2) word translation recognition accuracy and two directions of CI performance for SA and NSA groups (WM partialled out).*

L1-L2 Word Acc	CI	
	L2-L1	L1-L2
SA	.18	.49*
NSA	.09	.28

*$p < .05$, ** $p < .01$, *** $p < .001$, ~$.05 < p < .1$.

In sum, the pattern of results of these partial correlations was largely consistent with those reported in Section 4.2.3 for the SA group, but not for the NSA group. This suggests that for SA participants, lexical processing performance was related to CI performance in both directions. However, this was not the case for the NSA group.

4.2.8 Working Memory and Lexical Processing

To analyse the relationship between WM and lexical processing, we pooled SA and NSA groups together and computed the correlation between WM and word translation recognition accuracy (Word Acc) as well as WM and response times (RTs). Table 4.14 shows the correlation between WM and word translation recognition performance in both language directions. Word translation recognition accuracy in both language directions significantly correlated with WM ($p < .05$); however, only the response time in the direction of L2-L1 significantly correlated with WM. All these are moderate correlations.

Table 4.14. *Correlation between WM and two directions of word translation recognition for all participants.*

Correlation	Word translation recognition			
	L1-L2		L2-L1	
WM	Word Acc	RTs	Word Acc	RTs
	0.30*	-0.19	0.36*	-0.37**

*$p < .05$, ** $p < .01$, *** $p < .001$, ~$.05 < p < .1$.

A secondary analysis was conducted with the aim of further exploring the relationship between WM and lexical processing. Participants in SA and NSA groups were mixed together and were re-categorised into two new groups based on their WM resource availability (High-WM vs. Low-WM). This created High- and Low-WM groups composed of those participants whose WM was above (n = 26) or below the median (25.5, n = 24). The two new groups were matched in their vocabulary knowledge captured by LexTALE, $t(24) = .59$, $p = .56$. A 2 × (2) mixed factorial ANOVA was conducted with the between-subjects factor of group (High-WM vs. Low-WM) and the within-subjects factor of direction of word translation recognition accuracy (L2-L1 vs. L1-L2). A significant main effect of direction, $F(1, 48) = 36.88$, $p < .001$, $\eta_p^2 = .434$, indicated that overall, participants were more accurate at retrieving words in the L2-L1 direction than in the opposite direction (73.6% vs. 68.1%, respectively).

Table 4.15 provides the word translation recognition accuracy rates for the High-WM and Low-WM groups in both language directions. A significant main effect of group, $F(1, 48) = 4.68$, $p = .04$, $\eta_p^2 = .089$, confirmed that the High-WM group was more accurate in lexical processing in both language directions (accuracy rate: 74% vs. 68%, respectively). There was no significant group × direction effect, $F(1, 48) = 0.61$, $p = .113$, $\eta_p^2 = .051$.

Table 4.15. *Word translation recognition accuracy (%) in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for High-WM and Low-WM groups.*

Groups	L2-L1	L1-L2
High-WM	77%	70%
Low-WM	70%	66%

An analysis on the word translation recognition response times (L2-L1 vs. L1-L2) was also conducted. A significant main effect of direction, $F(1, 48) = 8.16$, $p = .006$, $\eta_p^2 = .145$, indicated that overall participants were faster at processing words from L2 to L1 than from L1 to L2 (981.2 vs. 1058.3, respectively).

Table 4.16 presents each group's mean response time, and again there was a significant main effect of group, $F(1, 48) = 12.62$, $p = .001$, $\eta_p^2 = .208$. The High-WM group was significantly faster than the Low-WM group in retrieving target words in both language directions (response time: 927.6 vs. 1112.0, respectively). No significant group × direction interaction was found, $F(1, 48) = 3.387$, $p = .537$, $\eta_p^2 = .008$.

Table 4.16. *Word translation recognition response times in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for High-WM and Low-WM groups.*

Groups	L2-L1	L1-L2
High-WM	880.65	974.45
Low-WM	1081.82	1142.07

4.2.9 Working Memory and Grammatical Accuracy

Previous studies have also observed positive correlations between WM and individuals' grammatical performance (e.g., Grey et al., 2015; Guntermann, 1995; Sagarra, 2017; Sagarra & Labrozzi, 2018; Verhagen & Leseman, 2016; Wright, 2013). Therefore, the author conducted *t*-tests to explore whether the High-WM group performed significantly better than the Low-WM group in terms of subject-verb number agreement accuracy in their L1-L2 interpreting output. The result was significant, $t(24) = 4.51$, $p < .001$, confirming that the High-WM group performed significantly better in attending to L2 form while processing the content at the same time, than their Low WM counterparts (73.6% vs. 53.5%, respectively).

In order to explore whether WM and SAE exert a joint positive effect on learners regarding subject-verb number agreement, we further divided the High-WM and Low-WM groups based on their SAE into High-WM-SA, High-WM-NSA, Low-WM-SA, and Low-WM-NSA groups. Participants with both high WM and SAE (i.e. the High-WM-SA group) exhibited the highest subject-verb number agreement rate in their L2 productions, and participants without SAE and low WM (i.e. the Low-WM-NSA group) demonstrated the lowest accuracy rate (Table 4.17). However, these differences were not significant.

Table 4.17. *Subject-verb number agreement (%) in Chinese-English (L1-L2) CI for High-WM-SA, High-WM-NSA, Low-WM-SA, and Low-WM-NSA groups.*

Groups	Subject-Verb Number Agreement
High-WM-SA	0.79
High-WM-NSA	0.67
Low-WM-SA	0.58
Low-WM-NSA	0.48

4.2.10 Working Memory and Consecutive Interpreting Performance

In investigating the relationship between WM and CI performance determined by the two individual assessors' mean ratings, the author conducted a 2 × (2) mixed factorial ANOVA with the between-subjects factor of group (High-WM vs. Low-WM) and the within-subjects factor of direction of interpreting (L2-L1 vs. L1-L2). A significant main effect of direction, $F(1, 48) = 36.88$, $p < .001$, $\eta_p^2 = .435$, indicated that overall, participants were more successful when interpreting from L1 to L2 than from L2 to L1 (5.46 vs. 4.60, respectively). There was also a significant main effect of group, $F(1, 48) = 15.59$, $p < .001$, $\eta_p^2 = .245$, which indicated that participants with higher WM performed significantly better than lower WM participants (5.77 vs. 4.30, respectively). There was also a significant group × direction interaction, $F(1, 48) = 9.96$, $p = .003$, $\eta_p^2 = .172$. Pairwise comparisons confirmed that the Low-WM group's interpreting performance was better for Chinese-English than English-Chinese, $t(24) = 8.24$, $p < .001$, whereas the High-WM group performed equally in the two directions of CI, $t(24) = 1.76$, $p = .09$, as shown in Table 4.18.

Table 4.18. *Consecutive interpreting performance in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for High-WM and Low-WM groups.*

Group	L2-L1	L1-L2
High-WM	5.51	6.02
Low-WM	3.69	4.90

Likewise, to explore whether WM and SAE exerted a joint positive effect on CI performance, the High-WM and Low-WM groups were further separated into High-WM-SA, High-WM-NSA, Low-WM-SA, and Low-WM-NSA groups based on their language contexts. The results indicate a similar pattern to the one found in grammatical processing: participants with both high WM and SAE outperformed the rest three groups in both CI directions, followed by the High-WM-NSA group (Table 4.19). However, participants in the Low-WM group exhibited comparable low CI scores regardless of their language contexts. Due to the limited number of participants, these differences were not significant.

Table 4.19. *Consecutive interpreting performance in English-Chinese (L2-L1) and Chinese-English (L1-L2) directions for High-WM-SA, High-WM-NSA, Low-WM-SA, and Low-WM-NSA groups.*

Group	L2-L1	L1-L2
High-WM-SA	6.21	6.17
High-WM-NSA	4.81	5.88
Low-WM-SA	3.75	4.80
Low-WM-NSA	3.62	5.00

4.2.11 Working Memory and Study Abroad Experience

To analyse whether language context influenced interpreting students' WM, a *t*-test was conducted to compare the two groups' L2 reading span scores. This result indicated that for the two groups of interpreting students with different learning contexts there was no statistical difference between their WM availability at the time of data collection, $t(24) = .29$, $p = .39$.

4.3 Summary

This chapter provides the data collected from 50 Chinese-English interpreting students who had been categorised into two groups according to their different learning contexts (SA group vs. NSA group).

The two groups were matched in terms of their age, age of onset learning the L2, course major, and the length of learning interpreting. The results from the LexTALE test and the LEAP-Q questionnaire suggest that the two groups were comparable in their English lexical knowledge and language proficiency. In terms of L2 activities per week, the SA group reported significantly more hours per week interacting with native speakers in English than the NSA group.

The results of the reading span task, word translation recognition accuracy and response times, as well as CI performance in both language directions were presented and correlated with each other. Generally speaking, the data indicate that the SA and

NSA groups did not differ statistically in terms of their WM; however, the two groups differed in the magnitude of word translation recognition asymmetry, grammatical accuracy, fluency in their L2 productions and their CI performance in the L2-L1 direction. Moreover, correlation and regression analyses showed that WM significantly correlated with both word translation recognition (accuracy and response time) and CI performance in both language directions. Furthermore, a stronger correlation between CI performance was identified in word translation recognition accuracy, rather than word translation recognition response time. The next chapter will interpret these results, integrate them with past findings in the research literature and explain their contribution to the field.

Chapter 5 Discussion

This chapter discusses how the results presented in Chapter 4 address the research questions for this study regarding the effect of study abroad experience (SAE) on interpreting students' linguistic performance, consecutive interpreting (CI) performance, and how participants' working memory (WM) interacted with their linguistic and CI performance in different language environments. The results lend partial support to some of the hypotheses, whereas others need to be rejected.

The first research question concerns whether participants with SAE will outperform their NSA counterparts in terms of linguistic performance, that is, L2 vocabulary knowledge, lexical processing performance, L2 oral fluency and grammatical accuracy. In addressing this research question, Hypothesis 1 predicts that interpreting students with SAE will perform better than NSA students. This hypothesis is partially supported by the results.

The second research question asks whether the SA context will have greater benefits than the L1 context for interpreting students' CI performance in both language directions. The corresponding Hypothesis 2 predicts that students with SAE will outperform their NSA counterparts in terms of overall CI performance. This hypothesis is also partially supported.

The third research question asks whether interpreting students with high WM resource availability and SAE will demonstrate superior linguistic and CI performance than the rest of the participants.

We hypothesised that WM resource availability is associated with learning interpreters' linguistic and CI performance, and that those participants with high WM as well as SAE will outperform their peers with low WM or without SAE. These hypotheses are supported. It was also predicted in Hypothesis 5 that the SA group will benefit from SAE and demonstrate larger WM resource availability than the NSA group. This hypothesis is not supported.

This chapter will discuss the results of the study in detail and is divided into three parts. Section 5.1 focuses on the effect of SAE on linguistic performance. Section 5.2 discusses the role of SAE on overall CI performance. Section 5.3 investigates the effect of WM on linguistic and CI performance for participants in different language contexts. Finally, Section 5.4 contains a summary of this chapter.

5.1 SAE and Linguistic Performance

Here, we consider differences between the groups in terms of their hours of L2 use and social contexts. According to their LEAP-Q responses, the SA and NSA groups differed in their weekly hours of interacting with native speakers in English; the SA group averaged 37.5 hours/week, and the NSA group reported 19.9 hours/week. Apart from that, the groups reported comparable hours spent watching English TV, listening to English radio/music, reading English materials and self-instruction in English, from 28 to 42 hours/week.

The first research question concerns the effect of SAE on linguistic performance by exploring whether interpreting students with SAE outperformed the students without SAE in terms of lexical knowledge (LexTALE), lexical processing (word translation recognition), L2 grammatical processing (L1-L2 CI performance) and oral fluency (L1-L2 CI performance).

The prediction that interpreting students with SAE will outperform their NSA counterparts in regard to the abovementioned tasks is partially supported. Even though no statistical difference was identified between the two groups in their L2 vocabulary knowledge at the time of testing, the SA group demonstrated less asymmetric switching costs than the NSA group in lexical processing. A speed-accuracy trade-off in word translation recognition was also found in the two groups, which refers to the situation where decisions are made accurately at the cost of speed or quickly at the expense of accuracy (Duckworth et al., 2018). In the present study, participants with SAE were faster but less accurate than the NSA participants in processing words in both language directions. When it comes to L2 oral grammatical accuracy and fluency measured through L1-L2 CI output, the SA group was more fluent and more accurate in subject-verb number agreement than the NSA group.

In sum, SAE exerted a beneficial effect on interpreting students' L2 grammatical accuracy and fluency, and partially on their lexical processing performance, but did not affect lexical knowledge.

The two groups' vocabulary knowledge is discussed below in Section 5.1.1, taking into consideration their different learning contexts, and Section 5.1.2 examines lexical processing, followed by the discussion of L2 grammatical accuracy in Section 5.1.3, and L2 fluency in Section 5.1.4.

5.1.1 Study Abroad Experience and Second Language Lexical Knowledge

Hypothesis 1 predicts that the SA environment could foster L2 vocabulary knowledge growth, and this is based on two assumptions: the first one stems from Interaction Theory which claims that learners who use their newly-acquired words more often are more likely to have solid word knowledge (Craik & Tulving, 1975). SAE is believed to provide learners with more opportunities to access and use novel L2 words than classroom instruction (Harley & Jean, 1999; Lo & Murphy, 2010). The second assumption is based on the Inhibitory Control model (Green, 1998) which suggests that learning the L2 in an L1 inhibition environment could reduce the L1 competition and thus promote L2 learning (Kroll et al., 1998). However, results of vocabulary task (LexTALE) do not support either of these assumptions, as the SA and NSA groups did not differ in terms of their L2 lexical knowledge. This finding is in line with studies which suggest that SAE does not make additional contributions to lexical knowledge development above that achieved through classroom instruction (e.g., Briggs, 2015; Collentine, 2004; Fitzpatrick, 2012; Grey et al., 2015; Serrano et al., 2011).

After careful consideration, we have identified three potential explanations for this null result concerning SAE and vocabulary knowledge. The first relates to the participants' ages of immersion. Studies reporting significant lexical gains as a result of target language immersion (e.g., Barik & Swain, 1976; Harley & Jean, 1999; Jimenez-Jimenez, 2010; Lo & Murphy, 2010), have mostly involved participants who were young children (e.g., Barik & Swain, 1976; Harley & Jean, 1999; Lo & Murphy, 2010), around grades two to ten, and thus they were much younger than the participants in our study. For late bilinguals, lexical knowledge, as a form of declarative knowledge, is very hard to acquire implicitly; thus, the acquisition is best obtained via the explicit instruction received in L2 classrooms (Kormos, 2006, p. 167).

The second possible explanation for the null result relates to participants' majors in their university studies. As all of the participants in the present study were students of master's degrees in interpreting and translation, their initial L2 level before SAE was quite high (on/above IELTS 6.5). According to previous studies examining the development of lexical knowledge, high-level learners improved slower than low-level learners due to the normal learning curve (B. Freed, 1995; Llanes & Muñoz, 2009; Milton & Meara, 1995; Zaytseva, 2016). Additionally, in previous studies, lexical gains in the basic 2000-word level as a result of target-language immersion were most significant, and when it came to words beyond the basic 2000 level, only slight increases were observed (Laufer, 1994, 1998; Lo & Murphy, 2010). This may explain why the lexical difference observed between the two groups failed to reach significance despite their different learning contexts. Moreover, since participants in the present study all majored in interpreting and translation, both SA and NSA groups had been encouraged to memorise novel L2 words in a diverse range of subjects during interpreting learning and practice (Corsellis, 2005; Han, 2015a; Liang et al., 2017; Niska, 2005), This is also likely to have diminished any additional potential benefit of novel word learning provided by the SAE over classroom instructions.

The third possible explanation lies in participants' L2 activities during their SAE. According to the questionnaire, the two groups of participants only differed significantly in their weekly hours of interacting with native speakers in English (see Table 4.1). Apart from that, the two groups' L2 reading and self-instruction time were comparable. Therefore, it can be concluded that in the present study, out-of-class interaction did not contribute to significant vocabulary knowledge gains. This finding is consistent with Briggs (2015), who reported that interactional exchanges offer limited vocabulary gains, because this type of contact does not provide learners with abundant opportunities to encounter and memorise novel lexical items. When coming across a novel word in a conversation, learners tend to guess the meaning of the unknown words from the context rather than interrupt the conversation flow and ask for a definition (Lafford, 2006). That is, interactions may constrain SA learners from getting opportunities for novel lexical learning and use (Briggs, 2015; Lo & Murphy, 2010).

This helps explain the lack of relationship between SAE and lexical gains in the present study. However, this result does not go against Interaction Theory but confirms that frequent interaction encourages learners to produce more target-like output, as SA participants demonstrated better L2 fluency as well as grammatical performance (Sections 5.1.3 and 5.1.4).

Based on the above three explanations, SA and NSA groups' comparable amount and variety of access to L2 novel words has probably diminished the benefit of L2 learning results brought by the L1 inhibitory environment (SA context). SA learners could not make full use of the L1 inhibition environment to boost their L2 word acquisition because they did not have significantly more access to novel L2 words. However, the benefits of L1 inhibition as a result of SAE manifest in improved lexical processing. Interpreters are a special subset of bilinguals who are required to keep their two languages constantly activated, even in a monolingual environment (García, 2019, p. 199; Grosjean, 1997; Schwieter & Ferreira, 2017); thus, SA participants in the present study experienced more constant language switching than their NSA counterparts, which benefited their lexical processing. This is evidenced by the lower language switching costs they experienced, as discussed in the following section.

In sum, SAE in the present study did not have an observable effect on interpreting students' L2 lexical knowledge.

5.1.2 Study Abroad Experience and Lexical Processing Performance

Both groups exhibited the direction-dependent translational asymmetry: word translation from the L2-L1 direction was faster and more accurate than in the opposite direction. These results lend support to the Inhibitory Control model (Green, 1998), which suggests that for unbalanced bilinguals, the dominant L1 is more active than the L2, which results in the situation that more cognitive effort is required to inhibit the L1 than the L2. However, compared with the NSA group, the SA group had significantly less direction-dependent asymmetry, which implies a more balanced and consistent lexical processing performance. Apart from direction-dependent asymmetry, when comparing the performance of SA and NSA groups, the results showed that there was a trade-off between word translation recognition accuracy and response times. The SA group was significantly faster but less accurate than the NSA group in tasks, indicating that SA participants tended to respond faster but at the risk of being less accurate, whereas the NSA group was more likely to prioritise accuracy by compromising speed.

