AN INVESTIGATION OF LEXICAL BUNDLES IN L1 SPEECHES AND L2 INTERPRETED LANGUAGES: A CORPUS-BASED STUDY

YI LIU

PhD

The Hong Kong Polytechnic University

2023

The Hong Kong Polytechnic University

Department of Chinese and Bilingual Studies

AN INVESTIGATION OF LEXICAL BUNDLES IN L1 SPEECHES AND L2 INTERPRETED LANGUAGES: A CORPUS-BASED STUDY

YI LIU

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

December, 2022

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

LIU YI (Name of student)

Abstract

In recent years, there has been increasing interest in formulaic language across a variety of linguistic disciplines from the fields of discourse analysis, language acquisition, language pathology, language pedagogy, and applied linguistics (Lu et al., 2018; Wray, 2002). Formulaic language is a long-recognized phenomenon and previous studies have identified its prevalent use in first language discourse (Erman & Warren, 2000; Nattinger & DeCarrico, 1992; Pawley & Syder, 1983). For example, Altenberg (1998) claimed that over 80 percent of words in the London Lund corpus are in formulaic sequences. Erman and Warren (2000) considered that over 50 percent of language could be formulaic. Formulaicity in the current study is operationalized via the concept of lexical bundles (LBs), first put forward by Biber et al. (1999b) who defined LBs as "bundles of words that show a statistical tendency to co-occur (1999b, p. 989)" and as "simply sequences of word forms that commonly go together in natural discourse" (1999b, p. 990). The present study attempts to extend earlier research on formulaic sequences in political speech by exploring use of lexical bundles used by L2 interpreters compared with L1 speakers and source text (ST) speeches.

Specifically, three issues are investigated. 1) The general distribution patterns of the use of LBs in L2 interpreted texts and L1 speeches in terms of their frequency. 2) The general distribution patterns of the use of LBs between L2 interpreting and L1 original speech in terms of their syntactic structures and discoursal functions. 3) The impact of source texts on L2 simultaneous interpreting (SI) regarding the use of LBs in interpreted texts.

A comparable and parallel interpreting corpus based on the United Nations Security

Council Meetings (UNSCCP corpus) is constructed. The comparable corpus includes two English components of L2 interpreted texts (abbreviated as L2I) and L1 original speech (abbreviated as L1O). The parallel corpus features one English component of target texts, which is half size of L2I, abbreviated as L2I (partial) and one Chinese component of source texts (abbreviated as STC).

Four-word lexical bundles are first retrieved and identified in L2I and L1O. These two groups of LBs are compared in terms of their frequency, syntactic structure, and discourse functions. The second section compares the relationship (including general patterns of equivalence, addition, and shifts) of LBs between ST and target texts (TT) in the parallel corpus. This step aims to investigate the link between source texts and target texts, which acts as a triangulation for studying ST interference, or the amount to which interpreters' use of lexical bundles is influenced by source texts.

Results from the comparable corpus indicate three points: 1) In terms of the general frequency L2 interpreters use LBs significantly more frequently than L1 English speakers, suggesting that L2 interpreters in political contexts depend more heavily on the idiom principle. Moreover, the distribution of functional bundles in L2I and L1O reveals divergent patterns, indicating that these two groups of speakers may apply distinct mechanisms in producing their speech. 2) In terms of the syntactic structure taxonomy, L2I contained most noun phrase (NP) / prepositional phrases (PP) and L1O comprised most verb phrases (VP), suggesting that L2I seems to feature more written language reflecting the formal speech style of political discourse in interpreting, while L1O is closer to spoken language. 3) In terms of discoursal function taxonomy, L2 interpreters used a greater number of different four-word bundle types across the three main functional categories than L1 speakers, according to the type counts of these

bundles. In each subcategory the results suggest that with the exception of ability bundles in the stance bundle group and introduction/transitional bundles in the discourse-organising bundle group, the type counts of L2I were mostly higher than those of L1O.

Regarding frequency counts, both L2I and L1O showed a similar pattern in the distribution of functional bundles. The majority of stance bundles used by L2 interpreters and L1 speakers, respectively, are used to express desire, obligation, intention, and epistemic stance. Therefore, it makes sense to surmise that lexical bundles of high frequency are to some extent pre-fabricated. Even though lexical bundles are not usually idiomatic, the fact that they are always useful suggests that they may be retained in memory unaltered and used for textual or interpersonal discourse functions.

The second part of the analysis examines the three types of translation relationships in which LBs are used in accordance with their distribution of discourse functions. The results showed that equivalence pattern accounts for 63% of all occurrences, followed by addition (29%) and shift (7%). Within the three major functional groups, stance bundles and referential bundles are mostly used as equivalent cases, while most discourse-organising bundles are used as addition in comparison to ST. The additional cases are mostly grammatical supplements when rendering Chinese ST into English TT. It is reasonable to infer that most bundles used by L2 interpreters correspond to the ST expressions. The addition and shift of LBs are used by interpreters to cope with the grammatical distinctions between Chinese and English.

The current research is the first attempt to carry out the exploration of the use of LBs

between L2 interpreting and L1 speakers as well as source texts and target texts. The emphasis on L2 interpreted texts contributes to the understanding of this underexplored field as most studies in corpus-based interpreting studies have focused on interpreting into L1. The current research adds to research on interpretese by presenting supportive evidence of normalisation, contributing to the knowledge on corpus-based interpreting studies. It also enhances knowledge of SI output by illustrating how L2 interpreters use LBs to form their speech.

Acknowledgements

I am deeply grateful to the following individuals, whose support and encouragement made it possible for me to undertake this research project and complete this journey.

First and foremost, I owe a debt of gratitude to my PhD supervisors, Prof. Li Dechao, for his wise counsel, unwavering encouragement, and tolerance. His vast knowledge and wealth of experience have been a constant source of inspiration, both in my academic research and daily lift.

I would like to express my deepest gratitude to my external examiners, Prof. Wang Binhua and Prof. Ren Wen, for their insightful review and valuable comments on my thesis. Their feedback has greatly contributed to the quality and rigor of my research.

I am also indebted to Dr. Cheung Kay Fan, Andrew, Dr. Liu Kanglong, and Dr. Wu Zhiwei for the academic support and assistance they provided over the course of my project.

My thanks also extend to Prof. Gert De Sutter of Ghent University, with whom I had the privilege of working during my PhD scholar visit. I would like to thank him for his insightful remarks and suggestions on my PhD work, as well as his wonderful assistance during my stay.

My gratitude goes as well to Prof. Daniel Gile, Prof. Haidee Kotze (Kruger), Prof. Kaisa Koskinen for their invaluable suggestions with regard to my research project at the 2022 CETRA Summer School. I also would like to thank Prof. Bart Defrancq, Prof. Sandra

Halverson, Prof. Marta Kajzer-Wietrzny, Prof. Silvia Bernardini, Prof. Lefer Marie-Aude, Prof. Granger Sylviane for their insightful ideas and invaluable suggestions during my presentation at the EST Congress in Oslo and CECL team at UCLouvain.

I would like to thank my colleagues and friends, Selena, Tina, Kiki, Yanjin, Jiaxin, Yanfang, and the great marvel team at AG518 for their emotional support and academic advice. I would also like to thank my friends met in the badminton GJD team 409, and those I met during my research stay in Ghent. I am grateful for their friendships, companionship, and the memories we shared.

To my families, I owe a debt of gratitude for their unconditional love and support. Their unwavering belief in me has been a constant source of strength and motivation.

Thank you all for helping me achieve this significant milestone in my academic career.

Table of Contents

List of Abbreviations

LBs: lexical bundles

SI: simultaneous interpreting (into B)

ST: source texts

TT: target texts

TU: translation universals

PP: prepositional phrase

NP: noun phrases

VP: verb phrases

List of Tables

List of Figures

List of Appendices

Chapter 1 Introduction

1.1 Research motivations

Formulaic language, prevalent in both written and spoken discourse (Altenberg, 1998; Erman & Warren, 2000) has attracted increasing attention in a wide range of disciplines (Wood, 2015; Wray, 2002). Formulaic sequences provide process advantages (Guz, 2014; Hyland, 2012; Schmitt, 2004; Simpson-Vlach & Ellis, 2010; Van Lancker-Sidtis & Rallon, 2004) through language production by allowing speakers to produce clusters of words stored holistically in their memory instead of assembling speech word by word (Arnon & Snider, 2010; Conklin & Schmitt, 2008; Grandage & Adolphs; Jiang & Nekrasova, 2007; Schmitt, 2004; Schmitt & Galpin, 2004; Wood, 2015; Wray, 2002; Wray & Perkins, 2000).

In recent years, studies on formulaic language have adopted a data-driven and frequency-based methodology to determine a certain kind of formulaic sequence. Formulaic sequences, therefore, have been further identified into lexical phrases (Nattinger & DeCarrico, 1992), formulas (Simpson-Vlach & Ellis, 2010), clusters (Carter & McCarthy, 2006), multi-word units or multi-word expressions (Adolphs, 2006; Moon, 1997), n-grams (Milton & Freeman, 1996), recurrent word combination (Altenberg, 1993), and lexical bundles (Biber et al., 1999a).

Specifically, formulaicity in the current study is operationalized via the concept of lexical bundles (LBs), defined as "bundles of words that show a statistical tendency to co-occur (Biber et al., 1999b, p. 989)" and as "simply sequences of word forms that commonly go together in natural discourse" (1999b, p. 990). It is commonly claimed

that LBs are the most frequent multi-word sequences in a corpus, and that they possess certain functions in discourse (Wood, 2015).

The use of lexical bundles is considered to be varied in different text types and registers (Breeze, 2013; Conrad & Biber, 2005; Hyland, 2008b; Shirazizadeh & Amirfazlian, 2021; Wang, 2017) and disciplines (Cortes, 2004; Hyland, 2008b). For example, Biber and Barbieri (2007) investigated the use of lexical bundles in both spoken and written discourse in a range of university registers. Their work suggested that lexical bundles are more frequently used in non-academic university registers than in instructional registers, and are more common in writing than in speech texts. Moreover, the use of lexical bundles is also related to language development (Chen & Baker, 2010; Cortes, 2004; Salazar, 2014). Many studies have concluded that L1 speakers tend to use more lexical bundles both in number and variation as compared to L2 speakers (Ädel & Erman, 2012; Chen & Baker, 2010; Pan et al., 2016).

Although there has been relatively little research interest in LBs in translation and interpreting studies, the investigation of lexical issues at multi-word level has attracted researchers' interest since the emergence of corpus-based translation studies. Many of these studies discuss lexical bundles or similar multi-word sequences (i.e., collocations, set phrases, idioms). In translation studies, the findings of several corpus studies (Lee, 2013; Xiao, 2010) support the hypothesis of normalisation in language pairs of English-Chinese and English-Korean in written translation. Kajzer-Wietrzny (2012) showed that interpreting subcorpora involving different language pairs tested confirmation of the normalisation hypothesis while the results generated from translation subcorpora are mixed among different language pairs. In addition, several scholars have identified LBs as a preferred technique in promoting fluency, improving readability, and

performing meta-discursive functions when assessing translators' stylistic preferences (Granger, 2014; Liu & Afzaal, 2021b; Shrefler, 2011).

In interpreting studies, the facilitation of "chunks" (including LBs) in simultaneous or consecutive interpreting has long been emphasized by interpreting practitioners, researchers, and trainers (Aston, 2018; Ferraresi & Miličević, 2017; Henriksen, 2007; Plevoets & Defrancq, 2018a). For instance, Henriksen (2007, p. 8) argued that "a great store of formulas diminishes the interpreter's production effort, insofar as these can be retrieved as single lexical items from memory." She also claimed that "the formula is the result of automatic language production" (2007, p. 13). Aston (2018) came up with a similar finding in that the use of recurrent formulaic phraseologies boosts the fluent output of interpreters as holistically stored formulaic phraseologies help interpreters reduce cognitive load when producing speech.

While the usefulness and benefits of applying formulaic expressions are agreed upon, the scarcity of extensive empirical studies suggests there are hindrances in these fields. Firstly, it is important to note that there is no systematic classification of units under investigation. Formulaic interpreting has been mostly limited to a small number of formulaic sequences such as idiomatic, collocational, and metaphorical expressions using various identification methods (Li & Halverson, 2022). Frequency-based formulaic sequence lexical bundles have been overlooked despite their omnipresent feature in language production (Altenberg, 1998; Erman & Warren, 2000). Secondly, previous studies have focused mainly on European languages. Genetically distant languages such as English and Chinese (Tsao, 1982) have rarely been discussed in this line of research with a few recent exceptions of corpus-based studies exploring the use of LBs in Chinese-English pairs through ST-TT descriptive data (Li & Halverson, 2020,

2022; Xu & Li, 2021). However, there has been limited interest in comparing the use of such formulae between interpreted language and L1 language in this language pair in examining how the two groups of speakers apply LBs as building blocks to construct their spoken output. Thirdly, the issue of directionality should also be taken into consideration. Some studies reveal that both professional interpreters and interpreting trainees have adopted different strategies when working in different directions (A-to-B and B-to-A interpreting; Bartłomiejczyk, 2006; Chang & Schallert, 2007; Wu & Liao, 2018). Most of the studies reviewed have focused on interpreting from one's B language into one's A language, perhaps because many international institutions (e.g., the United Nation, the European Union) favour interpreting into A language (Albl-Mikasa & Tiselius, 2021). Few studies have examined the interpreting products of B language, while retour interpreting from one's A language into one's B language is now widely recognized as an interpreting working mode that fulfils a genuine market demand and ensures interpreting quality (Albl-Mikasa & Tiselius, 2021).

To fill these research gaps, this research project carried out a comparison of the use of lexical bundles between L1 English speakers and L2 interpreters and ST-TT descriptive data in simultaneous interpreting (SI) mode from the perspective of textual features based on a relatively large-scale self-built comparable corpus.

Research has reported clusters to be genre-sensitive and to vary across genres (Biel et al., 2019; Cortes, 2004; Hyland, 2008b). The current study limits the corpora to the register of political debates and similar proceeding settings as political discourse is considered to be highly formulaic (Li & Li, 2015; Wu et al., 2021; Yusof, 2021). To limit the research scope, the current study focuses solely on investigating four-word sequences, which are the most studied, especially with respect to English (Ädel &

Erman, 2012; Biel et al., 2019; Chen & Baker, 2016) to facilitate comparison with previous studies.

1.2 Research questions

The aim of this study is twofold. Firstly, to contribute to the knowledge of how the two groups of speakers perform formulaicity in their speech, the distribution of lexical bundles in terms of structure and function used both in L1 speech and interpreted texts is examined. Secondly, to investigate ST interference (i.e., the degree to which interpreters' employment of lexical bundles is influenced by the source texts) the relationship between source texts and target texts regarding the use of LBs is identified.

RQ1: What are the general distribution patterns in the use of LBs in L2I and L1O in terms of frequency?

RQ1.1 Are interpreted language outputs more formulaic than native speech in terms of the use of LBs?

RQ 1.2 What are the most frequently used (top 50) LBs between the two sub-corpora?

RQ 1.3 What LBs are shared in the two groups?

RQ 1.4 Are there any patterns of overuse or underuse of the shared LBs between the two groups of speakers?

RQ2: What differences exist in the lexical bundles used by L2 interpreters and L1 English speakers based on their syntactic structures and discoursal functions?

RQ 2.1 What are the typical structural characteristics of lexical bundles used by L1 English speakers and professional L2 interpreters in the political register?

RQ 2.2 What are the typical discourse functions served by the lexical bundles used by the two groups of speakers?

RQ3: What is the impact of source texts on L2 SI regarding the use of LBs in interpreted texts?

RQ 3.1 What translation relationships do the target language LBs have to the corresponding parts of the STs?

RQ 3.2: How are these three types of translation relationships distributed in the dataset?

1.3 Data and methodology

To examine patterns in the use of lexical bundles (LBs) in the natural speech of L1 speakers and L2 interpreters, a comparable corpus of spontaneous speech is required. A parallel corpus was also built to determine whether and to what degree the employment of LBs by L2 interpreters is impacted by the source texts.

The comparable-parallel corpus (UNSCCP corpus) consisting of political debates in the United Nations Security Council (UNSC) was designed and built for the current study because of its openness, availability, and comparability. The comparable corpus included both simultaneously interpreted texts in English (interpreting into L2 language) and L1 original English speech, while the parallel corpus contained source texts in Chinese and target texts in English.

Following Biber et al. (1999), a lexical bundle is defined based on the frequency of recurring sequences of orthographic word units. This research project focused on four-word bundles as these are "the most researched length" in studies on lexical bundles and are often within "a manageable size for manual categorisation and concordance checks" (Chen & Baker, 2010, p. 32), as well as being the most studied in previous research. As the lexical bundles occur frequently and are distributed widely in different

texts (Biber, 2010, p. 170), frequency cut-off points should be set to identify lexical bundles. However, Biber et al. (2004) claimed that setting the cut-off point is somewhat arbitrary as identifying the significance of the frequency of lexical bundles is subjective.

In the present study, a relatively high frequency threshold of 40 occurrences per million words was chosen considering the high formulaicity of the corpus. The dispersion rate is another concern when identifying lexical bundles in ruling out idiosyncratic use by individual speakers or authors. It is stipulated that to be regarded as lexical bundles they should appear in at least eight different texts.

Corpus analysis software *Wordsmith 8.0* was used for the automatic retrieval of the four-word lexical bundles from the self-built corpora based on the criteria mentioned above. Following Chen and Baker (2010, 2016), overlapping bundles[1] were manually screened out to prevent inflated results via concordance analyses, and combined as appropriate (2016, p. 855). In addition, meaningless bundles such as *its role as the*, *peace and the Arab* were excluded as they do not process certain functions. It is believed that the scrutinized bundles genuinely reflect the frequency-related building blocks of discourse in mediated and non-mediated languages.

The identified bundles were then classified into different syntactic structural and discoursal functional groups based on the adapted taxonomy framework proposed by Biber et al. (2004). The high-frequency four-word bundles identified in the interpreted sub-corpora were then explored further in the parallel corpus containing source texts of Chinese speeches and target texts of English interpreting output, to investigate the

[1] The overlapping word sequences refer to two or more four-word bundles derived from the same longer bundle. For example, the five-word bundle "call on all parties to" derives bundles of *call on all parties* and *on all parties to*.

translation relationship (equivalence, addition, and shift) between source texts and target texts in the use of four-word bundles. This revealed the extent to which interpreters' use of lexical bundles is influenced by the source texts.

1.4 Outline of the study

The thesis is organised as follows:

The objective of the first chapter is to introduce the rationale and motivation for studying the use of lexical bundles of L2 interpreters. Lexical bundles (LBs), one of the most frequent multi-word sequences in corpora, have been extensively examined in applied linguistics, although the way they are employed by L2 interpreters has been rarely explored. The research questions to be answered are posed, followed by a brief introduction to the data and methods to be used. The dissertation's structure is then outlined.

Chapter two begins with a discussion of earlier research on formulaic language and then moves on to investigate the specific formulaic sequences known as lexical bundles. These investigations are conducted through the lens of applied linguistics, translation, and interpreting studies, highlighting the research gaps in this line of research.

Chapter three and chapter four introduces the self-built corpus for the current research and the methodology adopted to identify and analyse the use of LBs. The study compare the use of LBs through both comparable and parallel corpora. The examination of LBs is conducted from two perspectives: first, textual features including general distribution pattern, structural and functional analysis of the LBs identified from the comparable corpus of L1 speakers and L2 interpreters are analysed, and second, the translation

relationship of the use of LBs between ST and TT in the parallel corpus is examined.

Chapter five presents the study's findings and discusses them in light of the first two major research questions and the relevant literature. The primary objective is to compare the use of LBs between L2 simultaneous interpreting and L1 original speeches in terms of general distribution, structural and functional use, and to discuss how the two groups of speakers construct their language output using LBs.

Chapter six investigates and shows the LBs' three types of translation relationships (i.e., equivalence, addition, and shift) to the source texts in different discourse function groups to identify whether and to what extent source texts impact interpreters' usage of LBs. The results of this chapter reveal that equivalence pattern is found in most stance LBs and referential LBs used by L2 interpreters, but addition pattern is more common in discourse-organising bundles. Shift pattern only accounted for a small percentage of total occurrences.

Chapter seven summarizes the primary findings and conclusions of the current study, together with the scholarly contributions and significance of the research. The author also outlines the limitations of the current study as well as possible directions for further research.

Chapter 2 Literature Review

This chapter first reviews previous studies on formulaic language, then focuses on the research into the particular type of formulaic sequences known as lexical bundles in the fields of applied linguistics, translation, and interpreting studies. In addition, the author also reviews previous studies on corpus-based translation and interpreting studies. moreover, as the current study paid special attention to interpreting products of B language, research on the impact of working direction on simultaneous interpreting is also examined.

2.1 Formulaic language

Formulaic sequences are ubiquitous in natural language use. This long-recognized linguistic phenomenon has been extensively studied using a variety of research methodologies in the fields of corpus linguistics, psycholinguistics, discourse analysis, second language acquisition, and second language pedagogy (Altenberg, 1993; Arnon & Snider, 2010; Becker, 1975; Biber, 2009; Conklin & Schmitt, 2008; Firth, 1957; Hyland, 2008a; Raupach, 1984; Sinclair, 1991; Swinney & Cutler, 1979; Wray, 2002).

Wray (2002, p. 9) defines a formulaic sequence as "a sequence, continuous or discontinuous, of words or other elements, which is or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar". There are a few key points, not explicitly stated in this definition, that are fundamental to understanding the nature of formulaic sequences. Firstly, the notion of formulaic language is a complex one, and there a number of different types of formulaic sequences. Biber et al. (2004, p. 372) explain that studies on formulaic sequences mainly differ in the following ways:

- how such expressions are defined and identified, whether by perceptual importance, frequency or some other criterion;
- what exactly are the type of expressions under investigation, whether continuous sequences of words, frames or collocational frameworks, two-word collocations or longer sequences;
- how these formulaic sequences should be described, whether structurally or functionally or both.

The term "formulaic sequence" can thus be regarded as an umbrella term used to refer to different types of prefabricated word strings – including idioms such as *under the weather*, collocations such as *make an effort*, and multi-word units/expressions such as *in terms of* – which vary in terms of their idiomaticity, invariability, and structural completeness.

Secondly, this definition is based on the assumption that formulaic sequences are holistically stored and processed, something known as the holistic hypothesis (Jiang & Nekrasova, 2007). The holistic hypothesis posits that formulaic language facilitates fluent communication by allowing speakers to produce prefabricated chunks of words with specific meanings and functions that can be easily comprehended by readers or listeners, instead of constricting utterances word by word (Wood, 2015). In other words, formulaic sequences possess processing advantages by allowing speakers to lighten their cognitive burden during speech production (Conklin & Schmitt, 2012). Pawley and Syder (1983) note that in everyday conversation among L1 speakers, fluent speech contains a large number of strings of independent clauses. The presence of these memorized chunks in daily conversation means that a speaker is able to encode the required meanings in whole clauses and avoid hesitation. Arnon and Snider (2010)

carried out an experiment which found that L1 speakers reacted more quickly to higher frequency four-word sequences than to lower frequency combinations, indicating that frequent word strings are processed as wholes. Studies on the use of formulaic sequences involving L1 speakers seem unanimous in identifying faster processing speeds for frequently used sequences in contrast to less frequent ones: as if formulaic sequences are produced as single words (Conklin & Schmitt, 2008; Guz, 2014; Nattinger & DeCarrico, 1992; Schmitt et al., 2004; Underwood et al., 2004).

On the other hand, similar research on nonnative speakers has tended to yield mixed results. For example, Siyanova-Chanturia, Conklin, and Schmitt (2011) found that native speakers process idioms quicker than novel language phrases, while nonnative speakers demonstrated the opposite pattern: which may show that nonnative speakers encounter idioms less often. In contrast, Wood (2006, 2009, 2010) investigated the role of formulaic language in speech fluency among L2 learners, suggesting that increased use of formulaic language boosts the performance of speech fluency; while a small-scale experiment conducted by Boers, Eyckmans, Kappel, Stengers, and Demecheleer (2006) claimed to show that building a repertoire of formulaic sequences contributed to the improvement of L2 speakers' speech proficiency. It seems that whether formulaic sequences are processed as holistic units, thus facilitating speech fluency among L2 speakers, relates specifically to the type of formulaic sequences under investigation, and that idiomatic formulaic sequences, in particular, are difficult for nonnative speakers to process as holistic units.

In addition, the processing advantage of formulaic sequences may show differing effects among language learners at different proficient levels. Conklin and Schmitt (2012) argue that the use of formulaic sequences used by proficient learners shows

processing advantages similar to those of native speakers, but less proficient learners tended to perform formulaic sequences in a more word-by-word manner.

2.2 Lexical bundles

2.2.1 Defining lexical bundles

The current study follows a frequency-based approach to examine the use of a particular type of formulaic sequence, the *lexical bundle* (LB), a unit which is identified not intuitively but rather empirically through a corpus-driven methodology (Cortes, 2012, 2015). The origins of research into lexical bundles dates back at least to Altenberg (1993, 1998), who was one of the first researchers to adopt frequency as a primary selection criterion for identifying word combinations, and who employed a functional analysis to classify them. Later, the term "lexical bundle" (LB) was put forward by Biber and his colleagues in their *Longman Grammar of Spoken and Written English* (Biber et al., 1999b), where they define LBs as "bundles of words that show a statistical tendency to co-occur (1999b, p. 989)" or less technically as "simply sequences of word forms that commonly go together in natural discourse" (1999b, p. 990), and since then, research into LBs has flourished. In later work, Biber et al. (2004) claim that lexical bundles can be regarded as the basic building blocks of discourse both in written and spoken registers and that they play important functions in constructing discourse.

The use of an empirical methodology suggests that the length of lexical bundles tends to affect their identification as such. There are differences between three-word lexical bundles and longer expressions consisting of four or more words: in general, longer bundles are less common in natural discourse than shorter ones. Biber et al. (1999b, p. 992) pointed out that three-word bundles can be regarded as "a kind of extended

collocational association" but four-word, six-word or even longer bundles are "more phrasal in nature and correspondingly less common". In addition, the length of a bundle influences the "type of lexical items that make up the bundle, the grammatical group the bundle aligns with, and the communicative function of its use in a particular register" (Cortes, 2015, p. 204).

Frequency is the criterion most widely used to define lexical bundles. It is commonly believed that the most frequent multi-word sequences in a corpus are LBs, and that they occur across a range of texts or different disciplines. The high frequency of such word sequences suggest that they are stored and used as prefabricated chunks and they are more likely to be used unanalysed compared to lower frequency sequences (Cortes, 2015). The frequency cut-off used to identify lexical bundles in different studies is rather arbitrary (Altenberg, 1998; Biber et al., 2004; Hyland, 2008b). Frequency thresholds from 10 to 40 occurrences per million words have been employed in different studies (Biber et al., 1999b; Hyland, 2008b; Simpson-Vlach & Ellis, 2010). Researchers also set distribution criteria to identify LBs: in other words, LBs must generally appear in at least five texts in the corpus, or at least in a certain percentage of the texts therein, in order to be deemed as LBs (Biber & Barbieri, 2007; Hyland, 2008b). This procedure helps to eliminate the possibility of idiosyncrasy, which means that specific sequences may be often used more often by a particular user, or when a particular topic is discussed.

Although the use of a frequency-based method to identify LBs seems straightforward, there are still some challenges faced by researchers in targeting LBs in a corpus. Firstly, the appropriate size of corpus chosen to be studied remains undefined, although a threshold of one million words was suggested by Cortes (2015). The size of the corpus

also has implications for the length of lexical bundles, with two or three word strings able to be identified in smaller corpora if they occur frequently. Secondly, the quantitative criteria used ti to identify LBs such as frequency, and dispersion thresholds are commonly selected rather arbitrarily.

2.2.2 The structure of lexical bundles

In terms of internal structure, the majority of LBs are parts of phrases or clauses with embedded fragments rather than complete structural units (Biber et al., 1999b, pp. 993-1000). Biber et al. (1999b) found that only 15% of the LBs in conversation could be recognized as complete phrases or clauses, while less than 5% of LBs in academic prose are complete structural units. Although lexical bundles are structurally incomplete, they can connect two structural units, which means that the last word of the bundle is commonly the first element of the following structure (Biber et al., 2004). For example, lexical bundles such as *we hope that the*, *ready begins to continue to*, *is important that we*, begin with clauses or phrases with their last words forming another structural unit.

Although lexical bundles are often incomplete units, they have "strong grammatical correlates" which allow them to be grouped into different structural types (Cortes, 2015, p. 207). Biber et al. (2004, p. 381) proposed a framework for classifying LBs into structural types which represents a further refinement of the previous seminal work of (Biber et al., 1999b). The primary purpose of this taxonomy is to distinguish clausal bundles from phrasal bundles. Lexical bundles are classified into three broad types:

Type 1: lexical bundles incorporating verb phrase fragments, either beginning with a subject and followed by a verb phrase, or starting with a discourse marker and followed by a verb fragment, or starting directly with a verb phrase;

Type 2: lexical bundles incorporating dependent clause fragments; and

Type 3: lexical bundles incorporating noun phrase and prepositional phrase fragments.

These three broad structural types of lexical bundles are shown in Table 2.1 below, taken from Biber et al. 2004:

Table 2.1 Structural Types of LBs (Biber et al., 2004, p. 381)

1. Lexical bundles that incorporate *verb phrase* fragments
1 a. (connector +) 1st/2nd person pronoun + VP fragment Example bundles: *you don't have to, I'm not going to, well I don't know* 1b. (connector +) 3rd person pronoun + VP fragment Example bundles: *it's going to be, that's one of the, and this is a* 1c. Discourse marker + VP fragment Example bundles: *I mean you know ,you know it was, I mean I don't* 1d. Verb phrase (with non-passive verb): Example bundles: *is going to be, is one of the, have a lot of, take a look at* 1e. Verb phrase with passive verb: Example bundles: *is based on the, can be used to, shown in figure N* 1f. yes-no question fragments: Example bundles: are you going to, do you want to, does that make sense 1g. WH-question fragments: Example bundles: what do you think, how many of you, what does that mean
2. Lexical bundles that incorporate *dependent clause* fragments
2a. 1st/2nd person pronoun + dependent clause fragment Example bundles: *I want you to, I don't know if, I don't know why, you might want to* 2b. WH-clause fragments: Example bundles: *what I want to, what's going to happen, when we get to* 2c. *If*-clause fragments: Example bundles: if you want to, if you have a, if we look at 2d. (verb/adjective+) to-clause fragment Example bundles: *to be able to, to come up with, want to do is* 2e. *That*-clause fragments: Example bundles: *that there is a, that I want to, that this is a*
3. Lexical bundles that incorporate *noun phrase and prepositional phrase* fragments
3a. (connector +) Noun phrase with of-phrase fragment:

Example bundles: *one of the things, the end of the, a little bit of*

3b. Noun phrase with other post-modifier fragment:

Example bundles: *a little bit about, those of you who, the way in which*

3c. Other noun phrase expressions:

Example bundles: *a little bit more, or something like that, and stuff like that*

3d. Prepositional phrase expressions:

Example bundles: *of the things that, at the end of, at the same time*

3e. Comparative expressions:

Example bundles: *as far as the, greater than or equal, as well as the*

This structural taxonomy of lexical bundles has been foundational for later studies on lexical bundles. A number of scholars have modified this taxonomy slightly by grouping these structural types into different sub-groups in line with their research purposes (Cortes, 2004; Hyland, 2008b; Wang, 2017).

2.2.3 The function of lexical bundles

Apart from being incomplete in structure, the meaning of lexical bundles is usually not idiomatic; while by contrast idioms are often less frequent but are clearly prefabricated. The meaning of a lexical bundle can usually be derived from the combination of the meanings of the words that make it up. In lexical bundles such as *it is important that*, *hope that all the* or *give full play to*, the words in these expressions retain their own meaning and also contribute to the overall meaning of the lexical bundle when appearing in sequence in texts.