5.1.2.1 Direction-Dependent Asymmetry

The results of the SA and NSA groups revealed word direction-dependent asymmetry in their word translation recognition tasks, indicating that both groups were more accurate and faster from L2 into L1 than in the opposite direction. This result replicates previous research findings that forward lexical translation (L1-L2) is slower and more error-prone than backward lexical translation (L2-L1) (e.g., Costa & Santesteban, 2004; Linck et al., 2008; Meuter, 2009; Meuter & Allport, 1999; Olson, 2017; Peeters et al., 2014), which lends support to the Inhibitory Control model (Green, 1998). According to the model, both languages are activated in a non-selective manner, and L1-L2 word translation involves inhibiting the dominant L1 competitors to ensure that the intended L2 words are selected for output (see Green, 1998; Kroll et al., 2008). This process requires more cognitive effort than inhibiting the relatively weaker L2 in the other translation direction (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999; Olson, 2017; Peeters et al., 2014). This model explains why participants in the present study exerted more effort (realised as lower accuracy and longer response times) to inhibit their dominant, active L1 in order to complete the L1-L2 word translation recognition task than to inhibit the relatively weaker L2 when completing the task in the opposite direction. Note that in the present study, the SAE did not impact on this direction-dependent asymmetry, as the SA group also showed faster and more accurate performance in the L2-L1 than the other direction. This result is consistent with Meuter and Allport's (1999) suggestion that it is the relative strength of bilinguals' two languages that affects asymmetry in interpreting performance. In the present study, even though SAE helps attenuate L1 dominance and facilitate L2 processing (e.g., Baus et al., 2013; Linck et al., 2009; Sagarra & Labrozzi, 2018), for a group of late unbalanced bilinguals this attenuation appears to have been insufficient to override the dominance of the L1 in just two years' SAE. Therefore, the results of the present study show that both SA and NSA groups demonstrated the same direction-dependent asymmetry in their lexical processing tasks.

5.1.2.2 Study Abroad and Non-Study Abroad Groups

Even though the two groups demonstrated asymmetry in the same direction, the effect of different language contexts on participants' lexical processing was still observed. The SA group exhibited a smaller amount of asymmetry between the two language directions than the NSA group, replicating Schwartz and Kroll (2006) and Chmiel (2016) whose participants also had L2 immersion. In other words, the SA group in the present study was more balanced and consistent in terms of language switching costs during word translation recognition than the NSA group. Previous studies have suggested that factors such as participants' L2 proficiency (Costa, 2005; Costa & Santesteban, 2004; Meuter & Allport, 1999), language-switching habits (Christoffels et al., 2007), WM and learning contexts (Kroll et al., 1998; Kroll & Sunderman, 2003; Linck et al., 2008) impact on the language inhibitory process, and influence the switching costs between language directions. In the present study, participants in the SA and NSA groups were all interpreting students, and therefore, all were engaged in bilingual processing on a daily basis. Moreover, they were also comparable in their L2 proficiency, word knowledge and WM resource availability (Section 4.2.1). Therefore, the smaller degree of asymmetry observed in the SA group may be attributed to their relatively long-term exposure to standard English in daily life.

For the SA group, during their daily L2 immersion, the L2 environment helped them effectively attenuate the dominant L1 (e.g., Baus et al., 2013; Linck et al., 2009). So compared with the NSA group, SA participants were more at ease when making efforts to inhibit the L1 during L2 processing. Even though SA participants were in the L2 environment, their L1 was nevertheless stronger than their L2; therefore, the extra efforts required to inhibit the L2 could have been partially offset by their dominant L1. Consequently, the SA group exhibited more consistent and less asymmetric performance when processing the two languages compared to the NSA group. In contrast, when NSA participants' language environments coincided with their dominant L1, it intensified the difficulty of inhibiting the non-target but activated L1 during L2 processing. This contributed to the higher degree of asymmetry in the NSA group's performance. This also partially explains why the SA group demonstrated more balanced and consistent CI performance than the NSA group in both language directions, as discussed in Section 5.2.1.

Another salient difference between SA and NSA groups is that SA participants were faster than NSA participants in lexical processing in both language directions, but SA participants were less accurate than their NSA counterparts, indicating that the SA group was more inclined to give quick responses at the risk of making errors. This finding is consistent with DeKeyser (1991) and Tokowicz et al. (2004) who examined the effect of the L2 immersion experience and WM on the types of errors made in lexical processing. They found that when encountering unknown words, participants with longer SAE (exceeding one year) were prone to guess words' meanings at the risk of making errors, whereas individuals without SAE or with short SAE made more non-response choices. The present study is consistent with this explanation that SA learners tended to set a lower threshold for selecting a translation equivalent, and therefore, allowed less accurate translations to be given. This phenomenon reflects SA participants' need or desire to communicate in the absence of accurate words, which is believed to be fostered in an SA environment.

5.1.3 Study Abroad Experience and Grammatical Accuracy

Both groups met the emergence criteria, which indicates that all participants in the present study had acquired the linguistic knowledge of subject-verb number agreement at the time of data collection. Therefore, the errors made by participants should not be seen as a reflection of their inadequate knowledge about how to apply the third person singular markers, but as an indication that participants in the two different learning contexts varied in the accuracy with which they applied these markers when performing CI tasks, as shown by their differing accuracy rates.

Given that the SA group was more accurate in subject-verb number agreement than the NSA group, this result is in line with existing SLA evidence showing the superiority of SA over NSA learners in L2 grammatical processing (e.g., Grey et al., 2015; Howard, 2005; Isabelli & Nishida, 2005; Labrozzi, 2009; Llanes & Muñoz, 2009; Pérez-Vidal & Juan-Garau, 2011; Sagarra & Labrozzi, 2018; Song, 2015; Wright, 2013). Our results also support the assumption that SAE contributes to how participants use the grammatical structure that they have already acquired (Wright, 2013). Thus, we can now extend the influence of SAE on grammatical processing to the interpreting field.

We found that the two groups reported comparable amounts of English (L2) input in the form of watching English TV, listening to English radio/music, reading English materials and English self-learning. These types of input are important because they feed or nurture an innate linguistic system and aid its growth (B. Schwartz, 1993). Also, they provide evidence about what is acceptable within a language and what is not (Gass & Mackey, 2007). For instance, obtaining input is a way of determining that in English *he comes back* rather than *he come back* is grammatically correct. However, input alone is not sufficient for acquiring a language, unless the learners become involved in interactions (S. Zhang, 2009), which allows them to receive feedback or further information about the plausibility and correctness or incorrectness of their utterances in a real context (Gass & Mackey, 2007). During interactions, learners can form, test, modify and confirm or reject their linguistic hypotheses in a real language context which helps them integrate the learnt message into their linguistic system (Gass, 1988, 1997, p. 104; Izumi, 2003). And this is where the SA and the NSA groups differed in their social contexts in using the L2. In our study, SA participants reported spending more hours interacting with English speakers than their NSA counterparts. Frequent interaction is accompanied by more opportunities to produce output, which shifts the focus of learners' awareness from semantic processing to the syntax and morphology needed for grammatical processing (Gass, 1997, p. 105; Pica, 1994; Swain, 1995).

Through constant interaction and output during their SAE, the SA participants strengthened the weakly established grammatical knowledge base that they may have found difficult put into practice prior to their SAE. Consistent practice of applying grammatical rules into output helped them solidify the meaning-form connections and increase the automaticity of processing this grammatical structure (Izumi, 2003; Loschky & Bley-Vroman, 1993; Swain, 1985, 2005). Thus, the present study found that SAE triggers more accurate subject-verb number agreement. This is in line with the general consensus that intensive exposure to and use of appropriate linguistic forms leads to more efficient processing of those forms (Ellis, 2002; B. Freed, 1995; Paradis, 2004; Perani et al., 2003; Pliatsikas, 2010). But rote repetition without a context and without deeper cognitive processing does not appear to be very effective in increasing the efficiency of processing (Gile, 2009, p. 230). Therefore, compared with the NSA group, the higher grammatical accuracy rate of SA participants in the present study can be attributed to the greater amounts of interaction and output they experienced.

The finding that the SA group produced better grammatical processing performance is not consistent with studies which observed superior performance among the NSA groups, or with studies which proposed that SAE is unnecessary because classroom instruction is sufficient to promote grammatical processing (e.g., Collentine, 2004; Isabelli-García, 2010; Marqués-Pascual, 2011; Rothman & Iverson, 2007). Rothman and Iverson (2007) revealed that the two learning contexts (SAE and the classroom) both contribute to grammatical acquisition, however, their study lacked an NSA control group, and therefore it is unknown if SAE is more efficient than classroom instructions in boosting L2 learners' processing accuracy. The differing results of Collentine (2004) may stem from their short immersion time for the SA group, which was only one semester; moreover, the NSA group began the experiment with somewhat better overall L2 proficiency. Isabelli-García (2010) found no improvement in gender agreement among participant SA learners. This is not surprising given that, for late L2 learners whose L1 does not have grammatical gender, gender agreement is a very difficult task, both in traditional classroom settings and in SA conditions (e.g., Alarcón, 2009; Antón-Méndez et al., 2002; D. J. Davidson & Indefrey, 2009; Grey et al., 2015; Morgan-Short et al., 2010; Rossi et al., 2014; Sagarra & Herschensohn, 2010; Sagarra & Labrozzi, 2018). Thus, it appears that certain linguistic forms may be more likely to develop with SAE (Faretta-Stutenberg & Morgan-Short, 2018). Marqués-Pascual (2011) reported that SA students did not show greater gains than their NSA peers, but did find that SAE positively influenced L2 grammatical processing in highly proficient language learners. Given that our participants were all interpreting master's students with IELTS scores of 6.5 or higher, it is not surprising that they benefited from their SAE.

What can be concluded from the present study is that both SAE and NSA classroom instruction contribute to the development of subject-verb number agreement acquisition and processing. However, SA participants were more accurate in applying this grammatical structure than the NSA participants during their CI. This suggests that SAE helps promote grammatical processing automaticity through richer L2 interactions and output, during which participants are presented with more opportunities to put the linguistic knowledge they have acquired into real-world practice which seems to have helped them internalise these linguistic rules.

Automatic or efficient linguistic processing is considered to be beneficial as it requires fewer cognitive resources than controlled processing (Segalowitz, 2003). The more elements a learner is able to automatise, the more cognitive resources can be freed for other purposes (Antoniou et al., 2015; Christoffels et al., 2003; O'Brien et al., 2006; S. Zhang, 2009). For instance, if a language learner can process L2 grammatical structures automatically while using the language, then more cognitive resources can be allocated to processing semantic, pragmatic, and sociolinguistic levels of communication (Segalowitz, 2003). This advantage can be observed when participants undertake complex tasks such as CI (Section 5.2). Furthermore, automatic processing results in fast and accurate performance which is strongly associated with fluency.

5.1.4 Study Abroad Experience and Fluency

Fluency is the ability to speak quickly, accurately, and without inappropriate hesitation (Juan-Garau, 2018; S. Zhang, 2009). Results of the present study showed that the SA group was more fluent than the NSA group as captured by their syllables per minute and filled pauses in L1-L2 CI production. This finding is consistent with previous research that provides evidence of robust gains in L2 oral fluency as a result of SAE, but a lack of such gains after formal classroom instructions (e.g., Collentine, 2004; Llanes & Muñoz, 2009, 2013; Marqués-Pascual, 2011; Mora & Valls-Ferrer, 2012; Pérez-Vidal & Juan-Garau, 2011; Segalowitz et al., 2004; Wright, 2013, 2018, but see Segalowitz & Freed, 2004; Zaytseva, 2016).

As discussed earlier, the two groups were comparable in their hours and activities in the L2 input, but differed in their hours interacting with native speakers in the L2. Frequent interaction and output provide meaningful contexts for learners to use and solidify target words and rules (DeKeyser, 2007; Gass & Mackey, 2007; Izumi, 2003; McLaughlin, 1987, p. 134), which is an essential prerequisite for speedy access (Skehan, 1998, p. 60; S. Zhang, 2009) and automaticity (DeKeyser, 2001; Barbara Freed et al., 2003; Fuchs et al., 2001; Segalowitz, 2003; Swain, 1985).

For interpreting tasks which impose higher cognitive loads on participants than normal language tasks, efficient linguistic processing can be critical to alleviating participants' cognitive loads, and thus reducing dysfluencies. This is because one of the major causes of interpreting dysfluencies (i.e., filled pauses) relates to cognitive effort (B. Wang & Li, 2015).

In answering the first research question, the findings of the present study in relation to participants' lexical processing, grammatical processing and L2 fluency confirm the positive effect of SAE on participants' linguistic processing, but found that SAE had no effect on participants' lexical knowledge growth. Our results lend support to the Interaction Theory and the Inhibitory Control model. In terms of Interaction Theory, our study suggests that rich L2 input, interactions, and output effectively help consolidate the meaning-form connections, and thus promote automatic linguistic processing of acquired words and grammars. Moreover, interaction and output also play a crucial role in oral fluency. In regard to the Inhibitory Control model, the SA group exhibited less directional asymmetry in their lexical processing. This result confirms that for bilinguals, the processing of the target language is achieved by suppressing the unintended one. Our result is also in agreement with previous assumptions that L2 context can help attenuate the L1 and promote L2 processing. Additionally, our results advance our understanding of the SAE effect, suggesting that this effect also holds true for interpreting students who are required to keep their two languages co-activated, and deliberately switch between languages in their daily lives. Moreover, the findings in relation to lexical processing also confirm that SAE has an impact on participants' error patterns.

5.2 Study Abroad Experience on Consecutive Interpreting Performance

The second research question focuses on the effect of SAE on interpreting students' holistic CI performance as scored by two individual raters separately. In the present study, both English (L2) and Chinese (L1) source texts administered in CI tasks were on current, widely publicised global issues. They were suitable for general readers and not particularly difficult. The participants were likely familiar with the topics of the testing texts. Efforts were made to ensure that the source texts in L1 and L2 were comparable in regard to text type (monologic presentation), register (formal), context (speech), readability, length and speed. Therefore, the tasks completed by the participants were a clear reflection of their CI performance.

This section first discusses the correlation between the two groups' CI performance and their corresponding lexical processing performance, and then goes on to compare the CI performances of the SA and the NSA groups. The results (see Table 4.3.) suggest that L2-L1 word translation accuracy is closely associated with both groups in both CI directions, and this association is significantly stronger in the SA group than the NSA group. In terms of the CI performance comparison, both groups performed better in the direction of L1-L2 than L2-L1. The SA group scored higher than the NSA group in L2-L1 CI, while the two groups were comparable in L1-L2 CI. Participants in the SA group were also more consistent. This is shown by the fact that they demonstrated less asymmetry between the two interpreting directions.

5.2.1 Lexical Processing and Consecutive Interpreting Performance

The present results suggest that lexical processing performance correlates with interpreting students' CI performance, and this correlation is observable in both interpreting directions. These results are consistent with Christoffels, De Groot and Waldorp (2003) who found that lexical processing is an important factor in determining interpreting performance.

Even though some researchers argue that interpreting goes beyond the lexical level and involves comprehension and production of complete discourse (Gile, 2009, p. 160), the results of the present study confirm that linguistic processing on the lexical level is very important for bringing about a satisfactory interpreting performance. If retrieving an appropriate word for a concept during interpreting takes a long time, it is likely that the interpreting process will break down due to the loss of crucial cognitive resources and time (Christoffels et al., 2003). However, the difference between the present study and Christoffels et al. (2003) is that in their study, response times correlated with interpreting performance. But the data in the present study suggest that participants' lexical processing accuracy, rather than their response times, determined interpreting performance. This disparity may stem from the different interpreting modalities adopted by the two studies. Christoffels et al. (2003) examined SI, whereas the present study correlated lexical processing with the CI task. Compared with SI, CI imposes fewer time constraints on interpreters (R. Cai et al., 2015; Christoffels, 2004, p. 7; Gile, 2009, p. 165; Russell & Takeda, 2015; Tommola & Hyona, 1990), as SI requires interpreters to follow the pace set by the original language speaker. In the case of CI, prompt word translation is facilitative but not indispensable (De Bot, 2000), because there is an interval allowed between interpreters' comprehension and production. This suggests that participants in the present study had relatively more time to retrieve translation equivalents than did participants in Christoffels et al. (2003). This contributes to the finding in this study that there was a meaningful relationship between word translation recognition accuracy and CI performance instead of response times.

With regard to groups, according to Table 4.3, a stronger relationship was observed between lexical processing accuracy and CI performance in the SA group than in the NSA group. Furthermore, the direction of lexical processing was also related, such that L2-L1 word translation accuracy associated more strongly with CI performance than it did with L1-L2 word translation accuracy. In terms of L1-L2 word translation accuracy, a correlation was only observed with L1-L2 CI in the SA group. This suggests that in the present study, learners who were better at L2-L1 word translation were more capable of generating satisfactory CI performance, and this pattern tended to be more pronounced in the SA group. Moreover, this relationship, to some extent, is independent of participants' WM span. This is shown by the fact that when WM was partialled out (see Table 4.11 and Table 4.13), L2-L1 word translation recognition accuracy was still significantly correlated with CI in both directions in the SA group, but only in the L1-L2 CI direction for the NSA group.

These results indicate that members of the SA group, if they could understand the L2 words, were more capable of correctly comprehending the higher level of L2 and coming up with their corresponding translation equivalents. However, this was not the case in the NSA group. Given that reading for meaning involves more extensive brain activation than reading isolated words (Rimrodt et al., 2009), sentence comprehension is more than just word recognition. This finding seems to suggest that the SA group benefited from their SAE, and that this contributed to a higher-order language processing mechanism (in addition to word-level processing) and assisted them in obtaining a good level of L2 comprehension so long as they could capture the individual words.

In sum, the findings of the present study support Christoffels et al.'s (2003) conclusion that word translation correlates with interpreting performance. Their study revealed that response times in word translation are associated with SI performance, and the present study extends this by relating word translation recognition accuracy to CI performance. Moreover, the present study also demonstrates a stronger correlation between L2-L1 lexical processing accuracy and CI performance in both directions, and this correlation is more salient in the SA than in the NSA group.

5.2.2 Directionality in Consecutive Interpreting Performance

In the present study, the overall scores for CI performance for the SA and NSA groups revealed a significant difference in two interpreting directions. Participants of both groups were better at interpreting from Chinese to English (L1-L2) than the reverse. This directionality effect indicates that for our participants, CI in the L1-L2 direction was cognitively easier than in the L2-L1 direction. In the L1-L2 CI, the cognitive burden in the comprehension phase was lighter, likely because the source language was their native language. Even if they struggled with the expression in the output language, their understanding of the input language secured the success of the comprehension phase, which is considered to be the critical phase of CI (Gile, 2009, p. 176). Additionally, previous literature also suggests that the notetaking in CI is less cognitively demanding when the dominant L1 is the source language (Chabasse & Dingfelder Stone, 2015; Dam, 2004). However, when it comes to the L2-L1 direction, the cognitive burden falls on participants' comprehension phase; thus, participants are more subject to cognitive resource restrictions and saturation (Russell & Takeda, 2015). If participants failed to grasp the meaning of the input language, they could only resort to making up their own version or leaving gaps in their interpreting, which would negatively impact on their interpreting performance.

Therefore, the cognitive resources freed from the comprehension phase in the L1-L2 interpreting direction allowed participants to have more resources available to generate satisfactory L2 output (e.g., Barik, 1975; Chabasse & Dingfelder Stone, 2015; Kurz, 2008; S. Williams, 1995). This supports the directionality effect in this study.

It is also interesting to discover that participants in the present study were more successful at L2-L1 word translation than in the opposite direction (Section 5.1.2.1), while they were more successful when interpreting in the L1-L2 direction than the other way around. Based on the Inhibitory Control model (Green, 1998), L2 production is achieved by suppressing the L1. L1 inhibition requires additional effort than inhibiting the L2 (e.g., Bultena et al., 2015; Kroll et al., 2002; Kroll & Sunderman, 2003; Peeters et al., 2014), meaning that L2-L1 interpreting should consume fewer cognitive resources than interpreting in the other direction. Based on that, interpreting performance in the L2-L1 direction should be better than L1-L2, and this has indeed been observed in word translation in many studies (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999; Olson, 2017; Peeters et al., 2014). However, the contradictory result in interpreting direction once again suggests that the underlying difficulty of CI is not related to production-oriented operations but to comprehension (Gile, 2009, p. 176). If too much of an interpreter's cognitive resources are consumed by comprehension (listening and analysis), the resources spared for speech processing and speech production will be insufficient (Kurz, 2008). Therefore, we suggest that the cognitive burden required to inhibit the L1 during L1-L2 interpreting was lighter than the cognitive burden imposed by comprehending and analysing the L2 in L2-L1 CI. This directionality effect further indicates that, in CI tasks, an adequate understanding of the source language is more important than quickly retrieving the right words. That is, in interpreting, the advantage of a better and deeper understanding compensates for slower retrievability of words from the lexicon (De Bot, 2000).

Moreover, the interpreting directionality effect in the present study supports Zhang (2009), who reported that fast source language delivery speed led to better L1-L2 interpreting (SI) performance than the opposite direction. This is because the fast speed delivery of source text degrades interpreters' comprehension (Gerver, 1969; M. Liu, 2001; W. Zhang, 2009), especially when the source text is in their weaker language. In Zhang (2009), participants performed better in L2-L1 when delivery speed was slow (140 words/min), whereas all participants demonstrated better performance from the L1 to L2 when the delivery speed in the source language speeded up to 166 words/min. The delivery speed was set to 180 words/min in the present study, and that may explain the current directionality effect.

In sum, in terms of CI directions, the present study finds better performance in the L1-L2 direction than in the L2-L1 direction, and this directionality effect was independent of participants' language contexts. Our result further confirms Gile's (2009, p. 176) assumption that for CI, the comprehension phase places heavier cognitive pressure on interpreters than the production phase. With adequate source language comprehension, retrieving and producing corresponding equivalents in the L2 is not too effortful for our participants. However, it should be noted that participants in the present study were all interpreting master students with at least two years of interpreting training. Also, all were proficient L2 users, which implies that their L2 knowledge was sufficient for them to express what they had understood.

5.2.3 Study Abroad and Non-Study Abroad Groups

The overall scores for CI performance in the two groups differed in the L2-L1 direction, with the SA group attaining higher scores than the NSA group. However, the two groups were comparable in the L1-L2 direction. Moreover, the SA group exhibited more balanced interpreting performance in both interpreting directions, and the NSA group showed a significantly larger degree of asymmetry between the two directions. This result can be attributed to two factors.

Firstly, for the same reason that more balanced lexical processing was observed in the SA group (as discussed in Section 5.1.2.2), the SA environment helped inhibit participants' dominant L1, thereby alleviating the cognitive burden during L2 processing. Secondly, according to the interpreting Effort Model (Gile, 1995, 2009), the total processing resources required for an interpreting task are the sum of the processing resources of all sub-processes. When the SA group was processing the L2 in CI tasks, some cognitive burden may have been reduced by the L2 environment, and those freed resources could have been directed towards other components of interpreting (Chabasse & Dingfelder Stone, 2015; Christoffels et al., 2003; Gile, 2009, p. 159), such as memorising, notetaking and information retrieving, leading to better CI performance in the SA group than the NSA group. On the other hand, previous sections have shown that the SA group was more efficient in L2 processing than the NSA group; Therefore, it appears that more efficient L2 processing contributes to better interpreting performance. However, it is interesting to explore why the SA group demonstrated superior performance in the direction of L2-L1 interpreting but not in the opposite direction.

Psycholinguistic studies on multi-tasking show that language processing is slower when participants are engaged in more than one task at a time (Marcel Adam Just et al., 2008; Newman et al., 2007). This is also the case in interpreting. In CI, comprehension of the input language takes place concurrently with notetaking. As the incoming messages unfold, the WM, which is responsible for storing and processing the information, can analyse the corresponding sound signals and extract meaning. However, when the linguistic processing becomes too effortful, sound signals accumulate and are temporarily stored in the WM until participants are able to process the signals into meaningful segments or take notes. In such a situation, WM availability can be exceeded rapidly, and if this happens, either the incoming information cannot be attended to, or it is attended to at the expense of previously heard segments. In other words, while inefficient L2 processing in L1-L2 interpreting is only revealed as hesitation or grammatical mistakes in participants' L2 production (Gile, 2009, p. 223), in the L2-L1 interpreting direction, inefficient L2 processing can bring about non-comprehension which has far-reaching implications. This explains why in the present study, more efficient L2 processing in the SA group led to pronounced superiority in the L2-L1 CI direction, but this linguistic advantage was less noticeable in the L1-L2 direction. Additionally, with natural exposure, language constituents are heard and used regularly, so during the comprehension phase of CI, retrieving may be easier (Gile, 2009, p. 238), which also makes L2 comprehension less effortful for the SA group (Antoniou, 2010).

In answering the second research question, the findings of the present study suggest that both SA and NSA groups were better at CI in L1-L2 direction than the opposite direction. Moreover, the SA group demonstrated better CI performance than NSA group, and the contribution of better linguistic performance was more robust in the SA group's L2-L1 CI direction. The findings also show that the accuracy of L2-L1 lexical processing correlates more strongly with participants' CI performance than it does with L1-L2 lexical processing accuracy or lexical processing response times.

5.3 The Effect of Working Memory

The third research question aims to explore the role of WM on linguistic and CI performance among participants in different language environment.

Hypothesis 3 states that participants' WM resource availability will significantly correlate with linguistic and CI performance. In the present study, both groups' WM resource availability significantly correlated with their lexical and grammatical processing as well as their CI performance in both language directions. Moreover, participants with higher WM significantly outperformed their lower WM counterparts in both linguistic and CI performance, independently of language environment. This hypothesis was supported. According to WM multi-component model, proposed by Baddeley and Hitch (1974), WM involves the temporary storage and manipulation of information at the same time; therefore, with comparable linguistic knowledge and language proficiency, larger WM resource availability leads to better information storage and manipulation (Baddeley, 1996, 1998, 2006, 2010).

It is also hypothesised (in Hypothesis 4) that there will be a tendency for high WM participants with SAE to outperform the rest of the participants in linguistic and CI tasks. This hypothesis was supported. Participants with higher than median WM and SAE demonstrated better lexical and grammatical processing as well as higher and more balanced CI performance than participants with higher WM availability but no SAE and participants with SAE but lower than median WM availability.

Lastly, it was assumed that the language environment affects how WM is used in L2; thus, it was hypothesised (in Hypothesis 5) that SA participants would demonstrate larger WM resource availability than the NSA group. This hypothesis was not supported. The results showed that SA and NSA groups did not differ in their WM resource availability.

Section 5.3.1 discusses the association of WM with lexical processing; followed by a discussion of the correlation between WM and subject-verb number agreement in Section 5.3.2. In Section 5.3.3, the correlation between WM and CI performance is explored, followed by the correlation between WM and SAE in Section 5.3.4.

5.3.1 Working Memory and Lexical Processing

As predicted, WM correlated with participants' performance on word translation recognition tasks. The magnitude of these correlations ranged between .30 to .37 (Table 4.14). Even though moderately weak, the correlations are consistent with previous findings (e.g., K. Martin & Ellis, 2012; Sunderman & Kroll, 2009). In order to investigate the WM effect more closely, participants in the SA and NSA groups were combined and were re-categorised by a median split procedure based on WM availability into High-WM and Low-WM groups.

5.3.1.1 Direction-Dependent Asymmetry

The results showed that both High-WM and Low-WM groups exhibited word directional asymmetry: word translation from L2-L1 was faster and more accurate than translation in the opposite direction. This result again lends support to Green's (1998) bilingual Inhibitory Control model, confirming that bilinguals exert more cognitive effort in suppressing their dominant L1 than they do in suppressing L2 when both languages are co-activated. Therefore, our results support previous findings that lexical processing performance in L1-L2 word translation recognition is slower and less accurate than it is in the opposite L2-L1 direction (e.g., Costa & Santesteban, 2004; Linck et al., 2008; Meuter, 2009; Meuter & Allport, 1999; Olson, 2017; Peeters et al., 2014). Also, the existence of direction-dependent asymmetry confirms Christoffels et al.'s (2006) proposition that cognitive resources do not constitute a critical factor in modulating the magnitude of cross-language activation, as higher WM resource availability contributes to faster and more accurate performance in both directions. However, the asymmetry between the two directions is determined more by the relative strength of the two languages and the language context (Baus et al., 2013; De Bot, 2000; Linck et al., 2009; Meuter & Allport, 1999).

Therefore, our results suggest that the direction-dependent asymmetry exists regardless of participants' WM availability, and both High- and Low-WM groups were faster and more accurate in L2-L1 lexical processing than they were in the opposite direction.

5.3.1.2 High-WM and Low-WM groups

Results from the comparison of the High-WM and Low-WM groups showed that although the two groups had comparable lexical knowledge, the High-WM group was significantly better than the Low-WM group in terms of word translation recognition accuracy and response times in both directions. Our results support previous findings that learners with larger WM availability tend to exhibit faster and more accurate lexical processing performance (e.g., Kormos & Sáfár, 2008; Kroll et al., 2002; Linck et al., 2008, 2014; Michael et al., 2003; Sunderman & Kroll, 2009; Trude & Tokowicz, 2011). This phenomenon can be explained by the fact that participants with larger WM are more capable of inhibiting the activation of lexical competitors and are less hampered by interference from the non-target language (Linck et al., 2008; Michael et al., 2002, 2003; Michael & Gollan, 2005).

In sum, results from the present study support previous findings that WM has a robust, positive relationship with word translation recognition in both language directions. Our results also confirm previous findings that the executive function of WM facilitates lexical processing by inhibiting irrelevant language competitors.

5.3.2 Working Memory and Grammatical Accuracy

As hypothesised, in the present study, participants in the High-WM group were more capable of attending to subject-verb number agreement in their L2 production when performing L1-L2 CI tasks than participants in the Low-WM group. These results are consistent with previous studies reporting WM effects on L2 grammatical processing (e.g., Hartsuiker & Barkhuysen, 2006; Havik et al., 2009; Hopp, 2010; Keating, 2009; K. Martin & Ellis, 2012; McDonald, 2008; Roberts, 2012; Sagarra & Herschensohn, 2010; Sagarra & Labrozzi, 2018).

WM is responsible for attention allocation, and it also serves as a workspace to store and process information at the same time (Baddeley, 2003, 2006; K. Martin & Ellis, 2012). Similar to its effect on lexical processing, larger WM availability reduces the tendency for learners to be distracted by the intrusion of the unintended language or other distractors (Michael et al., 2002, 2003). This allows them to focus on target language processing. Furthermore, grammatical processing goes beyond the selection of individual words (K. Martin & Ellis, 2012), as grammatical agreement requires speakers to formulate the rest of the sentence while referring to what they have already uttered, so that they can make the upcoming parts grammatically consistent with what they have already produced (e.g., Badecker & Kuminiak, 2007; De Abreu & Gathercole, 2012; Hartsuiker & Barkhuysen, 2006; Jin, 2010; Kellogg, 2004; Miyake & Friedman, 1998). Therefore, for L2 learners who are not automatised in linguistic processing, attending to grammatical markers such as processing subject-verb number agreement demands a lot of their WM (e.g., Hilton, 2011, p. 20; Kormos, 2006, p. 167; Sagarra & Labrozzi, 2018; Sanz & Leow, 2011; Ullman, 2001).

In the present study, when participants were undertaking L1-L2 interpreting, firstly their L2 production was achieved by inhibiting the dominant L1 which drew on their

WM resources. Secondly, during the L2 processing, in order to make the verbs agree with the number of subjects in sentences, participants had to concurrently retrieve and maintain the subject while producing the verb and the remaining parts of a sentence (Ellis, 2011; Hartsuiker & Barkhuysen, 2006; McDonald, 2006). Even though participants recruited in our study had already acquired the target grammatical structure, English was nevertheless their L2. This implies that participants could not process English implicitly but had to draw on their declarative knowledge while speaking, and this controlled action is WM-demanding (e.g., Hilton, 2011, p. 146; Kormos, 2006, p. 167; Sanz & Leow, 2011; Ullman, 2001).

What makes interpreting tasks effortful is that WM resources are crucial but finite (Eriksson et al., 2015), and every non-automated element of the process requires a share of those resources. A sudden peak in requirements for one element would require additional resources to be diverted away from another element (Gile, 2009, p. 159), which would lead to resource shortages in the other elements and an impaired interpreting performance (Chabasse & Dingfelder Stone, 2015; Christoffels et al., 2003; Kurz, 2008). Note that participants in our study were not simply speaking their own minds but had to retrieve semantic information stored in their memory or notes on their notebooks while trying to produce sentences in the L2 in a grammatically correct manner, and this is particularly difficult. Therefore, it is not surprising that participants with high WM availability demonstrated better subject-verb number agreement than their low WM peers, independently of language contexts.

Critically, the results of the present study also indicate that within the High-WM group, participants with SAE were more accurate in their subject-verb number agreement than all other participants, followed by High-WM NSA participants, Low-WM SA participants, and Low-WM NSA participants. This result is consistent with previous studies (e.g., Labrozzi, 2009; Marqués-Pascual, 2011; Sagarra & Labrozzi, 2018), showing interactive effectiveness of WM and SAE regarding L2 grammatical processing. Large WM availability enables learners to cope effectively with rich input from the study abroad context and also allows them to pay more attention to language form and meaning during L2 processing (Lafford, 2006; Marqués-Pascual, 2011). However, the aforementioned interactions were not significant, perhaps due to limited statistical power related to sample size. To get a more statistically valid measurement, greater numbers of participants are needed. Therefore, this result should be interpreted with caution.

The findings of this study confirm that WM is strongly associated with grammatical processing. Comparisons of groups of high and low WM speakers with respect to accuracy rates for subject-verb agreement have been investigated by many studies involving speech-eliciting tasks, but very few of them are interpreting studies, and our results allow us to confirm that the effect of WM on subject-verb agreement also applies in the interpreting modality. Additionally, our results may be indicative of the superiority of the combined effects of high WM span and SAE over the rest.

5.3.3 Working Memory and Consecutive Interpreting Performance

A significant correlation between WM and both directions of CI performance was found in both the SA and the NSA groups. And when participants in the SA and NSA groups were mixed together and re-categorised on the basis of a median split on WM availability into High-WM and Low-WM groups, a more salient effect of WM on CI performance emerged. The High-WM group performed better than the Low-WM group in CI in both directions. This outcome again supports the findings of previous studies which have suggested that WM is positively correlated with interpreting performance (e.g., R. Cai et al., 2015; Christoffels et al., 2003; Injoque-Ricle et al., 2015; Macnamara & Conway, 2016; Timarová et al., 2014; Tzou et al., 2011; F. Wang, 2017). The results of our study also suggest that WM's effect on CI performance is independent of the participant's language context.

In the present study, the role of WM is more prominent when correlated with participants' L2-L1 CI performance than with their performance in the opposite direction, as the gap between the High- and Low-WM groups' mean scores were larger in the L2-L1 direction than in the opposite direction (L2-L1: 5.51 vs. 3.69; L1-L2: 6.02 vs. 4.90). This indicates that in CI, the L2-L1 direction is more WM-demanding than the L1-L2 direction, confirming the finding in Section 5.2.2 that listening and analysing the L2 imposes a heavier cognitive burden than producing a well-understood message in the L2.

These above findings contrast with findings indicating that there is no significant relationship between WM and interpreting performance (e.g., M. Liu et al., 2004; J.

Wang, 2016). This apparent contradiction may be explained by participants' different levels of interpreting experience. Both the abovementioned studies recruited professional interpreters, whereas the present study recruited interpreting students. Interpreting experience is a complex variable, as experience could add additional advantages to interpreting performance (Henrard & Van Daele, 2017), such as their superiority in adopting domain-specific skills (Christoffels et al., 2003; M. Liu et al., 2004; J. Wang, 2016). Compatibly, Dong and Xie (2014) also reported that interpreting experience could significantly contribute to cognitive enhancement. Therefore, it is hard to tell whether the superior interpreting performance of professional interpreters is the result of large WM or of abundant interpreting experience. In other words, for experienced interpreters, other factors may supersede the effect of WM (Kopke & Nespoulous, 2006), and this is why many researchers recruit student interpreters or untrained bilinguals to avoid 'contamination' by the strategies that professional interpreters have developed (e.g., Christoffels et al., 2003; F. Wang, 2017).

Participants in the present study were comparable in their ages, language proficiency, interpreting training length and word knowledge, and our results indicate that participants' WM resource availability, as measured by a reading span task, is strongly associated with their overall CI performance. Based on that, WM availability might be useful in grouping more-proficient and less-proficient interpreting students in interpreting training.

Our study also finds that High-WM participants with SAE demonstrated better CI performance than the rest of the participants. This manifested as better and more balanced CI performance in both language directions. Interestingly, our results also reveal that Low-WM participants from the SA and NSA groups demonstrated comparably low and inconsistent CI performance, which may indicate that without sufficient WM resources, SAE does not lend much help to participants' CI performance. These results suggest that WM and SAE exert a combined effect on CI performance. However, WM may make a higher contribution than SAE to interpreting performance. Even though more participants are needed if we want to obtain statistically valid measurements, the pattern is clear.

5.3.4 Working Memory and Study Abroad Experience

To investigate whether SAE influences how WM resources are used in the L2, this study also compared WM resource availability, as measured by reading span tasks, of the SA and NSA groups. The two groups did not differ in their WM availability at the time of data collection, which is consistent with the results from previous studies (e.g., Linck et al., 2008; Sunderman & Kroll, 2009; Tokowicz et al., 2004). In other words, the length of SAE does not link to WM resources used in the L2.