When analysing LBs, researchers normally assign functions to the most frequently used sequences and classify them into different categories. In this way, the saliency of a particular type of expressions across different registers or among different groups of users can be examined.

In their study using an inductive approach to investigate the language used in classroom teaching textbooks, (Biber et al., 2004, pp. 389-396) identify the three primary types of classroom language by functions as “stance expressions”, “discourse organisers”, and “referential bundles”. It is worth noting that, while a single bundle may perform more than one particular function even in a single occurrence, or serve a number of different functions according to the context, most bundles typically have one primary function. The structural and functional categories of bundles are closely related to each other:

Biber et al. (2004) found that bundles that contain complement clauses tend to perform a stance function, and those containing noun or prepositional phrases mainly serve referential functions. Table 2.2 below sets out Biber et al. 2004's functional taxonomy of LBs with examples:

Table 2.2 Functional taxonomy of LBs (Biber et al., 2004, pp. 384-388)

I. Stance expressions	II. Discourse organisers	III. Referential bundles
Express attitudes or assessments of certainty that frame some other proposition A. Epistemic stance *I don't know if,* *I think it was,* *are more likely to,* *the fact that the* B. Attitudinal/modality stance B1) Desire *if you want to,* *I don't want to* B2) Obligation/directive *you might want to,* *it is important to* B3) Intention/prediction *I'm not going to,* *it's going to be* B4) Ability *to be able to,* *can be used to*	Reflect relationships between prior and coming discourse A. Topic introduction/focus *what do you think,* *if you look at* *I would like to* B. Topic elaboration/ clarification *I mean you know,* *on the other hand*	Make direct reference to physical or abstract entities, or to the textual context itself A. Identification/focus *that's one of the, of the things that* B. Imprecision *or something like that,* *and stuff like that* C. Specification of attributes C1) Quantity specification *there's a lot of,* *how many of you* C2) Tangible framing attributes *the size of the, in the form of* C3) Intangible framing attributes *the nature of the,* *in the case of* D. Time/place/text reference D1) Place reference *in the United States* D2) Time reference *at the same time,*

		at the time of D3) Text deixis *shown in figure N,* *as shown in figure* D4) Multifunctional reference *the end of the,* *the beginning of the*

This functional taxonomy was quickly adopted by a number of studies of the use of LBs in academic registers (Biber & Barbieri, 2007; Cortes, 2004), modified and expanded by researchers to capture the features of different registers (Ädel & Erman, 2012; Chen & Baker, 2010; Li & Halverson, 2022; Simpson-Vlach & Ellis, 2010).

Li and Halverson (2022) employed this taxonomy into a parallel corpus of interpreted texts between Chinese and English to illustrate the functions of lexical bundles in target texts with complete equivalence in source texts. Li and Halverson retained the three main types of classification as with Biber et al. stance expressions, discourse organisers, and referential expressions – but modified the sub-types for the political discourse they investigated, adding the sub-types of People Deixis and Political Terms under referential expressions to capture instances occurring frequently in their parallel corpus.

Table 2.3 Functional Taxonomy of LBs (Li & Halverson, 2022, p. 11)

I. Stance expressions	1. 愿望类 我可以，我想，我要	1. Desire *I would like to*
	2. 责任 / 指令 我们要，我们会保持	2. Obligation / Directive *we need to ensure*
	3. 目的 / 将要发生的事	3. Intention / Prediction *the central government will*

	中央政府(要、会) 4. 会 / 能够 可以，能够，就会	 4. Ability *will / would be able to*
II. Discourse organisers	1. 引入 / 聚焦 这是个事实，确实，的确 您 怎么(看待、分析) 2. 详述 / 阐明 只要(/我们)(也、这)就是 说，(这)就是	1. Introduction / Focus *it is true that* *how do you see* 2. Elaboration / Clarification *as long as the* *that is to say*
III. Referential expressions	1. 识别/关注 [.....的]关系 2. 特征说明 很大，很多，充满，非 常 3. 地点指示	1. Identification / Focus *the relationship between the* 2. Specification of Attributes *have a lot of* 3. Place Reference *between / of the two sides*

	双边，双方，两岸，两地	
	4. 时间指示 这些年，这几年，最近几年	4. Time Reference *the past few years*
	5. 语篇指示 政府工作报告	5. Text Deixis *the government work report*
	6. 人的指示（新） 香港(民众、人民、人士)，港人	6. People Deixis (new) *people in Hong Kong*
	7. 政治术语（新） 一国两制 香港、香港特别行政区	7. Political Terms (new) *One Country Two Systems the Hong Kong SAR*

Hyland (2008b) modified this framework in an attempt to explore the frequency, forms and functions of LBs in a large corpus covering academic articles, theses and

dissertations from four different disciplines. Hyland put forward a new classification to better represent the use of LBs in academic writing, assigning each instance of an LB to one of three broad categories including research, text and participant-oriented type, with more detailed subcategories as shown in Table 2.4 below.

Table 2.4 Functional Taxonomy of LBs of Hyland (Hyland, 2008b, pp. 13-14)

Research-oriented functions help structure experience and activity of real world:	Location—*at the same time* Procedure—*the use of the* Quantification—a wide range of, Description—*the structure of the* Topic—*in the United States*
The textual functions are labelled by Hyland as text-oriented which deal with the meaning of text and its organisation:	Transition signals—*on the other hand* Resultative signals—*as a result of* Structuring signals—*in the present study* Framing signals—*in the case of*
The interpersonal functions are labelled participant-oriented: focusing on the writer or the reader:	Stance features—*may be due to* Engagement features—*as can be seen*

While previous studies mainly target written texts used in academic writing and classroom teaching, Wang (2017) proposed a classification which takes into consideration the features of spoken language, Four broad categories are distinguished, namely, real-world oriented, text-oriented and participant-oriented and others as shown in Table 2.5:

Table 2.5 Functional Classification of LBs (Wang, 2017, p. 198)

1. Real-world oriented: referring to real-world properties i. Time/place/personal reference, e.g. *at the end of, the rest of Europe* ii. Identification/descriptive attribute, e.g. *the first half of, the name of the* iii. Quantity specification, e.g. *a lot of er, a little bit of*
2. Text oriented: signaling the organisation of the speech and the elements of an argument. i. Transition signals: establishing logical links between elements, e.g. *on the other hand, so that we can* ii. Framing signals: situating arguments by specifying limiting conditions, e.g. *in the case of, on the basis of*
3. Participant oriented: focusing on the interaction between the speaker and the listener. i. Stance markers: expressing epistemic stance, e.g. *er it is not*, or the speaker's attitudinal/modality stance, e.g. *I don't know if, it has to be* ii. Engagement signals: addressing the hearer directly, often involving fragments of questions, e.g. *if you want to, what do you think*, or expressing agreement/ disagreement, e.g. *no no no no, yeah mhm hm yeah* iii. Procedure signals: indicating actions and the organisation of the lecture/seminar, e.g. *I would like to, you are going to* iv. Fillers: meaningless repetition of single words or sounds, e.g. *the the the the, of the of the*
4. Others A small number of instances whose functions cannot be easily classified into the above categories were put into a group labelled as "Other", e.g. *and so on and*, *a look at the*.

The function of a particular lexical bundle is not fixed across registers, but rather may need to be categorised under different functions in different registers. For example, the four-word bundle, *I would like to*, is assigned to the sub-group "topic introduction/focus" in Biber et al. (2004) in the register of classroom teaching and written texts, to the sub-group of "procedure signals" in Wang (2017) in the register of spoken academic English

as a Lingua Franca (ELF), and to the sub-group of "desire" in Li and Halverson (2022) in the register of interpreted political discourse. At the same time, the function of a lexical bundle is dependent on the context surrounding the sequence, and it may be necessary to examine the concordances of a particular lexical bundle in order to decide its function. In addition, any one bundle may perform more than one function in a certain corpus: in this case, in analysis a particular bundle will be matched to the function that it expresses most frequently (Biber et al., 2004; Cortes, 2015). Moreover, many bundles are found to be salient in specific registers, which would support the observation that some lexical bundles are register-bound.

2.3 Studies on lexical bundles across disciplines

2.3.1 Studies on lexical bundles in applied linguistics

Most studies on lexical bundles in the field of applied linguistics have followed the structural and functional framework proposed by Biber et al. (2004). The use of LBs has been proved to be pervasive in language use: LBs generally fall under several broad types of syntactic structures, and perform certain discourse functions in language production. It is demonstrated that the use of lexical bundles varies across text types, disciplines, spoken and written registers, and native and non-native speakers with varying levels of English proficiency. These inconsistent results of the use of LBs between native and non-native speakers have motivated this doctoral project, which attempts to compare the use of LBs between native speakers and L2 interpreters at professional level. Professional interpreters are defined as interpreting practitioners who possess interpreting related certificates or degree and have interpreting experience. The detailed description of professional interpreters is provided in Section 3.2 of Corpus Design. Professional interpreters allow this research to largely rule out the

impact factor of interpreters' language proficiency when addressing the second research question. That is to say, if the process advantages of LBs are not manifested among professional interpreters, the reason could be that they do not regard LBs as holistic units rather than they are not familiar with the use of these LBs.

Ever since Biber et al. (1999b) coined the term "lexical bundles" and examined the use of these structures in university teaching and textbooks, many studies have employed this framework to describing the syntactic structures and discourse functions of different types of lexical bundles. Most of these researchers have adopted a corpus-driven approach which is inductive in nature, and comparing the patterns of LBs across a variety of text types has become a frequent research topic (Breeze, 2013; Conrad & Biber, 2005; Hyland, 2008b; Shirazizadeh & Amirfazlian, 2021; Wang, 2017).

A substantial number of studies have focused on academic discourse (Hyland, 2008a, 2012; Hyland & Jiang, 2018; Le & Harrington, 2015; Shirazizadeh & Amirfazlian, 2021). For example, Biber and Barbieri (2007) investigated the use of LBs in both spoken and written discourse across a range of university registers and found that LBs are more frequently used in non-academic than in instructional registers in the university context. Nesi and Basturkmen (2006) investigated the cohesive role of lexical bundles used in university lectures. They argued that bundles within function groups of discourse organisation and referential bundles have the potential to perform cohesive roles. Native speakers of English may be able to recognize the functions of bundles, while non-native speakers are less likely to pick up this knowledge as they learn the language consciously rather than acquire it naturally.

Comparisons have also been made of the use of LBs in spoken and written registers,

with studies suggesting that LBs are more common in writing than in speech (Biber & Barbieri, 2007; Biber et al., 2004; Biber et al., 1999b; Conrad & Biber, 2005). Biber (2009) compared the uses of LBs between conversation and academic writing. His results suggest that the patterns of LBs differ substantially between the two registers: conversational texts tend to include more fixed sequences of function and content words; while academic writing is likely to contain formulaic frames incorporating invariable function words with variable content words.

In addition, the use of lexical bundles also demonstrates inter-disciplinary variation (Hyland, 2008b, 2012; Le & Harrington, 2015; Reppen & Olson, 2020) through the investigation of lexical bundles used in multifold disciplines such as the discourse of medicine, social sciences, natural sciences (Durrant, 2017; Wang, 2017), applied linguistics and pharmaceutical sciences (Ren, 2021), business (Yin & Li, 2021), education and psychology (Cao, 2021), and history and biology (Cortes, 2004).

For example, Ren (2021) applied the functional taxonomy proposed by Biber (2009) and Biber et al. (2004) compared the extent to which the lexical bundles are fixed when used in research articles in the disciplines of applied linguistics in contrast to the pharmaceutical sciences. Their findings reveal a higher level of variability in lexical bundles in articles in the discipline of applied linguistics compared to those in the pharmaceutical sciences. Wang (2017) explored the use of four-word bundles in terms of frequency, form, and function between two spoken genres (academic lectures and seminars) in four disciplines (Medicine, Social Sciences, Natural Sciences), again showing that the use of LBs varies according to genre and discipline.

While *inter*-disciplinary variation is widely acknowledged, *intra*-disciplinary variation

in the use of LBs has also been identified among different genres within the same discipline. For example. Shirazizadeh and Amirfazlian (2021) investigated the forms and functions of 4-word bundles in a corpus containing the three genres of textbooks, research articles and theses in the discipline of applied linguistics, revealing that the bundles used both vary significantly according to genre and at the same time show commonalities across different genres. Yin and Li (2021) approached both intradisciplinary and interdisciplinary variations by examining the structural and functional features of LBs. Their results indicate that the use of bundles displaying similar intradisciplinary features of similar distribution of bundles relating to "statistics" in business disciplines and lexical bundles can also serve as interdisciplinary mark to reveal distinctive intradisciplinary features across disciplines.

There are also other studies on the use of LBs which have explored issues regarding patterns of language development, specifically, the use of LBs between L1-English and L2-English learner writers (Ädel & Erman, 2012; Bychkovska & Lee, 2017; Chen & Baker, 2010; Lu & Deng, 2019; Pan et al., 2016; Salazar, 2014; Shin, 2018) and the proficiency levels of L2-English writers (Chen & Baker, 2016; Staples, Egbert, Biber, & McClair, 2013). A number of studies conclude that L1-English speakers tend to use more LBs in terms of both number and variation than L2-English speakers, and only few bundles are shared between L1 and L2 students (Ädel & Erman, 2012; Chen & Baker, 2010).

In contrast, there are also studies showing that L2 learners may use more LBs than L1 speakers. For example, Bychkovska and Lee (2017) used a structural and functional taxonomy compared the use of LBs between L1-English speakers and L1-Chinese learners, and showed that L2 learners use significantly more bundles compared to L1

speakers, and that the types of LBs used by the two groups of speakers differ both structurally and functionally. Therefore, the language proficiency levels and the native languages of L2 English speakers might also play a role in the use of LBs beside being L1 or L2 speakers.

Staples et al. (2013) examined the use of lexical bundles across three proficiency levels in the TOFEL iBT writing section. Their findings suggest that with the development of language proficiency, L2 learners would gradually be able to produce native-like expressions rather using memorized phrases. Similarly, Chen and Baker (2016) investigated the differences in the use of LBs among L1 Chinese learners of various English proficiency levels in essay writing, suggesting that the use of LBs by lower level learners rely more on verbs, especially the use of copula *be*, and colloquial quantifiers, features more common in conversation; while more proficient writers adopt a more impersonal tone through the use of more nominal components and academic or literate style bundles, more resembling academic prose.

In professional contexts, the research of Pan et al. (2016) compared the use of lexical bundles between L1-English and L2-English professionals writing and showed that L2 professional writers use more verb and clause fragments while L2 professional writers use more noun and prepositional phrases. In addition, L2 writers use LBs that are functionally different from that in L1 writers and tend to use specific bundles incorrectly such as using the four-word bundle “as well as the” as a conjunction rather than a preposition. Wei and Lei (2011) focused on the use of lexical bundles in the academic writing of advanced Chinse EFL learners and professional authors, singling out four-word bundles in a corpus of doctoral dissertations and published journal articles, and showing that learner writers depend more on LBs to formulate their articles

and employed less participant-oriented bundles than professionals. They also overuse passive structures and underuse anticipatory *it* structures of LBs than the professional ones, which might by account for the learners' preference for impersonality in their academic prose.

2.3.2 Studies on lexical bundles in translation

Compared to the large number of studies of LBs in linguistics, in the field of translation studies LBs have received only limited attention and remain relatively unexplored territory. A handful of studies that focus on lexical bundles or similar multi-word sequences (i.e. collocations, set phrases, idioms) focusing on the similarities and differences between translated and non-translated texts in the same language or across languages.

Studies of multi-word sequences in translation studies have gradually evolved from small-scale exploratory research into systematic quantitative analyses based on large-scale corpora. Analysis of these prefabricated phrases has proved to be effective in revealing different pattens across original and translated texts. The identification of these sequences in translation can also be considered as a way of improving language fluency.

Baker (2004) was one of the early scholars to introduce the study of multi-word phrases into translation studies. Although she did not use the term lexical bundles directly but opted for the terms "fixed" and "semi-fixed lexical phrases", she mainly studied four-word phrases. Although Baker stressed the unreliability of the word lists generated by the software available at that time, in this exploratory piece of research she was able to identify variation in lexical patterns between translated and non-translated texts in

English as well as variation in style in the work of different translators. She put forward the notion of "strategies of characterization" (Baker (2004, p. 180) as one possible reason to account for the differences she identified in the use of these lexical phrases, and proposed that higher frequency of certain lexical phrases in translated texts may be the result of direct carry over from features in the source texts. She also suggested that variation in the use of lexical phrases may be related to translation strategy or may simply represent favourite/quick expressions of the translators.

In later work Baker (2007), Baker conducted a pilot study on patterns of idiomaticity in translated and non-translated English. She argued that the relatively lower degree of idiomaticity in translations may be related to the phenomenon of normalisation, whereby translators tend to choose "safe, typical patterns of the target language" (p. 14), and by contrast avoid using idioms which are relatively opaque in meaning. It is a pity that only few examples were provided as explanation while no statistical data have been offered in the study. Although Baker's two studies were not able to offer conclusions or generalizations for the lexical patterns used in comparable corpus, they did initiate the study of multi-word units in translation studies and offered possible explanations to account for the differenced identified in this area of research.

Xiao (2010) also compared multi-word sequences in original and translated texts. Based on a one-million word corpus, he investigated idioms, word clusters and reformulation markers in translated Chinese in contrast to native Chinese, suggesting that word clusters are significantly more common in translated Chinese in terms of frequency, coverage and key clusters. The tendency to use more fixed and semi-fixed recurring patterns in translated Chinese can be explained as a strategy used by translators to improve fluency. Xiao attempted to verify whether the translation features he identified

in translated English could also be found in translated Chinese. His analysis of word clusters in this study reveals a relatively higher level of homogeneity in translated texts, which support the translation features of convergence and levelling out. And the overall higher frequency and coverage of word clusters used founded in in translated texts. The reformulation makers are regarded as a type of discourse markers which can enhance connectivity in discourse (Schourup, 1999), and they can be seen as markers of explicitness (Murillo, 2004). The findings regarding reformulation makers displayed a tendency to use stylistically simpler makers in translated Chinese than native Chinese, thus providing evidence of simplification in translation.

Shrefler (2011) is one of scholars in translation studies who explored this concept, examining the use of LBs in relation to the readability of translated texts of the Bible. Using both quantitative and qualitative analysis to compare two German Bible translations, he found that the translation incorporating more lexical bundles was more readable for modern readers. Shrefler employed Baker (2004) functional taxonomy in his study, and claimed in particular that lexical bundles falling into the category of discourse organisers may contribute to readability as they alert the readers to upcoming messages, while by contrast referential bundles might work *against* readability.

Lee (2013) applied lexical bundles as a discovery tool to identify textual features between translated texts vis-à-vis non-translated texts in Korean journalistic texts. He classified these bundles structurally and functionally using Biber et al. (2004) framework and extended the functional taxonomy by showing that singling out temporal and spatial phrases as a subcategory of referential bundles revealed variation in stylistic choices across translators. Lee's study connected function analysis and multi-word sequences in translation studies from a contrastive analysis perspective and

advocated to examine the use of LBs in concordances and collocates to explore how they are used in context, rather than simply relying on the overall distribution of LBs in a corpus. However, the relatively small corpus adopted, and initial bundle lists showed that this study represented only a beginning in using lexical bundles to identify significant differences between translated and non-translated texts.

The research of Ebeling and Ebeling (2017) compare the use of 3-word bundles in translated fictional texts and original English fictions in a multi-million word corpus with the aim of examine whether or to what extent translated language is different from original language. In extracting recurrent word-combinations, Ebeling and Ebeling used the threshold of distribution and threshold of recurrence to eliminate the idiosyncratic use of 3-word bundles by individual authors and translators as proposed by Biber et al. (1999b). They also proposed a functional classification of LBs based on the taxonomy of Biber et al. (2004) which comprises four main categories including evaluative, informational, modalising and organisational LBs as applied to translated texts. Based on the quantitative comparison of the 3-word bundle types, their findings suggested that translated texts contain more occurrences of frequent 3-word bundle types than original texts after the normalisation of frequency counts. They presumed that each translation can find its textual equivalents in source texts. Their research supported the translation universals of target language normalisation and source texts shining through tested in the previous translation studies (Laviosa, 2002; Xiao, 2010) in occurrences of comparison and spatial functions. The researchers raised concerns over the effect of the size of the corpus on the results of the analysis, and "justified excluding the identification of LBs in the source texts from their analysis on the grounds that translators strive to retain the "style" of the source texts in their target renderings. In other words, following the Hallidayan model of translation, one would expect each

target text to match its source text both ideationally and interpersonally, and hence analysing only the target text will still reveal features of the source text".

Based on a large corpus of EU documents, Biel (2018) investigated LBs in EU legal language and then expanded the investigation into more genres in Biel et al. (2019), which subsequently examined the use of LBs in a corpus-driven analysis of four institutional genres in EU and their corresponding Polish genres with the aim of investigating how the degree of formulaicity is influenced by the variables of genre. Their study adopted a relatively higher threshold of 40 occurrences per million words than that of 20 occurrences per million words used in Ebeling and Ebeling (2017). Their findings echoed those of previous research that LBs occur more in translated texts than in original texts with exception of only one genre. In addition, they claimed that translations and non-translations share very few LBs, which could be explained by the assumption that translations develop their own formulaic language, an indicator of a levelling out strategy. In addition, they argued that the results of the analysis on bundle usage could be affected by the different frequency thresholds used to identify LBs and the different genres under investigation.

LBs have also been used to compare translation styles. Liu and Afzaal (2021b) investigated the use of LBs in comparing the different stylistic features of both L1 translators and L2 translators in the self-built parallel corpora of the two translation versions of *Hongloumeng*, known in English as *Dream of the Red Chamber* or *The Story of the Stone*[2]. They have identified both three-word and four-word lexical bundles

[2] *Dream of the Red Chamber / The Story of the Stone*: a novel regarded as one of China's Four Great Classic Novels, written by Cao Xueqin in the mid Qing dynasty between the 1740s-1760s but not published till the 1790s, regarded as representing the pinnacle of Chinese fiction.

and then classified their functions based on Biber et al. (2004) framework. Their findings suggest that when a translator is translating into their L1, more and varied lexical bundles are used compared when a translator is translating into their L2, which can be accounted by different translation strategies adopted by translators when translating into L1 or L2 direction, and their language background, the translation Skopos, and the social, political and ideological background in which the work of translation was produced.

Lee (2022) examined the normalisation hypothesis through a multivariate exploratory analysis of the use of LBs between translated and non-translated English literary texts. When using the overall frequency of LBs, the normalisation is not supported. Nonetheless, when applying principal component analysis (PCA) and hierarchical cluster analysis (HCA) that take into account the relationships between individual texts, the use of LBs, and translated and non-translated group categories, normalisation occurred in a subset of discourse bundles. In addition, PCA reflect the differences between the types of LBs that translators favour. The researchers noted that the overall frequency analysis should provide information on general patterns in the data, but it risks overlooking the multidimensional nature of translation corpora and the relationships between source languages, translators, genres, and linguistic features.

The above-mentioned studies mainly employ comparable corpora to examine the use of multiword units in original compared with translated texts. Berūkštienė (2017), by contrast, used a parallel corpus to study the structural types of lexical bundles in court judgements of the Court of Justice of the European Union in English and how these structures are rendered in Lithuanian. The results of this study revealed that the high frequency of pre-formulated multi-word sequences in English court judgements manifest the high level of formality in legal texts. When examining the corresponding

translations of identified English bundles, it was found that the structures in source texts were rendered into a variety of different structures in target texts. This analysis shows that a lexical bundle can be treated as "a single cognitive unit for the purposes of establishing a translational equivalent" (p. 26).

2.3.3 Studies on lexical bundles in interpreting

While scholars working in corpus linguistics and translation studies have mainly applied frequency-based approaches to studying the use of formulaic expressions, interpreting scholars, trainers and practitioner paid more attention to the processing advantage of formulaic expressions by referring to them through terms such as "formulas" "recurrent formulaic phraseologies", "pat phrases", "chunks", "formulaic expressions" , and "formulae" (Aston, 2018; Gile, 1995/2009; Henriksen, 2007; Jones, 2014; Kajzer-Wietrzny & Grabowski, 2021; Plevoets & Defrancq, 2018b; Setton, 1999).

Early interpreting research on formulaic expressions generally rely on interpreters' own experience and a detailed examination of interpreted texts, and very few scholars have examined the topic through empirical studies. One of the reasons for lacking empirical investigations could be the unviability of large-scale interpreting corpus due to the significant efforts to develop a spoken corpus and the difficulty to collect raw materials especially in early days. Additionally, corpora of small size may not be may not be compatible with corpus software and scholars have relied on subjective and manual identification of formulaic sequences.

Collocation is among the first notion discussed in interpreting studies regarding multi-word sequences. The notion of collocation is related to the interpreting process and cognitively oriented studies. For example, Zanetti (1999) and Vandepitte (2001)

considered collocations as one of the interpreters' anticipation strategies as interpreters can recall the full expressions of the frequent use of fixed collocations and linguistic formulae when their initial parts appear in speech.

Henriksen (2007) was perhaps the first researcher to apply the term "formulaic language production" to simultaneous interpreting (SI), and to identify and categorise formulaic expressions both in the ST and the TT. Henriksen explained that formula in interpreting is "the result of automatic language production" and "related to the cliché and conveys recurrent ideas typical of the conference room (p. 13)". This description concerns prefabricated nature and the stereotypical meanings of such formulas while failing to mention their frequency of occurrence. Henriksen argued that formulaic language production enhances the homogeneity of the SI output as interpreters working in the same booth tend to borrow formulaic expressions from each other and claimed that formulaicity contributes to the "graduation standardization of the interpreting products (p. 2)" in an EU context. Henriksen also categorised the formulaic expressions in the source texts as those typical of the ritualised diplomatic conferences, those reflecting typical discourse of EU cooperation, and those stemming from legal/political/managerial discourse. She further pointed out that some formulaic expressions used by interpreters observed in target texts are idiomatic expressions. In addition, she argued that interpreters benefit from both recognizing and anticipating formulaic expressions from source texts and producing formulaic expressions in the target texts.

Aston (2018) has linked the uses of recurrent formulaic expressions with the fluent output of interpreters, arguing that the formulaic phraseologies are seemingly stored in memory as single lexical units, something that would help interpreters to reduce

cognitive load and hence boost fluency in production. Given the role played by recurrent formulaic expressions in reducing efforts both in production and reception in interpreting, he pointed out that teaching useful formulae for recurrent contexts has been long practiced in interpreting training. This study takes a pedagogical perspective to explore what are appropriate data for interpreting training in formulae identification and acquisition. His findings suggested that the training of simultaneous interpreters should stress the learning of phraseological units beyond the traditional emphases on the use of terminology.

These two above-mentioned studies supported the contention that the use of formulaic sequences helps to reduce the cognitive load in speech production, thus boosting speech fluency. However, the conclusions in Henriksen (2007) are based on the manual scrutiny of transcripts of only few interpreters, and the analysis of Aston (2018) is relatively intuitive based on a small-scale corpus with a selection of a few examples to support the argument. Quantitative examination of the relation between the use of formulaic sequences and speech fluency should be conducted to fill out the picture.

Focusing on the English-Chinese language pair, a number of scholars made use of the published *Parallel Corpus of Chinese EFL Learners*, or PACCEL with the transcribed data for the Test for English Major Band 8 (TEM8)[3] in China, to conduct empirical research on 4-word lexical bundles used by EFL learners. Major findings on the features of LBs used by EFL learners included that significantly more use of LBs significantly indicated higher interpreting quality by the metrics of test scores, fluency, and accuracy (Li, 2016, 2017), while incorrect LBs mostly occurred in terms of grammar and

[3] TEM8 is designed as an achievement to assess the overall English proficiency of senior undergraduate students majoring in English Language and Literature in China.

collocation due to a lack of target LBs stored in the long-term memory (Li and Zhao, 2019).

Moreover, Wang (2016) designed and carried out a 16-week experiment to test the effects of chunk training (stress on the use of chunks in SI) on fluency in Chinese-English SI working in both directions. With 30 participants allocated across experimental and control groups, a T-test of the experiment results showed that the group receiving chunk cognition training performed better in interpreting output in terms of indices of pauses, unfilled spaces, hesitations and repairs, suggesting that chunk training can significantly promote SI fluency among learners.

Shao (2018a, 2018b) studied the structure and function of 3-6 word lexical bundles extracted from a self-built small scale professional consecutive interpreting corpus of political discourse, using the London-Lund Corpus of Spoken English (LLC) as reference corpus. The findings of Shao's study showed that professional interpreters tend to use more but less varied lexical chunks compared to L1 speakers. Structurally, professional interpreters use more single clause constituents while L1 speakers produce more multiple clause constituents, and routinized and conventional chunks are used more often in this formal political interpreting setting.

Xu and Li (2021) used the classification of Biber et al. (2004) to analysed the use of 4-word lexical bundles among professional SI interpreters taken from a SI parallel corpus. The results showed that noun and prepositional phrase fragments made up the largest proportion of LBs used by interpreters followed by verb phrase fragments and dependent clause segments. Moreover, when comparing source and target texts, interpreters employ correspondence, addition, and substitution strategies to render the

identified bundles. Tang and Jiang (2022) compared the use of four-word lexical bundles between professional and trainee interpreters in consecutive interpreting mode, providing insights into how the use of lexical bundles affect interpreting proficiency. Their findings highlight the need for interpreter training programmes to focus on the development of lexical bundle use, as this can significantly impact the quality of the interpreting output.

Li and Halverson (2020) used the CICPPC corpus to carry out a quantitative analysis of LBs in consecutive interpreting. They chose the topic of LBs so as to explore the cognitive explanatory factors related to the phenomenon of explicitation. The analysis of ST-TT data suggested that recurring LBs demonstrated three addition textual patterns (including simple addition, repetitive addition, and quasi-repetitive addition), claiming that the explanation of how causal factors related to the cognitive complexity of bilingual processing and the ease of access of highly frequent chunks is feasible. In a later paper (2022), Li and Halverson (2022) extended their study to examine the discourse functions of 4-word bundles in the target texts and their relationships with their originals in the source texts. They put forward the notion of "constraints on formulaicity" (p. 16) as a generalization capturing the tension between frequency-driven selection and the need to establish a translation relationship, arguing that the translational relation of equivalence and addition are subject to lower constraints while shift may cause higher constraints.

In summary, interpreting scholars agree that the use of formulaic expressions can help to ease the cognitive load of interpreters and facilitate fluency in interpreting output. Based on the experience of interpreting practice and applying theories of modes of interpreting, early studies mainly rely on the manual examination of small corpora or

transcriptions of the discourse of a limited number of interpreters to illustrate how formulaic expressions are used by interpreters. More recently, empirical approaches have been employed using larger corpora of interpreting discourse to investigate the patterns of the prefabricated sequences. In particular, corpus-based interpreting studies have only recently begun to quantitative investigation on the role of LBs in interpreting settings. Current empirical studies on the use of LBs in interpreted texts have mainly focused on the comparison between source and target texts, while the use of these sequences between L1 and interpreted texts have not been fully explored. Additionally, although there are interpreting scholars using a structural and functional classification to describe the use of LBs in interpreting, a comprehensive dual analysis is still lacking.

2.4 Corpus-based interpreting studies as a research approach

2.4.1 Corpus linguistics as a research approach

The term "corpus linguistics" was coined by Jan Aarts in his co-edited book *Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research* (Aarts & Meijs, 1984). To some extent, corpus research can be seen as an expansion of quantitative research in sociolinguistics that began in the 1960s. Corpus linguistics explored the patterns of variation in a wider range of subjects, regarding spoken and written registers, in contrast to the relatively small range of varieties in quantitative sociolinguistics (Kendall, 2011). Biber (2010) summarised corpus linguistics as empirical in nature, analysing the actual patterns of use in natural texts based on both quantitative and qualitative analytical methods.

Corpus-based research largely relies on the representativeness of corpus (Biber, 1993; McEnery, Xiao, & Tono, 2006, pp. 13-21). Thus, the size and composition of corpus

are the top considerations in its design: in other words, the size of corpora should be large enough so that the linguistic features patterns can be representative, and the texts in a corpus should be able to represent the registers in the target domain of language use.

The two major approaches in corpus studies can be distinguished as “corpus-based” versus “corpus-driven” (Biber, 2009, 2010; McEnery et al., 2006). The term corpus-based refers to methods of analysing systematic patterns of variation and uses of pre-defined linguistic features, while the notion of corpus-driven is inductive in nature, constructing patterns merging from a corpus with minimal theoretical presumptions.

2.4.2 Corpus-based translation and interpreting studies

The corpus-based approach applied in translation studies has served as a new descriptive paradigm since the publication of the seminal paper by Baker (1993) in which Baker envisaged that corpus linguistics would provide the methodology for carrying out empirical investigations in translation studies. Two main types of corpus are used in corpus-based translation studies (CTS) and corpus-based interpreting studies (CIS), namely, comparable monolingual corpora, including a minimum of two subcorpora in the same language; or bilingual parallel corpora, including source texts and their corresponding target texts aligned to each other (Bernardini & Russo, 2017).

Baker (1993) suggested that the features of translated texts can best be explained through the use of comparable corpora with non-translated texts serving as a benchmark. Since then, the corpus-based approach has been adopted to examine relative frequencies of special words, collocations and lexico-syntactic structures (Bernardini, Ferraresi, &

Miličević, 2016; Olohan & Baker, 2000), revealing the "regularities of actual behaviour" (Toury, 1995, p. 265), in other words, patterns or universals. By contrast, the corpus-driven approach has enabled researchers to observe lexical features such as type-token ratio, the proportion of functional words, sentences length (Lanstyák & Heltai, 2012; Laviosa, 1998; Xiao, 2010).

Due to practical difficulties in compiling spoken corpora, the CIS appeared a bit later than CTS, originating from the highly influential paper of Shlesinger (1998b), where she referred to CIS as an offshoot of CTS. More than a decade later, a chapter written by Setton (2011) stated that "CIS is still a cottage industry". Setton (2011, p. 34) offered an overview of a number of corpus-based interpreting projects. In these projects, interpreting corpora created by researchers mostly feature authentic data with machine-readable transcripts (sometimes are tagged and indexed) covering a wide range of language combinations. Of these corpora, the European Parliament Interpreting Corpus (EPIC) is perhaps the most popular corpus in simultaneous interpreting mode in relation to European Parliament covering 23 languages (Bendazzoli, 2010), while the researchers in interpreting studies from Asia, particularly in China, has tended to draw more attention on consecutive interpreting (CI) based on the Chinese-English Interpreting Corpus of the Chinese's Premier's Annual Press Conferences (CEIPPC) (Wang & Tang, 2020).

In contrast to the general reference corpora containing hundreds of millions of words such as British National Corpus (BNC), and Corpus of Contemporary American English (COCA), the size of interpreting corpora is still quite limited. As interpreting corpora are mainly complied by individual researchers, it is unlikely to expect large interpreting corpora exceeding millions of words to emerge in the near future

(Bendazzoli, 2018). In addition, the difficulties of accessing authentic data, the significant amount of work involved in transcription and data annotation, and the lack of user-friendly and shared conventions for transcribing linguistic and paralinguistic features of orality features (Hu & Tao, 2013) all pose challenges for compiling interpreting corpora. Moreover, in annotating interpreting corpora, metadata information including the ethnographic features of the data (speaker, data, speed, mode of delivery, subject, timing, location, length), linguistic features (morphosyntactic and lexical features), paralinguistic features (disfluencies, prosody) all require great effort on the part of the compilers/researchers.

In relation to European languages at least, the parallel and comparable design of EPIC allows researchers to explore a wide range of investigations, including lexical patterns (Sandrelli, Bendazzoli, & Russo, 2010) and lexical varieties (Russo, Bendazzoli, & Sandrelli, 2006), interpreting universals (Bernardini et al., 2016), disfluencies and repairs (Bendazzoli, Sandrelli, & Russo, 2011), directionality (Monti, Bendazzoli, Sandrelli, & Russo, 2005) and many others. Furthermore, research on interpreting has also taken account of communication interactions and interpreters' strategies and norms through press conference interpreting data (Hu & Tao, 2013; B. Wang, 2012). Some research using small scale comparable corpora with PoS-annotation has investigated interpreter language, or "interpretese", by identifying specific lexical and morphosyntactic features. Future studies are needed on CIS to explore SI in cognitive, pragmatic, ethical, socio-cultural and ideological aspects of interpreting (Straniero Sergio & Falbo, 2012).

2.4.3 Interpreting as mediated spoken language

Translation has long been regarded as representing a distinct language variety that is different from non-translated language (Baker, 1993, 1996, 2004), something which has also been referred to as "third code" (Frawley, 1984), or "translationese" (Toury, 1995, p. 208). With the advent of corpus-based approaches to linguistic investigation, the distinctive features of translational texts can be then identified through the use of comparable corpora in larger-scale studies rather than intuitive finding generated from very small dataset. Baker (1993) suggests that all translations are likely to demonstrate certain linguistic features as a result of the process of translation affected by source language, which leads to the discussion of two overlapping concepts: translation universals and translationeses.

Similarly, studies of this nature have also been conducted in the field of interpreting studies, investigating the features of "interpretese". Ever since Shlesinger (1998a) pioneered Corpus-based Interpreting Studies (CIS) calling for more application of corpus approaches to verify existing intuitions regarding interpreting, a number of proposed translation universals found in translational texts are also tested in interpreting texts such as explicitation (Dayter, 2018; Gumul, 2006, 2021; Shlesinger, 1989, 1995; Tang & Li, 2017), simplification (Bernardini et al., 2016; Dayter, 2018; Kajzer-Wietrzny, 2012, 2015; Lv & Liang, 2019; Sandrelli & Bendazzoli, 2005), nominalization (Buendía, 2010; Defrancq & Collard, 2020; Kajzer-Wietrzny, 2012), levelling-out or convergence (Shlesinger, 2009), and source language interference (Dayter, 2018). The significance of CIS to explore the "psycholinguistic and pragmatic process of interpreting" was then repeatedly underlined (Setton, 2011, p. 34).

A number of lexical and grammatical features such as optional *that* (De Sutter & Lefer, 2020a; Kajzer-Wietrzny, 2018; Kruger & De Sutter, 2018; Olohan & Baker, 2000), lexical variety and density (Bernardini et al., 2016; Ferraresi, Bernardini, Petrović, & Lefer, 2018; Laviosa, 1998; Sandrelli & Bendazzoli, 2005) , part-of-speech (POS) (Dayter, 2018; Russo et al., 2006; Shlesinger, 2009), punctuation (May, 1997; Xiao & Dai, 2014), syntactic complexity measures (Liu & Afzaal, 2021a; Xu & Li, 2021) have been used to study translationese/interpretese among these proposed translation/interpreting universals. However, the use of formulaic expressions has been comparatively neglected and has only been mentioned in few studies relating to translation patterns of normalisation and levelling out (convergence), and explicitation.

Normalisation, also called "conventionalization" (Mauranen, 2007) is defined as "the tendency to conform to patterns and practices that are typical of the target language, even to the point of exaggerating them" by Baker (1996, pp. 176-177). As a result, there should be no frequency differences between translated texts and similar non-translated texts as translators tend to conform to the patterns in the target language; or even higher frequencies are detected in translated texts than in non-translated texts as translators overuse typical patterns of the target language. The normalisation hypothesis resembles Toury's law of growing standardization, which refers to the tendency to choose "more habitual options offered by a target culture (Toury, 1995, p. 268). Translators may resort to normalisation to meet the readers' linguistic expectations, using lexis or collocations that are typical in target texts (Baker, 2004; Øverås, 1998). Shlesinger (1998a) provided evidence for normalisation as interpreters completed unfinished sentences, offered grammatical rendition of ungrammatical source utterance, omitted false starts and self-corrections. Kajzer-Wietrzny (2012) explored nominalisation through an examination of trigrams and fixed expressions in the interpreted texts of two interpreters which

showed results that revealed an opposite tendency, i.e., away from normalisation. A text characterized by normalisation may make more frequent use of elements deemed standard in a given register such as lexical bundles or fixed expressions (Kajzer-Wietrzny, 2012, p. 53). Both the tendency to normalise and phraseology in general are share the common feature of conventionality, as suggested by Marco (2009). Xiao (2010) compared idioms, word clusters and reformulation markers in native Chinese and translated Chinese across different genres. His findings suggest that idioms are more common in native Chinese, while word clusters are used more frequently in translated Chinese, producing mixed evidence regarding the normalisation hypothesis. Levelling-out refers to the tendency of translational texts "to gravitate toward the center of a continuum" according to Baker (1996, p. 184), who found relatively lower level of variance in translated texts comparting to comparable non-translated texts. The term "levelling-out" was borrowed from Shlesinger (1989, p. 170), where it described the observed tendency of interpreted texts to gravitate towards the centre of the oral-literate continuum. Comparing the different conceptualization of "levelling out" in these two studies, Laviosa (2002, p. 72) opted for a different name of "convergence" to refer to "the relatively higher level of homogeneity of translated texts"; Olohan (2004, p. 100) also uses the term "convergence" to refer to the lower degree of variance in textual features in translated as compared to L1 texts. As Xiao and Dai (2014, p. 50) suggest, translated texts from the same or similar source languages may be similar to each other than to L1 language in the target language as a result of "translation language unique item under-representation and source language interference".

In the study of Xiao (2010), quantitative and qualitative differences in terms of the use of word clusters were found between native Chinese and translated Chinese texts showing that translated texts are more similar than non-translated texts in terms of the

use of word clusters, which implied that Chinese evidence supports the universal hypothesis of convergence or levelling out. As noted in Section 2.3.2 above, Henriksen (2007) has pointed out that the use of formulaic language enhances the homogeneity of the SI output as interpreters intentionally borrow formulaic expressions from each other when working in the same booth although the overall distribution of the use of such expressions are not provided.

The concepts of explicitation and implicitation were first put forward by Vinay and Darbelnet (1958/1995). They refer to explicitation as the "stylistic translation technique which consists of making explicit in the target language what remains implicit in the source language because it is apparent from either the context or the situation" (1995, p. 342). Later, Blum-Kulka (1986) formulated an explicitation hypothesis, postulating "an observed cohesive explicitness from source texts to target texts regardless of the increase traceable to differences between the two linguistic and textual systems involved" (1986, p. 19). According to Blum-Kulka, explicitation tends to be "translation-inherent" and may be regarded as a "universal strategy" in translation (1986, p. 21) rather than reflecting any particular differences between languages. According to this explicitation hypothesis, target texts tend to be more redundant than SL texts as they tend to feature an increased level of cohesive explicitness. As mentioned in 2.3.2, Li and Halverson (2020) examined the use of lexical bundles (LBs) between source texts and targeted texts in Chinese-English consecutive interpreting corpus, indicating increased levels of explicitation were realized through the use of LBs. They further identified possible causal factors as the cognitive complexity of bilingual processing and the ease of access of highly frequent chunks.

2.4.4 The influence of working direction on SI

When analysing interpreted speech, the issue of directionality – that is, whether one is interpreting from one's second/B language into one's first/A language, i.e., "into-A interpreting", or the reverse, i.,e., "into-B interpreting", also known as "retour interpreting" – has long been a contentious one, particularly when it comes to the discussion of interpreting strategies (Bartłomiejczyk, 2006; Chang & Schallert, 2007; Wu & Liao, 2018) and interpreting quality (Seleskovitch, 1999).

In Western European countries, simultaneous interpreters have traditionally interpreted only into their L1 language because they assumed that they could "express themselves naturally and idiomatically" only in their L1 language (Seeber, 2015, p. 84). Proponents of into-A interpreting claimed that into-B interpreting is "more cognitively demanding and more stressful" (Gumul, 2017, p. 312), as extra effort is required when producing interpreted renderings in the B language (Seleskovitch, 1999). On the other hand, in an Eastern European context, simultaneous interpreters have traditionally advocated for working from one's first language, claiming that listening to L1 language would allow them to fully understand the source texts, which may contribute to producing more complete and reliable interpreting output (Gumul, 2017; Lim, 2005; Mackintosh, 1999; Seleskovitch, 1999).

However, in earlier studies opinions on directionality on SI have mainly been based on "a mix of personal experience, ideology and tradition" (Gile, 2005a), while empirical studies on directionality are still scarce. The limited number of empirical studies that have been carried out on the influence of directionality on SI have reported mixed results. For example, Chang and Schallert (2007) examined the influence on

directionality in SI on the choice of strategies used by professional Chinese/English interpreters when coping with interpreting into both directions. Their results showed that interpreters who are used to working in both directions may have developed different strategies to mediate the different working modes of into-A and into-B interpreting. They also suggested that most Chinese interpreters are under higher levels of stress when interpreting into their B language (Chinese-English), and thus opted for meaning-based strategies such as generalizing, paraphrasing or omission to maintain interpreting quality. However, when interpreting into one's A language, interpreters may have more resources to process their output as they tend be more flexible in keeping with collocations and discourse patterns in the target language (Bartłomiejczyk, 2006). When interpreting into one's A language, interpreters may resort to strategies of addition, inferencing and transcoding (Bartłomiejczyk, 2006; Chang & Schallert, 2007; Tang & Li, 2017). In a study by Lin, Lv, and Liang (2018), directionality was identified as a salient factor influencing SI fluency in a Chinese and English language pair, since learner interpreters produce more fluent output when interpreting into their first language than the other way around, which seems to uncover the different demands of cognitive load of comprehension and production during SI.

Although the debates on which direction of interpreting in SI represents better practice have not reached a conclusion, there is agreement that there exist substantial differences between interpreting into the two directions (Bartłomiejczyk, 2006; Chang & Schallert, 2007; Dayter, 2020; Monti et al., 2005). The different interpreting strategies applied by interpreters, and the different level of cognitive constraints they experience, will lead to different language patterns manifested in their SI output.

The majority of existing studies on interpreting have concentrated on into-A interpreting (Kajzer-Wietrzny & Ivaska, 2020) and a large number of researchers, interpreting practitioners, and many international institutions (e.g. the UN, the EU) favour interpreting from one's second/B language into first/A language – L2-L1 interpreting (Albl-Mikasa & Tiselius, 2021; Bartłomiejczyk, 2006). However, into-B interpreting (L1-L2 interpreting) has been gradually recognized as an interpreting working mode which could ensure quality and fulfill genuine market demand (Albl-Mikasa & Tiselius, 2021; Bartłomiejczyk, 2015; Gile, 1995/2009).

In contrast, few studies so far have attempted to examine the linguistic features of into-B interpreting output. One of the few studies to do so, Dayter (2018), examined the simplification, explicitation, and normalisation features in interpreted texts of both directions, suggesting that different working directions may contribute to the contradictory results.

Discussion of directionality in SI also needs to take into account particular language pairs in relation to the features of languages involved and which language is serving as source versus target language (Bartłomiejczyk, 2006; Chang & Schallert, 2007). Previous investigation of linguistic features of interpreted texts has been confined to the European context, while the language pair of Chinese and English has rarely explored. English and Chinese are typologically distinct languages which differ from each other in terms of lexical features, syntactic rules and structures (Tsao, 1982), word order, logical form, and the encoding of thematic and case relations (Setton, 1999), and specific strategies may be required to overcome the differences between the two languages. For example, Dawrant (1996) indicated that when interpreting from Chinese to English, interpreters seemed to use strategies of waiting, linearity/segmentation to

cope with the word-order differences between the two language. The present study focuses on into-B interpreting between Chinese and English language pair.

2.5 Research gaps

It has been widely acknowledged that formulaic expressions are ubiquitous in our language production. Over the last three decades, an extensive body of research has applied a corpus-based approach to explore the use of multiple word strings in both spoken and written registers. Many of these studies have examined the use of lexical bundles (LBs), a particular type of frequency-based sequence within the broader category of formulaic expressions: LBs can be identified through specific frequency thresholds and dispersion requirements (Biber et al., 2004).

Research in applied linguistics has used the concept of LBs to compare the language development of L1-English and L2-English speakers, students and expert speakers, and speakers of different levels of L2-English language proficiency (Bychkovska & Lee, 2017; Chen & Baker, 2010, 2016; Cortes, 2004; Lu & Deng, 2019; Pan et al., 2016; Salazar, 2014; Shin, 2018; Staples et al., 2013). Scholars have also examined the discoursal functions of LBs in different registers and disciplines (Biber & Barbieri, 2007; Cao, 2021; Conrad & Biber, 2005; Durrant, 2017; Hyland, 2008b; Hyland & Jiang, 2018; Ren, 2021; Reppen & Olson, 2020). Most of these studies have followed the approach of applying structural and functional classification across corpora to generalize the features of bundles used by different groups of speakers.

Previous studies have shown that register is a salient factor influencing the use of LBs with regard to their syntactic structure and discoursal functions (Breeze, 2013; Conrad

& Biber, 2005; Hyland, 2008b; Shirazizadeh & Amirfazlian, 2021; Wang, 2017). Moreover, a number of studies support the observation that L1 speakers tend to use more lexical bundles in terms of both number and variety compared to L2 speakers (Ädel & Erman, 2012; Chen & Baker, 2010; Pan et al., 2016). There are only few bundles that are used both by L1 and L2 speakers. However, there are also studies arguing that L2 learner depend more on the use of lexical bundles (Bychkovska & Lee, 2017).

Although there has been relatively little research interest in LBs in the area of translation and interpreting studies, ever since the emergence of corpus-based translation studies, the investigation of lexical issues at the multi-word level has attracted significant interest from researchers. In terms of translated texts, many of these studies are concerned with lexical bundles or similar multi-word sequences (i.e., collocations, set phrases, idioms) with the aim of comparing the linguistic features of translated/interpreting and non-translated/interpreted language. Findings of several corpus studies (Lee, 2013; Xiao, 2010) supported the normalisation hypothesis in language pairs of English-Chinese and English-Korean in written translation; while in contrast, Kajzer-Wietrzny (2012) test of the normalisation hypothesis between interpreting subcorpora involving different language pairs came up with results that were not consistent among different language pairs. There are studies illustrating the general tendency that lexical bundles occur more in translational texts than in original texts, and that L1 speakers tend to use more and varied lexical bundles in their language production, which echoed the findings of corpus linguistic studies of lexical bundles. In addition, several scholars have identified applying LBs as a preferred technique to promote fluency, improve readability, and perform metadiscursive function when assessing translators' stylistic preferences (Granger, 2014; Liu & Afzaal, 2021b;

Shrefler, 2011).

On the other hand, scholars in interpreting studies have diverted from the frequency-based approach to its processing advantage of formulaic expressions. The facilitation of the use of "chunks" (including LBs) in simultaneous or consecutive interpreting has long been emphasized by interpreting practitioners, researchers and trainers (Aston, 2018; Ferraresi & Miličević, 2017; Henriksen, 2007; Plevoets & Defrancq, 2018a). While the usefulness and benefits of applying formulaic expressions in interpreting practice are acknowledged, the lack of extensive empirical studies indicates that there are hindrances in this research area.

Firstly, it is important to note that there is no systematic classification of units under investigation of formulaic expressions. The formulaic expressions in interpreted texts were mostly limited to a small number of formulaic sequences such as idiomatic, collocational, and metaphorical expressions due to various identification methods (Li & Halverson, 2022), while the frequency-based formulaic sequence, namely, lexical bundles has been overlooked despite its omnipresent feature in language production (Altenberg, 1998; Erman & Warren, 2000).

Secondly, previous studies reviewed focused mainly on European languages with a few very recent exceptions of corpus-based studies exploring the use of LBs in the Chinese-English pair using ST-TT descriptive data (Li & Halverson, 2020; Li & Halverson, 2022; Xu & Li, 2021). However, to the author's best knowledge, there has been no study comparing the use of such formulae between interpreted language and first language in this language pair to examine how the two groups of speakers apply LBs as building blocks to construct their spoken output.

Thirdly, the issue of directionality should also be taken into consideration. Some studies have revealed that both professional interpreters and interpreting trainees adopted different strategies when working in different directions: whether into-A or into-B interpreting (Bartłomiejczyk, 2006; Chang & Schallert, 2007; Wu & Liao, 2018). The strategies applied by interpreters may affect the use of lexical bundles to construct their interpreting output. Given that mixed results regarding the use of lexical bundles between L1 and L2 speakers, looking into the use of LBs in into-B interpreting products is also necessary when into-B interpreting has now widely recognized as an interpreting working mode especially in China.

In summary, the present research aims to fill the afore-mentioned research gaps by examining the use of lexical bundles of L2 interpreting between a typologically distinct language pair of Chinese and English through a frequency-based approach. It is hoped to improve the awareness and knowledge of how L2 interpreters apply LBs to facilitate their interpreting output

Chapter 3 Corpus and Methodology

A relatively large-scale corpus is the prerequisite for investigating the use of lexical bundles, a type of frequency-based formulaic sequence. In this chapter, a new comparable and parallel corpus named the United Nations Security Council comparable-parallel (UNSCCP) corpus is introduced to compare the use of lexical bundles (LBs) between English L1speakers and L2 interpreters and investigate the LBs' relationships to the source texts. Both comparable and parallel corpora will be adopted, using both corpus-driven and corpus-based approaches (see Section 2.4.2 above) to investigate LBs both at the macro and micro level and by carrying out quantitative and qualitative analysis.

3.1 Triangulation in corpus-based translation/interpreting studies

The current study examines interpreted texts using a corpus-based approach, involving the combination of both comparable and parallel corpora in its research design and both quantitative and qualitative methods of analysis to provide a comprehensive understanding of the data. The notion of triangulation, a methodology borrowed from the social sciences, was initially used in process-oriented translation studies with the aim to shed light on the nature of translation process (Alves & Gonçalves, 2003; Malamatidou, 2017; Shreve & Angelone, 2010). Saldanha and O'Brien (2014, p. 39) described triangulation as "cross-checking the results of one set of data provides with results from another set of data" and triangulation in translation studies refers to a combination of data and methods. According to Han (2018, p. 159), triangulation refers to "a process where researchers seek convergence and corroboration of results from different methods. The method of triangulation can be seen as "a mix of procedures to grasp complex phenomena" by using a combination of subjects, material,

strategies, methods, purposes, perspectives and investigators (Hansen, 2010, p. 207)."

From these definitions**,** applying triangulation indicates the use two or more data acquisition methodologies in one study with the aim to increase the validity and reliability of research findings. The rationale for triangulation used in the present study is complementarity, whereby the results generated from comparable corpus will be further elaborated/clarified through the use of a parallel corpus to facilitate the understanding of the interpreting process.

3.1.1 Combination of comparable corpus and a parallel corpus

A parallel corpus can be defined as "a corpus that contains source texts and their translations" (McEnery & Xiao, 2007, p. 133). In contrast, a comparable corpus can be defined as "a corpus containing components that are collected using the same sampling frame and similar balance and representativeness" (McEnery, 2003, p. 450) and also defined as a corpus "consisting of "separate collections of texts in the same language: one comprising original texts in the language in question and the other consists of translations in that language from a given source language or languages" by Baker (1995, p. 234). These two types of corpora "offer specific uses and opportunities" for contrastive and translation studies, as they provide new insights into the differences between source texts and translations, and between L1and non-L1texts, and can be used for comparative purposes, expanding our knowledge of universal features of translated and interpreted language (Aijmer, Altenberg, & Johansson, 1996).

These two types of corpora have their own advantages and disadvantages and therefore serve different purposes. In a parallel corpus, source and target texts can be used to study how the same content in source language is conveyed in another language

(Aijmer et al., 1996). as they have “the advantage of keeping meaning and function constant across the compared languages” (Altenberg & Granger, 2002, p. 9). However, a parallel corpus is insufficient for cross-linguistic contrasts as translated/interpreted texts cannot be free from the influence of translation-specific features (often referred to as translationese/interpretese) (Baker, 1993; Defrancq, 2018; Laviosa, 1997; Marco, 2018; McEnery & R. Xiao, 2007; Shlesinger, 1998a).

The use of a comparable corpus may overcome these limitations by incorporating translated/interpreted texts and non-translated/interpreted texts in the same sample frame. In a comparable corpus, specific linguistic features of translated/interpreted texts can be identified (Baker, 1996). However, a comparable corpus is less useful to study how messages in one language are expressed in another language during translation as they focus solely on the product rather than the process of translation. In order to confirm whether the observed differences result from translation/interpreting processes, the use of a parallel corpus is “indispensable to confirm that observed differences are indeed due to the translation process (vs. unrelated variables), to add explanatory power to “black box” observations when trying to infer decision-making process (translators/interpreters strategies and procedures) from corpus data and also, crucially, to link theory/description to teaching/practice” (Bernardini, 2011, p. 12). Therefore, comparable corpora can be a useful resource in combination with parallel corpora for translation/interpreting studies (Biel, 2016; Kenning, 2010; T. McEnery & Xiao, 2007; Shlesinger, 1998a; Zanettin, 2014).

In addition, the majority of existing parallel bilingual corpora tend to be relatively small and imbalanced (Altenberg & Granger, 2002) as they usually contain a single source text and its target text. In contrast, available comparable monolingual corpora tend to be larger in size and be more representative and balanced (Malamatidou, 2017) but the

comparability remains an issue that is hard to achieve to a near-perfect level (Baker, 2004; Bernardini et al., 2016; Laviosa, 1997). In discussion on the limitations of parallel and comparable corpora, it becomes apparent that the two types of corpora may complement each other in many ways, thus providing considerable advantages when used together. It has been acknowledged by a number of researchers that triangulation of corpora provides researcher with possibilities to study the same translation phenomenon from different aspects, using data in a complementary manner (Bernardini, 2011; McEnery & Xiao, 2007; Olohan, 2004), but little has been done in actually using combined corpora in translation/interpreting studies.

3.1.2 Combination of quantitative and qualitative analysis

Typically, the use of both qualitative and quantitative methods enables the researcher to compensate for the shortcomings of relying on any one single method. The methodological integration of qualitative and quantitative methods has been extensively discussed in the context of corpus-based interpreting studies (Davitti & Pasquandrea, 2014; Gile, 2005b; Hertog, Van Gucht, & de Bontridder, 2006; Hild, 2004; Pochhacker, 2006). While quantitative methods can identify relatively objective statistical patterns based on authentic data, qualitative methods allow researchers to investigate specific linguistic features in depth in particular contexts (Kenny, 2006). The current study adopts a method known as "corpus triangulation" which refers to "the combination of quantitative and qualitative corpus analysis techniques in the study of the same corpus" (Malamatidou, 2017, p. 66).

In terms of quantitative methods, both descriptive statistics, in the form of frequencies of the identified lexical bundles, and inferential statistics, i.e. log-likelihood test, are computed on the basis of the self-built corpora. The calculation of descriptive statistics

is the first step in the quantitative analysis, and these statistics are used to summarize linguistic features, identify patterns, and make comparisons and generalizations across different groups of speakers/writers (Mellinger & Hanson, 2016). The use of inferential statistics is a step towards studying variation across the corpora, and testing the significance of the differences observed between them. In addition, inferential statistics can enhance the quality of corpus-based language analysis, add insights and rigor, and allow for more focused analyses. The analysis will then move on to qualitative methods, conducting an analysis attentive to the frequently used lexical bundles in the contexts as revealed through concordance techniques. The combination of both quantitative and qualitative methods should provide deeper insights into the use of lexical bundles on the part of English L1speakers versus L2 interpreters.

3.2 Corpus design

3.2.1 Principles

As the purpose of the current study is to compare patterns of the use of lexical bundles (LBs) in the language of L1 speaker versus L2 interpreters in naturally occurring speech, a comparable corpus of spontaneous speech is needed. Additionally, in order to triangulate whether and to what extent the use of LBs of L2 interpreters is influenced by the source texts, a parallel corpus is also compiled. As previous studies (Biber & Barbieri, 2007; Cortes, 2004; Hyland, 2008b) have suggested, the pattern of LBs used in speech varies across different genres and subject areas, and empirical studies of translation and interpreting have also detected genre-related differences between translations/interpreted texts and non-translations/interpreted texts (De Sutter & Lefer, 2020b; Defrancq, De Clerck, & De Sutter, 2015; Delaere, De Sutter, & Plevoets, 2012; Teich, 2003). Thus, for the purposes of comparison, the speech in both the L1speaker

corpus and L2 interpreter corpus should belong to similar genres. Due to the difficulties of compiling spoken corpora (Cabrera Castro, 2017; Campoy & Luzón, 2007), the current study limited the corpus material to the genre of political discourse.

A comparable-parallel corpus (UNSCCP corpus) consisting of political debates in English was designed and built for the current study. The comparable corpus included both simultaneously interpreted texts in English (interpreting into L2 language)[4] and L1 original English speech, while the parallel corpus contained source texts in Chinese and target texts in English. Taking into account the availability and comparability of data, the speeches delivered by the delegates from the United Kingdom at the United Nation Security Council (UNSC) Meeting were chosen for the L1 English sub-corpus (L1O), and the simultaneous interpreting from Chinese into English of debates at the UNSC Meeting were selected for the interpreted English sub-corpus (L2I) in the comparable corpus. The speeches delivered by the delegates from China at the United Nation Security Council Meeting were chosen for the source texts (in Chinese) (STC) and their corresponding interpreting outputs (part of interpretered texts in L2I) (L2I partial) were target texts in the parallel corpus. The United Nation Audiovisual Library provides live broadcasts of all open meeting sessions. The data of the three sub-corpora all fall within the same time span of 2018-2019, when all speaking and interpreting took place on site, eliminating the influence of videoconferences and remote interpreting adopted in 2020 since the outbreak of the COVID-19 pandemic. In addition, speeches in more than 200 meetings were included in the corpus within these two years covering a wide range of general political, social and economic issues of public interest. All these sub-corpora share the same settings, topics, procedures and time frame so that

[4] Chinese interpreters are working both into and from Chinese (Lucía Ruiz Rosendo & Diur, 2021).

comparability can be ensured to a large extent.

The transcripts of speeches in the comparable-parallel corpus (UNSCCP) derive from the United Nations Security Council Debates corpus developed by Schoenfeld, Eckhard, Patz, Meegdenburg, and Pires (2019). This dataset contains verbatim records of speeches and their interpreting outputs in other five official languages[5] in the UN given at public meetings of the UNSC between 1995 and 2020 with metadata on the date of the speech, its speaker, and its position in the sequence of speeches at the meeting.

Apart from considerations of genre, the size of each corpus is another major concern. The ideal size of corpus for exploring the use of LBs is not straightforward. As the usual normalisation formulae tend to yield unreliable results, the comparison of lexical bundles identified in corpora of different sizes can be problematic (Cortes, 2015). This observation is also supported by Pan, Reppen, and Biber (2020) who showed that differences in the number of words across sub-corpora can have a significant influence on claimed differences in bundle use across groups. Similarly, Bestgen (2020) suggested that although identifying the normalised frequency threshold might allow researchers to compare the use of lexical bundles extracted from corpora of different sizes in an unbiased way, "the simplest and most correct way" (2020, p. 286) to avoid unfairness in the comparison of LBs between different corpora is to conduct the analysis in corpora of sizes as similar as possible. In determining the optimum size, in addition to the total number of words in the corpus, the number of texts and the length of each text also needs to be taken into consideration, as the lexical bundles are "(theoretically)

[5] In the provisions made in 1946, it was stipulated that English. France, Spanish, Russian and Chinese were to be the official languages. All these six languages are used as working languages. It meant that any of the five official languages could be used to deliver speeches, which would be interpreted into other five official languages (Lucía Ruiz Rosendo & Diur, 2021).

more likely to be repeated in a single long text than in a large corpus composed of many different short texts" (Pan et al., 2020, p. 218). It should be noted that the verbatim recordings of UNSC meetings do not contain the paralinguistic features common in spoken corpora, such as false starts, repetitions, filled pauses and hesitations. Since the aim of the present study is to investigate the use of formulaic sequences, the verbatim transcripts, which do not contain paralinguistic features are suitable.