In response to findings that WM resource availability does not benefit from SAE, Linck et al. (2008) argued that learners might need to adjust to the new language environment when they are initially immersed in an L2 country, and that this adjustment process consumes more cognitive resources at first. But they speculated that the resources would recover with time, as the immersion period gets longer and bilinguals have more experience managing the two languages in the L2 environment. The immersion period in Linck et al. was three months, which they deemed to be short. However, in the present study, the SAE lasted for 24 months on average, and this should have taken the SA participants past the initial costing stage proposed by Linck et al. Therefore, the longer immersion length of the present study rejects this assumption, indicating that SAE does not enhance WM resources per se.

The question is, does SAE bring no benefit to WM at all? Previous studies that investigated SAE and learners' WM suggest that exposure could facilitate linguistic processing and reduce the WM resources required to process the target language, thus eventually freeing the WM resources for other tasks (e.g., Antoniou et al., 2015; Christoffels et al., 2003; Orena et al., 2015). Gathercole and Baddeley point out that:

> …a tradeoff between processing and storage is necessary whenever a language processing task exceeds the limited resources available to the comprehender. The final principle is that there are important individual differences in WM capacity, and that these are due either to variation in the total capacity of resources available, or to the efficiency with which cognitive processes are executed (Gathercole & Baddeley, 1993, p. 222).

In other words, according to Gathercole and Baddeley, if learners can improve how efficiently they utilise the WM resources in a linguistic task by enhancing the efficiency of cognitive processes within that task, they can to some extent compensate for low WM availability and still achieve a satisfactory performance. This is compatible with the finding that large WM resource availability is important for linguistic and CI performance, as one must be cognisant of the fact that learners with larger WM availability are more likely to be able to utilise their WM in an efficient way (Tokowicz, 2014, p. 66). If resources are limited, the resolution of competition for resources will be too demanding and difficult to achieve. As a result, more errors will occur, especially in cognitively demanding tasks. This explains why participants with SAE combined with high WM have outperformed high WM learners without SAE and SA learners with low WM in previous SLA studies (Morgan-Short, 2007; Sagarra & Labrozzi, 2018; Sunderman & Kroll, 2009; Tokowicz et al., 2004; Wright, 2013) and the present interpreting study. SAE has promoted participants' linguistic processing automaticity, and thus reduced learners' WM burdens in linguistic and CI tasks.

In answering the third research question, the findings of the present study confirm that WM resource availability is associated with linguistic and CI performance, and this association is independent of language environment. Apart from that, High-WM participants with SAE demonstrated better linguistic and CI performance than the rest of the participants, implying that WM and SAE have a joint effect on participants' linguistic and CI performance. Specifically, WM seems to play a more important role in these aspects than language context does. In terms of the relationship between SAE and WM resources, SAE does not enhance learners' WM, but may help learners to utilise their WM resources more efficiently.

5.4 Summary

This chapter investigated the effects of SAE and WM on linguistic (lexicon, grammar and fluency) and CI performance in two groups of interpreting students with different language contexts. The results showed that interpreting students in different language contexts did differ in their linguistic and CI performance. In addition, students with high WM also distinguished themselves from the low WM students in linguistic and CI performance.

SAE exerted a positive effect on interpreting students' linguistic performance. This manifested in more balanced word translation recognition performance, higher grammatical accuracy and more fluent L2 output in the SA group. These results support Interaction Theory and the Inhibitory Control model, and are also consistent with the previous claim (e.g., Baus et al., 2013; Linck et al., 2009; Sagarra & Labrozzi, 2018) that the L2 context can help suppress the L1 and promote L2 processing. Our findings advance our understanding of the SAE effect by indicating that this effect also holds true to interpreting students who are required to constantly keep their two languages activated, and who deliberately switch between languages in their daily lives. Moreover, the findings from this task also confirm that SAE has an impact on participants' word lexical processing error patterns.

Another research question explored in the present study concerns the role of SAE on CI performance. The SA participants demonstrated better CI performance in the L2-L1 direction than their NSA counterparts. Both groups' CI performance was strongly correlated to their L2-L1 word translation recognition accuracy, indicating that L2-L1 lexical processing can be used as a predictor of participants' overall CI performance.

This study also examined whether WM resource availability, as assessed by the reading span task, correlated with interpreting students' linguistic and CI performance. The results show that participants with higher WM availability demonstrated faster and more accurate lexical processing, more accurate grammatical processing, and better CI performance. This implies that WM played a crucial role in participants' overall linguistic and CI performance, and the role of WM is independent of participants' language contexts. Specifically, High-WM participants with SAE demonstrated better linguistic and CI performance than the rest of the participants, indicating a combined effect from SAE and WM.

Chapter 6 Conclusion

This chapter concludes my investigation into whether SAE benefits interpreting performance, and how WM is involved. My conclusion consists of three sections. Section 6.1 presents the major findings and responds to each research question. Section 6.2 discusses the implications of this study, including the implications for Interaction Theory, the Inhibitory Control model, the WM multi-component model and the Efforts model, as well as the pedagogic implications for interpreting training and learning. Section 6.3 outlines the limitations of the thesis and offers some suggestions for future research.

6.1 Summary of Major Findings

In response to the research questions and hypotheses advanced in Chapter 3, the most relevant findings will be highlighted below.

(Q1) Compared with interpreting students without SAE, do interpreting students with SAE demonstrate better linguistic performance (i.e. better lexical and grammatical processing and fluency)?

Hypothesis predicted that SA students would be faster, more accurate, and more balanced (i.e. less asymmetry between the two directions of translation) in the word translation recognition task; they were expected to have higher subject-verb number agreement rate and be more fluent in their L2 production in Chinese-English (L1-L2) consecutive interpreting (CI) output. This prediction was based on Interaction Theory (Carroll, 1999; Long, 1981, 1996). SAE provides students with abundant opportunities to interact with native speakers in the target language, and according to Interaction Theory, frequent target-language interactions contribute to success in many areas of linguistic processing, such as comprehension, fluency, consolidation and the internalisation (automatisation) of existing linguistic knowledge. This claim is also supported by the Inhibitory Control model (Green, 1998), according to which the processing of the intended language is achieved by suppressing the unintended (but co-activated) one. Inhibiting the dominant L1 takes more effort than inhibiting the weaker L2, but SAE can effectively attenuate learners' dominant L1 and promote L2 processing (e.g., Baus et al., 2013; Linck et al., 2009; Sagarra & Labrozzi, 2018). The results of this study provide support for this Hypothesis 1. That is, although interpreting students with SAE were not more accurate in lexical retrieval, they exhibited more balanced word translation recognition performance. Also, they were more grammatically accurate overall, insofar as they exhibited higher subject-verb number agreement. Further, they were more fluent in their L2 production in the L1-L2 CI output.

(Q2) Compared with interpreting students without SAE, do interpreting students with SAE exhibit better CI performance?

Hypothesis 2 predicted that interpreting students with SAE would demonstrate better CI performance than their NSA counterparts, as measured by the assessments of two expert raters. This claim was based on the interpreting Effort Model (Gile, 1995, 2009), which asserts that the total cognitive resources required for an interpreting task is the sum of the cognitive resources of all sub-processes. If elements of linguistic processing become more efficient and automatised, then more resources can be redirected to other elements or sub-processes of interpreting, such as memorising, notetaking and information retrieving. This would presumably lead to better CI performance in the SA than the NSA group. However, it is noted that, despite the more balanced CI performance observed in the SA group, this higher linguistic processing efficiency led to pronounced superiority only in the L2-L1 CI direction, and was less obvious in the L1-L2 direction when compared with the NSA group. This confirms Gile's (2009, p. 176) assumption that for CI, the comprehension phase is more cognitively demanding than the production phase. In sum, the results of this study provide support for Hypothesis 2. That is, interpreting students with SAE outperformed those without this experience in overall CI performance, and the benefit of SAE was more salient in the L2-L1 direction.

(Q3) Does WM affect performance on linguistic and CI tasks?

Three hypotheses were proposed regarding the effect of WM in the present study. These hypotheses are based on the WM multi-component model (Baddeley, 2003, 2006), according to which WM is a cognitive resource in limited supply. WM is responsible for storing and manipulating the target information as well as inhibiting distractors, and these functions are important for both linguistic and CI performance.

Furthermore, WM is also a flexible resource which can be allocated among specific linguistic aspects so as to achieve an optimal processing result (Gathercole & Baddeley, 1993, p. 99; Kormos, 2006; Robinson, 2003). Therefore, efficient linguistic processing alleviates WM burdens during interpreting and thus contributes to CI performance.

Hypothesis 3 states that WM resource availability will significantly correlate with linguistic and CI performance. The results support this hypothesis, as WM correlated strongly with both groups' lexical and grammatical processing and bidirectional CI performance. Moreover, the High-WM group surpassed the Low-WM group in all the abovementioned tasks.

It was also hypothesised (in Hypothesis 4) that there would be a tendency for high WM participants with SAE to outperform the rest of the participants in linguistic (grammatical accuracy) and CI tasks. Our results suggest that within the High-WM group, participants with SAE were more accurate in their subject-verb number agreement than all other participants. This suggests that WM and SAE interact during L2 grammatical processing. A similar pattern was also found in CI performance; that is, high-WM participants with SAE performed better than the other participant subgroups, as indicated by their higher and more balanced bidirectional CI scores. Regarding Hypothesis 4, a tentative conclusion can be drawn that WM and SAE have an interactive and positive effect on linguistic and CI performance.

The fifth and last hypothesis concerns the relationship between WM and SAE. It predicts that WM resource availability will be larger in the SA group than in the NSA group. The results show that the two groups did not differ in their WM availability as assessed by the reading span task at the time of data collection. In other words, the length of SAE did not correlate with the availability of WM resources used in the L2. Therefore, this hypothesis is rejected.

In sum, the hypotheses related to Q3 confirm that WM plays a significant role in interpreting students' linguistic and CI performance, and this role is independent of participants' language environment. Apart from that, a trend could be observed that High-WM participants with SAE showed superior linguistic and CI performance than the rest of the participants, implying that WM and SAE have a joint effect on participants. Specifically, WM seems to have more impact on interpreting tasks than

SAE does. In terms of the relationship between SAE and WM resources, SAE does not increase learners' WM resource availability, but the experience of studying abroad may facilitate increased efficiency in the use of WM resources.

6.2 Implications of the Study

Theoretical Implications

This study used an interdisciplinary approach to investigate the effects of SAE and WM on interpreting students' linguistic and CI performance. The results make a contribution to addressing a research gap regarding the relationship between language environment, interpreting performance and WM. Firstly, the results lend support to Interaction Theory by showing that L2 contexts with rich and natural L2 input, interaction, and output can promote interpreting students' lexical and grammatical automaticity as well as their oral fluency. Importantly, these L2 benefits were not observed among participants who learnt the L2 in the L1 country with classroom instruction. Secondly, this study also contributes to the Inhibitory Control model by exploring lexical and CI translation asymmetry. The L2 context effectively helps resolve language competition by suppressing participants' L1, thus contributing to reduced language switching costs, even for interpreting students who have to constantly switch between languages and are required to deliberately keep their two languages activated in their daily lives. This study also offers insights into the WM multi-component model and the interpreting Effort Model. Efficient linguistic processing optimises WM resource allocation, which results in better CI performance.

Comparisons of bilinguals in different language learning contexts and with different WM capacities have been conducted in a number of second language acquisition studies using speech-eliciting tasks such as dialogues or storytelling. Few studies have investigated the effects of learning context and WM capacity on cognitively demanding tasks such as interpreting. The results of this study allow us to generalise about the effects of SAE and WM on the interpreting process (and specifically the understudied variant, CI). That is, this study indicates that SAE improves the grammatical accuracy and fluency of interpreting output in the L1-L2 direction. SAE also reduces overall interpreting direction-dependent asymmetry.

Pedagogical Implications

This study presents evidence that SAE may make linguistic processing (lexical and grammatical) and CI performance more efficient than domestic classroom instruction. This research has implications for several sub-disciplines of interpreting studies and second language acquisition, particularly in relation to the effects of linguistic and CI performance in adult interpreting students moving from foreign language instruction to study abroad settings. The lack of research on the impact of SAE on interpreting learners highlights the contribution this study makes to the field of interpreting as well as second language acquisition.

The present findings support the assumption that the influence of SAE on linguistic and CI performance is beneficial, as it underscores the significance of L2 input, interaction and output in linguistic processing. The findings of the present study highlight the importance of productive language use, as the two groups of participants enjoyed comparable amounts of L2 input, but were significantly different in the amounts of time spent interacting with native speakers.

One thing that can be gleaned from these findings is that language learners should be encouraged, not only to get rich target-language input, but also to interact with native speakers. This will create environments in which they can stretch themselves to make use of their linguistic knowledge to express their minds to the fullest extent and, in doing so, they will effectively promote the automatic processing of acquired linguistic rules. As has been reported in the literature and supported by the findings above, the more linguistic elements are automatised, the more WM resources can be freed for parallel cognitive processes during complex tasks such as CI.

This study also finds that WM resource availability significantly contributes to interpreting students' linguistic and CI performance, and that this cognitive resource plays a more significant role than SAE in modulating CI performance. Therefore, students who are undertaking CI training need to possess large WM resources, or they need to be able to allocate their WM resources efficiently. It is very difficult for an individual to develop their WM resources once they have reached maturity, but they can achieve more efficient WM allocation by enhancing their L2 processing efficiency.

Thus, the findings offer some guidance for L2 educators and interpreting educators in planning for effective training and choice of task mode for specific pedagogical purposes. The findings are also important for interpreting students who do not have SAE, and who therefore have relatively little contact with native speakers of their L2 language in their daily lives. Under such circumstances, more opportunities to interact with native speakers should be provided to interpreting students, for example through synchronous distance communication via the internet on a highly frequent basis to improve their L2 processing and thus enrich their CI training.

6.3 Limitations and Suggestions for Further Research

This study has several limitations which should be noted.

Firstly, the small sample size of each group makes it difficult to draw robust causal inferences regarding SAE and WM and their relationships to CI performance. A trend generalised from a small sample may not accurately reflect the situation for a whole population, and this limits generalisability.

Secondly, participants in this study were master's students majoring in Chinese-English interpreting and translation. They were at the end of their second year or the beginning of their third year of interpreting training. That is, their English learning was academically driven, and they were quite advanced. For bilinguals from other majors or levels of L2 proficiency or interpreting expertise, such as professional interpreters, the results may very well differ.

Thirdly, even though the use of expert raters is regarded as a suitable method for evaluating overall interpreting performance, and even though inter-rater reliability was high and thus deemed to be satisfactory for our purposes, we acknowledge that there also exist other objective methods for assessing interpreting performance.

In terms of research design, due to the lack of a pre-test before the SA group embarked on their abroad experience, the comparisons presented in this study correspond to their linguistic and CI performance during SAE. Thus, the current cross-sectional study does not provide empirical data that would test for whether SAE improved (or diminished) the SA groups' performance during the course of their interpreting studies. Perhaps, a longitudinal study would be better suited and would eliminate more variables and provide more solid results. This is a question that may be explored in future longitudinal research studies. Of course, a longitudinal study spanning three to five years would require more time and resources than would be feasible for a PhD thesis, but the results would advance the field.

We now turn to recommendations for future research. Firstly, the findings highlight the critical influence of SAE on interpreting students' linguistic and CI performance, and they highlight the need for more research. Indeed, to fully validate the results presented within this thesis, and the implications deriving from studying in the L2 environment, it would be ideal to also recruit participants with a broader spectrum of L2 proficiency levels and interpreting expertise as well as differing durations of SAE. It would also be useful to explore whether any benefits of SAE persist after returning to the home country. This would help to further elucidate what it is about this experience that accounts for divergent performances between SA and NSA learners. This is particularly pertinent during uncertain times such as the current COVID-19 pandemic, when study abroad programs are disrupted and many students may be locked down and obliged to communicate with target-language speakers online.

Secondly, the present study shows that SAE contributes to improvements in interpreting students' subject-verb number agreement accuracy rates in their L2 output. Future research may examine the impact of SAE on accuracy in processing a wider range of grammatical structures to determine whether the observed pattern generalises across to other aspects of grammar. Additionally, SAE is a complex and high-order variable that could encompass many different aspects, such as language experiences, cultural interactions and emotions. This thesis only focuses on the language development brought about by such an environment. Future studies could also take participants' psychological state such as anxiety and pressure into consideration when they are in an environment with different culture from their own country. And it is interesting to know how these emotions affect their linguistic performance.

Thirdly, this study provides tentative conclusion regarding the joint effects of SAE and WM. More evidence is needed to confirm this finding.

In conclusion, the present thesis adopted an interdisciplinary approach to the examination of interpreting studies. The results demonstrate that SAE and WM contribute to interpreting students' linguistic and CI performance. Importantly, regarding CI, the influence of L2 environment was more salient in the L2-L1 than in the L1-L2 direction, indicating that L2 comprehension may be more malleable (or more susceptible to language environment) than production, even after a sustained period of L2 exposure and usage. This study also highlights the salient role of WM in linguistic processing and interpreting among students studying in different language contexts.

References

Abutalebi, J., Annoni, J.-M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., & Khateb, A. (2008). Language control and lexical competition in bilinguals: An event-related fMRI study. *Cerebral Cortex*, *18*(7), 1496–1505.

Alarcón, I. (2009). The processing of gender agreement in L1 and L2 Spanish: Evidence from reaction time data. *Hispania*, *92*(4), 814–828.

Alloway, R. G., & Alloway, T. P. (2013). Working memory: An introduction. In T. P. Alloway & R. G. Alloway (Eds.), *Working memory: The connected intelligence*. Psychology Press.

Anderson, J. R. (2015). *Cognitive psychology and its implications* (8th ed.). Worth.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, *30*(3), 221–256.

Antoniou, M. (2010). *One head, two languages: Speech production and perception in Greek-English bilinguals* [Doctoral dissertation, Western Sydney University]. Western Sydney University ResearchDirect.

Antoniou, M., Wong, P. C., & Wang, S. (2015). The effect of intensified language exposure on accommodating talker variability. *Journal of Speech, Language, and Hearing Research*, *58*(3), 722–727. https://doi.org/10.1044/2015_JSLHR-S-14-0259

Antón-Méndez, I., Nicol, J. L., & Garrett, M. F. (2002). The relation between gender and number agreement processing. *Syntax*, *5*(1), 1–25.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders*, *36*(3), 233–240. https://doi.org/10.1016/S0021-9924(03)00022-4

Arnett, C. (2013). Syntactic gains in short-term study abroad. *Foreign Language Annals*, *46*(4), 705–712. https://doi.org/10.1111/flan.12052

Atkins, P. W. B., & Baddeley, A. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, *19*(4), 537–552. https://doi.org/10.1017/S0142716400010353

Autin, F., & Croizet, J.-C. (2012). Improving working memory efficiency by reframing metacognitive interpretation of task difficulty. *Journal of Experimental Psychology: General*, *141*(4), 610.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 5–28.