To meet these requirements, a comparable corpus was built consisting of L2 interpreted English and L1 English of similar size of 200,000 words; and a parallel corpus was built consisting of two sub-corpora of around 100,000 words of interpreted texts in English (speeches in 2019 meetings in L2I) and around their corresponding source texts in Chinese. In order to deal with possible complications from the relatively small size of the corpora, conservative standards (see Section 3.3 below) are applied when identifying lexical bundles (Bestgen, 2018).

Table 3.1 depicts the distribution of the comparable corpus and the parallel corpus. in the comparable corpus, the L2I corpus contains 199,735 words across 359 texts, whereas the L1O corpus contains 200,278 words across 278 texts. In the parallel corpus, the STC contains 154,449 Chinese characters, while the L2I (partial) contains 104,109 words across 194 texts.

Table 3.1 Comparable and parallel corpus breakdown

		words	**texts**
Comparable	**L1O**	200,266	278
	L2I	199,735	359
Parallel	**STC**	154,449	194
	L2I (partial)	104,109	194

3.2.1 A special setting: The United Nations Security Council

Lucía Ruiz Rosendo and Diur (2021) present an overview of conference interpreting at the United Nations (UN). The United Nations Security Council (UNSC) is one of the six main bodies of the UN, charging with maintaining international peace and security. The UN has stipulated English, French, Spanish, Russian, Arabic and Chinese as its six official languages, which means that since 1947 speeches can be delivered in any of these five official languages and will be simultaneously interpreted into the other four languages. As a matter of fact, the UN was one of the earliest organisations to adopt simultaneous interpreting.

As stated in Section 2.4.4., most researchers and interpreters favour interpreting from one's second or B language into one's first or A language (Bartłomiejczyk, 2006; Kajzer-Wietrzny & Ivaska, 2020). However, due to the fact that interpreters who are able to interpret from Chinese and Arabic to their L1 language are very limited. Thus, Chinese and Arabic booths are working in both directions (Lucía Ruiz Rosendo & Diur, 2021). This accounts for the motive of investigating the interpreting into L2 direction for the current study as the retour interpreting is the mainstream service provided when concerning the language of Chinese.

Since the 1970s, the majority of staff and freelance interpreters who are working at the UN have received formal training in conference interpreting schools and need to pass a Language Competitive Examination (low pass rate of less than twenty percent) or an accreditation examination (Lucia Ruiz Rosendo & Marie, 2017). Throughout their careers, their performance will be reviewed by respective chiefs of each language interpreting section (Baigorri-Jalón & Travieso-Rodríguez, 2017). Therefore, it is safe

to conclude that the interpreters working in the UN are professional and experienced and the quality of their interpreting output is assured to be of a high caliber. According to the interview with an interpreter working in the Chinese at the UN, https://www.sohu.com/a/229395792_176673, there are currently 26 interpreters (April, 2018) working in the Chinese booth, providing simultaneous interpreting in both directions.

The UNSC debates represent the everyday high-level diplomatic politics of the UN and it is commonly understood that political discourse tends to be "formulaic, institutionalized, and authoritative" (B. Wu et al., 2021, p. 2). The Security Council holds almost daily meetings at the UN headquarters, and the delegates representing China, one of the five permanent members of the Security Council, attend and make statements basically at every meeting, providing ample data for interpreting analysis. The speeches delivered in the UN tend to be characterized by typical features of written language such as "dense chunks, low-frequency vocabulary, lexical density, complex syntactical structures and a high level of internal cohesion and coherence" (Lucía Ruiz Rosendo & Diur, 2021, p. 120). Wu et al. (2021, p. 9) summarize three distinctive features of the UNSC speeches delivered by Chinese delegates. First, they are "homogeneous in terms of topic and structure". Second, the speeches are "written to be read out" rather than spontaneously delivered, imposing an increased processing load on interpreters. Thirdly, the source speeches delivered by Chinese delegates at the UNSC meetings tend to be delivered at very high speech rate, amounting to around 190 characters per minute (cpm), which is way faster than the ideal speech rate of 150-180 cpm for interpreting (Li, 2010). These aforementioned features pose great challenges for SI interpreters to working from Chinese as source language to English as target language as they are required to interpreting all the message with precision. Being

familiar with formulaic expressions may facilitate simultaneous interpreters to deal with the interpreting difficulties including high information density and fast speech (Wu et al., 2021).

Despite the scale of the data and the social and political importance of their content, the UN corpora have been a surprisingly rare object of research within translation and interpreting studies and other disciplines. “Interpreting scholars may also be reluctant to study the interpreted products of Chinese UN speeches because both the statements and the corresponding SI renditions are likely to be based on prepared written texts (Wu et al., 2021, p. 3). In some cases, interpreters in the Chinese booth will receive the speech manuscripts in advance when Chinese was first made an official language at the UN (L. Liu, 1988). However, according to a recent news report (Bao, 2020) **https://www.chinanews.com.cn/gn/2020/06-09/9207261.shtml**, the Chinese interpreter Mr. Zhou Yuqiang claimed that the interpreters at the UN are required to interpret impromptu speeches especially concerning the disputed topics such as human rights, disarmament when delegates improvising their statements. Meanwhile, interpreters in the UN are regularly confronted with speakers reading out prepared presentation in the UN meetings (interpreters may or may not have the access to the manuscript) as explained in the following section. Thus, examining SI renditions based on prepared speech can provide insight into a practice that is common in international organisations (Gile, 1995/2009; Monacelli, 2009). Additionally, investigating how the speech delivered by Chinese delegates are interpreted by UN interpreters may contribute to the knowledge of how these political discourses are mediated and how formulaic expressions such as lexical bundles are used to construct their arguments.

3.2.2 SI with texts

As the UN has developed into 193-membership from the original 51 founding Member States, there are growing number of speakers/delegates giving speeches at meeting sessions. Accordingly, the speaking time of a delegate at UN has been gradually reduced to meet the prescribed length of sessions, adhering to a stricter schedule (Baigorri-Jalón & Travieso-Rodríguez, 2017; Diur, 2015). The imaginable consequence is that the speed of speeches has increased and unscripted have generally been replaced by scripted speeches as delegates seek to fit as much information as possible into their limited time slot (Lucía Ruiz Rosendo & Diur, 2021). Therefore, the UN interpreters are delivering their SI output with text, which requires the combination of simultaneous interpreting skills and sight translation skills. According to a survey of 50 professional conference interpreters (Cammoun, Davies, Ivanov, & Naimushin, 2009), most interpreters (92%) have acknowledged the usefulness of the texts. However, it worth mentioning that the interpreters do not always have access to the written texts (Bartłomiejczyk & Stachowiak-Szymczak, 2021). Moreover, even if the texts are given to interpreters, the speakers may not follow the texts completely and produce incongruent items.

When undertaking SI with the use of written texts, interpreters are required to mediate across three different channels: the speech delivered by the speaker in the source language in the auditory channel; the source text provided in the visual channel; and the interpreting output produced by the interpreter in the target language in the auditory channel (Chmiel, Janikowski, & Lijewska, 2020). In this type of SI, even more multitasking is required, and self-monitoring is more important as interpreters need to avoid source language interference from both the auditory and visual channels on target

language output. While this kind of interpreting with the aid of written texts is common SI practice in the context of international meetings (Setton & Motta, 2007), only a small number of studies have examined interpreted texts of this kind and analysed how written texts presence may affects SI performance.

Gile (1995/2009) and Shlesinger (1995) proposed that interpreting with text may facilitate interpreters in boosting their interpreting output quality, especially when interpreting under extreme conditions such as high speech rate and difficult accents uttered by speakers. These presumptions have been supported by the empirical study of Yang, Li, and Lei (2020), who reported that the presence of written texts has significantly improved the interpreting quality of interpreting trainees under fast-speech rate condition between Chinese and English. Chmiel et al. (2020) investigated how conference interpreters cope with multi-sensory input in the task of SI with text. Their results showed that participants in the study tent to focus more on the written modality (texts of source language) than auditory modality. They also reported that "no evidence for a facilitation effect for congruent items, and identified an impeding effect of the presence of the visual text for incongruent items" (p. 37). This might be explained by the fact that "a multimodal facilitation effect during SI with text is contingent on the synchronicity of the two signals (e.g., text and speech)" while "asynchrony of signals on different channels is expected to increase cognitive load" (Seeber, 2016).

Chapter 4. Analytical Framework

The analytical framework presented in this chapter lays out a methodology for identifying and categorising four-word lexical bundles (LBs) in the corpus. This framework includes specific steps for retrieving four-word bundles from the WordSmith 8.0 corpus tool, as well as manual screening criteria. Then, LBs will be examined from two perspectives: firstly, textual features will be analysed including structural and functional analysis of the lexical bundles identified from the comparable corpus; and then the translation relationships will be examined using the parallel corpus. Thus, the analytical framework also includes the classification framework for syntactic structure and discoursal functions, which has been adapted for the current research. Additionally, this chapter explains how to compare the translation relationships of four-word LBs between source texts and target texts. A second rater is recruited to achieve inter-rater reliability. The log-likelihood test is used to determine whether the differences in the frequencies of the identified four-word bundles of different sub-corpora are statistically significant.

4.1 Identification of LBs

Lexical bundles (LBs) were defined as "a recurring sequence of three of more words (Biber et al., 1999b, p. 990) Following the definition given in Biber et al. (1999), it can be seen that LBs are defined based on the frequency of recurring sequences of orthographic word units. Frequency thresholds are usually set to identify bundles that are frequent enough to be considered representative in a corpus. LBs can consist of a different number of words ranging from two to six and even more such as *one of*, *in order to*, *I don't know if*, *in the case of the*. However, four-word bundles in English are "the most researched length" for the studies on lexical bundles (Ädel & Erman, 2012;

Bychkovska & Lee, 2017; Chen & Baker, 2016; Durrant, 2017; Li & Halverson, 2022; Pan et al., 2020; Shahriari, 2017). The current study also chose four-word bundles as the research subject for the following reasons. Firstly, in a relatively large corpus, four-word bundles are more manageable than two-word and three bundles (Cortes, 2004), and four-word bundles generally incorporate two or three word bundles (as in the case of *as a result of* which incorporate *as a result*). Secondly, four-word bundles tend to perform a clearer range of structures and functions than three-word bundles (Hyland, 2008b). Thirdly, five-word and six-word bundles are usually phrases and occur much less commonly in natural speech, and they incorporate four-word LBs in most occurrences (as in the case of *at the same time we* include *at the same time*). The analysis of four-word bundles can also be compared to the previous studies as they seem to be most often studied (Cortes, 2013; Hyland, 2012; Hyland & Jiang, 2022; Hyland & Jiang, 2018; Xia, Ai, & Pae, 2022).

As lexical bundles occur frequently and are distributed widely across different text types, frequency cut-off points need to be set to identify lexical bundles (Biber, 2010). Biber et al. (2004, p. 376) claimed that setting the cut-off point is "somewhat arbitrary" as the question of how significant is the frequency of lexical bundles is rather subjective. The frequency thresholds used in previous studies varied depending on the size of the corpus used (Lu, Kisselev, Yoon, & Amory, 2018). Given the relatively small size of the corpora used in the present study and highly formulaic nature of political discourse, a relatively high-frequency cut-off of 40 times per million words is set, which means that the raw occurrence is set at eight in the current study.

In addition, dispersion rate[6] is another concern when identifying lexical bundles to rule out idiosyncratic uses by individual speakers or authors and the local repetitions concerning the immediate topic of the discourse (Biber et al., 2004; Chen & Baker, 2016; Pan et al., 2016). The dispersion criterion is rather arbitrary, resulting in diverse practices in the literature. Three to ten texts are frequently used as a criterion for four-word bundles (Ädel & Erman, 2012; Biber & Barbieri, 2007; Chen & Baker, 2016; Cortes, 2004, 2015), but percentages are also occasionally employed (Hyland, 2008b). In the current study, the L1 English sub-corpus and L2 interpreted sub-corpus in the comparable corpus contain 278 and 359 texts respectively, it is stipulated that the lexical bundles should at least appear in eight different texts to be identified as lexical bundles.

Apart from the frequency and dispersion criteria, Simpson-Vlach and Ellis (2010) and Salazar (2014) suggested calculating a Mutual Information (MI) score to supplement the frequency-based criteria for identifying LBs, as a method based solely on frequency may filter out some "distinctively useful but lower frequency phrases whose component words are highly unlikely to occur together by chance will not make it to the top of the frequency-ordered list" (Simpson-Vlach & Ellis, 2010, p. 493). MI is "a statistical measure of the coherence of a phrase, or of the relative "stickiness" of the words which make up a phrase", one which was initially used for two-word collocations (Wood, 2015, p. 123). High MI n-grams indicate that those word strings show greater coherence, while a lower MI score means that the words may simply be co-occurring by chance.

[6] According to Pan et al. (2016, p. 63), "dispersion thresholds are important to ensure that bundles are not restricted to a few texts or authors. A dispersion requirement ensures that the bundles are typical of the entire corpus, not just a few texts".

However, Biber (2009) expressed concerns over the use of MI in identifying lexical bundles since the score does not consider the order of the words in strings, and MI is known to preferentially identify low-frequency content words. Hyland (2012) also criticized the validity of applying MI in strings of three or more words for potentially ignoring the order of words. Considering four-word LBs is chosen as the research object and a large number of bundles are derived from the corpora, the MI method is not included in the current research.

The corpus analysis software, WordSmith Tools 8.0 (Scott, 2020) was used for the automatic retrieval of the 4-word lexical bundles from the self-built corpora based on the criteria mentioned above. Two sets of raw four-word bundles retrieved from the comparable corpus including L1 speech and L2 interpreted speech were created (as shown in Appendix 1 and 2). Further refinement was then carried out on these lexical bundles (see the section below). It is believed that the scrutinized bundles (screened bundles based on exclusion criteria, see below) can genuinely reflect the frequency-related building blocks of discourse in mediated and non-mediated languages.

4.1.1 Exclusion criteria

Next, manual interventions were carried out in order to screen out meaningless sequences. Meaningless sequences, referring to bundles that are "completely devoid of any identifiable meaning" (Salazar, 2014, p. 49) such as *its role as the*, *peace and the Arab,* were deleted as these sequences can be considered noise as they do not present clear meaning and thus do not process certain functions.

Overlapping word sequences which refer to two or more 4-word bundles that are

derived from the same longer bundle also deserve manual intervention. According to Chen and Baker (2010), there are two major types of overlapping bundles. One is "complete overlap" (p. 32), which refers to two four-word bundles that are derived from a single five-word bundle. For example, the four-word bundles of *call on all parties* and *on all parties to* both occur 15 times, deriving from the five-word bundle of *call on all parties to* in L1O. The other type of overlap is "complete subsumption," which refers to "a situation where two or more four-word bundles overlap and the occurrences of one of the bundles subsume those of the other overlapping bundle(s)" (p. 32). For example, *I would like to* occurs 235 times, while *would like to thank* occur 37 times, both of which occur as a subset of the five-word bundle *I would like to thank.* To prevent inflated results, these overlapping bundles are grouped together and considered as one unit, following the practice of Chen and Baker (2010, 2016) via concordance analyses manually. This step is to minimize the excessive repetitiveness with the final list.

4.1.2 Topic-specific/context-dependent LBs

Moreover, context-dependent bundles (also known as topic-related bundles) referring to bundles incorporating proper nouns (e.g., *by the United nations, Democratic republic of the)* are also considered problematic and excluded in analysis in some studies (Ädel & Erman, 2012; Bychkovska & Lee, 2017; Chen & Baker, 2010; B. Wu et al., 2021). For example, Chen and Baker (2016) warn that context-dependent bundles are not "building blocks" (2016, p. 855) which show distinct discourse features but likely a result of the topics and/socio-geographical contexts. However, the context-dependent bundles referring to organisations (e.g., *Community of West African States*), countries (e.g., *Democratic Republic of the Congo*), or people (e.g., *United Nations High Commissioner*, *members of the council*) could be linked to the features of language use

in UN setting. When analysing the use of lexical bundles in a similar political discourse of European Union documents, Jablonkai (2010) found that context-dependent bundles referring to the European Union made up the second most prominent group of bundles, implying that this type of bundle are paramount in the discourse of international organisations. Therefore, the current study retained the context-dependent bundles in the analysis and classified them functionally in the sub-group of referential groups as shown in the coding scheme in Section 3.4.2 below.

4.2 Classification of LBs

4.2.1 Structural categories of LBs

The present study classified these four-word bundles into three broad categories in terms of structure on the basis of the framework proposed by Biber et al. (2004) introduced in Section 2.2.2 above. This four-category taxonomy includes verb phrase (VP), clausal fragments, noun phrase (NP) and preposition phrase (PP), and other (bundles which do not fit into the above mentioned four categories). This kind of structural annotation tends to be able to be applied unambiguously and was therefore performed by the author.

Table 4.1 illustrates all the categories with examples.

Table 4.1 Structural categorisation of lexical bundles

Structure	**Examples**
1. Lexical bundles that incorporate verb phrase	*I would like to;* *will be able to*
2. Lexical bundles that incorporate dependent clause fragments	*It is important that;* *I think it is;*

	take this opportunity to; *we hope that the*
3. Lexical bundles that incorporate noun phrase and prepositional phrase	*by the United Nations;* *for the people of;* *at the same time;* *political settlement of the*

4.2.2 Adapted analytical framework based on functional categories of LBs.

As introduced in Section 2.2.3, a small number of scholars have proposed functional classification (Biber et al., 2004; Cortes, 2004; Hyland, 2008b; Li & Halverson, 2022) that could be applied across different text types. The current study adopts a classification of functions based mainly on the framework of Li and Halverson (2022) which in turn is based on Biber et al. (2004). The category of context-dependent in this study was incorporated under the broad category of referential expressions.

Table 4.2 Functional categorisation of lexical bundles

Function	**Sub-category**	**Example**
Stance expressions	Desire	*I would like to* *I hope that we*
	Obligation/Directive	*We need to ensure* *Urge all parties to*
	Intention	*I will be brief* *Look forward to the*
	Ability	*will be able to*
	Action	*pay tributes to the* *play a constructive role*
	Epistemic stance	*I don't know if* *I think it is*

Discourse organisers	Introduction/transitional signals Elaboration/Clarification	*With regard to the* *This is way the* *as well as the*
Referential expressions	Identification/Focus Specification of Attributes Time reference Context-dependent	*the relationship between the* *a large number of* *the past few years* *Democratic Republic of the* *Congo community of west Africa* *report of the Secretary-General* *Stabilization Mission in Mali*

4.3 Lexical bundles analysis in ST-TT comparison

The four-word lexical bundles were retrieved and compared in terms of their structure and function in the two sub-corpora in the comparable corpus. Subsequently, the high-frequency four-word bundles identified in the interpreted sub-corpora were further investigated in the parallel corpus containing the source texts of Chinese speeches and the target texts of English interpreting output. This study aims to investigate the relationship between lexical bundles in sources texts and their corresponding target renditions. This step involves examining how the meaning and form used in the source language are transferred to the target language during the interpreting process. This serves as a method of triangulation to explore the extent of source text interference on interpreters' use of lexical bundles. In other words, the current study examine how the interpreters; use of lexical bundles is influenced by the source texts.

The current study employs Toury (2012) concept of "coupled pairs" to analyse the patterns of lexical bundles and the related strategies used by simultaneous interpreters

from the target language. According to Toury (2012, p. 117), “the units of comparative would always emerge as couple pairs of target- and source-text segment, ‘replacing’ and ‘replaced’ segments, respectively.”

When analysing the relationship between the “couple pairs”, both “form-based” and “meaning-based features (Isham, 1994, pp. 206-208; Li & Halverson, 2022) were included in the analytic framework to depict the relationships between TT lexical bundles and the corresponding Chinese STs. A form-based relationship is dependent on word-for-word interpreting, also known as “direct transmission” or “transcoding” (Dam, 2001, p. 27), while a meaning-based relationship relies on “semantic comparison and contrast” (Li & Halverson, 2022, p. 7).

In other words, two levels of source-target correspondence are discussed in the current study. Firstly, formal correspondence. This refers to the degree to which the formal features of the source language are reflected in the target language. These features include grammatical structures, vocabulary, and idiomatic expressions. A formally equivalent interpreting attempts to reproduce the source text as closely as possible in the target language.

The second level is textual correspondence. This refers to the degree to which the content and meaning of the source text are preserved in the target language. A textually equivalent interpreting attempts to convey the same message as the source text, using equivalent expressions and idioms that are appropriate in the target language.

The analysis translation relationships of target lexical bundles applied the adapted analytic framework proposed by Li and Halverson (2022) and Xu and Li (2021), which

included three general patterns of equivalence, addition and shifts.

Specifically, equivalence refers to the lexical bundles that are interpreted directly "through reliable equivalents stored in the experienced interpreter's memory" (Setton, 1999, p. 80). In the current study, equivalence refers to the lexical bundle used in TT corresponds word-for-word with a unit in the ST. The second category is addition, which is treated as a strategy "when the interpreter decides to add, by way of explanation, something the original speaker did not say because the interpreter thinks the interpreting may otherwise not be clear for the audience (Bartłomiejczyk, 2006, p. 160). In the current study, addition refer to the translation relationship that the "lexical bundle is used without corresponding to any ST unit" (Li & Halverson, 2022, p. 8). The interpreter makes the information or logical relationships in the original text explicit by adding lexical bundles in their interpreting output. Shift refers to the translation relationships that the LB used in SI is triggered by source language but they are subject to deviation at lexical/semantic/pragmatic/grammatical level from that in the source language as opposed to an equivalent translation relationship (Xu & Li, 2021, p. 590).

The analytic framework of translation relationships is shown in Table 4.3 with examples.

Table 4.3 The translation relationships in the use of LBs

Translation relationships	**Examples**
equivalence	中国将继续与国际社会一道 *China will continue to work with the international community* 为推动也门问题的解决发挥建设性作用 *play a constructive role in promoting the settlement of the Yemen issue.*

addition	我们要加快发展，ø消除族群暴力滋生的土壤。 *We must accelerate development in West Africa with a view to eliminating this breeding ground for intercommunal violence and conflicts, which in essence are a struggle for development resources.* 中方高度重视安理会改进工作方法，提高权威和效率，ø更好地履行《联合国宪章》赋予的职责。 China is very committed to improving the Council's working methods and enhancing its authority and effectiveness *so that it can* better discharge the duties mandated for it by the Charter of the United Nations
shift	中方希望马里和平协议各方巩固当前的良好势头 *We hope that the signatory parties will consolidate the current positive momentum* 中方支持任何有助于稳定利比亚局势、 推动利比亚问题政治解决进程的努力。 *China supports any and all efforts to help stabilize the situation on the ground and promote a political settlement process for the Libyan question.*

Note. "ø" means no corresponding ST

4.4 Inter-rater reliability

To achieve inter-rater reliability, an outsider examiner was recruited. The author (rater A) and rater B (a PhD scholar in linguistics) classified identified bundles respectively into different function groups based on the coding scheme introduced in Section 3.4. The two raters first coded 15% of the identified bundles in the L1 English corpus. A moderate rate (89.8%) of error identification agreement was achieved in the preliminary rounds due to certain disagreements among raters, after which each disagreement was resolved through negotiation.

4.5 Log-likelihood test

The frequency differences of the identified four-word bundles in the two corpora were tested for statistical significance using Rayson's (n.d.) log-likelihood test (Rayson & Garside, 2000). The calculation of log-likelihood is a probability statistic often used in corpus analysis (Baker, 2010; Bychkovska & Lee, 2017; Chen & Baker, 2016; De Cock & Granger, 2021; J. Ebeling & Ebeling, 2018; Salazar, 2014), which can be used to compare differences in occurrence in the two corpora and is therefore useful for examining the language use of the two populations under investigation. The calculation produces a value for log-likelihood (LL). The token frequencies for each structural and functional category and subcategory in the two sub-corpora of the comparable corpus were computed using the UCREL log-likelihood calculator (http://ucrel.lancs.ac.uk/llwizard.html). A High LL indicates a significant difference between the relative frequencies of a lexical bundle based on the sizes of the two corpora. Specifically, A LL score of 3.84 or greater is significant at the $p < .05$ level; a score of 6.63 or greater is significant at $p < .01$; a score of 10.83 or greater is significant at $p < .001$; and a score of 15.13 or greater is significant at $p < .0001$. The results of the log-likelihood tests can be used to identify statistically significant cases of overuse and underuse of bundles in the two corpora (Salazar, 2014).

Chapter 5 Distribution Patterns of LBs Used in L2I and LIO

This chapter describes the results of the study and discusses them with reference to the first two general research questions and the relevant literature. The overriding objectives are (1) to compare the use of LBs between L2 simultaneous interpreting (SI) and L1 original speech in terms of general distribution, structural and functional use and (2) to explore how the two groups of speakers use these formulaic expressions to construct their language outputs. The chapter begins by describing the extracted lexical bundles based on the selection and exclusion criteria. Then, to address RQ 1.1, it presents the frequency of identified four-word bundles in L2I and L1O. Next, the chapter examines the convergent and divergent use of LBs between the L2I and L1O, with a specific comparison of the top 50 most frequently used LBs. Then, to address RQ 1.3 and RQ 1.4, the shared LBs used by bot sub-corpora are investigated, with an exploration of the patterns of overuse or underuse of the shared LBs. The chapter goes on to describe the structural distribution of the two sub-corpora and discusses the results in relation to RQ 2.1. Finally, this chapter compares the discoursal functions of the identified bundles in L2I and L1O and discusses the results, addressing RQ 2.2.

5.1 Overall distribution of extracted lexical bundles in L2I and L1O

According to the selection criteria discussed in Section 3.3, the frequency threshold was set at 40 times per million words (raw occurrence of eight) and the dispersion rate was set at eight texts. Using the corpus analysis software WordSmith Tools 8.0, this study first generated two lists of four-word bundles from the sub-corpora L2I and L1O. The L2I extracted 13,436 four-word sequences of 755 types, amounting to 27% of the total words. From the L1O sub-corpus, 4251 four-word sequences of 281 types were extracted. Then, manual interventions were performed to exclude meaningless

sequences (e.g., *nations as the main, peace and the Arab*) and overlapping word sequences, as explained in Section 3.3.1.

The overlapping word sequences refer to two or more four-word bundles that are derived from the same longer bundles. For example, the six-word sequence *United Kingdom will continue to support* derived four-word bundles of *United Kingdom will continue*, *Kingdom will continue to*, and *will continue to support*. The five-word sequence *call on all parties to* generated bundles of *call on all parties* and *on all parties to*. No seven-word or longer bundles were spotted in either sub-corpora. In addition, in the United Nations Security Council (UNSC) meetings, some frequently discussed entities have names and titles that form sequences of more than four words, such as *United Nations high commissioner for human right*s and *United Nations multidimensional integrated stabilization mission*.

To prevent inflated results, this study employed Chen and Baker's strategy of manually checking bundles via concordance analyses and combining them as appropriate (2016). After screening out overlapping bundles, 606 and 214 types of four-word sequences were identified as LBs in the two sub-corpora, with frequencies of 10,873 and 3,219 in L2I and L1O, respectively, accounting for 22% and 6% of total words in the two sub-corpora, as shown in Table 5.1 and Table 5.2.

As introduced in Section 4.2.2, the next step was to examine the controversial context-dependent LBs in the list to determine the contribution of this type of bundle to the total number of identified bundles. The context-dependent bundles mainly referred to organisations, countries, people and contexts. For example, some of the frequently used four-word bundles (or four-word bundles derived from longer bundles) referring to

countries or organisations include *Democratic republic of the (Congo)*, *the organization for the prohibition (of chemical weapons)*, and *the office for the coordination (of Humanitarian Affairs)*. As for bundles referring to people, some examples are *members of the council*, *members of the United (Nations)*, and *(special) envoy of the secretary-general*. Four-word bundles such as *international peace and security*, *for peacekeeping to peacebuilding*, and *unity and territorial integrity* were also categorised as context-dependent bundles because they are closely related to the topics of the meetings in UNSC.

There are 325 and 102 types of context-dependent bundles in L2I and L1O, with frequencies of 6,393 and 1,559, respectively. These context-dependent bundles take up a large proportion of the total identified four-word bundles in both groups (59% in L2I and 49% in L1O), which echoes the findings of Jablonkai (2010), who claimed that context-dependent bundles predominate in the discourse of international organisations.

The frequency differences of the identified four-word bundles in the two corpora were tested for statistical significance using Rayson's (n.d.) log-likelihood test (Rayson & Garside, 2000), and the algorithm is calculated using the log-likelihood and effect-size calculator on http://ucrel.lancs.ac.uk/llwizard.html. A log-likelihood score of 3.84 or greater is significant at the $p < .05$ level; a score of 6.63 or greater is significant at $p < .01$; a score of 10.83 or greater is significant at $p < .001$; and a score of 15.13 or greater is significant at $p < .0001$.

Table 5.1 Lexical Bundle Frequency Summary in L2I

L2I-Eng	Normalised frequency	% of total words

Type	755			
(-manual)	604		-20.0%	
content-dependent	325			
(-content-dependent)	279		-53.8%	
Token	13436	67269		27
(-manual)**	10847	54307	-19.3%	22
content-dependent	6393	41392		17
(-content-dependent)**	4454	22300	-58.9%	9

****$p < 0.0001$*** $p < 0.001$, ** $p < 0.01$, *$p < 0.05$

Table 5.2 Lexical Bundle Frequency Summary in L1O

	L1O-Eng	**Normalised frequency**		**% of total words**
Type	281			
(-manual)	213		-24.2%	
content-dependent	102			
(-content-dependent))	110		-48.4%	
Token	4251	21227		8
(-manual)**	3197	15964	-24.8%	6
content-dependent	1559	7785		6
(-content-dependent)**	1638	8179	-48.8%	3

****$p < 0.0001$*** $p < 0.001$, ** $p < 0.01$, *$p < 0.05$

5.1.1 Frequency description of identified four-word bundles

Among the 604 types of bundles, totaling 10,847 cases, identified (after manual screening of meaningless and overlapping bundles) in the L2I, the most frequent bundle was *as soon as possibl*e, which appeared 163 times in 122 texts. In addition, there were eight bundles that appeared over 100 times: *as soon as possible, countries of the region, the Democratic Republic of (the Congo), I would like to, peace and stability in, play a constructive role, regional and subregional organizations,* and *international peace and security.* In total, there were 29 bundles whose frequency exceeded 50 occurrences in L2I. In contrast, only 213 bundles were found in L1O, with a total frequency of 3,197, which is a little more than one third of the number in L2I. The bundle most frequently used by L1 speakers was *I would like to,* with a frequency of 235; this was the only bundle that occurred more than 100 times in the L1O. Five bundles were used more than 50 times: *I would like to, representative of the secretary, Republic of the Congo, international peace and security, would also like to*.

The data demonstrate that the four-word bundles occurred significantly more frequently ($p < 0.001$) in L2I than in L1O in terms of raw frequency, the manually screened bundles, and the list of bundles excluding context-dependent types. These results indicate that, in the political discourse of the UNSC, L2 interpreters tend to rely more heavily on lexical bundles in their speaking than L1 English speakers. As L2I interpreters are interpreting from their L1 to the L2 language, the language in L2I includes features of both L2 language production and interpreting language production. These results contradict studies on lexical bundles in second-language acquisition that suggest L1 speakers tend to use more lexical bundles than non-native speakers (Ädel & Erman, 2012; Chen & Baker, 2010; Pan, Reppen, & Biber, 2016; Salazar, 2014).

However, the results in the current study conform with findings in translation studies that show that translated texts contain more formulaic sequences than original texts (Baker, 2007; Biel, Koźbiał, & Wasilewska, 2019; Ebeling & Ebeling, 2017; Xiao, 2010). It seems that the language outputs of L2 interpreters in L2I more closely resemble translated texts than L2 language production in terms of using lexical bundles.

5.1.1.1 Effect of normalisation

The fact that L2 interpreters extensively use formulaic expressions in four-word lexical bundles (LBs) to construct their outputs may suggest that the translation pattern of normalisation (Mauranen, 2007) exists in the political discourse in the UNSC setting. Normalisation, also called 'conventionalisation' (Mauranen, 2007), is defined as "the tendency to conform to patterns and practices that are typical of the target language, even to the point of exaggerating them" (Baker (1996, pp. 176-177). Marco (2009) argued that phraseological use could be regarded as an indicator of normalisation, which is also supported by Lee (2022) and Xiao (2010) when analysing the use of four-word bundles in language pairs of English-Chinese and English-Korean in written translation.

As the occurrence of lexical bundles is regarded as an omnipresent feature in language production (Altenberg, 1998; Erman & Warren, 2000), the higher frequencies of LBs in the interpreted texts as opposed to non-interpreted texts reflect interpreters' overuse of this particular formulaic expression in the target texts. This normalisation pattern in interpreting might be seen as a strategy of interpreters to conform to the patterns in the target language. The normalisation pattern also resembles Toury's law of growing standardization, which refers to the tendency to choose "more habitual options offered

by a target culture" (Toury, 1995, p. 268). It is reasonable to postulate that interpreters in the current study may resort to normalisation to meet the readers' linguistic expectations. Resorting to fixed lexical phrases can be closely associated with "idiomaticity and fluency" (Xiao, 2010, p. 3).

As Baker (2004) claimed, if translators prefer fluency as an overall strategy when translating into English, this preference would be reflected in greater usage of recurring, familiar lexical phrases in the language. Frequent use of recognisable, fixed or semi-fixed lexical phrases conveys the impression of fluent speech in a text. The fact that L2 interpreters, in the current study, rely on a large number of LBs might be evidence of a "preferred strategy" that translators tend to adopt, according to the normalisation universal hypothesis in translation studies (Baker, 2004, p. 182).

Moreover, the normalisation hypothesis confirmed in this study may imply that interpreters are likely to employ safe, typical patterns of the target language and shy away from creative or playful uses. It might be postulated that LBs used by L2 interpreters provide them with convenient, ready-made sections of speech that convey the naturalness and fluency in their speech output that interpreters actively seek. In this sense, it is reasonable to discover that interpreters make extensive use of LBs in their interpreter output.

Apart from the possible influence of bilingual language production and interpretive use of language in interpreting, Kajzer-Wietrzny (2012) examined normalisation through the use of lexical bundles between different language pairs in interpreted texts, arguing that the results are mixed due to the different source-target language combinations. It seems that the use of LBs is affected to some extent by the interpreting language pairs. Although the current study focuses on only Chinese-English language pairs, which are

rarely studied, the use of LBs in Chinese source texts may also need to be considered, which will be discussed in the next chapter.

5.1.1.2 Effect of cognitive load of SI

The overuse of lexical bundles (LBs) by L2 interpreters might be related to interpreters' techniques for coping with heavy cognitive load during simultaneous interpreting (SI). SI is regarded as an extreme multitasking situation with an intense cognitive load. In the interpreting process, more structural information taxes limited cognitive resources. Consequently, SI interpreters might resort to fixed expressions to reduce their cognitive load quickly when dealing with plain information, due to the need to simultaneously produce and comprehend the message (Jiang & Jiang, 2022).

The Effort Model proposed by Gile (2009) modelled SI as a process involving a set of operations on "successive speech segments" (168), which consists of Listening and Analysis (Effort L), the Short-term memory (Effort M), the Speech production (Effort P), and a Coordination (Effort C). Each effort competes for limited processing resources, and certain efforts may be compromised due to the increased cognitive load of other efforts.

Formulaicity is assumed to decrease the cognitive load on interpreters (Plevoets & Defrancq, 2018b), as the occurrence of formulaic sequences is considered a sign of decreasing the process demands of language use (Conklin & Schmitt, 2012; Underwood, Schmitt, & Galpin, 2004). The use of high-frequency bundles (a sub-type of formulaic sequences) suggests, to some extent, that these sequences were both stored and used as prefabricated chunks (Biber, Conrad, & Cortes, 2004). In interpreting studies, the

processing of formulaic chunks is considered ‘a holistic process’ (Plevoets & Defrancq, 2018b, p. 6) that contributes to the automation of cognitive operations (Eyckmans, 2007; Van Rietvelde, Eyckmans, & Bauwens, 2010). In the current study, the reliance of L2 interpreters on LBs might be viewed as a way to decrease the cognitive load in language outputs while leaving more cognitive processing capability for other Effort modes.

While research into the effects of using LBs on interpreters’ performance is scarce, the facilitation of “chunks” (including LBs) in simultaneous interpreting has long been understood by practitioners, researchers and trainers (Aston, 2018; Ferraresi & Miličević, 2017; Henriksen, 2007; Plevoets & Defrancq, 2018a) to mediate the heavy cognitive load experienced by SI interpreters. Several scholars have identified LBs as a preferred technique for promoting fluency, improving readability and performing metadiscursive functions when assessing translators’ stylistic preferences (Granger, 2014; Liu & Afzaal, 2021; Shrefler, 2011).

Similarly, interpreters are trained to adopt the strategy of chunking/segmentation to save processing capacity for memory (Cheung, 2012; Donato, 2003; Gile, 2009), with the aim of consuming the least processing capacity (X. Li, 2015; Shlesinger, 2003; Wu & Liao, 2018), especially when dealing with long and complex sentences (Kader and Seubert, 2014). This linearity technique requires interpreters to divide original sentences into shorter segments and reformulate them, especially when dealing with long, complex sentences (Kader & Seubert, 2014). This interpreting process might explain why L2 interpreters use significantly more four-word LBs than L1 speakers in the same setting of UNSC.

5.1.1.3 Excessive use of context-dependent bundles

The large number of context-dependent bundles used by L2 interpreters in this study might also be related to the strategy used by professional interpreters during SI. Although the number of context-dependent bundles accounted for relatively similar proportions—59% in L2I and 49% in L1O—the raw frequencies of these bundles are quite different, amounting to 6,393 and 1,559 in L2I and L1O, respectively. The considerably larger number of context-dependent bundles used by L2 interpreters might signal a strategy of presenting familiar information while buying time to process their interpreting outputs. Aston (2018, p. 88) described what he called "topic-related terms" as ready-made expressions used by interpreters.

In addition, as pointed out by Riccardi (2005, p. 760), professional interpreters will resort to their repertoire of "ready-made" phrases to interpret "recurrent, stereotypical parts of a conference", which generally refer to welcomes, greetings, thanks and the introduction of an agenda. In addition, Wu and Liao (2018) suggested that interpreters may undergo the process of repetition, as defined by Donato (2003), or repair, as defined by Bartłomiejczyk (2006), which refer to the repetition of what has been said by using formulated terms. The tendency to use similar terms may also explain the excessive use of four-word lexical bundles by interpreters in the current study.

Moreover, the set dispersion criterion (identified four-word bundles should appear in at least eight texts) excluded only a limited number of bundles in both L2I (90 types of 886 frequency) and L1O (19 types of 180 frequency). It seems that the limit decrease of LBs, in terms of both type and token, when eliminating LBs that occur in less than eight texts, might suggest that interpreters tend to develop their own formulaic

repository, which is shared in the interpreter booth. This postulation conforms with the linguistic convergence detected in the European Parliament (EP) discourse. Defrancq (2018) argued that interpreters in the EP and the members of the EP (MEP) constitute a discourse community. Defrancq and Plevoets (2022a) and Defrancq and Plevoets (2022b) further claimed that interpreters and MEPs develop a common linguistic repertoire, based on the analysis of three- and four-word bundles.

5.1.2 The 50 most frequently used bundles (Top 50)

In human languages, Zipf's law asserts that the words with the highest frequency make up the most linguistic tokens (Baayen, 2001; Bestgen, 2020; McEnery & Gabrielatos, 2006; Zipf, 1935/1965). This law has been recognized in the analysis of natural languages, whether consisting of a few thousand words or several tens of millions. The frequency of words decreases "as a power function to their rank in the frequency list, with the most frequent word occurring approximately twice as often as the second most frequent word, which occurs twice of the fourth most frequent word, and so forth" (Ellis, 2012, p. 38; Zipf, 1935/1965). Several studies on formulaic expressions also conform to Zipf's distribution (Bestgen, 2020; Ellis, 2012; Tichý, 2021). The results of the current study suggest that the frequencies of the two sub-corpora also follow a Zipfian curve, as shown in Figure 5.1 and Figure 5.2.

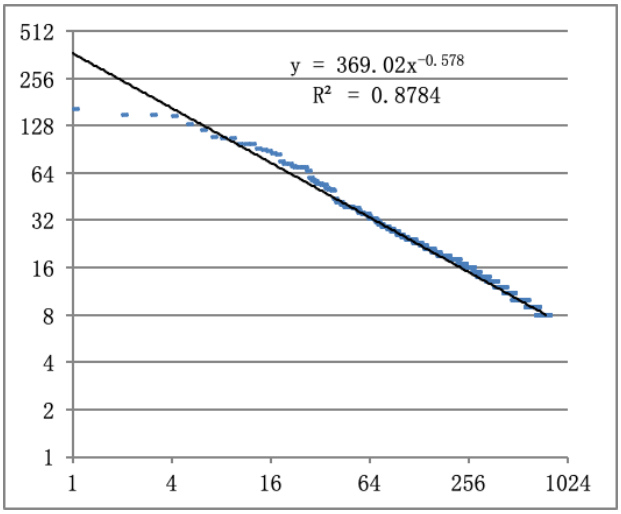

Figure 5.1 L2I number of four-word bundles frequencies (illustration of Zipf's law)

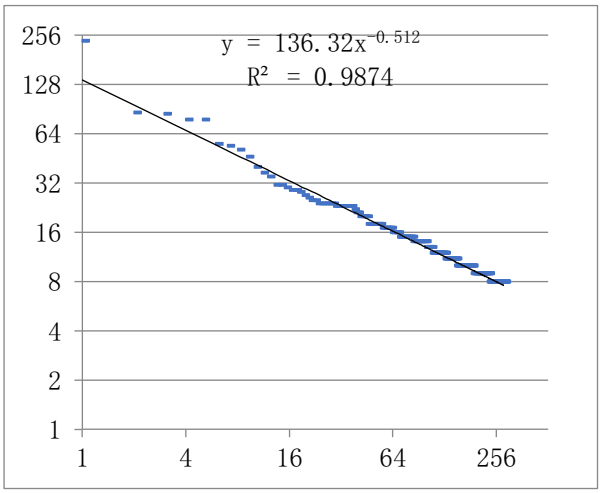

Figure 5.2 L1O number of four-word bundles frequencies (illustration of Zipf's law)

Thus, this study first focused on the 50 most frequently used bundles, as shown in Table 5.3. The frequency of the 50 most frequently used bundles in L2I is 3,361, which accounts for 31% of the total frequency, while in L1O the total frequency of the 50 most frequently used bundles is 1,524, or 48% of the total frequency. It seems that the 50 most frequently used words constitute a large proportion of the total word frequency. It is reasonable to observe the patterns of these 50 bundles in each list as representative of the overall lists for easier operation and clear demonstration.

Table 5.3 Most frequent 50 four-word bundles in the LI2 and L1O

Four-word bundles in L2I				Four-word bundles in L1O			
Bundles	**Freq.**	**per million**	**N. of Texts**	**Bundles in**	**Freq.**	**per million**	**N. of Texts**
as soon as possible	163	816	122	I would like to	235	1,173	131
countries of the region	151	756	78	**special representative of the (secretary-general)**	86	429	56
(the) Democratic Republic of the (Congo)	149	746	23	**democratic republic of the (congo)**	78	389	19
I would like to	130	651	101	**international peace and security**	55	275	44
peace and stability in	121	606	87	I would also like (to)	54	270	50
play a constructive role	108	541	99	**use of chemical weapons**	46	230	19
regional and subregional organizations	107	536	75	**United Kingdom welcomes the**	40	200	34
international peace and security	107	536	76	**members of the council**	35	175	33
united nations and the	98	491	70	with regard to the	31	155	28
with a view to	97	486	82	**United Nations and the**	31	155	29
charter of the united	91	456	72	**with the united nations**	30	150	22
international community should continue	90	451	74	take this opportunity to	29	145	28
we hope that the	88	441	78	**sexual violence in conflict**	29	145	15
at the same time	84	421	74	it is important that	28	140	25
China will continue to	76	381	68	will be able to	27	135	21
Special Representative of the Secretary-General	73	365	55	it is vital that	26	130	23
China stands ready to	73	365	73	**international humanitarian law and**	25	125	20
on the basis of	71	355	61	as well as the	25	125	23
economic and social development	70	350	55	**United Kingdom will continue**	24	120	23
unity and territorial integrity	69	345	69	situation on the ground	24	120	22

stability and development in	69	345	64	role to play in	24	120	24
purposes and principles of	60	300	48	for the people of	24	120	21
humanitarian situation in Syria	58	290	31	**(the office for the) coordination of Humanitarian Affairs**	24	120	23
political settlement of the	57	285	44	**Charter of the United (Nations)**	24	120	19
with the international community	55	275	55	**United Nations high commissioner (for human rights)**	23	115	20
ready to work with	55	275	55	**the organization for the prohibition of Chemical Weapons (OPCW)**	23	115	18
principles of the charter	54	270	46	**members of the security (council)**	21	105	19
a political solution to	54	270	44	at the same time	21	105	20
China would like to	51	255	48	**women and peace and**	20	100	17
united nations peacekeeping operations	50	250	26	**report of the secretary**	20	100	20
it is important to	50	250	38	continue to support the	20	100	20
it is necessary to	44	220	33	**by the united nations**	20	100	20
as well as the	42	210	39	**United Nations multidimensional integrated (stabilization mission)**	18	90	16
an important role in	42	210	40	per cent of the	18	90	16
through dialogue and consultation	40	200	39	on behalf of the	18	90	18
relevant Security Council resolutions	40	200	38	it is clear that	18	90	15
with the United Nations	39	195	32	**weapons of mass destruction**	17	85	11
should continue to support	39	195	37	pay tribute to the	17	85	16

peace and security in	39	195	28	**the Office of the United (Nations)**	17	85	16
minister for foreign affairs	39	195	37	**members of the United (Nations)**	17	85	17
countries in the region	39	195	27	look forward to the	17	85	15
work with the international	38	190	38	it is very good	17	85	14
relevant united nations resolutions	38	190	23	I think that we	17	85	14
china is ready to	38	190	37	I think it is	17	85	15
a political settlement of	38	190	35	very much welcome the	16	80	13
situation on the ground	36	180	28	I also want to	16	80	15
maintaining international peace and	36	180	5	as we have heard	16	80	14
Special Envoy of the (Secretary-General)	36	180	31	about the importance of	16	80	15
regional peace and stability	35	175	31	**why the United Kingdom**	15	75	14
it is imperative to	35	175	31	**United Nations assistance mission**	15	75	14
Total	3,362			Total	1,494		

First, there were 32 content-dependent bundles (in bold) in L2I and 24 content-dependent bundles in L1O in the top 50 list. This phenomenon echoed the total frequency trend in which content-dependent bundles occupy a large proportion of the total bundles identified in both sub-corpora.

Second, as shown in Table 5.3, out of the 50 most frequently used bundles, only ten are used by both L2I and L1O (highlighted in red). Of these ten, five are context-dependent bundles. This finding suggests that the language repository possessed by L2 SI interpreters and L1 speakers is different. SI interpreters are known to process their

output under great stress, attending to both listening and speaking, while L1 speakers do not experience this strain. The fact that only a small number of bundles are shared by both groups of speakers reveals different mechanisms for L1 speakers and L2 interpreters in constructing sentences. These results echo Biel et al. (2019), who found consistently few bundle overlaps between translations and non-translations. To confirm this assumption, in the next section, the log-likelihood calculation was conducted on shared bundles addressing all the identified bundles.

5.1.3 Patterns of overuse or underuse of shared LBs

In addition to the analysis of the 50 bundles most frequently used by the two groups of speakers, the author checked the shared bundles used in the two sub-corpora. As shown in Table 5.4, 106 bundles were used in L2I and LIO. Regarding L2I, shared bundles make up 14% of the total extracted bundles in type number and 21% in token number. As for L1O, the number of shared bundles comprises 38% of the total extracted bundles in type number and 50% in token number. It appears that shared bundles constitute a small portion of L2I, whereas L1 speakers use many bundles, which partially confirms the trend shown in the top 50 bundle lists. It may be plausible to conclude that the common bundles have a moderately significant influence in constructing the UNSC discourse. This conclusion is in line with data from the European Parliament corpus, where only 12% of LB types are shared in interpreted texts and 15% in L1 speeches. (Wu, 2021).

The log-likelihood ratio of the bundles identified by both L2 interpreters is calculated based on the equation developed by Rayson, Berridge, and Francis (2004); Rayson and Garside (2000) with the list of bundles used by L1 speakers as the reference. All told,

32 items show significant differences in occurrences of four-word bundles between the two sub-corpora. There are 28 types of four-word bundles (highlighted in red) in L2I that were used significantly more than those in L1O, while four bundles (highlighted in blue) used in L1O were used significantly more than those in L2I.

In addition, the log-likelihood is calculated based on the raw data lists of four-word bundles extracted using Wordsmith 8.0. Thus, the list still contained overlapping bundles. If the overlapping bundles are combined (i.e., *I would like to*, *would like to thank* can be combined into one bundle; *would also like to* and *I would also like* can be combined; *Republic of the congo* and *democratic republic of the* can be combined; *envoy of the secretary* and *special envoy of the* can be combined; *Group of five for* and *five for the sahel* can be combined), then even fewer bundles were used differently by the two groups of speakers. Of the 32 items with significant differences, the majority are content-dependent bundles with only 11 items (highlighted in bold) excepted.

The four bundles that are overused in L1O are *I would like to*, *would also like to*, *I would also lik*e and *take this opportunity to*. Three of them refer to the expression *I would (also) like to*, which was the most frequently used four-word bundle employed by L1 speakers.

It can be seen from Table 5.4 that only a few bundles were used by both L2 interpreters and L1 speakers. Additionally, based on the log-likelihood results, it is safe to conclude that the L2 interpreters tend to use different bundles from those used by L1 speakers.

Table 5.4 Shared four-word bundles in L2I and L1O with log-likelihood calculation

Word	Freq. in Corpus 1_L2I	Freq. in Corpus _ L1O	Log-likelihood	Sig.	
I would like to	130	235	30.36	**0.000**	*** -
representative of the secretary	69	86	1.82	**0.177**	-
special representative of the	73	84	0.74	**0.389**	-
Republic of the Congo	148	78	22.23	**0.000**	*** +
Democratic Republic of the	149	78	22.77	**0.000**	*** +
international peace and security	107	55	17.13	**0.000**	*** +
would also like to	16	54	21.68	**0.000**	*** -
I would also like	16	51	19.13	**0.000**	*** -
use of chemical weapons	31	46	2.90	**0.089**	-
would like to thank	66	37	8.35	**0.004**	** +
members of the council	24	35	2.03	**0.154**	-
with regard to the	31	31	0.00	**0.992**	+
United Nations and the	98	31	36.74	**0.000**	*** +
with the united nations	39	30	1.20	**0.273**	+
take this opportunity to	11	29	8.35	**0.004**	** -
as well as the	42	25	4.41	**0.036**	* +
situation on the ground	36	24	2.45	**0.118**	+
charter of the United	91	24	41.79	**0.000**	*** +
United Nations high commissioner	13	23	2.79	**0.095**	-
prohibition of chemical weapons	12	23	3.49	**0.062**	-
organization for the prohibition	13	23	2.79	**0.095**	-
nations high commissioner for	13	23	2.79	**0.095**	-
for the prohibition of	12	23	3.49	**0.062**	-
also like to thank	12	22	2.96	**0.085**	-
members of the security	20	21	0.02	**0.883**	-
at the same time	84	21	40.64	**0.000**	*** +
report of the secretary	14	20	1.05	**0.306**	-
continue to support the	37	20	5.19	**0.023**	* +
by the united nations	18	20	0.10	**0.752**	-
United Nations Multidimensional integrated	18	18	0.00	**0.994**	+
nations multidimensional integrated stabilization	18	18	0.00	**0.994**	+

multidimensional integrated stabilization mission	18	18	0.00	**0.994**	+
integrated stabilization mission in	18	18	0.00	**0.994**	+
weapons of mass destruction	17	17	0.00	**0.994**	+
office of the United	10	17	1.82	**0.178**	-
United Nations assistance mission	18	15	0.28	**0.596**	+
support the United Nations	19	15	0.48	**0.487**	+
parties to the conflict	19	15	0.48	**0.487**	+
on all parties to	15	15	0.00	**0.994**	+
it is important to	50	15	19.98	**0.000**	*** +
for the United Nations	10	15	0.99	**0.319**	-
established pursuant to resolution	24	15	2.12	**0.145**	+
deeply concerned about the	10	15	0.99	**0.319**	-
call on all parties	15	15	0.00	**0.994**	+
United Nations hybrid operation	9	14	1.08	**0.298**	-
stabilization mission in Mali	12	14	0.15	**0.700**	-
peace and security in	39	14	12.34	**0.000**	*** +
nations hybrid operation in	9	14	1.08	**0.298**	-
implementation of the peace	17	14	0.30	**0.585**	+
hybrid operation in Darfur	9	14	1.08	**0.298**	-
for his briefing and	15	14	0.04	**0.847**	+
committee established pursuant to	25	14	3.17	**0.075**	+
will continue to support	19	13	1.15	**0.284**	+
peace and reconciliation in	17	13	0.55	**0.460**	+
envoy of the Secretary	33	13	9.05	**0.003**	** +
on the implementation of	12	12	0.00	**0.995**	+
implementation of the agreement	31	12	8.74	**0.003**	** +
African union and the	26	12	5.32	**0.021**	* +
would like to begin	8	11	0.47	**0.494**	-
stabilization mission in the	19	11	2.18	**0.140**	+
like to begin by	9	11	0.20	**0.659**	-
work of the united	21	10	4.02	**0.045**	* +
will continue to work	20	10	3.42	**0.064**	+
security council and the	27	10	8.16	**0.004**	** +
on the ground and	20	10	3.42	**0.064**	+
on the council's agenda	10	10	0.00	**0.995**	+
nations assistance mission in	11	10	0.05	**0.822**	+

high commissioner for refugees	12	10	0.19	**0.665**	+
group of five for	21	10	4.02	**0.045**	* +
for the implementation of	10	10	0.00	**0.995**	+
five for the Sahel	21	10	4.02	**0.045**	* +
by the security council	27	10	8.16	**0.004**	** +
all the parties to	12	10	0.19	**0.665**	+
works agency for Palestine	18	9	3.08	**0.079**	+
special envoy of the	35	9	16.48	**0.000**	*** +
relief and works agency	19	9	3.68	**0.055**	+
refugees in the near	18	9	3.08	**0.079**	+
Palestine refugees in the	18	9	3.08	**0.079**	+
other members of the	10	9	0.06	**0.814**	+
joint comprehensive plan of	10	9	0.06	**0.814**	+
is the only way	27	9	9.47	**0.002**	** +
international committee of the	10	9	0.06	**0.814**	+
for Palestine refugees in	18	9	3.08	**0.079**	+
economic community of west	17	9	2.52	**0.112**	+
comprehensive plan of action	10	9	0.06	**0.814**	+
community of west African	17	9	2.52	**0.112**	+
committee of the red	10	9	0.06	**0.814**	+
agency for Palestine refugees	18	9	3.08	**0.079**	+
West Africa and the	15	8	2.18	**0.140**	+
united nations relief and	19	8	4.64	**0.031**	* +
united nations office for	9	8	0.06	**0.804**	+
thank special representative of	16	8	2.74	**0.098**	+
state in Iraq and	18	8	3.97	**0.046**	* +
root causes of the	10	8	0.23	**0.633**	+
revitalized agreement on the	11	8	0.48	**0.487**	+
resolution of the conflict	10	8	0.23	**0.633**	+
peace and stability in	121	8	119.15	**0.000**	*** +
on the basis of	71	8	57.88	**0.000**	*** +
nations relief and works	19	8	4.64	**0.031**	* +
mission in Mali minusma	10	8	0.23	**0.633**	+
Islamic state in Iraq	18	8	3.97	**0.046**	* +
is deeply concerned about	9	8	0.06	**0.804**	+
continue to work with	23	8	7.61	**0.006**	** +
at the heart of	9	8	0.06	**0.804**	+

assistance mission in Afghanistan	9	8	0.06	**0.804**	+
agreement on the resolution	11	8	0.48	**0.487**	+

5.2 Structural distribution of identified lexical bundles

This part of the analysis focuses on the structural distribution of identified four-word bundles in L2I and L2O with the aim of further understanding the convergent and divergent uses of these formulaic expressions. As introduced in Section 3.4.1, the present study follows the structural taxonomy developed by Biber et al. (2004), classifying these bundles into three categories, including verb phrase (VP), clausal fragments, as well as noun phrase (NP) and preposition phrase (PP) fragments. It should be noted that the context-dependent bundles were excluded in this part of the analysis.

5.2.1 Structural distribution of four-word bundles in L2I and L1O corpus

Type and token frequencies and proportions of the three structural types in the two corpora are shown in Table 5.5 and Table 5.6. Specifically, in L2I, the NP and PP categories occupied the largest proportion, at 46% in terms of type and 43% in terms of token, followed by VP with 37% in terms of type and 33% in terms of token. Clausal phrases made up the smallest proportion, at 18% in terms of type and 24% in terms of token. Thus, the distribution patterns of three structural categories are the same for both token and type of identified bundles in L2I. This trend of structural distribution matches that in Wu (2021) study of interpreted texts in European Parliament discourse, in which NP/PP-based bundles accounted for 49% of the bundle types, followed by VP-based bundles for 33% and dependent clause fragments for 16%.

In L1O, by contrast, the distribution of structural classification of identified bundles showed a different pattern for type and token. Specifically, in terms of bundle types, L1 speakers used more varied types of clausal fragments, which accounted for the largest proportion at 39%, followed very closely by NP and PP categories at 35%. The remaining 29% consisted of verb phrases. In terms of tokens, however, VP constituted the largest proportion, at 37%, followed by clausal phrases at 33% and NP and PP at 29%. These differences result from the excessive use of the VP *I would like to*, with 235 occurrences, which accounts for 14% of the total frequency of all identified bundles, excluding context-dependent ones. This relatively large percentage might explain the inflation of the two sets of data. The following analysis focuses mainly on token distribution. This trend differs from the L1 English speech in the European Parliament discourse (Wu, 2021), in which the NP/PP-based bundles accounted for the largest proportion, at 64%, followed by VP-based bundles at 20% and dependent clause bundles at 15%. The difference may be explained by the literate nature of EP speeches, as most of them are read-out.

Table 5.5 Type distribution of the structures of four-word bundles in L2I and LIO

	L2I		**L1O**	
	Type	**%**	**Type**	**%**
VP	102	37	29	26
Clausal	50	18	43	39
NP/PP	127	46	38	35
Total	279	100	110	100

Table 5.6 Token distribution of the structures of four-word bundles in L2I and LIO

	L2I		**L1O**	
	Token	**%**	**Token**	**%**
VP	1471	33	612	37
Clausal	1055	24	544	33

NP/PP	1928	43	482	29
Total	4454	100	1638	100

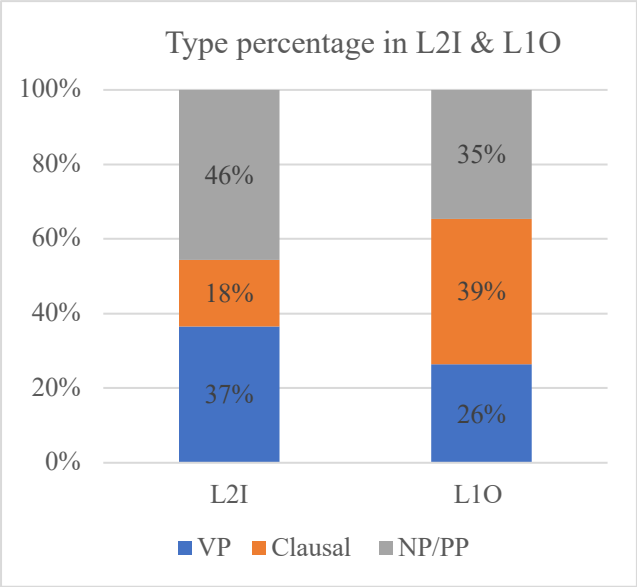

Figure 5.3 Distribution of type percentage in terms of syntactic structures in L2I and L1O

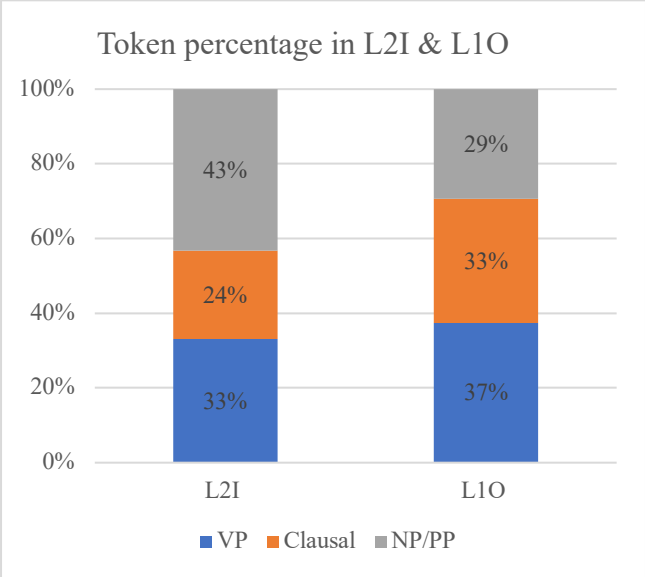

Figure 5.4 Distribution of token percentage in terms of syntactic structures in L2I and L1O

Table 5.7 shows the five most frequently used four-word bundles in the three structural categories. According to standardised frequency statistics (frequency/million words), the frequency of these four-word bundles was greater than 40 times/million words, which means they were used very commonly in the L2I sub-corpus.

Although it has been argued that lexical bundles tend to be neither idiomatic nor structurally complete (Biber & Barbieri, 2007; Biber et al., 2004), these top five four-word bundles in each structural group are important building blocks in discourse. Four-word bundles within this list of VP and clausal fragments tend to function as discourse frames for expressing new ideas by forming a kind of pragmatic 'head' for larger phrases and clauses. Bundles of NP/PP type bridge two phrases together.

5.2.1.1 Contextual use of key VP-based four-word bundles

In terms of the VP sub-group, two out of the top five bundles included the combination of pronoun + verb phrase *would (also) like (to)* (e.g., *I would like to*, *China would like to*, *I would also like*) in both L2I and L1O. There were no four-word bundles including a passive-voice verb. It might be inferred that passive voice is not commonly used by L2 interpreting and L1 speakers in the UNSC meetings. The total frequencies of the top five VP bundles are very close (356 for L2I and 353 for L2O).

Examples 1 to 9 illustrate how VP are used in the concordance in L2I and L1O. In Examples 1, 3[1], 4, 5, 6 and 9[1], the identified four-word bundles in the structure of *pronoun (I or we)/noun (China) + verb* serve as the subject and predicate in these four sentences.

In Examples 2 and 3[2], the identified four-word bundles (*work with the international* and *pay tribute to the*) are part of a to-clause acting as the objects in the sentences. The four-word bundle *will be able to* is a (verb) to-clause fragment, which is the predicate verb fragment in Example 7. In Example 8, the four-word bundle *continue to support the* consists of both predicate (the intransitive verb *continue*) and object (infinitive *to support the*).