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology*, *8*(2), 234–238.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208.

Baddeley, A. (2006). Working memory: An overview. In S. J. Pickering & G. D. Phye (Eds.), *Working memory and education* (pp. 1–31). Academic Press.

Baddeley, A. (2009). *Memory*. Psychology Press.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), 136–140.

Baddeley, A., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158.

Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–89). Academic Press.

Baddeley, A., & Hitch, G. J. (2000). Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged? *Journal of Experimental Child Psychology*, *77*(2), 128–137.

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory Language*, *56*(1), 65–85.

Bajo, M., Padilla, F., & Padilla, P. (2000). Comprehension processes in simultaneous interpreting. In A. R. A. Chesterman, N. Gallardo San Salvador, & Y. Gambier (Eds.), *Translation in context: Selected papers from the EST Congress* (pp. 127–142). John Benjamins.

Bardovi-Harlig, K. (1992). The use of adverbials and natural order in the development of temporal expression. *IRAL: International Review of Applied Linguistics in Language Teaching*, *30*(4), 299. https://doi.org/10.1515/iral.1992.30.4.299

Barik, H. C. (1975). Simultaneous interpretation: Qualitative and linguistic data. *Language and Speech*, *18*(3), 272–297. https://doi.org/10.1177/002383097501800310

Barik, H. C., & Swain, M. (1976). Primary-grade French immersion in a unilingual English-Canadian setting: The Toronto study through grade 2. *Canadian Journal of Education/Revue Canadienne de l'education*, *1*(1), 39–58.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570-585.

Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, *2*, 324. https://doi.org/10.3389/fpsyg.2011.00324

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*(3), 402–409.

Becker, M., Schubert, T., Strobach, T., Gallinat, J., & Kühn, S. (2016). Simultaneous interpreters vs. professional multilingual controls: Group differences in cognitive control as well as brain structure and function. *NeuroImage*, *134*, 250–260. https://doi.org/10.1016/j.neuroimage.2016.03.079

Bhandari, R. (2017). *A World on the Move: Trends in Global Student Mobility* (Higher Education Policy for Minorities in the United States) [Press release]. IIE Center for Academic Mobility Research and Impact. http://hdl.handle.net/10919/83142

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*(3), 89–129.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*(2), 102–118.

Bock, K., & Levelt, W. (1994). Language production—Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition*, *10*(3), 299–314. https://doi.org/10.1017/S1366728907003082

Borovsky, A., Burns, E., Elman, J. L., & Evans, J. L. (2013). Lexical activation during sentence comprehension in adolescents with history of specific language impairment. *Journal of Communication Disorders*, *46*(5–6), 413–427. https://doi.org/10.1016/j.jcomdis.2013.09.001

Briggs, J. G. (2015). Out-of-class language contact and vocabulary gain in a study abroad context. *System*, *53*, 129–140. http://dx.doi.org/10.1016/j.system.2015.07.007

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press.

Brooks, P. J., Kempe, V., & Sionov, A. (2006). The role of learner and input variables in learning inflectional morphology. *Applied Psycholinguistics*, *27*(2), 185–209.

Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, *30*(2), 225–251.

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 137.

Bühler, H. (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. *Multilingua*, *5*(4), 231–235.

Bultena, S., Dijkstra, T., & Van Hell, J. (2015). Language switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism: Language and Cognition*, *18*(3), 453–469.

Bygate, M. (2002). Speaking. In R. B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 27–38). Oxford University Press.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles (research article). *PLoS ONE*, *5*(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Cai, R., Dong, Y., Zhao, N., & Lin, J. (2015). Factors contributing to individual differences in the development of consecutive interpreting competence for beginner student interpreters. *The Interpreter and Translator Trainer*, *9*(1), 104–120.

Carpenter, P. A., Miyake, A., & Just, M. A. (1995). Language comprehension: Sentence and discourse processing. *Annual Review of Psychology*, *46*(1), 91–120.

Carroll, J. B. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals*, *1*(2), 131–151. https://doi.org/10.1111/j.1944-9720.1967.tb00127.x

Carroll, S. E. (1999). Putting ‘input’in its proper place. *Second Language Research*, *15*(4), 337–388.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386–404.

Chabasse, C., & Dingfelder Stone, M. (2015). Capacity management in interpretation: Efforts, directionality, and language pair considerations. In D. Andres & M. Behr (Eds.), *To know how to suggest...: Approaches to teaching conference interpreting* (pp. 75–102). Frank & Timme GmbH.

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268.

Chen, L., Shu, H. U. A., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition*, *10*(2), 161–174. https://doi.org/10.1017/s136672890700291x

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology*, *32*(5), 867. https://doi.org/10.1037/0012-1649.32.5.867

Chiaro, D., & Nocella, G. (2004). Interpreters' perception of linguistic and non-linguistic factors affecting quality: A survey through the World Wide Web. *Meta: Journal Des Traducteurs/Meta: Translators' Journal*, *49*(2), 278–293.

Chincotta, D., & Underwood, G. (1998). Simultaneous interpreters and the effect of concurrent articulation on immediate memory: A bilingual digit span study. *Interpreting*, *3*(1), 1–20.

Chmiel, A. (2016). Directionality and context effects in word translation tasks performed by conference interpreters. *Poznan Studies in Contemporary Linguistics*, *52*(2), 269–295. https://doi.org/10.1515/psicl-2016-0010

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT press.

Christoffels, I. (2004). *Cognitive studies in simultaneous interpreting*. PrintPartners Ipskamp.

Christoffels, I., & De Groot, A. M. B. (2004). Components of simultaneous interpreting: Comparing interpreting with shadowing and paraphrasing. *Bilingualism: Language and Cognition*, *7*(3), 227–240. https://doi.org/10.1017/S1366728904001609

Christoffels, I., & De Groot, A. M. B. (2005). Simultaneous interpreting: A cognitive perspective. In J. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 454–479). Oxford University Press.

Christoffels, I., De Groot, A. M. B., & Kroll, J. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, *54*(3), 324–345. https://doi.org/10.1016/j.jml.2005.12.004

Christoffels, I., De Groot, A. M. B., & Waldorp, L. J. (2003). Basic skills in a complex task: A graphical model relating memory and lexical retrieval to simultaneous interpreting. *Bilingualism: Language and Cognition*, *6*(3), 201–211. https://doi.org/10.1017/s1366728903001135

Christoffels, I., Firk, C., & Schiller, N. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208.

Chun, D. M., & Payne, J. S. (2004). What makes students click: Working memory and look-up behavior. *System*, *32*(4), 481–503.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3–42.

Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, *26*(2), 227–248. https://doi.org/10.1017/S0272263104262040

Collentine, J., & Freed, B. (2004). Learning context and its effects on second language acquisition: Introduction. *Studies in Second Language Acquisition*, *26*(2), 153–171. https://doi.org/10.1017/S0272263104262015

Conway, A. R., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354.

Cook, V. (2003). Changing the first language in the L2 user's mind. *Effects of the Second Language on the First*, *1*, 1–18.

Corsellis, A. (2005). Training Interpreters to Work in the Public Services. In M. Tennent (Ed.), *Training for the new millennium: Pedagogies for translation and interpreting* (pp. 153–173). John Benjamins.

Costa, A. (2005). Lexical access in bilingual production. In A. M. B. De Groot & J. Kroll (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). Oxford University Press.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283-1296.

Costa, A., Kovacic, D., Franck, J., & Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Bilingualism: Language and Cognition*, *6*(3), 181–200. https://doi.org/10.1017/S1366728903001123

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory*, *50*(4), 491–511.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, *169*, 323–338.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268.

Crain, S., & Lillo-Martin, D. (1999). *An introduction to linguistic theory and language acquisition*. Wiley-Blackwell.

Dam, H. V. (2004). Interpreters' notes: On the choice of language. *Interpreting*, *6*(1), 3–17.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 561.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*(1), 1–18.

Daneman, M., & Hannon, B. (2007). What do working memory span tasks like reading span really measure? In N. Osaka, R. H. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 21–42). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198570394.003.0002

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433.

Darò, V., & Fabbro, F. (1994). Verbal memory during simultaneous interpretation: Effects of phonological interference. *Applied Linguistics*, *15*(4), 365–381.

Davidson, D. E. (2007). Study abroad and outcomes measurements: The case of Russian. *The Modern Language Journal*, *91*(2), 276–280.

Davidson, D. J., & Indefrey, P. (2009). An event-related potential study on changes of violation and error responses during morphosyntactic learning. *Journal of Cognitive Neuroscience*, *21*(3), 433–446.

De Abreu, P. M. J. E. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory*, *19*(5), 529–537.

De Abreu, P. M. J. E., & Gathercole, S. E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, *104*(4), 974. https://doi.org/10.1037/a0028390

De Bot, K. (2000). Simultaneous interpreting as language production. In B. Englund Dimitrova & K. Hyltenstam (Eds.), *Language processing and simultaneous interpreting* (Vol. 40, pp. 65–88). John Benjamins.

De Bruin, A., Roelofs, A., Dijkstra, T., & FitzPatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: FMRI evidence from trilingual speakers. *NeuroImage*, *90*, 348–359.

De Groot, A. M. B. (2000). A complex-skill approach to translation and interpreting. In S. Tirkkonen-Condit & R. Jääskeläinen (Eds.), *Tapping and mapping the processes of translation and interpreting* (Vol. 37, pp. 53–68). John Benjamins. https://doi.org/10.1075/btl.37.06gro

De Groot, A. M. B., Borgwaldt, S., Bos, M., & Van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language*, *47*(1), 91–124.

De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology Section A*, *53*(2), 397–428.

De Groot, A. M. B., & Starreveld, P. A. (2015). Parallel language activation in bilinguals' word production and its modulating factors: A review and computer simulations. In J. W. Schwieter (Ed.), *Cambridge handbooks in language and linguistics* (pp. 389–415). Cambridge University Press.

DeKeyser, R. (1991). The semester overseas: What difference does it make? *ADFL Bulletin*, *22*(2), 42–48.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, *19*(2), 195–221.

DeKeyser, R. (2001). Automaticity and automatisation. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). Cambridge University Press.

DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

DeKeyser, R. (2010). Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals*, *43*(1), 80–92. https://doi.org/10.1111/j.1944-9720.2010.01061.x

Dewey, D. P. (2008). Japanese vocabulary acquisition by learners in three contexts. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *15*, 127–148.

Dewi, R. (2017). Lexical complexity in the introductions of undergraduate students' research articles. *Exposure Journal*, *6*(2), 161–172.

Di Biase, B., & Kawaguchi, S. (2002). Exploring the typological plausibility of Processability Theory: Language development in Italian second language and Japanese second language. *Second Language Research*, *18*(3), 274–302. https://doi.org/10.1191/0267658302sr204oa

Dijkstra, T., Grainger, J., & Van Heuven, W. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496–518.

Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language Cognition*, *5*(3), 175–197.

Dijkstra, T., & Van Heuven, W. (2012). Word recognition in the bilingual brain. In M. Faust (Ed.), *The handbook of the neuropsychology of language.* (pp. 451–471). Blackwell Publishing.

Dong, Y., & Cai, R. (2015). Working memory and interpreting: A commentary on theoretical models. In Z. Wen & M. B. Mota (Eds.), *Working memory in second language acquisition and processing* (pp. 63–81). Multilingual Matters.

Dong, Y., Liu, Y., & Cai, R. (2018). How does consecutive interpreting training influence working memory: A longitudinal study of potential links between the two. *Frontiers in Psychology*, *9*, 875–875. https://doi.org/10.3389/fpsyg.2018.00875

Dong, Y., & Xie, Z. (2014). Contributions of second language proficiency and interpreting experience to cognitive control differences among young adult bilinguals. *Journal of Cognitive Psychology*, *26*(5), 506–519.

Donovan, C. (2005). Teaching simultaneous interpretation into B: A challenge for responsible interpreter training. *Communication & Cognition. Monographies*, *38*(1–2), 147–166.

Donovan, C. (2003). Entrance exam testing for conference interpretation courses: How important is it? *International Journal of Interpretation and Translation*, *1*(2), 17–44.

Doughty, C. J. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, *13*(4), 431–469.

Dowens, M. G., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience*, *22*(8), 1870–1887.

Duckworth, R. A., Potticary, A. L., & Badyaev, A. V. (2018). On the origins of adaptive behavioral complexity: Developmental channeling of structural trade-offs. In D. M. Fragaszy (Ed.), *Advances in the study of behavior* (Vol. 50, pp. 1–36). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/S0065345417300189

Duez, D. (1982). Silent and non-silent pauses in three speech styles. *Language and Speech*, *25*(1), 11–28.

Duperron, L. (2006). Study abroad and the second language acquisition of tense and aspect in French: Is longer better. In S. Wilkinson (Ed.), *Insights from study abroad for language programs* (pp. 47–51). Heinle Boston.

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*(2), 143–188.

Ellis, N. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. P. Leow (Eds.), *Implicit and Explicit Language Learning: Conditions, Processes, and Knowledge in SLA and Bilingualism* (pp. 35–47). Georgetown University Press.

Ellis, N., & Sagarra, N. (2010). Learned Attention Effects in L2 Temporal Reference: The First Hour and the Next Eight Semesters. *Language Learning*, *60*(2), 85–108. https://doi.org/10.1111/j.1467-9922.2010.00602.x

Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science*, *33*(4), 547–582.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102–134). Cambridge University Press.

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron*, *88*(1), 33–46.

Fabbro, F., & Darò, V. (1995). Delayed auditory feedback in polyglot simultaneous interpreters. *Brain and Language*, *48*(3), 309–319.

Færch, C., Haastrup, K., & Philipson, R. (Eds.). (1984). *Learner language and language learning*. Multilingual Matters Limited.

Faretta-Stutenberg, M., & Morgan-Short, K. (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research*, *34*(1), 67–101. https://doi.org/10.1177/0267658316684903

Fedorenko, E., Piantadosi, S., & Gibson, E. (2012). Processing relative clauses in supportive contexts. *Cognitive Science*, *36*(3), 471–497.

Fernández, E. I. (2013). Unpacking delivery criteria in interpreting quality assessment. *Humanidades*, *10*, 169–182.

Fernández, E. M., & Cairns, H. S. (2011). *Fundamentals of psycholinguistics*. John Wiley & Sons.

Fernández-García, M., & Martínez-Arbelaiz, A. (2014). Native speaker–non-native speaker study abroad conversations: Do they provide feedback and opportunities for pushed output? *System*, *42*(1), 93–104. https://doi.org/10.1016/j.system.2013.10.020

Ferreira, A., & Schwieter, J. W. (2017). Translation and cognition: An overview. In A. Ferreira & J. W. Schwieter (Eds.), *The handbook of translation and cognition* (pp. 3–17). Wiley Online Library.

Ferreira, A., Schwieter, J. W., & Gile, D. (2015). The position of psycholinguistic and cognitive science in translation and interpreting. In A. Ferreira & J. W. Schwieter (Eds.), *Psycholinguistic and cognitive inquiries into translation and interpreting* (Vol. 115, pp. 3–15). John Benjamins.

Ferreira, V., & Slevc, L. R. (2007). Grammatical encoding. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 453–469). Oxford University Press.

Fitzpatrick, T. (2012). Tracking the changes: Vocabulary acquisition in the study abroad context. *The Language Learning Journal*, *40*(1), 81–98.

Fodor, J. (2008). The Modularity of mind: An essay on faculty psychology. In J. E. Adler & L. J. Rips (Eds.), *Reasoning—Studies of human inference and its foundations* (pp. 878–914). Cambridge University Press.

Freed, B. (1995). Language learning and study abroad. In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 3–34). John Benjamins.

Freed, B., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, *26*(2), 275–301. https://doi.org/10.1017/S0272263104262064

Freed, B., So, S., & Lazar, N. A. (2003). Language learning abroad: How do gains in written fluency compare with gains in oral fluency in French as a second language? *ADFL Bulletin*, *34*(3), 34–40. https://doi.org/10.1632/adfl.34.3.34

French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, *29*(3), 463–487. https://doi.org/10.1017/S0142716408080211

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology A*, *50*(1), 119–148.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Fromkin, V. (2012). *An introduction to language* (7th Australian and New Zealand ed.). Cengage Learning Australia.

Fu, R. (2012). 符荣波　著：英汉双向交替传译中译语停顿的对比研究 [Pausing in two-way E-C consecutive interpreting: A contrastive analysis of trainee interpreter's performance]. *外语教学与研究（外国语文双月刊）[Foreign Language Teaching and Research]*, *44(3)*, 437–447.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256.

García, A. M. (2019). *The Neurocognition of Translation and Interpreting* (Vol. 147). John Benjamins. https://doi.org/10.1075/btl.147

Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics*, *9*(2), 198–217.

Gass, S. (1997). *Input, interaction, and the second language learner*. Erlbaum.

Gass, S., & Glew, M. (2018). Second language acquisition and bilingualism. In J. Altarriba & R. R. Heredia (Eds.), *An introduction to bilingualism: Principles and processes*. Routledge.

Gass, S., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (Vol. 175199, pp. 180–206). Lawrence Erlbaum Associates.

Gathercole, S. E. (2007). Working memory and language. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 757–769). Oxford University Press.

Gathercole, S. E., & Baddeley, A. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*(3), 259.

Genesee, F. (1985). Second language learning through immersion: A review of US programs. *Review of Educational Research*, *55*(4), 541–561.

Gerver, D. (1969). The effects of source language presentation rate on the performance of simultaneous conference interpreters. In F. Pöchhacker & M. Shlesinger (Eds.), *The interpreting studies reader* (pp. 52–67). Routledge.

Gerver, D. (1976). Empirical studies of simultaneous interpretation: A review and a model. In R. W. Brislin (Ed.), *Translation: Applications and research* (pp. 165–203). New York Gardner Press.

Giezen, M. R., & Emmorey, K. (2016). Language co-activation and lexical selection in bimodal bilinguals: Evidence from picture–word interference. *Bilingualism*, *19*(2), 264.

Gile, D. (1995). *Basic concepts and models for interpreter and translator training*. John Benjamins.

Gile, D. (1998). Observational studies and experimental studies in the investigation of conference interpreting. *International Journal of Translation Studies*, *10*(1), 69–93.

Gile, D. (2005). Teaching conference interpreting. In M. Tennent (Ed.), *Training for the new millennium: Pedagogies for translation and interpreting* (pp. 127–151). John Benjamins.

Gile, D. (2009). *Basic concepts and models for interpreter and translator training* (Rev. ed.). John Benjamins.

Godfrey, L., Treacy, C., & Tarone, E. (2014). Change in French second language writing in study abroad and domestic contexts. *Foreign Language Annals*, *47*(1), 48–65. https://doi.org/10.1111/flan.12072

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech*, *15*(2), 103–113.