These examples show that the four-word bundles classified as verb phrases in terms of their structure can play the roles of subject, predicate or object in these sentences. In addition, there can be more than one four-word bundle in one sentence, as shown in Examples 3 and 9.

(1) With regard to the Council's work this month, ***I would like to*** reaffirm our positions on the following three issues. (L2I)

(2) China will endeavour to ***play a constructive role*** in promoting a solution to the Palestinian question. (L2I)

(3) ***China would like to***[1] ***work with the international***[2] community, engage in close communication, display mutual confidence and mutual respect, and strengthen cooperation in pursuing continued progress towards a suitable solution to the Korean peninsula issue. (L2I)

(4) ***We call on the*** international community to step up its mediation efforts in order to impel the Yemeni parties to cease hostilities and to resume the negotiations for a political settlement so that peace, security and stability can be restored in Yemen as soon as possible. (L2I)

(5) ***I would like to*** begin by addressing the situation in the Lake Chad basin. (L1O)

(6) While this was not discussed by the Committee, ***I would also like*** to take this opportunity to say a few words about the current situation in the Sudan with regard to the ongoing protests. (L1O)

(7) We hope that the pledging conference to be held in Geneva on 26 February ***will be able to*** give a boost to these important efforts (L1O)

(8) The United Kingdom will ***continue to support the*** sides in their efforts to achieve a settlement, and we stand ready to play our part in any future talks. (L1O)

(9) In closing, ***I would like to***[1] ***pay tribute to the***[2] ongoing efforts of the African. Union and African subregional organizations to prevent unconstitutional changes of Government. (L1O)

5.2.1.2 Contextual use of key clausal fragments-based four-word bundles

The total frequencies of clausal fragments in L2I are much higher than that in L1O. As shown in Table 5.6, the total frequency of clausal fragment in L2I in the top five bundles is also higher than that of L1O (342 in L2I and 118 in L2O). Among the top five bundles of clausal fragments, the structure of *anticipatory it + verb /adjective phrase* (*it is important/vital/clear that*, *it is important to*) accounted for four out of the five in L1O, while L2I used more *first-person pronoun + dependent clause fragments* (*we hope that the*, *China will continue to*, *China stands ready to*). In this top-five list, both L2I and L1O used one *(verb/adjective) + to-clause fragment* in the clausal group (*ready to work with* and *take this opportunity to*). No bundle in the clausal fragment group was shared between the two corpora.

The structure of *anticipatory it + verb/adjective phrase* is one of the most common patterns in academic writing (Biber, Johansson, Leech, Conrad, & Finegan, 1999). As

shown in Examples 16, 17, 18 and 19, the bundles of anticipatory-it pattern are used to introduce extraposed structures and function to express the speakers' evaluation without identifying their sources explicitly. In L2I, the top five clausal segments contain more bundles starting with personal pronoun (e.g., *we*) or nouns (e.g., *China*), as shown in Examples 10, 11 and 12.

(10) **We hope that the** Mission will continue to enhance contacts and communication with the Government and FARC and make greater efforts to enhance trust and dispel doubts. (L2I)

(11) **China will continue to** follow developments closed and support all efforts to stabilize the situation on the ground, promote the political process for settling the Libyan question and help all parties to unite in combating terrorism. (L2I)

(12) **China stands ready to** enhance its contribution to the early restoration of peace, stability and prosperity in Libya. (L2I)

(13) We are **ready to work with** all parties to engage in assessing innovations, support the improvement of the Council's working methods with concrete actions, and encourage the Council to conduct its work in a more effective, transparent, democratic and impartial manner and to play an important and constructive role in maintaining international peace and security. (L2I)

(14) ***It is important to*** increase funding and technical support to developing countries to assist them in coping with climate change and improving their ability to adapt to it. (L2I)

(15) I would also like to ***take this opportunity to*** say a few words about the current situation in the Sudan with regard to the ongoing protests. (L1O)

(16) From the perspective of my Government, and I am sure that I speak for many in the Chamber and beyond, I certainly make that commitment because ***it is***

important that we work together to find common solutions, and the youth have an important role to play. (L1O)

(17) ***It is vital that*** we follow up the success of the Berlin Conference and ensure that the international community strengthens its commitment to tackling the problems of the region. (L1O)

(18) When we consider instability in regions across the world, ***it is clear that*** meeting the aspirations of young people is crucial in defending and promoting human rights, resolving conflict and sustaining peace. (L1O)

(19) I think **it is very good** that the Security Council has been able to be united on this very important issue. (L1O)

5.2.1.3 Contextual use of key NP/PP-based four-word bundles

As with the clausal fragments, the numbers of total frequencies for the NP/PP group in L2I are much higher than that in L1O. The total frequency of the top five bundles in the NP/PP group is 457 in L2I and 110 in L2O. Specifically, three out of five bundles in this list in L1O are prepositional phrases with an embedded *of*-phrase (*for the people of*, *on behalf of the*, and *about the importance of*), as shown in Examples 26, 28 and 29, while only one bundle of this type appears in L2I (*on the basis of*) in Example 23. The *noun phrase + of* structure is one of the most common patterns of four-word bundles in academic writing (Biber et al., 1999).

In these top-five bundles in the NP/PP group, only one bundle (*at the same time*) in Examples 22 and 27 is shared in both corpora. The bundle *at the same time* is also shared in the corpora of argumentative essays written by US-based L1-English and L1-Chinese ESL undergraduate students (Bychkovska & Lee, 2017) and the corpus of

interpreted political discourse (Li & Halverson, 2020), which suggests this particular bundle is used frequently in English writing and speaking regardless of topic. The top-five lists of L2I and L1O also support the argument put forward previously that very few bundles are shared between the two corpora.

(20) The international community should continue to seek a political solution to the situation in Syria as a matter of priority, continue to count on the United Nations as the main channel for mediation and work towards the comprehensive, just and careful resolution of the situation on Syrian ***as soon as possible***. (L2I)

(21) ***With a view to*** promoting the implementation of the agreement by the Government and FARC, China also hopes that the Government and the National Liberation Army will overcome the current difficulties and reach agreement without delay on an extension of the ceasefire. (L2I)

(22) ***At the same time***, a comprehensive and balanced approach should be put in place to promote non-proliferation and the use of science and technology for peaceful purposes. (L2I)

(23) The status of Jerusalem must be decided by the parties concerned ***on the basis of*** final-status negotiations. (L2I)

(24) Secondly, ***it is important to*** commit to the peaceful resolution of issues by political and diplomatic means. (L2I)

(25) Finally, ***with regard to the*** economic situation, the Council needs to continue to protect the Libyan people from economic hardship, including by supporting the restoration of the economy and the delivery of service across the country. (L1O)

(26) Unfortunately ***for the people of*** Syria, that could not be further from the reality. (L1O)

(27) ***At the same time***, it is important to recognize that Iran has legitimate security interests in the region. (L1O)

(28) Russian is not authorized to carry out an investigation ***on behalf of the*** Security Council. (L1O)

(29) Secondly, we very much agree with the Special Envoy ***about the importance of*** the rule of law and human rights. (L1O)

Table 5.7 The top five most frequent LBs in the three structural types in L2I and L1O

	L2I			L1O		
Structure	**Four-word bundles**	**Freq.**	**Freq./ million**	**Four-word bundles**	**Freq.**	**Freq./ million**
VP	I would like to	130	651	I would like to	235	1173
	play a constructive role	108	541	I would also like	54	270
	China would like to	51	255	will be able to	27	135
	work with the international	38	190	continue to support the	20	100
	we call on the	29	145	pay tribute to the	17	85
Total Freq.		**356**			**353**	
Clausal	we hope that the	88	441	take this opportunity to	29	145
	China will continue to	76	381	it is important that	28	140
	china stands ready to	73	365	it is vital that	26	130
	ready to work with	55	275	it is clear that	18	90
	it is important to	50	250	it is very good	17	85
Total Freq.		**342**			**118**	
NP/PP	as soon as possible	163	816	with regard to the	31	155
	with a view to	97	486	for the people of	24	120
	at the same time	84	421	at the same time	21	105
	on the basis of	71	355	on behalf of the	18	90
	as well as the	42	210	about the importance of	16	80

Total Freq.	**457**	**110**

5.2.2 Discussion of results

5.2.2.1 Comparison between L2I and L1O

The results generated from both L2I and L1O contrast with the findings of Wang (2017), in which clausal phrase fragments were the most common structure in spoken academic lectures and seminars, as L2I contained the most NP/PP phrases and L1O contained the most VP in terms of token number. This difference in the findings might be explained by the different genres examined in these two studies.

As shown in Figure 5.4, the proportion of VP in total frequency was 37%, which is the largest proportion among all three categories. The prevalent use of VP in L1O echoes the finding in 4.1 that *I + Verb* collocations occurred more frequently in the top 50 four-word bundle list of L1O.

Many studies of LBs in written language across genres demonstrate that NP and PP are used most frequently (Biber & Barbieri, 2007; Conrad & Biber, 2005; Pan et al., 2016; Shirazizadeh & Amirfazlian, 2021), indicating that written and spoken language are composed of different structural bundles. Using more NP and PP means that more information is integrated into a sentence (Pan et al., 2016). Therefore, it can be inferred that academic written language contains more information than that of spoken language. In the current study, L1 speakers tended to rely more on VP than L2 interpreters in their speech output, which means that L2I seems to include more features of written language in interpreting, reflecting the formal speech style of political discourse, while L1O

remains closer to the features of spoken language. This differences between L2I and L1O might be explained by the oral-literate continuum put forward by Shlesinger (1989), who argues that simultaneous interpreters tend to interpret oral texts with more literate features (Baker, 1996).

According to Biber et al. (2004), VP-based bundles and clause fragments are found to be more common in spoken registers (classroom teaching and conversation) than in written registers (textbooks and academic essays). In the current study, as shown in Figures 4.3 and 4.4, the verb phrases account for a similar percentage in L2I (33%) and L1O (37%), whereas the percentage of clausal fragments of L2I is 24% and 33% in L1O. The percentage of VP-based bundles and clause fragments in L1O is larger than that on L2I, indicating that the L1 speech is more oral than interpreted texts.

5.2.2.2 Comparison between L2I and interpreted corpora of previous studies

Xu and Li (2021) found that NP/PP phrases were used mostly by L2 interpreters (44%) as opposed to other structure types in their LegCo corpus, which aligns with the results of the current study (46%). The LegCo corpus is also based on political discourse and included the interpreted texts of the meetings in the Chief Executive's Question and Answer Sessions. Similarly, these speeches were also simultaneously interpreted by L2 interpreters working from Chinese to English. It should be noted, however, that their source speeches were in Cantonese, one of major dialects in Chinese (Y.-S. Lee, Vakoch, & Wurm, 1996), while in the current study the source speeches are in Mandarin.

However, in the LegCo corpus, the second largest proportion were clause fragments, accounting for 33%, followed by VP at 23%; by contrast, in the current study, VP made up 31% and clausal fragments constituted 22%. It seems that L2 interpreters used more VP in the UNSC setting than that of LegCo, but the difference was around 10% in each category.

The total size of TT in the LegCo corpus was 82,922 words in seven texts, which is less than half of the size of the L2I in the current study. In addition, when they set the same cut-off point of 40 times per million, the standardised frequency (per million) of extracted four-word bundles is 13,181, which is much less than that the total standardised frequency of identified bundles (after manual screening) in the current study (54,307). These contrasting results further suggest the relatively high level of formulaicity in the interpreted texts in the UNSC setting.

Li and Halverson (2020) retrieved four-word lexical bundles based on a cut-off point of normalised frequency of 40-50 times per million, like that of the current study, from the CICPPC (the Chinese-English Interpreting Corpus of Premier Press Conference), a parallel consecutive interpreting corpus. Altogether, they identified 349 types of four-word bundles with the normalised per million frequency of 27,775, which is also much less than the number of 54,307 in the current study. Similarly, in CICPPC, the NP bundles constituted the largest proportion (56%), followed by VP bundles (30%) and clausal fragments (15%), which accords with the distribution pattern of L2I in the current study. Both CICPPC and the sub-corpus of L2I in UNSCCP contain the interpreted texts in English of formal political speeches in Chinese delivered by Chinese officials/delegates. The language style of the source texts tends to be consistent, but the CICPP is based on consecutive interpreting, while the current study is based on

simultaneous interpreting. It seems that NP/PP phrases are the most commonly used lexical bundle structures in these three corpora, all which are based on political discourse.

Table 5.8 lists the top five LBs in each type by frequency in the UNSCCP, CICPPC and LegCo. Focusing on the bundles of VP structures, it is surprising to see that *I would like to* is the mostly frequently used VP bundle in all three corpora. In the CICPPC corpus, *will be able to* and *we will be able* are likely to be combined into one five-word bundle, *we will be able to*. The four-word bundles of *are you going to* and *you are going to* in LegCo both include *you*, indicating the interactive function between the speaker and the other half of the interaction. Interestingly, both CICPPC and LegCo include question and answer sessions, but only LegCo demonstrated the interaction between the speaker and the listener using the most frequent bundles of verb phrases.

As for the clausal fragment group, no bundle was shared among the three corpora. In the NP/PP group, UNSCCP and CICPPC share one bundle, *at the same time*, and UNSCCP and LegCo share one bundle, *as soon as possible*. Both CICPPC and LegCo maintained context-dependent bundles that accounted for four out of five in the structure of NP/PP (*of the Taiwan straits*, *sides of the Taiwan*, *one country two system*, and *Belt and Road Initiative*). It might be reasonable to infer that context-dependent bundles are also commonly used in these two corpora, as in the current study. In general, with the exception of *I would like to*, very few bundles are shared among the top five frequent bundle lists across these three corpora, despite the similar political settings.

Table 5.8 The top five frequent LBs extracted from interpreted texts in the three structural types in UNSCCP, CICPPC and LegCo.

	Lexical bundles		
Structure	**UNSCCP**	**CICPPC.**	**LegCo**
VP	I would like to	I would like to	I'd/would like to
	play a constructive role	will be able to	are you going to
	china would like to	we will continue to	not be able to
	work with the international	we are going to	do not want to
	we call on the	we will be able	you are going to
Clausal	we hope that the	would like to ask	when it comes to
	china will continue to	would like to know	we are talking about
	china stands ready to	it is true that	it is not just
	ready to work with	to press ahead with	so that is why
	it is important to	as long as we	we have to look
NP/PP	as soon as possible	at the same time	one country two system
	with a view to	between the two sides	in the policy address
	at the same time	of the Taiwan straits	members of the public
	on the basis of	sides of the Taiwan	as soon as possible
	as well as the	between the two countries	Belt and Road Initiative

5.3 Functional distribution of identified lexical bundles in L2I and L1O corpus

This section analyses the discourse function distribution of the identified four-word bundles in L2I and L2O. As explained in Section 3.4.2, the present study adapted the functional taxonomy framework of Li and Halverson (2022), based on the classification of Biber et al. (2004), to classify these bundles into three broad categories: stance expressions, discourse organisers and referential expressions. It should be noted that the context-dependent bundles were excluded in this analysis.

5.3.1 Discourse function distribution of four-word bundles in L2I and L1O

The four-word bundles identified in L2I and L1O were categorised into the sub-groups of the functional framework. Biber et al. (2004) asserted that LBs can serve multiple functions depending on the context, but most bundles should have a primary function.

Therefore, each four-word bundle was assigned to its most commonly used category by examining the concordance of each bundle.

Table 5.9 shows the distribution of bundle functions in terms of type counts across the three categories. L2I used 124 different types of stance bundles, accounting for 44% of the total types, followed by 85 types of referential bundles for 30% and 70 types of discourse organising bundles for 25%. In L1O, there are 49 types in the function of discourse category, making up 49% of the total, more than in stance expressions (34 types, for 34%) and referential bundles (19 types, for 17%).

Table 5.10 presents the total frequency counts of the four-word bundles used in L2I and L1O, organised by functions. According to the log-likelihood calculation, L2 interpreting used significantly more varied four-word bundles in the three categories than the L1 speakers, which is not surprising, as the total number of lexical bundles in terms of types for L2I is already significantly higher than that of L1O.

The distribution of the frequency of identified four-word bundles shows different patterns, as depicted in Figures 4.5 and 4.6. L2I used 46% stance bundles, with a frequency of 2,046, followed by 28% discourse organising bundles (frequency of 1,225) and 27% referential bundles (frequency of 1,183). In contrast, LIO used roughly the same number of stance bundles, with a frequency of 715 at 44%, and discourse organising bundles at 43% with a frequency of 711. Only 13% of the bundles are classified as referential bundles, with a frequency of 212.

Table 5.9 Distribution of the functional categories/subcategories in the two corpora (type)

	Function	L2I		L1O	
		NO.	%	NO.	%
Stance expressions	Desire	15	5	8	7
	Obligation/Directive	17	6	1	1
	Intention	10	4	6	5
	Ability	0	0	3	3
	Action	79	28	12	11
	Epistemic stance	3	1	7	6
	Total	**124**	**44**	**37**	**34**
Discourse organisers	Introduction/transitional signals	28	10	32	29
	Elaboration/Clarification	42	15	22	20
	Total	**70**	**25**	**54**	**49**
Referential expressions	Identification/Focus	3	1	2	2
	Specification of Attributes	80	29	16	15
	Time reference	2	1	1	1
	Total	**85**	**30**	**19**	17
	Total	**279**	**100**	**110**	**100**

Table 5.10 Distribution of the functions of four-word bundles in the two corpora (token)

	Function	L2I		L1O	
		NO.	%	NO.	%
Stance expressions	Desire	431	10	352	21
	Obligation/Directive	241	5	12	1
	Intention	322	7	75	5
	Ability	0	0	45	3
	Action	1018	23	141	9
	Epistemic stance	34	1	90	5
	Total	**2046**	**46**	**715**	**44**
Discourse organisers	Introduction/transitional signals	658	15	441	27
	Elaboration/Clarification	567	13	270	16
	Total	**1225**	**28**	**711**	**43**
Referential expressions	Identification/Focus	33	1	23	1
	Specification of Attributes	964	22	181	11
	Time reference	186	4	8	0
	Total	**1183**	**27**	**212**	**13**
	Total	**4454**	**100**	**1638**	**100**

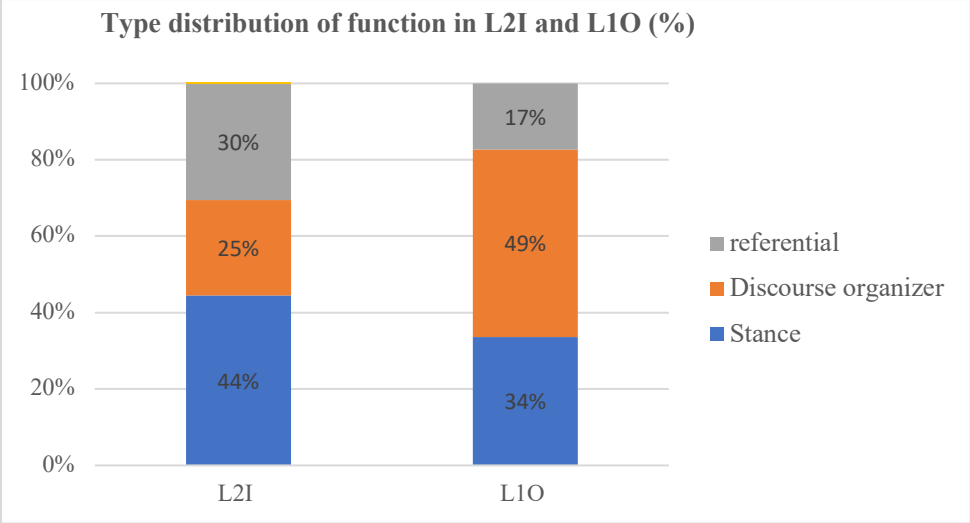

Figure 5.5 Distribution of type percentage in terms of discoursal function in L2I and L1O

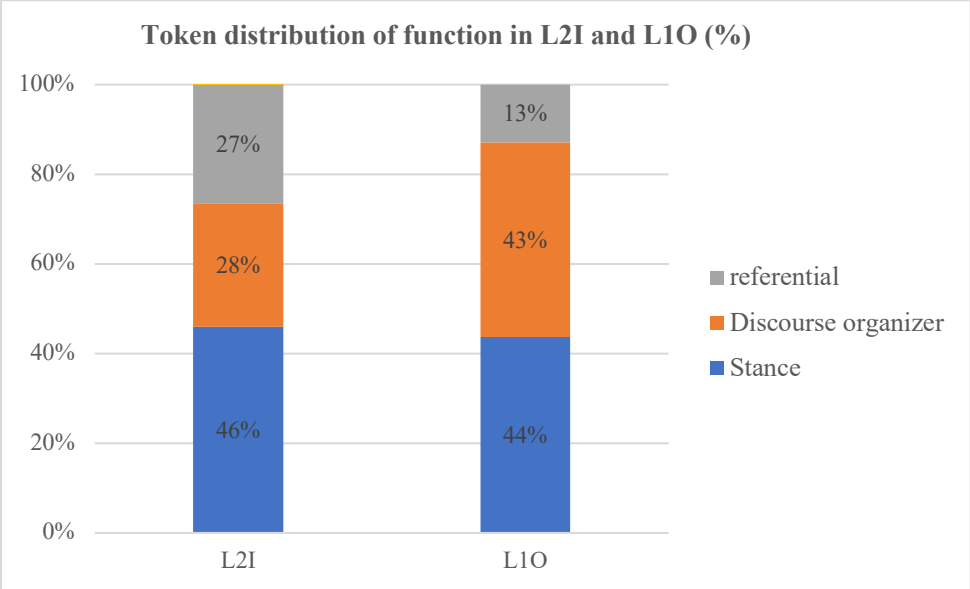

Figure 5.6 Distribution of token percentage in terms of discoursal function in L2I and L1O

5.3.1.1 Stance bundles

Regarding stance expressions, L21 used 124 types (44%) of four-word bundles and L1O used 37 types (34%). In terms of frequency count, stance expression bundles appeared 2,046 times in L2I, accounting for 46% of the total bundle frequency, making

them the most frequently used bundles among the three categories. On the other hand, in L1O, stance expression bundles appeared 715 times, constituting 44% of total frequency in L1O, making them the most frequently used bundle types in L1O.

The first subcategory under stance bundles is desire bundles, which express wishes and expectation. There are 15 types of desire bundles used by L2 interpreters, including expressions such as *I would like to, we hope that the, I wish to thank* and *China is willing to*, while eight types of desire bundles (i.e., *I would (also) like to, I also want to,* and *welcome the efforts of*) are identified in L1O. In terms of frequency counts, there are 431 desire bundles in L2I, accounting for 10% of total frequency in L2I, and 352 desire bundles in L1O, making up 21% of total frequency in L1O. Desire bundles often include personal expressions of stance, "which frame self-motivated wishes and desires" (Biber et al., 2004, p. 390). Table 5.11 lists all desire bundles in both L2I and LO. The desire bundles in the current study mostly take the structure of *first pronoun + would like/hope*, used to start an utterance. *I would like to* is the most frequently used bundle in this subcategory, both in L2I and L1O, which "provides a frame for the interpretation of the following proposition" (Biber et al., 2004, p. 389) and functions to express the speaker's attitude toward the information in the proposition (Staples, Egbert, Biber, & McClair, 2013), as shown in Examples 30 and 32. *I would like to* is also reported as a frequency bundle in spoken registers in conversation and classroom teaching (Biber et al., 2004).

A closer look of desire bundles shows similar levels of variety between L2I and L1O, as desire bundles in L2I contain words like *would like, hope, wish* and *willing,* while L1O contains *would like, want, welcome, hope,* and *should like*. This may suggest that although L2 interpreters use more types of desire bundles, their flexibility in expressing

desire is relatively similar to that of L1 speakers. The L2 interpreters, working under the heavy cognitive load of SI, tend to use similar lexical bundles repeatedly. The desire bundles seen in the current study are similar to those used in European parliament L1 speech and interpreting; those also contains bundles such as *I hope that we, I very much welcome,* and *we want to do*. (Wu, 2021, p. 65).

(30) ***I would like to*** reaffirm our positions in the following three issues. (L2I)

(31) ***We hope that the*** Mission will continue to enhance contacts and communication with the Government and FARC and make greater efforts to enhance trust and dispel doubts. (L2I)

(32) At this point, ***I would like to*** mention the sad history sexual exploitation and abuse that Haiti has witnessed and to say that we will work with the Secretary-General to ensure the implementation of the zero-tolerance policy throughout the whole of the United Nations. (L1O)

(33) We also ***welcome the efforts of*** the Chair of the new Ceasefire and Transitional Security Arrangements Monitoring Mechanism to increase the timeliness of reporting. (L1O)

Table 5.11 Stance bundles in L2I and L1O

L2I			L1O		
Desire	**Freq.**	**Texts**	**Desire**	**Freq.**	**Texts**
I would like to	130	101	I would like to	235	131
we hope that the	88	78	I would also like (to)	54	50
China would like to	51	48	I also want to	16	15
China hopes that the	33	30	welcome the efforts of	11	11
we hope that all	20	20	would therefore like to	10	10
I wish to thank	19	19	I would therefore like	9	9
I would also like to	16	16	I hope that we	9	9
hope that the parties	13	13	I should like to	8	8
we would like to	10	10			
would first like to	9	9			

China is willing to	9	9	
China hopes that all	9	9	
I would first like	8	8	
hope that the mission	8	8	
China wishes to thank	8	8	

Obligation/directive bundles are used to call for actions. There are 17 obligation/directive bundles in L2I, the second largest proportion in this category in terms of bundle type. Meanwhile, there is only one obligation/directive bundle in L2I (*urge all parties to*). Regarding tokens of obligation/directive bundles, the frequency is 241 in L2I and 12 in L1O. Most bundles in this subcategory contain the modal verbs *should* or *must,* as shown in Table 5.12, such as *should continue to support, we should continue to, we must continue to,* and *should be given to*, indicating that speakers are about to carry out certain actions (see Example 35) or directing other members to carry out the actions expressed by the speakers (see Example 34). Obligation bundles are common in both spoken and written university registers (Biber & Barbieri, 2007) and EP register (Wu, 2021), but are not supported in the present study. It is interesting to see that the second-person pronoun *you* constantly appears in obligation bundles in university teaching, directing the listener to take certain actions (Biber et al., 2004), while *you* does not appear in the UNSC register. Instead, the plural form of the first person pronoun, *we,* is frequently used. This difference could imply that speakers, who are typically classroom instructors, are more authoritative than their listeners, who are typically students, whereas the delegates of the UN Security Council, whom interpreters serve, are of equal standing.

(34) The international community ***should continue to support*** the capacity-building of the Afghan National Defense and Security Forces, enhance the country's independent self-defense capabilities and work together to effectively

address the threats posed by terrorism, transnational crime and narcotics trafficking. (L2I)

(35) ***We must continue to*** maintain our confidence and patience so as to help to facilitate the political process in South Sudan. (L2I)

(36) As always, we ***urge all parties to*** promote non-violence and to engage. constructively towards a two-State solution as the only basis for a sustainable settlement to the conflict.

Table 5.12 Obligation/directive bundles in L2I and L1O

L2I			L1O		
obligation	**Freq.**	**Texts**	**obligation**	**Freq.**	**Texts**
should continue to support	39	37	urge all parties to	12	11
we should continue to	25	19			
we must continue to	16	13			
should be given to	16	13			
council should continue to	16	15			
should fully respect the	14	14			
should work together to	13	13			
council should remain united	13	13			
community should fully respect	13	13			
should respect the sovereignty	11	11			
should continue to help	11	9			
should abide by the	11	11			
should focus on the	10	10			
should remain united and	9	9			
should strengthen coordination and	8	8			
should continue to strengthen	8	8			
should adhere to the	8	8			

The third subcategory of stance expressions is intention bundles, which express intention or future commitment. Table 5.13 lists intention bundles for both L2I and L1O. For L2I, 10 bundle types (frequency of 322) have been classified in this subcategory,

while in L1O there are six bundle types (frequency of 75). Examples 37–40 show the contextual use of some intention bundles in both corpora. In most cases, these expressions describe joint actions, with the subjects of the county (*China*, or *we*, which also refers to the country which the speaker is representing) used to announce the future actions of the country. Only one bundle in L1O (*I will be brief*) expresses the speaker's personal intention to perform some future action.

Table 5.13 Intention bundles in L2I and L1O

L2I			L1O		
Intention	**Freq.**	**Texts**	**Intention**	**Freq.**	**Texts**
China will continue to	76	68	continue to support the	20	20
China stands ready to	73	73	look forward to the	17	15
ready to work with	55	55	will continue to work	10	9
China is ready to	38	37	look forward to working	10	10
we will continue to	24	24	I will be brief	10	10
will continue to play	16	16	look forward to hearing	8	8
will continue to strengthen	15	15			
that the mission will	9	9			
will continue to take	8	8			
we look forward to	8	8			

(37) ***China will continue to*** support the work of Special Representative of the Secretary-General Chambas and UNOWAS going forward. (L2I)

(38) ***We look forward to*** seeing the next round of peace talks convened without incident so that consensus can be reached on the framework document for the political negotiations and an inclusive solution arrived at as soon as possible.

(39) ***We look forward to*** the upcoming strategic review of MUNUSMA, which will be crucial to establishing the future direction of peacekeeping support to Mali. (L1O)

(40) To that end, ***we will continue to*** work with our friends and allies to coordination an international response. (L1O)

Very few bundles have been categorised into the ability subcategory, with zero occurrences in L2I and three in L1O (*will be able to, have been able to, and has been able to*), as shown in Table 5.14. Examples 41 and 42 show that the ability bundles can indicate both future actions (*will be able to*) and past actions (*have/has been able to*).

(41) Secondly, the parties should take concrete steps to ensure that refugees ***will be able to*** return home on a voluntary basis and with access to accurate information. (L1O)

(42) They took it again, but, sadly, foreign fighters ***have been able to*** re-establish themselves there. (L1O)

Table 5.14 Ability bundles in L1O

L1O		
Ability	**Freq.**	**Texts**
will be able to	27	21
have been able to	9	9
has been able to	9	9

Table 5.15 presents action bundles in L2I, and Table 5.17 lists action bundles in L1O. In L2I, the subcategory of action contains 79 types of bundles, which take the largest proportion within stance expressions at 28% of total bundle types. Action bundles also account for the largest proportion in L2O, accounting for 11% of total bundle types with 12 types. The frequency of action bundles in L2I is 1,018, accounting for 23% of total bundle number—the most frequently used bundle function by L2 interpreters—while the frequency count is 141 in L1O, occupying 9% of total bundle numbers. This action subcategory is added to the current study to accommodate bundles that express

a concrete action and do not fit in other subcategories of stance expressions, such as *play a constructive role, work with the international, we call on the* and *listened carefully to the*. Many action bundles contain transitive verbs (such as *play, meet, resolve, find, achieve*) or intransitive verbs (such as *call on, listen to, focus on, work with*) to indicate different actions. There are also action bundles following the structure of *be + adjective* (*China is deeply concerned, we are ready to*). Present tense is common in action bundles in both corpora (*play a constructive role*, *China stands ready to*, *work with the international*), while there are also action bundles using present perfect tense (*China has always supported*, *has played an important*, *has taken note of*) and very few in simple past tense (*listened attentively to the, listened carefully to the*). Among these action bundles in both corpora, there are only two bundles—*are held to account* and *set out in the*—that used passive-voice verbs, shown in Example 46, while the rest of the action bundles employ active verbs. In Examples 43–45, the action bundles all refer to the action of the subject (*China* or *I*), serving as part of the objects in the sentences.