Gor, K., & Long, M. (2009). Input and second language processing. In W. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (2nd ed., pp. 445–472). Academic Press.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81.

Grey, S., Cox, J. G., Serafini, E. J., & Sanz, C. (2015). The Role of individual differences in the study abroad context: Cognitive capacity and language development during short-term intensive language exposure. *Modern Language Journal*, *99*(1), 137–157. https://doi.org/10.1111/modl.12190

Griffiths, R. (1992). Speech rate and listening comprehension: Further evidence of the relationship. *TESOL Quarterly*, *26*(2), 385–390. https://doi.org/10.2307/3587015

Grosjean, F. (1980). Temporal variables within and between languages. In H. W. Dechert & M. Raupach (Eds.), *Towards a cross-linguistic assessment of speech production* (pp. 39–53). Lang.

Grosjean, F. (1992). Another view of bilingualism. *Advances in Psychology*, *83*, 51–62.

Grosjean, F. (1997). The bilingual individual. *Interpreting*, *2*(1–2), 163–187.

Grosjean, F. (2013). Speech perception and comprehension. In F. Grosjean & P. Li (Eds.), *The Psycholinguistics of Bilingualism*. John Wiley & Sons. http://ebookcentral.proquest.com/lib/uwsau/detail.action?docID=1120620

Gumul, E., & Lyda, A. (2007). The time constraint in conference interpreting: Simultaneous vs. consecutive. *Research in Language*, *5*, 165–183.

Guntermann, G. (1995). The peace corps experience: Language learning in training and in the field. In B. Freed (Ed.), *Studies in bilingualism* (Vol. 9, pp. 149–170). John Benjamins. https://doi.org/10.1075/sibil.9.10gun

Hale, S., & Campbell, S. (2002). The interaction between text difficulty and translation accuracy. *Babel*, *48*(1), 14–33.

Han, C. (2015a). *Building the validity foundation for interpreter certification performance testing* [Doctoral dissertation]. Macquarie University.

Han, C. (2015b). (Para) linguistic correlates of perceived fluency in English-to-Chinese simultaneous interpretation. *International Journal of Comparative Literature and Translation Studies*, *3*(4), 32–37.

Han, C. (2018). Quantitative research methods in translation and interpreting studies. *The Interpreter and Translator Trainer*, *12*(2), 244–247.

Han, C., Chen, S., Fu, R., & Fan, Q. (2020). Modeling the relationship between utterance fluency and raters' perceived fluency of consecutive interpreting. *Interpreting*. https://doi.org/10.1075/intp.00040.han

Han, C., & Riazi, M. (2017). Investigating the effects of speech rate and accent on simultaneous interpretation: A mixed-methods approach. *Across Languages and Cultures*, *18*(2), 237–259. https://doi.org/10.1556/084.2017.18.2.4

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, *19*(3), 379–400.

Harley, B., & Jean, G. (1999). Vocabulary skills of French immersion students in their second language. *Zeitschrift Für Interkulturellen Fremdsprachenunterricht*, *4*(2), 1–23.

Harley, B., & Swain, M. (1984). The interlanguage of immersion students and its implication for L2 teaching. In A. Davies, C. Criper, & A. Howatt (Eds.), *Interlanguage* (pp. 291–311). Edinburgh University Press.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*(1), 25–38. https://doi.org/10.1017/S0272263100010457

Hartsuiker, R. J., & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*(1–3), 181–204.

Havik, E., Roberts, L., Van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, *59*(1), 73–112.

Hawkins, R., & Liszka, S. (2003). Locating the source of defective past tense marking in advanced L2 English speakers. *Language Acquisition Language Disorders*, *30*, 21–44.

Henrard, S., & Van Daele, A. (2017). Different bilingual experiences might modulate executive tasks advantages: Comparative analysis between monolinguals, translators, and interpreters. *Frontiers in Psychology*, *8*, 1870. https://doi.org/10.3389/fpsyg.2017.01870

Hernandez, A. E., & Meschyan, G. (2006). Executive function is necessary to enhance lexical processing in a less proficient L2: Evidence from fMRI during picture naming. *Bilingualism: Language and Cognition*, *9*(2), 177–188.

Hilton, H. E. (2011). What is implicit and what is explicit in L2 speech? Findings from an oral corpus. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 145–157). Georgetown University Press.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*(4), 901–931. https://doi.org/10.1016/j.lingua.2009.06.004

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, *17*, 5–27. https://doi.org/10.11431/secondlanguage.17.0_5

Howard, M. (2005). Second language acquisition in a study abroad context: A comparative investigation of the effects of study abroad and foreign language instruction on the L2 learner's grammatical development. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 495–530). De Gruyter.

Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm. *Memory & Cognition*, *35*(7), 1675–1684.

Hulstijn, J. H. (1997). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition*, *19*(2), 131–143.

Ife, A., Vives Boix, G., & Meara, P. (2000). The impact of study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics*, *4*(1), 55–84.

Injoque-Ricle, I., Barreyro, J. P., Formoso, J., & Jaichenco, V. I. (2015). *Expertise, working memory and articulatory suppression effect: Their relation with simultaneous interpreting performance. 11*(2), 56. https://doi.org/10.5709/acp-0171-1

Isabelli, C. A., & Nishida, C. (2005). Development of the Spanish subjunctive in a nine-month study-abroad Setting. In D. Eddington (Ed.), *In Selected Proceedings of* *the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 78–91). Cascadilla Press.

Isabelli-García, C. (2010). Acquisition of Spanish gender agreement in two learning contexts: Study abroad and at home. *Foreign Language Annals*, *43*(2), 289–303.

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, *24*(2), 168–196.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *Tesol Quarterly*, *34*(2), 239–278.

Jared, D., & Kroll, J. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, *44*(1), 2–31.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.

Jimenez-Jimenez, A. F. (2010). A comparative study on second language vocabulary development: Study abroad vs classroom settings. *Frontiers*, *19*, 105–124.

Jin, Y. (2010). *Is working memory working in consecutive interpreting?* [Doctoral dissertation]. University of Edinburgh.

Jin, Y. (2017). Consecutive interpreting. In C. Shei & Z. Gao (Eds.), *The Routledge handbook of Chinese translation* (pp. 321–335). Routledge.

Juan-Garau, M. (2014). Oral accuracy growth after formal instruction and study abroad: Onset level, contact factors and long-term effects. In C. Pérez–Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 87–109). John Benjamins.

Juan-Garau, M. (2018). Exploring oral L2 fluency development during a three-month stay abroad through a dialogic task. In C. S. A. Morales-Front (Ed.), *The Routledge handbook of study abroad research and practice*. Routledge.

Juffs, A. (1998). Main verb versus reduced relative clause ambiguity resolution in L2 sentence processing. *Language Learning*, *48*(1), 107–147.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*(2), 137–166. https://doi.org/10.1017/S0261444810000509

Just, Marcel A, & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *47*(2), 310.

Just, Marcel Adam, & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149.

Just, Marcel Adam, Keller, T. A., & Cynkar, J. (2008). A decrease in brain activation associated with driving when listening to someone speak. *Brain Research*, *1205*, 70–80.

Kállay, É., & Visu-Petra, L. (2014). Indicators of cognitive and emotional functioning in first-year master students in conference interpreting: A focus on individual differences. *Cognition, Brain, Behavior*, *18*(3), 209-228.

Kao, C.-W. (2019). Developing second language literacy: Taiwanese college students' error types in focused feedback effectiveness. In B. L. Reynolds & M. F. Teng (Eds.), *English literacy instruction for Chinese speakers* (pp. 245–264). Springer.

Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2019). The language experience and proficiency questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*, 1–6. https://doi.org/10.1017/S1366728919000038

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, *59*(3), 503–535. https://doi.org/10.1111/j.1467-9922.2009.00516.x

Kellogg, R. T. (2004). Working memory components in written sentence generation. *The American Journal of Psychology*, *117*(3), 341–361.

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *The Quarterly Journal of Experimental Psychology*, *68*(8), 1665–1692.

Kirchhoff, H. (1976). Simultaneous interpreting: Interdependence of variables in the interpreting process, interpreting models and interpreting strategies. In F. Pöchhacker & M. Shlesinger (Eds.), *The interpreting studies reader* (pp. 111–119). Routledge.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40.

Kopke, B., & Nespoulous, J.-L. (2006). Working memory performance in expert and novice interpreters. *Interpreting*, *8*(1), 1–23.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, *49*(2), 303–342.

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum Associates.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, *11*(2), 261–271. https://doi.org/10.1017/S1366728908003416

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Addison-Wesley Longman.

Kroll, J., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, *128*(3), 416–430.

Kroll, J., & Dijkstra, T. (2002). The bilingual lexicon. In R. B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 301–321). Oxford University Press.

Kroll, J., Michael, E. B., & Sankaranarayanan, A. (1998). A model of bilingual representation and its implications for second language acquisition. In A. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 365–395). Lawrence Erlbaum Associates.

Kroll, J., Michael, E. B., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, *18*(2), 137–171. https://doi.org/10.1191/0267658302sr201oa

Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149-174.

Kroll, J., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals—The development of lexical and conceptual representation. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 104–129). Blackwell.

Kroll, J., & Tokowicz, N. (2005). Models of bilingual representation and processing: Looking back and to the future. In A. M. B. De Groot & J. Kroll (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 531–553). Oxford University Press. http://ebookcentral.proquest.com/lib/uwsau/detail.action?docID=281414

Kurz, I. (1993). Conference interpretation: Expectations of different user groups. *The Interpreters' Newsletter*, *5*, 13–21.

Kurz, I. (2001). Conference interpreting: Quality in the ears of the user. *Meta: Translators' Journal*, *46*(2), 394–409.

Kurz, I. (2008). The impact of non-native English on students' interpreting performance. In G. Hansen, A. Chesterman, & H. Gerzymisch-Arbogast (Eds.), *Efforts and models in interpreting and translation research* (pp. 179–192). John Benjamins.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433.

Labrozzi, R. (2009). *Processing of lexical and morphological cues in a study abroad context* [Doctoral dissertation]. The Pennsylvania State University.

LaBrozzi, R. (2012). The role of study abroad and inhibitory control on processing redundant cues. In K. Geeslin & M. Dıaz–Campos (Eds.), *Selected proceedings of the 14th Hispanic Linguistics Symposium* (pp. 228–241). Cascadilla.

Lafford, B. (2006). The effects of study abroad vs. Classroom contexts on Spanish SLA: Old assumptions, new insights and future research directions. In C. A. Klee & T. L. Face (Eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 1–25). Cascadilla.

Lara, A. R. (2014). *Complexity, accuracy and fluency development through study abroad programmes varying in duration*. [Doctoral dissertation]. Universitat Pompeu Fabra.

Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, *25*(2), 21–33.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, *19*(2), 255–271.

Laver, J. D. (1973). The detection and correction of slips of the tongue. In V. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 132–143). Mouton.

Leeman, J. (2003). Recasts and second language development beyond negative evidence. *Studies in Second Language Acquisition*, *25*(1), 37–63. https://doi.org/10.1017/S0272263103000020

Leonard, K. R., & Shea, C. E. (2017). L2 speaking development during study abroad: Fluency, accuracy, complexity, and underlying cognitive factors. *Modern Language Journal*, *101*(1), 179–193. https://doi.org/10.1111/modl.12382

Levelt, W. (1989). *Speaking: From intention to articulation*. MIT Press.

Levelt, W. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*(23), 13464–13471.

Levelt, W., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75.

Li, C., & Thompson, S. A. (2008). Chinese. In B. Comrie (Ed.), *The world's major languages* (pp. 703–723). Routledge. http://ebookcentral.proquest.com/lib/uwsau/detail.action?docID=401871

Li, X. (2013). Are interpreting strategies teachable? Correlating trainees' strategy use with trainers' training in the consecutive interpreting classroom. *Eut Edizioni Università Di Trieste*, *18*(18), 105–128.

Liang, J., Fang, Y., Lv, Q., & Liu, H. (2017). Dependency distance differences across interpreting types: Implications for cognitive demand. *Frontiers in Psychology*, *8*, 1–10. https://doi.org/10.3389/fpsyg.2017.02132

Lin, Y., Lv, Q., & Liang, J. (2018). Predicting fluency with language proficiency, working memory, and directionality in simultaneous interpreting. *Frontiers in Psychology*, *9*, 1543–1543. https://doi.org/10.3389/fpsyg.2018.01543

Linck, J. A., Hoshino, N., & Kroll, J. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon*, *3*(3), 349–374.

Linck, J. A., Kroll, J., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, *20*(12), 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861–883.

Linck, J. A., & Weiss, D. J. (2011). Working memory predicts the acquisition of explicit L2 knowledge. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 101–114). Georgetown University Press.

Liter, A., Heffner, C. C., & Schmitt, C. (2017). The interpretation of plural morphology and (non-)obligatory number marking: An argument from artificial language learning. *Language Learning and Development*, *13*(4), 451–480. https://doi.org/10.1080/15475441.2017.1324307

Liu, H., Bates, E., & Li, P. (1992). Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics*, *13*(4), 451–484. https://doi.org/10.1017/S0142716400005762

Liu, M. (2001). *Expertise in simultaneous interpreting: A working memory analysis* [Doctoral dissertation]. University of Texas at Austin.

Liu, M. (2008). How do experts interpret? Implications from research in Interpreting Studies and cognitive science. In G. Hansen, A. Chesterman, & H. Gerzymisch-Arbogast (Eds.), *Efforts and models in interpreting and translation research: A tribute to Daniel Gile* (Vol. 80, pp. 159–177). John Benjamins.

Liu, M., & Chiu, Y.-H. (2011). Assessing source material difficulty for consecutive interpreting. In R. Setton (Ed.), *Interpreting Chinese, interpreting China* (pp. 135–156). John Benjamins.

Liu, M., Schallert, D., & Carroll, P. J. (2004). Working Memory and Expertise in Simultaneous Interpreting. *Interpreting*, *6*(1), 19–42. https://doi.org/10.1075/intp.6.1.04liu

Llanes, À., & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, *37*(3), 353–365. https://doi.org/10.1016/j.system.2009.03.001

Llanes, À., & Muñoz, C. (2013). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, *63*(1), 63–90. https://doi.org/10.1111/j.1467-9922.2012.00731.x

Lo, Y. Y., & Murphy, V. A. (2010). Vocabulary knowledge and growth in immersion and regular language-learning programmes in Hong Kong. *Language and Education*, *24*(3), 215–238.

Long, M. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences*, *379*(1), 259–278.

Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*(2), 126–141.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press. https://doi.org/10.1016/B978-012589042-7/50015-3

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based learning. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–167). Multilingual Matters.

Ma, Y. (2017). *Task complexity and grammatical development in ESL* [Doctoral dissertation, Western Sydney University]. Western Sydney University ResearchDirect.

Macías, M. P. (2006). Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. *Interpreting*, *8*(1), 25–43.

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, *21*(4), 557–587.

Mackey, A. (Ed.). (2007). *Conversational interaction in second language acquisition: A series of empirical studies*. Oxford University Press.

Mackintosh, J. (1985). The Kintsch and Van Dijk model of discourse comprehension and production applied to the interpretation process. *Meta: Journal Des Traducteurs/Meta: Translators' Journal*, *30*(1), 37–43.

Macnamara, B. N., & Conway, A. R. A. (2016). Working memory capacity as a predictor of simultaneous language interpreting performance. *Journal of Applied Research in Memory and Cognition*, *5*(4), 434–444. https://doi.org/10.1016/j.jarmac.2015.12.001

Magiste, E. (1986). Selected issues in second and third language learning. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neurolinguistic perspectives* (pp. 97–122). Lawrence Erlbaum Associates.

Maitreyee, R., & Goswami, S. P. (2009). Language proficiency questionnaire: An adaptation of LEAP-Q in Indian context. *Student Research at A.I.I.S.H. Mysore*, 214–232. https://www.aiish.ac.in/images/pdf/08-09slp.pdf#page=219

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940. https://doi.org/10.1044/1092-4388(2007/067)

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, *129*(3), 361.

Marqués-Pascual, L. (2011). Study abroad, previous language experience, and Spanish L2 development. *Foreign Language Annals*, *44*(3), 565–582. https://doi.org/10.1111/j.1944-9720.2011.01148.x

Martin, K., & Ellis, N. (2012). The roles of phonological short-term memory and working memory in l2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*(3), 379–413. https://doi.org/10.1017/S0272263112000125

Martin, R. C., & Slevc, L. R. (2014). Language production and working memory. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 437–450). Oxford University Press.

Massaro, D. W. (1975). *Understanding language*. Academic Press.

Matthei, E., & Roeper, T. (1985). *Understanding and producing speech*. Universe Pub.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401.

McDonald, J. L. (2008). Grammaticality judgments in children: The role of age, working memory and phonological ability. *Journal of Child Language*, *35*(2), 247–268. https://doi.org/10.1017/S0305000907008367

McLaughlin, B. (1987). *Theories of second language learning*. Edward Arnold.

Mead, P. (2002). Exploring hesitation in consecutive interpreting: An empirical study. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century* (pp. 73–82). John Benjamins.

Mead, P. (2005). Methodological issues in the study of interpreters' fluency. *The Interpreters' Newsletter*, 39–64.

Mellinger, C. D., & Hanson, T. A. (2019). Meta-analyses of simultaneous interpreting and working memory. *Interpreting*, *21*(2), 165–195. https://doi.org/10.1075/intp.00026.mel

Meuter, R. F. (2009). Language selection and performance optimisation in multilinguals. In L. Isurin, D. Winford, & K. De Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 27–51). John Benjamins.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*(1), 25–40.

Michael, E. B., Dijkstra, T., & Kroll, J. (2002). *Individual differences in the degree of language nonselectivity in fluent bilinguals*. Poster presented at the annual meeting of the Psychonomic Society, Kansas City.

Michael, E. B., & Gollan, T. H. (2005). Being and becoming bilingual. In J. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). Oxford University Press.

Michael, E. B., Tokowicz, N., & Kroll, J. (2003). *Modulating access to L2 words: The role of individual differences and language immersion experience*. Fourth International Symposium on Bilingualism, Tempe.

Miller, G. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. (Partial reprint Psychological Review 63 (1956) pp.81-97). *Psychological Review*, *101*(2), 343–352. https://doi.org/10.1037/0033-295X.101.2.343

Miller, G., Galanter, E., & Pribram, K. H. (1962). Plans and the structure of behavior. *The American Journal of Psychology*, *75*(1), 161. https://doi.org/10.2307/1419559

Milton, J., & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITL-International Journal of Applied Linguistics*, *107*(1), 17–34.

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, *67*(1), 224–237.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Lawrence Erlbaum Associates.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*(4), 621-640. https://doi.org/10.1037//0096-3445.130.4.621

Miyake, A., & Shah, P. (1999). Models of working memory: An introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.

Mizuno, A. (2005). Process model for simultaneous interpreting and working memory. *Meta: Journal Des Traducteurs/Meta: Translators' Journal*, *50*(2), 739–752.

Montero, L., Serrano, R., & Llanes, À. (2017). The influence of learning context and age on the use of L2 communication strategies. *The Language Learning Journal*, *45*(1), 117–132. https://doi.org/10.1080/09571736.2013.853824

Moon, J., & Jiang, N. (2012). Non-selective lexical access in different-script bilinguals. *Bilingualism: Language and Cognition*, *15*(1), 173–180.

Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, *46*(4), 610–641. https://doi.org/10.1002/tesq.34

Moreno, S., Bialystok, E., Wodniecka, Z., & Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *Journal of Neurolinguistics*, *23*(6), 564–579.

Morgan-Short, K. (2007). *A neurolinguistic investigation of late-learned second language knowledge: The effects of explicit and implicit conditions* [Doctoral dissertation]. Georgetown University.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, *60*(1), 154–193.

Moser-Mercer, B. (1978). Simultaneous interpretation: A hypothetical model and its practical application. In D. Gerver & H. W. Sinaiko (Eds.), *Language interpretation and communication* (pp. 353–368). Plenum Press.

Moser-Mercer, B., Frauenfelder, U., Casado, B., & Künzli, A. (2000). Searching to define expertise in interpreting. *Language Processing and Simultaneous Interpreting. Interdisciplinary Perspective*, 107–131.

Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, *29*(4), 578–596.

Muñoz, M., & Marquardt, T. (2003). Picture naming and identification in bilingual speakers of Spanish and English with and without aphasia. *Aphasiology*, *17*(12), 1115–1132. https://doi.org/10.1080/02687030344000427

Napier, J. (2015). Comparing signed and spoken language interpreting. In H. Mikkelson & R. Jourdenais (Eds.), *The Routledge handbook of interpreting* (pp. 129–143). Routledge.

Newman, S. D., Keller, T. A., & Just, M. A. (2007). Volitional control of attention and brain activation in dual task performance. *Human Brain Mapping*, *28*(2), 109–117.

Niska, H. (2005). Evaluation of training programmes. In M. Tennent (Ed.), *Training for the new millennium: Pedagogies for translation and interpreting* (pp. 35–64). John Benjamins.

Norris, D., Butterfield, S., Hall, J., & Page, M. (2018). Phonological recoding under articulatory suppression. *Memory & Cognition*, *46*(2), 173–180. https://doi.org/10.3758/s13421-017-0754-8

Obler, L. K. (2012). Conference interpreting as extreme language use. *International Journal of Bilingualism*, *16*(2), 177–182. https://doi.org/10.1177/1367006911403199

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, *27*(3), 377–402. https://doi.org/10.1017/S0142716406060322

Odlin, T. (2003). Crosslinguistic influence. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 436–486). Blackwell.

Odlin, T. (2012). Crosslinguistic influence in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (p. wbeal0292). Blackwell Publishing. https://doi.org/10.1002/9781405198431.wbeal0292

Olson, D. (2017). Bilingual language switching costs in auditory comprehension. *Language, Cognition and Neuroscience*, *32*(4), 494–513.

Orena, A. J., Theodore, R. M., & Polka, L. (2015). Language exposure facilitates talker learning prior to language comprehension, even in adults. *Cognition*, *143*, 36–40. https://doi.org/10.1016/j.cognition.2015.06.002

Padilla, F., Bajo, M., & Macizo, P. (2005). Articulatory suppression in language interpretation: Working memory capacity, dual tasking and word knowledge. *Bilingualism: Language and Cognition*, *8*(3), 207–219.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. John Benjamins.

Park, H. I., Solon, M., Henderson, C., & Dehghan–Chaleshtori, M. (2020). The roles of working memory and oral language abilities in elicited imitation performance. *The Modern Language Journal*, *104*(1), 133–151.

Parodi, T., Schwartz, B., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals. *Linguistics*, *42*(3), 669–706. https://doi.org/10.1515/ling.2004.022

Pavlacic, P. (2018). Raising the profile in China of Australia's excellence in the delivery of English language training. English Australia. https://internationaleducation.gov.au/International-network/Australia/InternationalStrategy/EGIProjects/Documents/Raising%20the%20Profile%20in%20China%20in%20English%20Language%20Training%20FINAL.pdf

Pavlović, N. (2007). Directionality in translation and interpreting practice: Report on a questionnaire survey in Croatia. *Revue Internationale d'interprétation et de Traduction/International Journal of Interpretation and Translation*, *5*(2), 79–99. https://doi.org/10.1075/forum.5.2.05pav

Peeters, D., Runnqvist, E., Bertrand, D., & Grainger, J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 284.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, *19*(3), 170–182. https://doi.org/10.1002/hbm.10110

Pérez-Vidal, C., & Juan-Garau, M. (2011). The effect of context and input conditions on oral and written development: A Study Abroad perspective1. *IRAL, International Review of Applied Linguistics in Language Teaching*, *49*(2), 157–185. https://doi.org/10.1515/iral.2011.008

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, *44*(3), 493–527.

Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition*, *6*(2), 186–214.

Pienemann, M. (1998). *Language processing and second language development: Processability theory* (Vol. 15). John Benjamins.

Pliatsikas, C. (2010). *Grammatical processing in second language learners of English*. University of Reading.

Poarch, G. J., & Van Hell, J. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *111*(3), 419–438.

Pöchhacker, F. (1993). On the Science of Interpretation. *The Interpreters' Newsletter*, 52–59.

Pöchhacker, F. (2004). *Introducing interpreting studies*. Routledge.

Pöchhacker, F. (2016). *Introducing interpreting studies* (2nd edition). Routledge.

Pöchhacker, F., & Zwischenberger, C. (2010). *Survey on quality and role: Conference interpreters' expectations and self-perceptions*. Electronic version:< http://aiic.net/page/3405/survey-on-quality-and-role-conference-interpreters-expectations-and-self-perceptions/lang/1.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production* (Vol. 20). John Benjamins.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, *15*(1), 36–57.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). *The myth of cognitive decline: Non-linear dynamics of lifelong learning. Topics in Cognitive Science, 6 (1), 5–42*.

Ransdell, S., Barbier, M.-L., & Niit, T. (2006). Metacognitions about language skill and working memory among monolingual and bilingual college students: When does multilingualism matter? *International Journal of Bilingual Education and Bilingualism*, *9*(6), 728–741.

Ratner, N. B., & Gleason, J. B. (2004). Psycholinguistics. In L. R. Squire (Ed.), *Encyclopedia of Neuroscience* (pp. 1199–1204). Academic Press. https://doi.org/10.1016/B978-008045046-9.01893-3

Rennert, S. (2010). The impact of fluency on the subjective assessment of interpreting quality. *The Interpreters' Newsletter*, *15*, 101–115.

Rimrodt, S., Clements-Stephens, A., Pugh, K., Courtney, S., Gaur, P., Pekar, J. J., & Cutting, L. (2009). Functional MRI of sentence comprehension in children with dyslexia: Beyond word recognition. *Cerebral Cortex*, *19*(2), 402–413.

Roberts, L. (2012). Individual differences in second language sentence processing. *Language Learning*, *62*, 172–188. https://doi.org/10.1111/j.1467-9922.2012.00711.x

Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 631–678). Blackwell.

Rossi, E., Kroll, J., & Dussias, P. E. (2014). Clitic pronouns reveal the time course of processing gender and number in a second language. *Neuropsychologia*, *62*, 11–25.

Rothman, J., & Iverson, M. (2007). Input type and parameter resetting: Is naturalistic input necessary? *IRAL-International Review of Applied Linguistics in Language Teaching*, *45*(4), 285–319. https://doi.org/10.1515/IRAL.2007.013

Russell, D. (2005). Consecutive and simultaneous interpreting. In T. Janzen (Ed.), *Topics in signed language interpreting: Theory and practice* (Vol. 63, pp. 135–164). John Benjamins.

Russell, D., & Takeda, K. (2015). Consecutive interpreting. In H. Mikkelson & R. Jourdenais (Eds.), *The Routledge handbook of interpreting* (pp. 96–111). Routledge.

Sagarra, N. (2017). Longitudinal effects of working memory on L2 grammar and reading abilities. *Second Language Research*, *33*(3), 341–363.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, *120*(8), 2022–2039.

Sagarra, N., & Labrozzi, R. (2018). Benefits of study abroad and working memory on L2 morphosyntactic processing. In C. Sanz & A. Morales-Front (Eds.), *The Routledge handbook of study abroad research practice*. Routledge.

Santilli, M., Vilas, M. G., Mikulan, E., Caro, M. M., Munoz, E., Sedeno, L., Ibanez, A., & Garcia, A. M. (2019). Bilingual memory, to the extreme: Lexical processing in simultaneous interpreters. *Bilingualism Language and Cognition*, *22*(2), 331–348. https://doi.org/10.1017/S1366728918000378

Sanz, C., & Leow, R. P. (2011). Introduction. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: conditions, processes, and knowledge in SLA and bilingualism* (pp. 1–6). Georgetown University Press.

Sanz, C., Lin, H.-J., Lado, B., Stafford, C. A., & Bowden, H. W. (2016). One size fits all? Learning conditions and working memory capacity in ab initio language development. *Applied Linguistics*, *37*(5), 669–692.

Schwartz, A. I., & Kroll, J. (2006). Bilingual lexical activation in sentence context. *Journal of Memory Language Cognitive Processes*, *55*(2), 197–212.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, *15*(2), 147–163.

Schwieter, J. W., & Ferreira, A. (2017). Bilingualism in cognitive translation and interpreting studies. In A. Ferreira & J. W. Schwieter (Eds.), *The handbook of translation and cognition* (pp. 144–164). Wiley Online Library.

Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (Vol. 27, p. 382). Blackwell.

Segalowitz, N., & Freed, B. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, *26*(2), 173–199.

Segalowitz, N., Freed, B., Collentine, J., Lafford, B., Lazar, N., & Díaz-Campos, M. (2004). A comparison of Spanish second language acquisition in two different learning contexts: Study abroad and the domestic classroom. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 1–18.

Seleskovitch, D. (1999). The teaching of conference interpretation in the course of the last 50 years. *Interpreting*, *4*(1), 55–66.

Serrano, R., Llanes, À., & Tragant, E. (2011). Analyzing the effect of context of second language learning: Domestic intensive and semi-intensive courses vs. study abroad in Europe. *System*, *39*(2), 133–143.

Signorelli, T., Haarmann, H. J., & Obler, L. K. (2011). Working memory in simultaneous interpreters: Effects of task and age. *International Journal of Bilingualism*, *16*(2), 198–212. https://doi.org/10.1177/1367006911403200

Signorelli, T., & Obler, L. K. (2012). Working memory in simultaneous interpreters. In J. Altarriba & L. Isurin (Eds.), *Memory, language, and bilingualism: Theoretical and applied approaches* (pp. 95–125). Cambridge University Press.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Song, Y. (2015). *Second language processing of derivational and inflectional morphology in English* [Doctoral dissertation]. Georgetown University.

Starreveld, P. A., De Groot, A. M. B., Rossmark, B. M., & Van Hell, J. G. (2014). Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism: Language and Cognition*, *17*(2), 258–276.

Stavrakaki, S., Megari, K., Kosmidis, M. H., Apostolidou, M., & Takou, E. (2012). Working memory and verbal fluency in simultaneous interpreters. *Journal of Clinical and Experimental Neuropsychology*, *34*(6), 624–633.

Sunderman, G., & Kroll, J. (2009). When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics*, *30*(1), 79–99. https://doi.org/10.1017/S0142716408090048

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in Second Language Acquisition*, *15*, 165–179.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honor of HG Widdowson* (pp. 125–144). Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–82). Cambridge University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Heinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Lawrence Erlbaum Associates.

Szego, J. (2020). *Australia's Chinese lesson—The nation's urgent need to engage with international students from China* (Press Release NARRATIVE #4). Scanlon Foundation Research Institute. https://scanloninstitute.org.au/publication/australias-chinese-lesson-nations-urgent-need-engage-international-students-china

Tannenbaum, P. H., Williams, F., & Wood, B. S. (1967). Hesitation phenomena and related encoding characteristics in speech and typewriting. *Language and Speech*, *10*(3), 203–215.

Tanner, D., Nicol, J., Herschensohn, J., & Osterhout, L. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. In A. K. Biller, E. Y. Chung, & A. E. Kimball (Eds.), *BUCLD 36: Proceedings of the 36th annual Boston university conference on language development* (pp. 594–606). Cascadilla Press.

Tedick, D. J. (1998). Immersion education: International perspectives by Robert Keith Johnson, Merrill Swain. *The Modern Language Journal*, *82*(4), 585–587. https://doi.org/10.2307/330232

Thierry, G., & Wu, Y. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535.

Thomas, M., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, *43*(1), 44–66.

Timarová, S. (2008). Working memory and simultaneous interpreting. In P. Boulogne (Ed.), *Translation and its others: Selected papers of the CETRA research seminar in translation studies 2007* (pp. 1–28). KULeuven.

Timarová, S., Čeňková, I., Meylaerts, R., Hertog, E., Szmalec, A., & Duyck, W. (2014). Simultaneous interpreting and working memory executive control. *Interpreting*, *16*(2), 139–168.

Timarová, S., Čeňková, I., Meylaerts, R., Hertog, E., Szmalec, A., & Duyck, W. (2015). Simultaneous interpreting and working memory capacity. In A. Ferreira & J. W. Schwieter (Eds.), *Psycholinguistic and cognitive inquiries into translation and interpreting* (pp. 101–126). John Benjamins.

Tissi, B. (2000). Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. *The Interpreters' Newsletter*, *10*(4), 103–127.

Tokowicz, N. (2014). *Lexical processing and second language acquisition*. Routledge.

Tokowicz, N., & Kroll, J. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes*, *22*(5), 727–779.

Tokowicz, N., Michael, E. B., & Kroll, J. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, *7*(3), 255–272. https://doi.org/10.1017/S1366728904001634

Tommola, J., & Hyona, J. (1990, April 15-21). Mental load in listening, speech shadowing and simultaneous interpreting: A pupillometric study [Conference session]. The World Congress of Applied Linguistics, Thessaloniki, Greece. https://files.eric.ed.gov/fulltext/ED327061.pdf

Toudic, D. (1997). *National report on the training of translators and interpreters in France* (Thematic Network Project in the Area of Languages Sub-Project 7: Translation and Interpreting, pp. 50–64). Université de Rennes. http://www.celelc.org/projects/Past_Projects/TNP_Languages/TNP1_resources/SP7FinalRecs.pdf

Trude, A. M., & Tokowicz, N. (2011). Negative transfer from Spanish and English to Portuguese pronunciation: The roles of inhibition and working memory. *Language Learning*, *61*(1), 259–280.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Tzou, Y.-Z. (2008). *The roles of working memory, language proficiency, and training in simultaneous interpretation performance: Evidence from Chinese-English bilinguals*. Texas A&M University.

Tzou, Y.-Z., Eslami, Z. R., Chen, H.-C., & Vaid, J. (2011). Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: Evidence from Mandarin–English speakers. *International Journal of Bilingualism*, *16*(2), 213–227. https://doi.org/10.1177/1367006911403197

Ullman, M. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, *30*(1), 37–69.

Van Hell, J., & De Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica*, *128*(3), 431–451.

Van Heuven, W., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458–483.

Van Heuven, W., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The*

Quarterly Journal of Experimental Psychology, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Greenwood Publishing Group.

Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology*, *141*, 65.

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*, *128*(3), 442.

Vigliocco, G., & Vinson, D. P. (2007). Semantic representation. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 195–215). Oxford university press.

Votaw, M. C. (1992). A functional view of bilingual lexicosemantic organization. *Advances in Psychology*, *83*, 299–321.

Wang, B., & Li, T. (2015). An empirical study of pauses in Chinese-English simultaneous interpreting. *Perspectives*, *23*(1), 124–142.

Wang, F. (2017). 王非 著：*口译过程研究—记忆机制与信息加工模型 [Interpreting process: Memory mechanism and information processing model]*. 科学出版社 [China Science Publishing and Media].

Wang, J. (2016). The relationship between working memory capacity and simultaneous interpreting performance: A mixed methods study on professional Auslan/English interpreters. *Interpreting*, *18*(1), 1–33.

Wen, H., & Dong, Y. (2019). How does interpreting experience enhance working memory and short-term memory: A meta-analysis. *Journal of Cognitive Psychology*, *31*(8), 769–784. https://doi-org.ezproxy.uws.edu.au/10.1080/20445911.2019.1674857

Wiedenhof, J. author. (2015). *A grammar of Mandarin*. John Benjamins.

Wilhelm, O., Hildebrandt, A. H., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, *4*, 433.

Williams, J., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, *53*(1), 67–121. https://doi.org/10.1111/1467-9922.00211

Williams, S. (1995). Research and bilingualism and its relevance for interpreting. *Hermes, Journal of Linguistics*, *15*, 143–154.

Wright, C. (2013). An investigation of working memory effects on oral grammatical accuracy and fluency in producing questions in English. *TESOL Quarterly*, *47*(2), 352–374. https://doi.org/10.1002/tesq.68

Wright, C. (2018). Effects of time and task on L2 Mandarin Chinese language development during study abroad. In C. Sanz & A. Morales-Front (Eds.), *The Routledge handbook of study abroad research practice*. Routledge.

Wu, Y., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition*, *129*(2), 418–425.

Xie, Z. (2018). The influence of second language (L2) proficiency on cognitive control among young adult unbalanced Chinese-English bilinguals. *Frontiers in Psychology*, *9*, 412. https://doi.org/10.3389/fpsyg.2018.00412

Xie, Z., & Dong, Y. (2017). Contributions of bilingualism and public speaking training to cognitive control differences among young adults. *Bilingualism: Language and Cognition*, *20*(1), 55–68.

Xu, H. (2010). Pauses in conference consecutive interpreting from English into Chinese: An empirical study. *Foreign Languages Research*, *1*, 66–73.

Yager, K. (1998). Learning Spanish in Mexico: The effect of informal contact and student attitudes on language gain. *Hispania*, 898–913.

Yu, W., & Van Heuven, V. (2017). Predicting judged fluency of consecutive interpreting from acoustic measures: Potential for automatic assessment and pedagogic implications. *Interpreting*, *19*(1), 47–68.

Zaytseva, V. (2016). *Vocabulary acquisition in study abroad and formal instruction: An investigation on oral and written lexical development* [Doctoral dissertation]. Universitat Pompeu Fabra.

Zaytseva, V., Pérez-Vidal, C., & Miralpeix, I. (2018). Vocabulary acquisition during study abroad. In C. Sanz & A. Morales-Front (Eds.), *The Routledge handbook of study abroad research practice*. Routledge.