Table 5.15 Action bundles in L2I

L2I					
Action	**Freq.**	**Texts**		**Freq.**	**Texts**
play a constructive role	108	99	play a positive role	10	10
China stands ready to	73	73	meet each other halfway	10	10
work with the international	38	38	have played an important	10	10
we call on the	29	28	has taken note of	10	10
played an important role	25	25	has been made in	10	10
listened carefully to the	23	23	has actively participated in	10	10
resolve their differences through	20	20	give full play to	10	9
listened attentively to the	20	20	focus on the following	10	10
we call on all	19	17	find a political solution	10	9
play an active role	18	17	create conditions conducive to	10	10
China has always supported	18	17	commends the efforts of	10	10

taken note of the	17	14	China welcomes those developments	10	10
support the work of	17	17	adhere to the purposes	10	10
strengthen communication and coordination	17	16	achieve lasting peace and	10	10
promote the political process	16	16	supports the efforts of	9	9
alleviate the humanitarian situation	16	13	resolve differences through dialogue	9	9
solution is the only	15	15	promote the implementation of	9	8
address the root causes	15	14	promote a political settlement	9	9
support the efforts of	14	13	playing a constructive role	9	9
prevent violence against civilians	14	14	pay close attention to	9	9
attaches great importance to	14	13	is deeply concerned about	9	9
provide constructive assistance to	13	13	China welcomes the efforts	9	9
China is deeply concerned	13	13	China supports the efforts	9	8
respect the leadership of	12	12	China is concerned about	9	7
maintain international peace and	12	11	China has noted that	9	9
ease the humanitarian situation	12	12	China has been closely	9	9
China has always been	12	12	build a community of	9	9
abide by the purposes	12	12	been closely following the	9	9
work with all parties	11	11	bears the primary responsibility	9	9
we stand ready to	11	11	adhere to the principle	9	8
take this opportunity to	11	11	support African countries in	8	6
respect the sovereignty and	11	11	strengthen its security capacity	8	8
play an important role	11	11	respect the views of	8	7
make the following points	11	11	remain seized of the	8	8
is an important country	11	11	remain committed to the	8	8
create favourable conditions for	11	11	play a positive and	8	8
continue to contribute to	11	11	leverage the role of	8	8
China expresses its appreciation	11	10	I also thank the	8	7
China appreciates the efforts	11	11	has always been committed	8	8
achieve peace and stability	11	11	fully respect the leadership	8	8
we are ready to	10	10	emphasize the following points	8	8
takes note of the	10	9	continue to advance the	8	7
support the Syrian government	10	8	China has taken note	8	8
support the role of	10	10	China abstained in the	8	8
support the mediation efforts	10	9	cease all settlement activities	8	8
speak with one voice	10	10	advance the political settlement	8	8

respect the sovereignty of	10	10

Table 5.16 Action bundles in L1O

L1O		
Action	**Freq.**	**Texts**
pay tribute to the	17	16
very much welcome the	16	13
set out in the	15	13
deeply concerned about the	15	13
call on all parties	15	15
preventing sexual violence in	13	9
have a responsibility to	12	10
join others in thanking	9	9
I also thank the	9	9
are held to account	9	8
thank him for his	8	8
I join others in	8	8
call on the government	8	8

(43) China stands ready to continue to ***play a constructive role*** in achieving lasting peace, stability and sustainable development in South Sudan. (L2I)

(44) China would like to ***work with the international*** community, engage in close communication, display mutual confidence and mutual respect and strengthen cooperation in pursuing continued progress towards a suitable solution to the Korean peninsula issue. (L2I)

(45) In closing, I would like to ***pay tribute to the*** ongoing efforts of the African Union and African subregional organizations to prevent unconstitutional changes of Government. (L1O)

(46) It is essential that the election calendar be honoured and that key dates ***set out in the*** electoral timetable are met, including the completion of the voter register

list by 5 April, the audit of the electoral list by 25 May, convocation of the electoral process in June and the registration of presidential candidates in July.

Table 5.17 listed epistemic bundles in L2I and L1O, which are used to "convey knowledge, thoughts, beliefs, viewpoints, awareness, and facts" (Wu, 2021, p. 64). In L2I, the epistemic stance subcategory only has three bundle types with a frequency of 34, including *we believe that the*, *China believes that the*, and *has always believed that*, which all feature the word of *believe*, expressing the speaker's certainty. In L1O, there are seven type bundles with a frequency of 90 in this category, with the majority containing the verb *think* to express the speaker's uncertainty. In L2I, most epistemic bundles express the stance of the country with the subject of *we* or *China*, while most epistemic bundles in L1O express the speaker's stance. Examples 47–49 demonstrate the contextual use of epistemic bundles.

Table 5.17 Epistemic bundles in L2I and L1O

L2I			L1O		
Epistemic	**Freq.**	**Texts**	**Epistemic**	**Freq.**	**Texts**
we believe that the	13	11	I think that we	17	14
China believes that the	13	12	I think it is	17	15
has always believed that	8	8	think that it is	12	10
			I think we all	12	11
			I think that the	11	8
			believe that it is	11	11
			I am sure that	10	9

(47) ***We believe that the*** matter under consideration is extremely complex and delicate. (L2I)

(48) ***China believes that the*** parties concerned should strictly comply with their obligations under the Chemical Weapons Convention and, in line with the relevant provisions of the Convention, carry out a comprehensive, impartial and objective investigation and deal with the issues concerned within the framework of the OPCW. (L2I)

(48) ***I think that we*** accept that the Myanmar authorities are deeply concerned about development issues in Rakhine state as a whole, and the Council believes that it is one aspect that will need to be addressed. (L1O)

(49) But ultimately, ***I am sure that***, as others have done, or will do, during the course of this debate, the solution to long-term peace and stability lies not with the military but in a peace process that is Afghan-led and Afghan-owned. (L1O)

5.3.1.2 Discourse organising bundles

The second broad category is discourse organising bundles, which mainly serve the functions of introduction/focus and topic elaboration/clarification. There are 70 bundle types (25%) in L2I and 45 bundle types (49%) in L1O. In terms of frequency count, discourse organising bundles appeared 1,225 times in L2I, accounting for 28% of total frequency, and 715 times in L1O, making up 43% of total frequency. Although more types of bundles in this broad category are identified in L2I, discourse organising bundles in L1O accounted for a greater percentage in L1O due to the total number differences of identified bundles between the two corpora in the current study.

Topic introduction/focus bundles mainly serve as transitional signals to link segments of sentences or to introduce new topics. However, the boundaries between topic introduction/focus and topic elaboration/clarification bundles are sometimes blurred, as

these two subcategories include many inconsistencies between the author and the second coder during the first round of coding.

Table 5.18 presents topic introduction/focus bundles in both L2I and L1O. In L2I, there are 28 bundle types, such as *with a view to, at the same time,* and *on the issue of*. In L1O, there are 32 bundle types in this subcategory, which is the only subcategory where the number in L1O exceeded that of L2I. The frequency count of topic introduction/focus bundles was 658 times in L2I and 441 times in L1O. These results show a greater diversity in L1O than in L2I, but L2 interpreters tend to use topic introduction bundles repeatedly, which might suggest that interpreters tend to rely more on the fixed expressions of lexical bundles to connect segments within a sentence, but they use a relatively smaller repertoire of topic introduction bundles. This type of bundle serves the discourse function of identifying a new topic, as shown in Examples 50–54 and 56. As in Example 55, the bundle *on behalf of the* introduces the involved party. They are often impersonal, as they do not use first- or second-person pronouns, with few exceptions (*let me begin/start/conclude by, if we are to, as we have heard/said, as we heard from*) in L1O.

Table 5.18 Topic introduction/focus bundles in L2I and L1O

L2I			L1O		
Topic introduction	**Freq.**	**Texts**	**Topic introduction**	**Freq.**	**Texts**
with a view to	97	82	with regard to the	31	28
at the same time	84	74	take this opportunity to	29	28
it is important to	50	38	it is important that	28	25
it is necessary to	44	33	it is vital that	26	23
it is imperative to	35	31	at the same time	21	20
with regard to the	31	30	on behalf of the	18	18
on the issue of	29	25	it is clear that	18	15
is the only way	27	26	it is very good	17	14

there is a need	22	11	as we have heard	16	14
as the main channel	21	20	about the importance of	16	15
that is acceptable to	20	18	there can be no	15	14
so as to achieve	18	18	it is important to	15	14
so that they can	17	17	let me begin by	14	13
on the question of	17	14	it is very important	12	12
so that it can	16	16	it is essential that	12	11
so as to create	15	15	as a result of	12	10
it is essential to	15	13	if we are to	11	10
so as to ensure	13	12	let me conclude by	10	10
when it comes to	10	9	had to say about	10	10
as chair of the	10	10	there needs to be	9	8
so as to reach	9	9	let me start by	9	9
as one of the	9	9	it is right that	9	9
act in accordance with	9	9	is very good to	9	9
only way to resolve	8	8	is the only way	9	9
on the one hand	8	8	as we heard from	9	8
last but not least	8	8	this is the first	8	8
cooperation in order to	8	8	let me say that	8	8
as to reach a	8	8	it is crucial that	8	8
			it is also important	8	8
			but it is also	8	8
			as we have said	8	8
			as this is the	8	8

(50) ***With a view to*** promoting the implementation of the agreement by the Government and FARC, China also hopes that the Government and the National Liberation Army will overcome the current difficulties and reach agreement without delay on an extension of the ceasefire. (L2I)

(51) ***At the same time***, however, the security situation in northern and central Mali remains grim. (L2I)

(52) China calls upon the Security Council to remain united ***on the issue of*** Syria, speak with one voice and create favourable conditions for substantive progress in Syria's political process at an early date. (L2I)

(53) Finally, ***with regards to the*** economic situation, the Council needs to protect the Libyan people from economic hardship, including by supporting the restoration of the economy and the delivery of service across the country. (L1O)

(54) I think ***it is important that*** we take a moment to pay tribute to all the peacekeepers who have given their lives in service of the United Nations, and to all the brave women and men who are serving now in support of the United Nations peacekeeping operations. (L1O)

(55) ***On behalf of the*** United Kingdom, I too would like to express our gratitude to the Polish presidency for organizing today's important discussion, and to thank our briefers for their statements this morning. (L1O)

(56) ***Let me conclude by*** saying that, in seeking to improve United Nations peacekeeping, the Secretary-General has sought commitment and consensus. (L1O).

There are also bundles in this type indicating cause-and-effect relations, a reason, a purpose or an effect, such as *so as to achieve, as that they can, cooperation in order to.* These examples provide evidence to support the argument that lexical bundles are "stored as unanalysed units in the mental lexicon" (Biber et al., 2004, p. 392).

(57) We hope that the members of the Council will continue to engage in constructive consultations ***so as to achieve*** consensus at an early date. (L2I)

(58) China calls on all parties to assume a responsible attitude, commit to patiently engaging in dialogue and ***cooperation in order to*** preserve the seriousness and authority of the JCPOA, and resolve issues through dialogue and negotiation. (L2I)

(59) The facility was also seriously damaged and at least six people, including three medical staff, were injured ***as a result of*** the attack. (L1O)

(60) ***For the sake of*** the Syrian people and for preventing the future uses of chemical weapons, we call on Russia to persuade its Syrian friends to get rid of their chemical weapons and comply fully with the Chemical Weapons Convention. (L1O)

The second subcategory of discourse organising bundles are elaboration/clarification bundles. Table 5.19 shows this type of bundle in both L2I and L1O. In L2I and L1O, respectively, there are 42 bundle types and 22 bundle types identified as elaboration/clarification bundles. In terms of frequency counts, elaboration/clarification bundles appeared 567 times in L2I and 270 times in L1O. This type of bundles serves the functions of elaboration, clarification or modification. The bundle *as well as the*, a common bundle in both sub-corpora is used for explicit comparison and contrast as shown in Examples 62 and 63. However, it may be argued that *as well as the* in this sentence serves the function of introducing a new topic, but it can also be seen as the elaboration of the current topic. This is a typical example showing that the boundary between the two subcategories is not clear-cut.

Bundles of prepositional phrases (such as *on the basis of*, *for foreign affairs of*, *with the rest of*, *for the people of*) are one of the major groups of bundles in this subcategory used to modify an object.

Table 5.19 Elaboration/clarification bundles in L2I and L1O

L2I			L1O		
Elaboration/clarification	**Freq.**	**Texts**	**Elaboration/clarification**	**Freq.**	**Texts**
on the basis of	71	61	as well as the	25	23
as well as the	42	39	situation on the ground	24	22
the rest of the	23	23	for the people of	24	21
a political settlement to	20	18	on the importance of	14	12
within the framework of	18	18	on the ground in	14	14

organizations such as the	16	16	on the implementation of	12	12
for the maintenance of	16	15	by the government of	12	10
voted in favour of	15	15	as set out in	12	12
with the government of	14	14	on the situation in	11	11
with the aim of	14	12	for the council to	11	11
on the part of	14	12	vote in favor of	10	8
role in that regard	13	12	our support for the	10	9
with the purposes and	12	12	on the ground and	10	10
on the implementation of	12	11	on the council's agenda	10	10
is at the core	12	12	for the sake of	10	9
for the resumption of	12	12	for the implementation of	10	10
for the prohibition of	12	12	with the support of	9	9
with the parties concerned	11	11	full implementation of the	9	9
on the resolution of	11	11	for the government of	9	9
on the parties concerned	10	9	on the basis of	8	8
on the council's agenda	10	9	doing on the ground	8	8
on assuming the presidency	10	10	at the heart of	8	8
on an equal footing	10	10			
for the month of	10	8			
for the implementation of	10	8			
for foreign affairs and	10	10			
for convening today's meeting	10	10			
at the core of	10	10			
with the countries of	9	9			
under the leadership of	9	8			
together with the rest	9	9			
position on the issue	9	9			
developments on the ground	9	9			
at the heart of	9	9			
at a critical juncture	9	9			
with the relevant parties	8	8			
with the relevant council	8	8			
with the countries concerned	8	8			
with all the parties	8	8			
with all parties to	8	8			
with all parties in	8	8			
on the following three	8	8			

(61) The international community should continue to hep Mali to enhance its capacity for self-development and governance, ***on the basis of*** respecting Mali's independence, sovereignty and territorial integrity. (L2I)

(62) However, the country still faces multiple challenges, including difficulties in their political process, a fragile security environment and rampant terrorist activity, ***as well as the*** large-scale movements of refugees and migrants. (L2I)

(63) We acknowledge both the right of Palestinians to freedom of assembly and protest, ***as well as the*** right of Israelis to security. (L1O)

(64) The United Kingdom comments the Under-Secretary-General's efforts to start a meaningful dialogue between the United Nations and the Syrian regime in order to improve the humanitarian situation ***for the people of*** Syria. (L1O)

5.3.1.3 Referential bundles

It should be noted that context-dependent bundles are also categorised as referential bundles, but they are not included in this part of analysis. If context-dependent bundles were added, referential bundles would certainly constitute the largest proportion of both sub-corpora, which aligns with the results of studies of the EP discourse (Wu, 2021) and CIPPC corpus (Li & Halverson, 2022). All told, there are 85 types of referential bundles with a total frequency of 1,183 in L2I and 19 types of bundles with a frequency of 212 in L1O. The number of referential bundles in L2I, both in type and token, is over four times that of L1O. Referential bundles generally function to "identify an entity or single out some particular attribute of an entity as especially important" (Biber et al., 2004, p. 393). The discussion of discoursal functions of bundles in this section concerns three major subcategories included under referential bundles: identification/focus, specification of attributes and time reference. The context-

dependent bundles referring to country, organisation, people, or topic also belong to the referential bundles, but they are not discussed in this section.

Regarding identification/focus bundles, there are only three bundle types of 33 frequencies (*between the two sides, both the symptoms and, cooperation between the two*) in L2I and two bundles (*are in need of*, *is one of the*) that signal local emphasis in L1O with a frequency of 23, as shown in Table 5.20. In Example 65, the bundle *between the two sides* identifies the people of Palestine and the Arab people. As in Example 68, the bundle *is one of the* identifies part of the entity as noteworthy and is used to introduce a discussion by stating the main point first and providing details afterwards.

In Example 66, the bundle *both the symptoms and* identifies the symptoms and the root cause. Ten out of eleven occurrences of *both the symptoms and* are followed by the expression *and (the) root causes of*. It seems that L2 interpreters frequently use the expressions *both symptoms and* and *the root causes of*. Although *both the symptoms and* does not include references to country, organisation, people, or topic and thus is not classified as a context-dependent bundle, this expression is closely related to the topics discussed in the UNSC. It might be reasonable to infer that the corresponding source text is also a formulaic expression that has been used frequently by the source text speakers (the Chinese delegates), and thus interpreters are using this fixed expression to interpret the same formulaic expressions in the source text.

Table 5.20 Identification/focus bundles in L2I and L1O

L2I			L1O		
Identification/focus	**Freq.**	**Texts**	**Identification/focus**	**Freq.**	**Texts**
between the two sides	12	9	are in need of	13	13
both the symptoms and	11	11	is one of the	10	10
cooperation between the two	10	8			

(65) The Council and the international community should uphold justice by taking swift action to respond to the legitimate aspirations of the people of Palestine and the Arab people, intensifying efforts to promote peace and negotiation and facilitating the early resumption of talks ***between the two sides*** in order to achieve the two-State solution that will bring about peaceful coexistence between the two countries. (L2I)

(66) It is vital to pay equal attention to development and peace, achieve peace through development, facilitate development through peach, address ***both the symptoms and*** the root causes of conflicts and lay a solid basis for sustainable reconciliation and peace by strengthening development capacities. (L2I)

(67) Two thirds of the population ***are in need of*** humanitarian assistance in what is a man-made disaster. (L1O)

(68) This ***is one of the*** most serious challenges to the international non-proliferation regime we have ever faced. (L1O)

The second subcategory of referential bundles identifies specific attributes of nouns in most cases. In L2I, there are 80 types of bundles, accounting for 29% of the bundle type, which is the most varied bundle type used by L2 interpreters; in L1O, there are 16 bundle types, making up 15% of the total bundle types. As for frequency count, specification of attributes bundles appeared 964 times, accounting for the 22% of the total frequency, while the number in L1O is 181, making up 11% of the total frequency number. Table 5.21 lists specification of attributes bundles in L2I and L1O. Most bundles in this subcategory are incomplete in structure.

Table 5.21 Specification of attributes bundles in L2I and L1O

L2I			L1O		
Specification of attributes	**Freq.**	**Texts**	**Specification of attributes**	**Freq.**	**Texts**
situation on the ground	36	28	role to play in	24	24
efforts should be made (to)	24	22	per cent of the	18	16
situation in the country	23	23	parties to the conflict	15	14
constructive role in the	22	21	progress that has been	13	12
a shared future for	22	22	solution to the conflict	12	11
solution to the issue	21	20	important role to play	11	11
constructive role in achieving	21	21	a number of speakers	11	9
community with a shared future	19	18	a vital role in	10	10
root causes of conflict	19	17	refugees in the near	9	9
rights and interests of	19	16	comprehensive plan of action	9	9
interests of the country	19	18	causes of the conflict	9	9
stability in the country	18	15	root causes of the	8	8
concerns of all parties	18	18	participation of women in	8	8
progress has been made	16	16	efforts to ensure that	8	8
joint force of the	16	16	all of us to	8	8
a positive role in	16	16	agreement on the resolution	8	8
solution to the question	14	10			
norms governing international relations	14	14			
efforts to improve the	14	14			
by the purposes and	14	14			
interests of all parties	13	13			
future of the country	13	13			
an objective and impartial	13	13			
objective and impartial position	12	12			
main channel of mediation	12	11			
effective implementation of the	12	11			
a solution to the	12	10			
a political solution is	12	12			
solution through dialogue and	11	11			
solution that is acceptable	11	9			
settlement is the only	11	11			
responsibility for maintaining international	11	11			

positive and constructive role	11	10
objective and impartial investigation	11	9
efforts to promote peace	11	10
agreement on the resolution	11	11
acceptable to all parties	11	10
root causes of the	10	10
resolution of the conflict	10	10
proper settlement of the	10	8
offices of the united	10	8
office of the united	10	8
joint comprehensive plan of	10	8
issues through dialogue and	10	9
interference in the internal	10	9
cooperation with regional and	10	10
China's position on the	10	10
assistance to the country	10	10
a proper solution to	10	8
a positive and constructive	10	9
settlement of the question	9	8
root causes of conflicts	9	9
primary responsibility for the	9	8
parts of the country	9	9
parties in the country	9	9
measures taken by the	9	9
its support for the	9	9
efforts to find a	9	9
efforts of all parties	9	9
best of its ability	9	9
a joint effort to	9	9
settlement to the issue	8	8
settlement of hotspot issues	8	8
responsibility of the international	8	8
responsibility for the maintenance	8	8
process in the country	8	8
measures to prevent violence	8	8
legitimate concerns of all	8	8
joint efforts of the	8	8
efforts to promote the	8	8

efforts to implement the	8	8	
efforts of the international	8	8	
efforts of the government	8	8	
efforts made by the	8	8	
concerted efforts of the	8	8	
best of our ability	8	8	
an important country in	8	8	
addressing the root causes	8	8	
a large number of	8	8	
a community of a	8	8	

As shown in Examples 69–72, some specification bundles identify abstract characteristics such as *role of the united, interests of the country,* and *the principle of land*.

(69) The world's counter-terrorism efforts should respect the sovereignty of the countries affected and their ownership in the fight against terrorism, adhere to the purposes and principles of the Charter of the United Nations and make full use of the leading ***role of the United*** Nations and its Security Council. (L2I)

(70) All Libyan parties should place the overall ***interests of the country*** and the vital interests of its people above all else and jointly take active measures to promote the peace process. (L2I)

(71) The international community must remain committed to the relevant United Nations resolutions, the ***principle of land for*** peace and the Arab Peace Initiative, and must step up its efforts to see negotiations resumed so as to achieve a comprehensive, just and lasting peace through dialogue and negotiation without delay. (L2I)

(72) The Council has played ***a vital role in*** shining a spotlight on the situation and encouraging action by the authorities on the ground. (L1O) Some specification bundles are used to modify the features of the noun within the bundle.

In Example 73, for the bundle *situation in the country*, the prepositional phrase *in the country* is used to modify *situation*. In some cases, they are tangible framing bundles that contain concrete words (such as *role*) with metaphorical meaning. In Example 76, the four-word bundle of *constructive role in the* is structurally incomplete; the complete structure should be *play a constructive role in the process*. Both the adjective *constructive* and the prepositional phrase *in the process* are used to modify the noun *role*. In Example 77, the noun *causes* is modified by the premodifier of *root* and postmodifier *of conflict*.

(73). Actions taken by the Security Council on Syria's humanitarian issues should not only help ease the overall humanitarian ***situation in the country***, but also help consolidate the momentum for a ceasefire in Syria and be conducive to a bigger picture of a political settlement to the issue. (L2I)

(74) As we all agreed, there can be no purely military ***solution to the conflict.***

(75) All ***parties to the conflict*** bear a responsibility for that, but the Government alone is responsible for the bureaucratic impediments – such as the $4,000 work permit fees which continue to hinder the efforts of the relief agencies to save lives.

(76) China hopes that the Myanmar and Bangladesh will continue to take appropriate steps and that the Security Council and the international community will play a ***constructive role in the*** process. (L2I)

(77) Secondly, we should give Africa effective assistance in addressing the ***root. causes of conflict***. (L2I)

Some of the specification of attributes bundles specify quantities or amounts such as *a large number of, percent of the,* and *a number of speakers,* as shown in Examples 78–80.

(78) As a result of the efforts of the parties concerned some ceasefire agreements have been reached and ***a large number of*** civilians evacuated through the corridor. (L2I)

(79) As President Santos Calderon said in Cartagena last week, less than 10 ***per cent of the*** time set out for the peach agreement to be delivered has passed. (L1O)

(80) As ***a number of speakers*** have noted, more than half the world's population is below the age of 30. (L1O)

The third subcategory of referential bundles refers to times in the text itself, but these are used only in one bundle both in L2I (*at an early date*) and L1O (*over the past few*) with a frequency of 23 and eight, respectively, as shown in Table 5.22. Examples 81 and 82 demonstrate how they are used in the context.

Table 5.22 Time reference bundles in L2I and L1O

L2I			L1O		
Time reference	**Freq.**	**Texts**	**Time reference**	**Freq.**	**Texts**
as soon as possible	163	122	over the past few	8	8
at an early date	23	22			

(81) China is ready, together with other members of the Council, to continue to play its role in promoting the peace process in Colombia and to achieve comprehensive peace and stability in that country ***at an early date***.

(82) We have been concerned by the number of Palestinian deaths in the West Bank ***over the past few*** weeks.

5.3.2 Discussion of results

5.3.2.1 Comparison between L2I and L1O

To answer the research question, the identified four-word bundles were analysed with reference to the adapted functional framework (Biber et al., 2004) between L2I and L1O. Drawing on the analyses in the previous section, the divergent use of lexical bundles in terms of their discoursal functions between L2I and L1O has been identified. The distribution patterns of the four-word bundles used by L2 interpreters and L1 speakers in terms of their discoursal functions differ in type and token counts.

Regarding the type counts of four-word bundles, L2 interpreters used more varied bundles in the three broad functional categories than L1 speakers. Zooming in on each subcategory, the type counts of L2I outnumbered that of L1O in most of the subcategories, except for ability bundles in the stance bundle group and introduction/transitional bundles in the discourse organiser bundle group. While L2I used more varied bundle types in the stance bundles group (44%), L1O used more varied types of bundles in the discourse organiser group. It seems that L1 speakers tend to use more varied linguistic devices to establish logic links or connections between different segments of the sentences.

As for the frequency counts, the distribution of functional bundles has demonstrated a similar pattern in both L2I and L1O. Both L2 interpreters and L1 speakers resort more often to stance bundles to express their desire, obligation, intention and epistemic stance, accounting for the largest proportion in both L2I (46%) and L1O (44%), with frequencies of 2,046 and 715, respectively. The second largest proportion is discourse organising bundles, with a frequency of 1,225 in L2I (28%) and 711 (43%) in L1O.

The figure shows that L1 speakers use similar amounts of stance (715) and discourse organising (711) bundles in their spoken utterances. Examining each subcategory, four subcategory bundles were used more than 500 times in L2I. Specifically, action bundles and specification bundles were used most often, with a frequency of over 900 times, followed by introduction/focus bundles at a frequency of 658 and elaboration bundles at a frequency of 567 in L2I. L1 speakers use introduction bundles more often, with a frequency of 441, followed by desire bundles, with a frequency of 352. The elaboration bundles were used 270 times and the specification of attributes bundles were used 181 times.

In general, L2 interpreters rely more on the use of lexical bundles to construct their utterances than L1 speakers, given the significantly higher frequency in L2I than in L1O, except for ability and time reference bundles. This might be explained by the heavy cognitive load interpreters experience and their frequent use of chunking techniques. The simultaneous interpreting from the L1–L2 direction is believed to be impose a heavy cognitive load, as SI interpreters must produce a corresponding output instantaneously when listening to the source speech, with a time-lag of only a few seconds (Christoffels and De Groot, 2004; Gerver, 1976; Kroll and De Groot, 2005; Liang et al., 2019; Seeber, 2011). Thus, SI interpreters may use short chunks in the target language to maintain simultaneity with the source language, especially when dealing with long and complex sentences (Kader and Seubert, 2014). The strategy of chunking is described as “syntactic restructuring” in simultaneous interpreting (Riccardi, 2022, p. 378), referring to the operation of breaking long utterances of complex syntactic structure into shorter segments. Because producing a long target utterance may cause excessive cognitive load, SI interpreters often resort to this strategy (Chernov, 2004) to help them to save their cognitive resources (Cheung, (2012); Donato,

2003; Gile, 2009; Li, 2015; Shlesinger, 2003; Wu and Liao, 2018). The chunking technique might explain the frequent use of four-word bundles identified in the current study.

In addition, stance expressions take up a similar proportion of the total frequency in L2I and L1O and the percentage of the other two broad categories does not differ dramatically. Thus, it might be safe to infer that L2 interpreters and L1 speakers resort to similar groups of bundles to construct their speech output in the same UNSC setting, although the two groups of speakers share only a small number of lexical bundles as shown in Section 4.1.1.3; this supports the argument that the four-word bundles 'have important discourse functions that fit the context and purposes of the registers in which they are common' (Conrad & Biber, 2005, p. 69). Wang (2017) argued that academic lectures are a form of speech event, and as such may necessarily reflect characteristics of stereotyped speech. Likewise, in the current study, it is reasonable to see the similar distribution of functional bundles in L2I and L1O, given the homogenous political setting of the UNSC in the two sub-corpora.

Stance bundles are found to be most common in spoken register (both conversation and classroom teaching), demonstrating communicative priorities in speech output (Biber et al., 2004). Referential bundles, the most prevalent functional category in academic writing registers and a very typical occurrence in an academic spoken register (classroom instruction), are primarily used to provide information (Biber et al., 2004). If communicative and informational priorities are two ends of a continuum, both L2I and L1O in the current study would be located at the communicative end and the L1O would be at the informational end, which implies that interpreted texts are more informational than the UNSC discourse.

The greater density of introduction/focus bundles in L1O might indicate that L1 speakers rely on a set of formulaic expressions to make connections between propositions clear when condensing informational content to be delivered in the political speech setting, which aligns with the findings of Wang (2017, p. 200), who argued that speeches in academic lectures have a higher density of procedure bundles because the "heavy content to be delivered in the lecture setting may prompt the use of a set of routine phrases so as to make connections between propositions clear to the audience".

On the other hand, L2 interpreters also used introduction/focus bundles more often but with fewer types. This chunking technique might explain why L2 interpreters use more lexical bundles in their interpreting output. However, it is possible for interpreters to use introduction/focus bundles to link these chunks together, but L1 speakers may employ more varied expressions to achieve the same goal.

5.3.2.2 Comparison between L2I and interpreted corpora of previous studies

To further illustrate the distribution of functional bundles in the L2I, this section compared the functional patterns found in the current study with two previous studies using similar functional classification frameworks based on interpreted political discourse (Li & Halverson, 2022; Wu, 2021). There are several reasons to compare these studies. First, all three studies focus on four-word lexical bundles to assess formulaicity in spoken corpora. Second, the retrieval and screening processes of the studies are similar in terms of frequency cut-off points and dispersion rate. Third, these three studies are all based on political discourses sharing similar speech topics. Fourth, the functional classification framework adopted in these three studies is based on the

taxonomy proposed by Biber et al. (2004), with minor adaptations for each individual study. However, it should also be mentioned that the corpus sizes differ in each study, as do the source-language pairs and the interpreting working modes.

Wu (2021) examined the simultaneous interpreting and L1 English speeches from the European Parliament Interpreting Corpus (EPIC). The simultaneous interpreted texts were interpreted from Spanish and Italian, containing 37 interpreted texts with a total of 18,611 words. The L1 English speech sub-corpus contains 81 English speeches with a total of 40,711 words. The paralinguistic features, including repetitions, hesitations, fillers and words half-produced, were removed as processes from the current study. The topics in Wu's corpora includes politics, justice and health in the EP context, which can be regarded as political discourse. The functional taxonomy is also adapted from the framework proposed by Biber et al. (2004), including stance organisers, discourse organisers and referential bundles. Wu added a subcategory of subject-specific bundles, referring to organisation, institution or documents, which is like the context-dependent bundles in the current study. Her findings showed that stance bundles, referential bundles and subject-specific bundles almost equally take up one third of the corpus, with discourse organisers accounting for 8% of the bundle types. On the other hand, the L1 speech used the largest proportion of subject-specific bundles at 50%, followed by 24% for referential bundles, 20% for stance bundles and 6% for discourse organisers.

Li and Halverson (2022) adopted the Chinese-English Interpreting Corpus of Premier Press Conference (CICPPC), consisting of source Chinese texts and target English consecutive interpreted texts of 147,108 words. Similarly, paralinguistic features such as repetitions or hesitations filters were not included in the corpora. They added one broad functional category—special conversational functions—based on the framework

of Biber et al. (2004). However, it seems that no examples of this type of bundle were given in their study. Their findings show that referential bundles occupy around half of the total bundles both in terms of type (48%) and token (52%). Focusing on taken counts, stance expression accounted for the second largest group, at 25%, followed by discourse organisers at 18%.