Zhang, S. (2009). The role of input, interaction and output in the development of oral fluency. *English Language Teaching*, *2*(4), 91–100.

Zhang, T., Van Heuven, W., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science*, *22*(10), 1237–1242.

Zhang, W. (2009). Working memory and simultaneous interpreting performance: A report on empirical research among Chinese interpreters of English. *Foreign Language and Literature*, *25*(4), 128–134.

Zheng, B., & Xiang, X. (2017). The psycholinguistics of Chinese translation. In C. Shei & Z. Gao (Eds.), *The Routledge handbook of Chinese translation* (pp. 291–306). Routledge.

Appendix A Lexical Processing Stimuli

Stimuli	Misleading	Stimuli	Misleading
淀粉	stench	farce	脸
侵略	enter	taboo	纹身
契约	contact	empty	阿姨
智力	encouraging	maritime	婚期
大陆	content	discard	忽视
高原	highground	flaw	流动
武器	war	radical	种族的
珊瑚虫	carols	consult	袭击
寄生虫	parasol	liberate	蓄意的
掠夺者	prey	objective	主观的
臭氧层	stinksphere	enterprise	奖品
潮湿的	dump	meadow	寡妇
贬值	appreciate	decent	百分比
寒冷的	chilli	complaint	完成
肖像	painter	territory	恐怖分子
自治权	anatomy	celebrity	庆祝
作证	terrify	battery	黄油的
生态系统	biosystem	committee	诺言
领土的	terrorist	soybean	酱油
审美的	anaesthetic	peninsula	青霉素
评论家	commander	universal	制服
捐助者	donation	unique	独角兽
草坪	loan	preserve	精确的
海军的	native	touchy	火炬
透明的	translucent	sesame	相似的
葡萄糖	grape sugar	colonise	克隆
折扣	break	asset	评估
管弦乐队	concert	hilarious	歇斯底里的
新陈代谢	metabolist	impromptu	不及时的
气压计	pressuremeter	inject	引擎
肥料	fertiliser	transaction	交易
收入	revenue	splash	溅

请愿	petition	complete	完整的
过剩的	surplus	auspicious	吉利的
丛林	jungle	stink	臭味
病态的	morbid	intensive	集中的
潮汐	tide	provoke	煽动
破产	bankruptcy	estimate	估计
旋风	whirlwind	qualified	有资格的
峡谷	gorge	violate	违反
机密的	confidential	diligent	勤勉的
通风	ventilation	fossil	化石
微风	breeze	grant	授予
方言	dialect	permit	许可
刺客	assassin	bizarre	奇异的
誓言	oath	inflation	通货膨胀
没收	confiscate	taste	味道
湿润的	moist	atom	原子
学科	discipline	nationality	国籍
消过毒的	sterile	iceberg	冰山
兽皮	hide	comparable	可比较的
候选人	candidate	dehydrate	脱水
凡人	mortal	paraphrase	释义
加速度	acceleration	smash	粉碎
逃跑	flee	assure	保证
审查	audit	valid	有效的
绿洲	oasis	uprising	起义
审判	trial	frugal	节俭的
污染物质	pollutant	defend	辩护
茄子	eggplant	pollen	花粉

Appendix B Material for Reading Span Tasks

(Modified from Harrington & Sawyer, 1992)

Harrington, M., & Sawyer, M. (1992). L2 Working Memory Capacity and L2 Reading Skill. *Studies in Second Language Acquisition*, *14*(1), 25–38. https://doi.org/10.1017/S0272263100010457

No.	Stimuli	Make sense?
1.	He played baseball all day at the park and got a sore leg.	J
2.	The clerk in the department store put the presents in a dustbin.	F
3.	His younger brother played guitar in a rock and roll band.	J
4.	Suddenly the taxi opened its door in front of the bank.	J
5.	The last thing he did was to take a nice hot bath.	J
6.	Her best memory of America was the Tower of London bell.	F
7.	At the very top of the tall tree sat a small bird.	J
8.	She took a deep breath and reached into the darkness.	J
9.	China is famous for its butter and cheese.	F
10.	He overslept and missed all of the morning economics class.	J
11.	Popular foods in the winter are watermelon and ice cream.	F
12.	There was nothing left to do except leave and lock the door.	J
13.	In order to attend the dinner she needed to buy a dress.	J
14.	The woman screamed and slapped the rude man in the face.	J
15.	She leaned over the candle and her hair caught on fire.	J
16.	The drinks were all gone and all that remained was the beer.	F
17.	The hunting knife was so sharp that it cut his right hand.	J
18.	She soon realized that the man forgot to leave the room key.	J
19.	The saw that he brought was not strong enough for the lock.	J
20.	The first driver out in the morning always picks up the mail.	J
21.	All that remained in the lunch box was one salted nut.	J
22.	The boat engine would not run because it was full of oil.	F
23.	The letter said to come to the market to claim the prize.	J
24.	It was a very simple meal of salted fish and boiled rice.	J

25.	They decided to take an afternoon break by the large rock.	J
26.	He wanted to leave his bags and jacket in the hotel room.	J
27.	There were so many people that I could easily find a seat.	F
28.	He opened the bottom drawer and pulled out a shirt.	J
29.	The skiing was so wonderful that he got bored and started to read a book.	F
30.	It is impolite to eat spaghetti with hands.	J
31.	The season that people often associate with love is spring.	J
32.	The letter was lost because it did not have a postage stamp.	J
33.	People in southern China always like to eat noodles.	F
34.	At night the prisoners escaped through a hole in the wall.	J
35.	She made full use of her time and became a top student in her class.	J
36.	There are two birds barking outside the window.	F
37.	Drinking a lot of water could prevent us from getting sunstroke.	J
38.	There are 48 hours per day.	F
39.	She has no passport so she cannot go overseas.	J
40.	A three-hour sleep at night is far from enough for mankind.	J
41.	Be careful of the vast on the table.	J
42.	We should be treated with respect and kindness.	J

Appendix C Consecutive Interpreting Materials

Chinese-English

虽然中国已经在新能源汽车这一领域实现了快速发展但是接下来质量取代数量应该是行业可持续性发展的重中之重。尤其是在汽车的安全以及电池质量这两方面,都是需要经过严密的监督和检查的。

由于政府出台了购车补贴税费减免等扶持政策,近两年来我国新能源汽车产业呈现爆炸式增长。中国汽车工业协会数据显示,今年 1-10月全国累计销售电动汽车已超过 17 万辆,同比增长 290%。

随着世界各国努力摆脱对化石燃料的依赖并大量利用风能和太阳能技术,全世界今年使用的可再生能源超过了煤炭的使用量。国际能源总署执行主任在一份声明中说道:“我们正在目睹世界能源市场向可再生能源的转变,在其他领域也是如此,现在可再生能源的增长的中心正在移向新兴市场国家。”

这些数据只反映出了现在能够生产的能源总量,而不是已经生产出的能源总量。国际能源总署表示,在近五年内,可再生能源已是增长最快的发电来源,未来将会“迅速赶上和煤炭之间的差距。”近年来,可再生能源的成本大幅下降,国家对这种新技术的投资也直线上升。

English-Chinese

Over my lifetime, I have seen very significant social changes. Probably one of the most significant, is the rise of artificial intelligence. In short, I believe that the rise of powerful AI, will be either the best thing, or the worst, ever to happen to humanity. We do not yet know which. But we should do all we can, to ensure that its future development benefits us, as well as our environment.

The progress in AI research and development is swift, and perhaps we should all stop for a moment, and focus our research, not only on making AI more capable, but on maximizing its societal benefit. Everything that civilization has to offer, is a product of human intelligence, and I believe there is no real difference between what can be achieved by a biological brain, and what can be achieved by a computer.

It therefore follows that computers can, in theory, emulate human intelligence, and exceed it. But we don't know. So we cannot know if we will be infinitely helped by AI, or ignored by it and side-lined, or probably destroyed by it. Indeed, we have concerns that clever machines will be capable of undertaking work currently done by humans, and swiftly destroy millions of jobs.

While primitive forms of artificial intelligence developed so far, have proved very useful, I fear the consequences of creating something that can match or surpass humans. In fact, humans are limited by slow biological evolution, couldn't compete, and would be superseded, then AI would bring great disruption to our economy. some scientists, however, believe that humans can command the rate of technology for a decently long time, and that the potential of AI to solve many of the world's problems will be realised.

Appendix D LexTALE

LexTALE stands for *Lexical Test for Advanced Learners of English*. It is intended for cognitive researchers studying participants with an advanced level of English as a second language in an experimental setting.

What is it?

The LexTALE is a quick and practically feasible test of vocabulary knowledge for medium to highly proficient speakers of English as a second language. It consists of a simple unspeeded visual lexical decision task. In contrast to other vocabulary or proficiency tests, it has been designed to meet the needs of cognitive researchers. It is quick, easy to administer, and free, and yet it is a valid and standardized test of vocabulary knowledge. It has also been shown to give a fair indication of general English proficiency.

Quick

On average, the LexTALE takes about 3.5 minutes to complete. It comprises only 60 trials, making it a practically feasible addition to any psycholinguistic experiment.

Easy

The LexTALE can either be administered online, or implemented in any experimental software (download the item list and instructions for implementation). The LexTALE can also be downloaded in Praat, Presentation, and Matlab format.

Valid

In a large-scale study (Lemhöfer & Broersma, 2012) on Dutch and Korean advanced learners of English, the LexTALE was evaluated 1) as a measure of English vocabulary knowledge, 2) as an indicator of general English proficiency, and 3) as an indicator of performance on two psycholinguistic experimental paradigms. LexTALE scores were found to be good predictors of vocabulary knowledge (as measured by L1-L2 and L2-L1 translation), to give a fair indication of general English proficiency (as measured by two thorough and extensive proficiency tests, the TOEIC and the Quick Placement Test), and to correlate well with experimental word recognition data (from lexical decision and progressive demasking experiments). See Validity for more information about LexTALE in relation to other tests.

Better than self-ratings

As many bilingual studies use in-house questionnaires including language history questions or proficiency ratings, the predictive power of the LexTALE was compared to that of self-ratings (Lemhöfer & Broersma, 2012). Self-ratings were assessed separately for writing, reading, listening and speaking proficiency. They turned out to be significant predictors of some of the translation accuracy and general proficiency variables, but not as consistently as the LexTALE. Moreover, self-ratings were not significantly related to the experimental word recognition data at all, whereas the LexTALE was.

Free

We would like to invite everybody to use the LexTALE for their research purposes. We would appreciate it if you let us know, just for our information. In your publications, please refer to www.lextale.com and to Lemhöfer and Broersma (2012) for updates on this reference, please keep an eye on this website.

Dutch and German versions

Apart from the standardized and validated English version of the LexTALE, there are also a German and a Dutch version of LexTALE available. Although they are not yet validated or tested for their equivalence with the English version, they were developed in parallel to the English version. Both can be done online, and downloaded as item list and in Praat Presentation, and Matlab format.

Appendix E Language Experience and Proficiency Questionnaire (LEAP-Q)

(Modified from Marian et al., 2007)Automatic citation updates are disabled. To see the bibliography, click Refresh in the Zotero tab. Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940. https://doi.org/10.1044/1092-4388(2007/067)

Language Experience and Proficiency Questionnaire (LEAP-Q)

Last Name	fdd	First Name		Today's Date	
Age		Date of Birth		Male ☐	Female ☐

(1) Please list all the languages you know **in order of dominance**:

1	**2**	**3**	**4**	**5**

(2) Please list all the languages you know **in order of acquisition** (your native language first):

1	**2**	**3**	**4**	**5**

(3) Please list what percentage of the time you are *currently* and *on average* exposed to each language.
(*Your percentages should add up to 100%*):

List language here:					
List percentage here:					

(4) When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original was written in another language, which is unknown to you.
(*Your percentages should add up to 100%*):

List language here					
List percentage here:					

(5) When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each language? Please report percent of total time.
(*Your percentages should add up to 100%*):

List language here					
List percentage here:					

(6) Please name the cultures with which you identify. On a scale from zero to ten, please rate the extent to which you identify with each culture. (Examples of possible cultures include US-American, Chinese, Jewish-Orthodox, etc):

List cultures here					
	(click here for scale)	(click here for scale)	(click here for scale)	(click here for scale)	(click here for scale)

(7) How many years of formal education do you have? ______ ______________________________
Please check your highest education level (or the approximate US equivalent to a degree obtained in another country):

☐ Less than High School ☐ Some College ☐ Masters
☐ High School ☐ College ☐ Ph.D./M.D./J.D.
☐ Professional Training ☐ Some Graduate School ☐ Other:

(8) Date of immigration to the USA, if applicable ___ ______________________________
If you have ever immigrated to another country, please provide name of country and date of immigration here. ________________ ______________________________

(9) Have you ever had a vision problem ☐, hearing impairment ☐, language disability ☐, or learning disability ☐ ? (Check all applicable). If yes, please explain (including any corrections):
______________________________ ______________________________

English

This is my (please select from pull-down menu) language.

All questions below refer to your knowledge of .

(1) Age when you…:

began acquiring :	*became fluent* in :	*began reading* in :	*became fluent reading* in :

(2) Please list the number of years and months you spent in each language environment:

	Years	Months
A country where is spoken		
A family where is spoken		
A school and/or working environment where is spoken		

(3) On a scale from zero to ten, please select your *level of **proficiency*** in speaking, understanding, and reading from the scroll-down menus:

Speaking	(click here for scale)	Understanding spoken language	(click here for scale)	Reading	(click here for scale)

(4) On a scale from zero to ten, please select how much the following factors contributed to you learning :

Interacting with NSs	(click here for pull-down scale)	Language tapes/self instruction	(click here for pull-down scale)
Interacting with family	(click here for pull-down scale)	Watching TV	(click here for pull-down scale)
Reading	(click here for pull-down scale)	Listening to the radio	(click here for pull-down scale)

(5) Please rate to what extent you are currently exposed to in the following contexts:

Interacting with NSs	(click here for pull-down scale)	Listening to radio/music	(click here for pull-down scale)
Interacting with family	(click here for pull-down scale)	Reading	(click here for pull-down scale)
Watching TV	(click here for pull-down scale)	Language-lab/self-instruction	(click here for pull-down scale)

(6) In your perception, how much of a foreign accent do you have in ?

(click here for pull-down scale)

(7) Please rate how frequently others identify you as a non-native speaker based on your accent in :

(click here for pull-down scale)

Appendix F Participants' Consent Form

Consent Form – General (Specific)

Project Title: The Effect of Study Abroad Experience and Working Memory on Chinese-English Consecutive Interpreting Performance

I hereby consent to participate in the above named research project.

I acknowledge that:

- I have read the participant information sheet (or where appropriate, have had it read to me) and have been given the opportunity to discuss the information and my involvement in the project with the researcher/s

- The procedures required for the project and the time involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.

I consent to:

[Insert tick box option for each specific activity e.g.

☐ *Participating in an interview*

☐ *Having the interview audio recorded*

☐

I consent for my data and information provided to be used for this project.

I understand that my involvement is confidential and that the information gained during the study may be published but no information about me will be used in any way that reveals my identity.

I understand that I can withdraw from the study at any time without affecting my relationship with the researcher/s, and any organisations involved, now or in the future.

Signed:

Name:

Date:

This study has been approved by the Human Research Ethics Committee at Western Sydney University. The ethics reference number is: H12405

What if I have a complaint?

If you have any complaints or reservations about the ethical conduct of this research, you may contact the Ethics Committee through Research Engagement, Development and Innovation (REDI) on Tel +61 2 4736 0229 or email humanethics@westernsydney.edu.au.

Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

Appendix G Participant Information Sheet

Participant Information Sheet – General (Specific)

Project Title: The Effect of Intensified Language Exposure on Consecutive Interpreting Performance

Project Summary: This project aims to explore the relationship between intensive second language exposure and interpreting performance.

You are invited to participate in a research study being conducted by Ruiyuan Wang, School of Humanities and Communication Arts under the Supervision of Jing Han, Bruno Di Biase, Mark Antoniou, School of Humanities and Communication Arts.

The research is to explore the impact of intensive second language exposure on interpreting performance.

How is the study being paid for?

You will be given a voucher in the value of $25 as an appreciation for your time.

What will I be asked to do?

Firstly, you will be asked to finish a working memory task, and then a word translation recognition task. The last task will be a consecutive interpreting performance task.

Participation will take no more than 50 minutes in total.

What benefits will I, and/or the broader community, receive for participating?

You will be given a voucher in the value of $25 as an appreciation for your time. You will also get a good experience in how the research conducts.

By exploring the correlation between intensive language exposures and interpreting performance, this study may contribute to interpreting trainers when they design curricula and lectures for interpreting training programs.

Will the study involve any risk or discomfort for me? If so, what will be done to rectify it?

This study involves no foreseeable discomfort, and if you become distressed you may withdraw at any stage or avoid answering questions that appear intrusive without any penalty. Your decision to withdraw will not prejudice your future relationship with School of Humanities and Communication Arts, Western Sydney University.

How do you intend to publish or disseminate the results?

It is anticipated that the results of this research project will be published and/or presented in a variety of forums. In any publication and/or presentation, information will be provided in such a way that the participant cannot be identified, except with your permission.

No information obtained in connection with this study will be used to identify you. When the study results are reported, participants' identities will be protected, and identifying information such as your name, will not be revealed. All data obtained from the study will be stored in password-protected computers accessible only by research staff.

Will the data and information that I have provided be disposed of?

Please be assured that only the researchers will have access to the raw data you provide and that your data will not be used in any other projects. Please note that minimum retention period for data collection is five years post publication. The data and information you have provided will be securely disposed of.

Can I withdraw from the study?

Participation is entirely voluntary and you are not obliged to be involved. If you do participate you can withdraw at any time without giving reason.

If you do choose to withdraw, any information that you have supplied will be directly deleted and will not be used in this study.

Can I tell other people about the study?

Yes, you can tell other people about the study by email or orally. *For example:*

Providing them with the Chief Investigator's contact details. They can contact the Chief Investigator to discuss their participation in the research project and obtain a copy of the information sheet.

What if I require further information?

Please contact Ruiyuan Wang should you wish to discuss the research further before deciding whether or not to participate.

Ruiyuan Wang, Principal Researcher, Ph. D. student

Mobile number: +61 4 0689 4914

Email: 17945006@student.westernsydney.edu.au

What if I have a complaint?

If you have any complaints or reservations about the ethical conduct of this research, you may contact the Ethics Committee through Research Engagement, Development and Innovation (REDI) on Tel +61 2 4736 0229 or email humanethics@westernsydney.edu.au.

Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

If you agree to participate in this study, you may be asked to sign the Participant Consent Form. The information sheet is for you to keep and the consent form is retained by the researcher/s.

This study has been approved by the Western Sydney University Human Research Ethics Committee. The Approval number is *[enter approval number once the project has been approved]*.