The results of these three studies of the use of lexical bundles in political discourse indicate that context-dependent (subject-specific in EPIC) bundles constitute a sizeable component of the data. However, as stated in Section 4.1.1, many context-dependent bundles were not included in the analysis of the functional classification. The context-dependent bundles, also known as subject-specific bundles, were classified as a dependent function group within EPIC, and they have also been incorporated into the referential bundles within CICPPC. Because this kind of bundle was classified differently in the three investigations, the distribution pattern looked different.

Looking at raw counts of the bundles types across function categories in EPIC (Wu, 2021, p. 85), there are similar numbers of stance expressions (91) and referential expressions (97) in interpreted texts, and the discourse organisers (22) constitute less than one quarter of the other two categories. In English speech, referential expressions have 54 type, followed by stance expressions at 44 types and discourse organisers at 14 types. Compared to the current study, these results show different patterns, as the discourse organisers are used more often in UNSC discourse, taking up the 43% in L1O and 28% in L2I. In addition, bundles functioning as discourse organisers (84%) are also used more frequently in CICPPC with the type counts higher than that of stance expressions (68%) (Li & Halverson, 2022). This difference might be explained by the source-target language pair involved in the three studies, as the EPIC is based on

interpreted texts from Spanish and Italian and the UNSC and CICPPC are interpreted from English to Chinese. Alternatively, it might be influenced by the interpreting directionality, as the practice in EPIC would be interpreting into their L1 language direction, while the interpreters in the other two corpora are working into their L2 language directions. The cross-linguistic influence on the use of lexical bundles should be further explored in future studies.

In EPIC, stance bundles and referential bundles accounted for the same percentage, which also differs from the present analysis, as the percentage of type counts for stance bundles in both L2I and L1O are greater than 30%. This difference may be explained by the frequent use of desire bundles, such as *I would like to*, in the present study. In addition, the frequent employment of action bundles, which is unique to the present study, might account for a notable number of stance bundles.

5.4 Comparison of the relationship between structural and functional categories

According to Biber et al. (2004), there is a strong relationship between syntactic structural type and discoursal functional for lexical bundles. They found that most stance bundles consist of dependent sentence fragments and verb phrase fragments, whereas most referential bundles consist of noun phrase or prepositional phrase fragments.

In the following section, the numbers of bundle types with different structures in the three broad functional categories are compared to find out whether the relationship is also supported in L2 interpreting and L1 speech.

As demonstrated in Figures 4.7 and 4.8, the structural distributions of stance bundles and referential bundles in both L2I and L1O exhibit a consistent pattern. Most attitude bundles in L2I and L1O consist of verb phrases and clausal fragments, with a higher proportion of verb phrases in L1O. As for referential bundles, they are dominated by noun phrases and prepositional phrases in L2I (89%) and L1O (69 %).

In contrast, L2 interpreters employ noun phrases and preposition phrases as discourse organisers, while L1 speakers employ a comparable number of NP/PP and clausal fragments in discourse organising bundles. This difference might be explained by the excessive use of the structure of anticipatory *it is + adj* (*it is important/vital/clear/crucial that*) by L1 speakers in introduction/focus bundles to introduce new topics. It seems that L2 interpreters diverge from L1 speakers in their use of formulaic language in political discourse in their use of discourse organising bundles.

The results of this section suggest there is a strong relationship between the structural and functional classifications of four-word bundles.

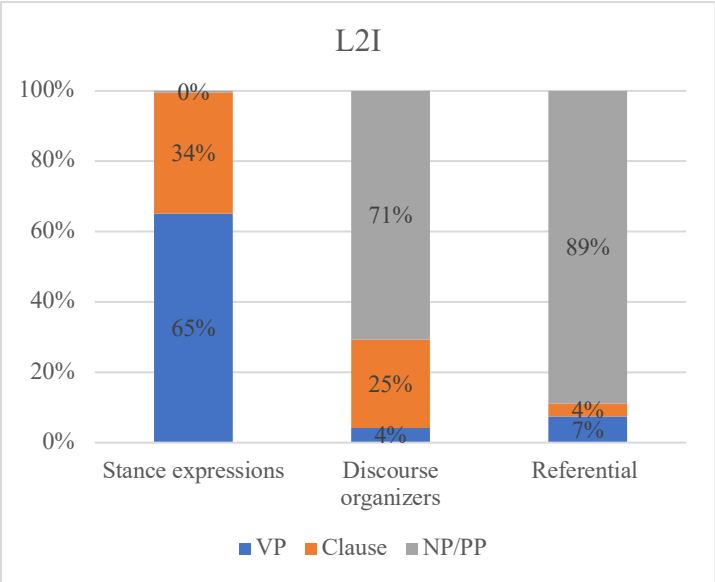

Figure 5.7 Structural-functional relationship of LBs in L2I

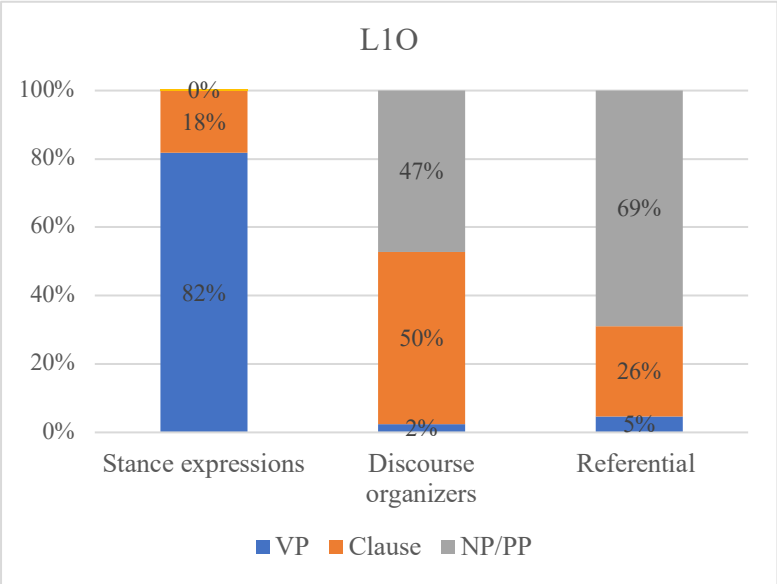

Figure 5.8 Structural-functional relationship of LBs in LIO

5.5 Summary

This study used frequency-based inquiry to highlight the commonalities and differences in the use of LBs in L2I and L1O with the aim of examining how L2 interpreters and L1 speakers use formulaic expressions to construct their language production. Specifically, this study explored the different use of four-word bundles in L1 spoken English and L2 interpreted English in a political register via a corpus-based approach. Previous studies (Biber, 2009; Biber & Barbieri, 2007; Biber et al., 2004; Cortes, 2004; Hyland, 2008b) have shown that the use of LBs is different across genres and groups of speakers. The current study reveals that L2 interpreters use LBs substantially more than L1 English speakers, suggesting that L2 interpreters in political discourses rely more heavily on the idiom principle. In addition, the distribution of functional bundles in L2I and L1O exhibits divergent patterns, suggesting that these two sets of speakers may employ different devices when constructing their utterances.

Lexical bundles are one kind of formulaic expression based on frequency. It has been claimed that frequent word sequences are not necessarily pre-fabricated, which means that they are "stored and retrieved whole from memory" (Wray, 2002, p. 9). There is a lack of adequate experimental methods for validating the nature of pre-fabricated formulaic expressions. Previous studies have instead relied on linguistic features of word sequences to identify "pre-fabricated sequences based on intuition or perceived salience, rather than on evidence from actual linguistic production and comprehension" (Biber & Barbieri, 2007, p. 283). In practice, pre-fabricated sequences are typically limited to word combinations that are relatively fixed in form and not compositional in meaning—i.e., to some extent, idiomatic (Howarth, 1998; Hudson, 1998). Extremely common word sequences do not fit these criteria; hence, they have often been ignored in previous research.

However, in-depth analyses of the discourse roles played by lexical bundles, which are defined based on frequency, show that these word combinations can also be seen as pre-fabricated units. Lexical bundles provide a discourse framework for new assertions rather than expressing new propositional meaning. Although they do not always have an idiomatic meaning, they do provide crucial discourse functions such as position expression, discourse organisation and referential framing.

Therefore, it is reasonable to hypothesize that lexical bundles of high frequency are to some extent pre-fabricated. While lexical bundles are not typically idiomatic, the fact that they are constantly functional suggests that they are possibly retained intact in memory and employed for textual or interpersonal discourse functions as unanalysed language chunks. Of course, it should be pointed out that further experimental research

is needed to determine whether L1 speakers or interpreters actually comprehend and produce lexical bundles as unanalyed pre-fabricated sequences.

In spite of the fact that they are psycholinguistic in nature, lexical bundles are thought to be conducive to the understanding and production of speech (Biber and Barbieri, 2007). It is indisputable that these sequences are frequent in both the L2 interpreting output and L1 speech production. The findings of the investigation of the discoursal function of the identified lexical bundles have confirmed the assertion that these sequences are always functional.

Although researchers are acknowledging the significance of formulaic language for fluency (Van Lancker-Sidtis and Rallon, 2004; Wray and Perkins, 2000), there is little consensus on how formulaic sequences can be effectively introduced in interpreting training and quality assessment. The description carried out in this chapter facilitates our understanding of the convergent and divergent use of lexical bundles between L2 interpreters and L1 speakers. The next chapter explores the relationship between source text and target text in terms of the use of lexical bundles to determine if the differential use of lexical bundles in L2 interpreting is due to the influence of source texts during SI interpreting.

Chapter 6 The Translation relationships Entailed in LBs Between the ST and the TT

This chapter examines and illustrates the three types of translation relationships in which lexical bundles (LBs) are used according to the distribution of their discourse functions in an attempt to determine whether and to what extent source texts (STs) influence interpreters' use of LBs. To specifically examine the association between discourse functions and translation relationships, the chapter focuses on the 71 high-frequency LBs with a raw frequency of 16 or above identified in L2I. These 71 LBs were distributed across all the sub-functional groups except for epistemic stance bundles and identification/focus bundles, in which the frequencies were less than 16. Thus, six LBs in these two sub-groups have also been included in this part of the analysis. Overall, 77 LBs encompass all the sub-functional groups, which should be suitable for the requirements of the research (see Appendix 3).

Of note, the parallel corpus included the Chinese STs (STC) and the corresponding English target texts (L2I partial), with the latter consisting of all the 194 interpreted texts that were delivered in 2019 as part of L2I (including interpreted texts from 2018 and 2019). The composition of the parallel corpus is presented in Table 6.1. The lack of corpus data is the main reason for the smaller size of the parallel corpus compared to the comparable corpus, as not all of the interpreted speeches that were collected for the comparable corpus had STs. In addition, the L2I partial contains all of the 279 identified four-word bundles (excluding context-dependent bundles) in L2I, thus indicating that the L2I partial is sufficiently representative to support the current research. As there are three types of correspondence, the translation relationships tended to occur repeatedly within a single bundle. The concordances of the 77

investigated bundles satisfy the requirements for demonstrating the general patterns. Table 5.2 shows the functional distribution of the 77 LBs of 1,223 frequency under investigation.

Table 6.1 The parallel corpus breakdown

	words	**texts**
STC	154,449	194
L2I (Partial)	104,109	194

Table 6.2 Functional distribution of lexical bundle tokens in L2I (partial)

Function		**token**		**type**	
		NO.	%	NO.	%
Stance expressions	Desire	170	14	7	9
	Obligation/Directive	68	6	5	6
	Intention	131	11	6	8
	Ability	0	0	0	0
	Action	177	14	13	17
	Epistemic stance	13	1	3	4
	Total	**559**	**46**	**34**	**44**
Discourse organisers	Introduction/transitional signals	283	23	15	19
	Elaboration/Clarification	90	7	7	9
	Total	**373**	**30**	**22**	**29**
Referential expressions	Identification/Focus	17	1	3	4
	Specification of Attributes	182	15	16	21

Time reference	92	8	2	3
Total	**292**	**24**	**21**	**27**
Total	**1223**	**100**	**77**	**100**

Sketch Engine was used to examine the parallel concordances in both the ST and the TT. The scope of the comparison was set as the sentence level (the ST sentence versus the corresponding TT sentence). The comparison was based on the semantic meaning[7] of a single substantive word in one LB, as well as of the LB itself as a whole in the TT.

The correspondence type of equivalence means that the LBs in the TT were equivalent to the messages in the ST, which was the most common translation relationship in the use of four-word bundles in the ST and the TT (63%). The large percentage of the equivalence pattern may provide evidence to support the prefabricated nature of LBs, as the equivalence in the use of LBs between the ST and the TT suggested that these LBs may have been stored as fixed expressions in the interpreters' minds and can be uttered without requiring significant cognitive effort.

Klaudy and Károly (2005) indicated that the process of explicitation, which may be required or optional, was frequently related to the addition of elements: "In the case of obligatory explicitation, the translator has no choice: if no explicitation is performed, the TT sentence will be ill-formed" (p. 16) and "[in] the case of optional explicitation, the translator is faced with a choice: s/he may produce a well-formed target language sentence even without carrying out explication" (p. 17). In this situation, explicitation

[7] Semantic meaning in the current study refers to the literal meaning of the ST or the TT. If the implicit meaning of the ST is expressed by interpreters via an LB in the TT, the translation relationship will be classified as addition.

is typically not required due to linguistic differences, but due to differences in language use, discourse structure and background information instead. The discussion of addition may be relevant to the debate about explicitation, but this issue will not be addressed here. This type of LB is mainly used for grammatical supplementation: The deconstruction of LBs contributes to matching the output and omissions in the STC with LB components by taking their meaning and grammatical functions into account. In other words, a grammatical supplement is the explicit expression of one or more non-spoken components of LBs in the TT; this may be mandatory or optional. They can be further divided into the addition of a subject, the addition of a modal verb, and the addition of prepositional phrases.

The shift pattern occurs at the lexical level, but it is important to note that the instance itself has been identified at the level of LBs. There are several types of LB shift, the first being a shift in personal references. The noun 'China' is shifted to the pronoun 'we' when using LBs such as *we hope that the* and *we hope that all* as the rendition of 中方希望. The second is a shift in the part-of-speech. The propositional phrasal LB *for the maintenance of* is the rendition of the verb phrase 维护, while the third type is a shift in tense. The passive tense four-word bundle *to which (priority)* should be given is the rendition of 应重点. The fourth type of shift is a shift from a more specific concept to a more general concept: The ST version of 中非 ('between China and Africa') has been interpreted as *between the two sides* in the LB.

The discussion of translation relationships is based on the functional classification of LBs. Examples of the realisation of each correspondence type are provided in each sub-category of the functional taxonomy.

The frequency of the three correspondence patterns according to the discourse function is presented in Table 6.3. It is notable that equivalence was the most common pattern overall, accounting for 63% of occurrences, followed by addition at 29% and shifts at 7%. These results are consistent with the findings in the study by Li and Halverson (2022), despite their findings having been based on consecutive interpreters' performances in the Chinese-English Interpreting Corpus of Premier Press Conference (CICPPC) (token: equivalence 56.8% > addition 36.3% > shift 6.9%), as well as in the study by Xu and Li (2021), which focused on simultaneous interpreting (SI) between the Cantonese and English language pair (token: equivalence 65% > addition 19% > shift 16%).

With regard to the three main categories of discourse functions, the three correspondence patterns were distributed differently, as shown in Table 6.3. The frequency of equivalence was still highest in the stance bundles (31%) and the referential bundles (18%). However, the frequency of addition in discourse organisers (16%) was greater than was that of equivalence (15%). The proportion of shift across the three broad bundle types was always the smallest. In the following section, the relationship between discourse functions and the translation relationship is discussed in more depth. Examples of the realisation of each type of correspondence are presented according to each sub-category of the three broad functional groups.

Table 6.3. Correspondence patterns by discourse functions (token frequency)

Function	**Equivalence**	**%**	**Addition**	**%**	**Shift**	**%**

Stance expressions	374	31	136	11	48	4
Discourse organisers	178	15	193	16	2	0
Referential expressions	223	18	28	2	40	3
Total	775	63	358	29	90	7

6.1 Stance expressions

With regard to the sub-categories in the stance bundles, as shown in Figure 6.1, the prevalence of the equivalence pattern occurred in four out of five sub-categories in the stance bundles,[8] with the exception being the obligation/directive bundles.

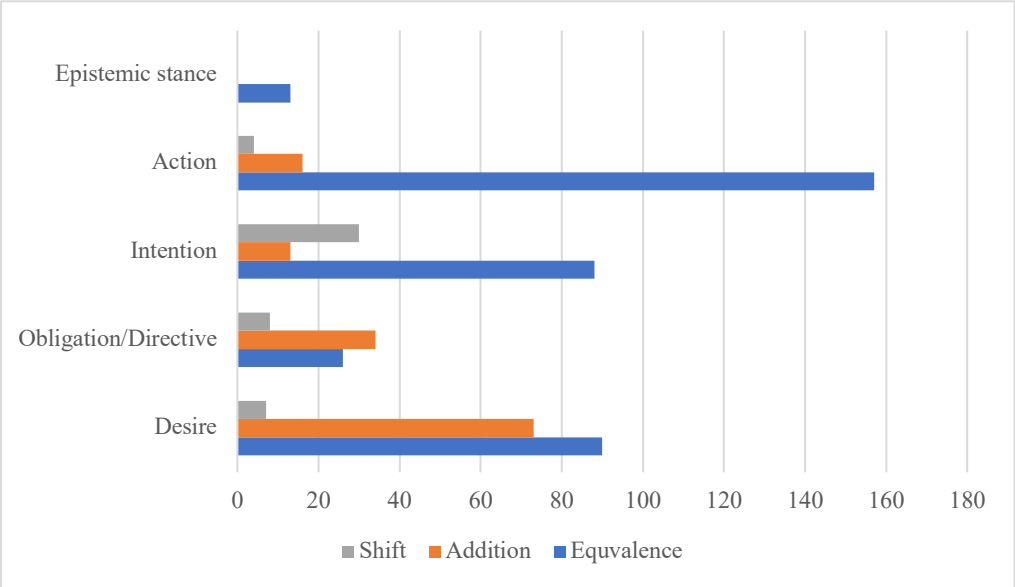

Figure 6.1 The relationship between correspondence pattern of stance LBs

6.1.1 Desire bundles

[8] The sub-category of ability bundles was not included in the analysis as there were no occurrences of such bundles in L2I.

Table 6.4 presents the frequency of the desire bundles in L2I (partial) and their correspondence patterns (absolute token frequency). The equivalence pattern accounted for 53%, followed by addition at 43% and shift at 4%.

Table 6.4 Correspondence pattern in desire bundles

Lexical bundle	Freq. in L2I(partial)	Equivalence	Addition	Shift
I would like to	61	43	18	0
we hope that the	48	13	31	4
China would like to	26	14	12	0
China hopes that the	10	9	1	0
we hope that all	10	4	5	1
I wish to thank	6	4	1	1
would also like to	9	3	6	0
Total	170	90	74	6
		53%	44%	4%

6.1.1.1.1 Equivalence in desire bundles

The majority of the desire bundles (53%) were typically equivalent to the ST. These types of bundles were frequently used by speakers and interpreters at the start of utterances to express the speakers' desire to do something. The most frequently used desire bundle *I would like to* was used as the direct interpreting of 我想/我愿 in the ST in most cases (43), as shown in Example 1. Similarly, the four-word bundle *China hopes that the* was interpreted as 中方希望 in nine out of 10 occurrences (see Example 2).

Example 1

ST: 今天 是 中国 农历 新年 ，首先 **我 想** 祝愿 总统 阁下 以及 各位 同事 新春 快乐 ，同时 也 祝愿 世界 各 国 人民 和平、 安宁 。

Jīntiān shì zhōngguó nónglì xīnnián, shǒuxiān wǒ xiǎng zhùyuàn zǒngtǒng géxià yǐjí gèwèi tóngshì xīnchūn kuàilè, tóngshí yě zhùyuàn shìjiè gèguó rénmín hépíng, ānníng.

BT: Today is the Chinese Lunar New Year. First of all, **I would like to** wish His Excellency the President and colleagues a happy Chinese New Year. I also wish peace and tranquility to the people of all countries in the world.

TT: Today is the eve of the Chinese new year. First of all, **I would like** to convey my best new year greetings to you, Sir. At the same time, I wish peace and tranquility to the people of the world

Example 2

ST: **中方 希望** 国际 社会 加强 合作 ，共同 应对 安全 挑战。

Zhōngfāng xīwàng guójì shèhuì jiāqiáng hézuò, gòngtóng yìngduì ānquán tiǎozhàn.

BT: **China hopes that the** international community will strengthen cooperation and jointly address security challenges.

TT: **China hopes that the** international community will strengthen its cooperation and respond jointly to the region's security challenges.

6.1.1.1.2 Addition in desire bundles

The translation relationship of addition was also common in desire bundles. The four-word bundle *I would like to* was a complete addition, as none of the meaning in this expression was expressed in the ST. The reason that the interpreter added this LB in the interpreting output was for the purpose of grammatical supplementation. As shown in Example 3, there was no subject in the ST, but the corresponding rendition in English required a subject in order to be grammatically correct.

Example 3

ST: 第四 ，🇨🇦 加大 对 沿岸 国家 发展 援助。

Dì sì, jiā dà duì yán'àn guójiā fāzhǎn yuánzhù

BT: Fourth, 🇨🇦 increase development assistance to coastal countries.

TT: Fourthly, **I would like to** point out the importance of increasing development assistance to coastal States.

In the United Nations Security Council (UNSC) setting, 我感谢…的发言/通报（'I thank … for his/their briefings') is frequently used by Chinese delegates at the beginning of their speeches, as shown in Example 4. While the direct translation of 我感谢 is *I thank*, this expression is often interpreted as *I would like to thank* to express an extreme degree of politeness that is only used in formal settings. Moreover, the desire bundles often occurred at the beginning of a sentence; by uttering the four-word bundle *I would like to* compared to the two-word expression *I thank*, the interpreter may have had more time to wait for more information to be produced without being silent in front of the audience. The four-word bundle *China would like to* was also often added

as a buffer in front of other notional verbs to imply either tentativeness or extreme politeness, as shown in Examples 5.

Example 4

ST:我∅感谢 格里菲斯 特使 、 洛科克 副秘书长 和 比 斯利 执行 主任 的 通报 。

Wǒ gǎnxiè gé lǐ fēi sī tèshǐ, luò kē kè fù mìshū zhǎng hé bǐ sī lì zhíxíng zhǔrèn de tōngbào.

BT: **I** thank Special Envoy Griffith, Under-Secretary-General Lowcock and Executive Director Beasley for their briefings.

TT: At the outset, **I would like to** thank Special Envoy Griffiths, Under-Secretary-General Lowcock and Executive Director Beasley for their briefings.

Example 5

ST: 中方∅对 推进 维和 行动 改革 有 以下 看法。

Zhōngfāng ∅ duì tuījìn wéihé xíngdòng gǎigé yǒu yǐxià kànfǎ

BT: **China** has the following views on promoting the reform of peacekeeping operations.

TT: **China would like to** share the following views on promoting PKO reform.

The third type of addition was the addition of a personal reference (a subject in this case) in the TT when the subject had been omitted in the ST. As shown in Example 6, the subject of 希望 (‘hope’) was omitted because the subject of 中方 appeared in the previous sentence. In Chinese, it is common for two or more verbs or verb phrases to be connected by simply placing them in a parallel structure without the use of

connecting words when describing two or multiple actions in a row. However, as this sentence structure is not idiomatic in English, the interpreter added the subject *we* in their interpreting output. Moreover, the reason for adding *we* instead of *China* is that *we* is a plural form that implies a more general meaning, thus decreasing the risk of making a mistake. Furthermore, by using a plural form as the subject, the interpreter can use the original form of the verb as a predicate without making an extra effort to ensure subject-verb agreement.

Example 6

ST: 中方 呼吁 委内瑞拉 朝野 各 方 在 宪法 和 法律 框架 内, 通过 对话 协商 寻求 政治 解决 方案 。**希望** 国际 社会 做 真正 有利于 委内瑞拉 国家 稳定 、 经济 发展 、 民生 改善 的 事情 , 在 尊重 委内瑞拉 主权 前提 下 , 向 委内瑞拉 提供 建设性 帮助 , 推动 有关 问题 尽快 平稳 解决 。

Zhōngfāng hūyù wěinèiruìlā cháoyě gè fāng zài xiànfǎ hé fǎlǜ kuàngjià nèi, tōngguò duìhuà xiéshāng xúnqiú zhèngzhì jiějué fāng'àn. Xīwàng guójì shèhuì zuò zhēnzhèng yǒu lìyú wěinèiruìlā guójiā wěndìng, jīngjì fāzhǎn, mínshēng gǎishàn de shìqíng, zài zūnzhòng wěinèiruìlā zhǔquán qiántí xià, xiàng wěinèiruìlā tígōng jiànshè xìng bāngzhù, tuīdòng yǒuguān wèntí jǐnkuài píngwěn jiějué.

BT: China calls on all parties in the ruling and opposition parties in Venezuela to seek a political solution through dialogue and consultation within the constitutional and legal framework. **Hoped that** the international community will do things that are truly conducive to Venezuela's national stability, economic development, and improvement of people's livelihood. On the premise of

respecting Venezuela's sovereignty, it will provide constructive assistance to Venezuela and promote the smooth resolution of relevant issues as soon as possible.

TT: China calls upon the Venezuelan Government and opposition parties to seek a political solution through dialogue and consultation within the constitutional and legal framework. **We hope that the** international community will take actions that are truly conducive to the stability and economic development of Venezuela and the improvement of the livelihoods of people in the country. Under the premise of respecting the sovereignty of Venezuela, we should provide constructive assistance to the country to promote a smooth resolution of relevant issues as far as possible.

6.1.1.1.3 Shift in desire bundles

The shift pattern only occurred in 4% of the desire bundles, and all the shifts were shifts in personal reference when the noun *China* was shifted to *we* in the four-word bundles *we hope that the* and *we hope that all*, as shown in Example 7. The reason was similar to the addition of *we* in Example 6, and the use of a more general expression was likely to have been a strategy that the interpreter adopted for risk management. In addition, using a noun in the plural form as a subject makes it easier to process the subsequent sentences, particularly for interpreters whose L1 language is not English.

Example 7

ST: **中方 希望** 马里 和平 协议 各 方 巩固 当前 的 良好 势头 ，不断 加强 互信 ，推进 协定 各 条款 的 落实 ，共同 致力于 马里 国家 发展 和 建设 。

Zhōngfāng xīwàng mǎlǐ hépíng xiéyì gè fāng gǒnggù dāngqián de liánghǎo shìtóu, bùduàn jiāqiáng hùxìn, tuījìn xiédìng gè tiáokuǎn de luòshí, gòngtóng zhìlì yú mǎlǐ guójiā fāzhǎn hé jiànshè.

BT: **China hopes that** all parties to the Mali peace agreement will consolidate the current good momentum, continuously strengthen mutual trust, promote the implementation of various provisions of the agreement, and work together for the national development and construction of Mali.

TT: **We hope that the** signatory parties will consolidate the current positive momentum, continue to strengthen mutual trust and implement all provisions of the agreement so as to jointly engage in Mali's development and reconstruction.

6.1.2 Obligation/directive bundles

Table 6.5 presents the frequency of the obligation/directive bundles in the L2I (partial) and their correspondence patterns (absolute token frequency). The addition pattern accounted for 51%, followed by equivalence at 34% and shift at 15%.

Table 6.5 Correspondence pattern in obligation/directive bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
should continue to support	19	9	8	2
we should continue to	22	1	20	1
we must continue to	12	5	7	0
should be given to	7	0	0	7
council should continue to	8	8	0	0

	68	23	35	10
		34%	51%	15%

6.1.1.2.1 Equivalence in obligation/directive bundles

Obligation/directive bundles generally contained model verbs such as *should* and *must*, which are the translations of 应 and 要 in the Chinese TT. The equivalence pattern accounted for 34% of the obligation/directive bundles, which were literal translations of the ST, as shown in Examples 8-9. The four-word bundle *should continue to support* was the translation of 应继续支持, and the bundle *we must continue to* was a direct translation of 我们要继续.

Example 8

ST: 国际 社会 **应 继续 支持** 联合国 驻 南 苏丹 特派团 履职 。

Guójì shèhuì yīng jìxù zhīchí liánhéguó zhù nán sūdān tèpài tuán lǚ zhí.

BT: The international community **should continue to support** the UN Mission in South Sudan in fulfilling its mandate.

TT: The international community **should continue to support** the Mission in carrying out its mandate.

Example 9

ST:**我们 要 继续** 促进 青年 全面 发展 ，支持 青年 一 代 为 世界 和 平 与 发展 作出 重要 贡献 。

Wǒmen yào jìxù cùjìn qīngnián quánmiàn fāzhǎn, zhīchí qīngnián yīdài wèi shìjiè hépíng yǔ fāzhǎn zuòchū zhòngyào gòngxiàn.

BT: **We must continue to** promote the all-round development of youth and support the young generation to make important contributions to world peace and development.

TT: **We must continue to** promote the comprehensive development of youth and support the younger generation in making important contributions to world peace and development.

6.1.1.2.2 Addition in obligation/directive bundles

In contrast to the other four sub-categories in the stance bundles, the addition pattern occurred most frequently in the obligation/directive bundles. There were two types of addition in the obligation/directive bundles, one being the addition of personal references and the other the addition of modality. One important reason was that most of the bundles in this type incorporated the word *continue*, which implied that they were to be followed by the structure of *subject* + *verb* fragments and the subject in the ST sentence had been omitted, as illustrated in Example 6. In addition, as a verb often follows the expression *continue to*, interpreters tend to add *should* or *must* in their utterances.

Example 10 demonstrated the addition of a personal reference (the subject *we*) in the TT. In Example 11, the four-word bundle *we should continue to* in the TT has had both a person reference (*we*) and modality (*should*) added, in contrast to the expression in the ST (继续).

Example 10

ST: 应 不断 完善 联合国 同 非盟 等 区域 和 次区域 组织 合作 伙伴 关系 ，加强 在 冲突 预防 、 危机 管理 和 冲突 后 重建 等 领域 工 作 。

Yìng bùduàn wánshàn liánhéguó tóng fēi méng děng qūyù hé cì qūyù zǔzhī hézuò huǒbàn guānxì, jiāqiáng zài chōngtú yùfáng, wéijī guǎnlǐ hé chōngtú hòu chóngjiàn děng lǐngyù gōngzuò.

BT: **should continue to** improve the partnership between the United Nations and regional and sub-regional organizations such as the African Union, strengthen the fields of conflict prevention, crisis management and post-conflict reconstruction.

TT: **We should continue to** improve cooperation and partnership between the United Nations and regional and subregional organizations, such as the African Union, and scale up efforts in conflict prevention, crisis management and post-conflict reconstruction.

Example 11

ST： 第二 ，继续 支持 非盟 等 区、 次区域 组织 的 努力 ，帮助 维 护 索马里 的 和平 与 安全 。

Dì èr, jìxù zhīchí fēi méng děng qū, cì qūyù zǔzhī de nǔlì, bāngzhù wéihù suǒmǎlǐ de hépíng yǔ ānquán.

BT: Second, **continue to** support the efforts of the African Union and other regional and sub-regional organizations to help maintain peace and security in Somalia.

TT： Secondly, **we should continue to** support the efforts of the AU and other regional and subregional organizations that help maintain peace and security in Somalia.

6.1.1.2.3 Shift in obligation/directive bundles

Most of the shift patterns in the obligation/directive bundles occurred in the four-word bundle *should be given to*, which were categorised as a shift in tense and a shift in part-of-speech. As shown in Example 12, 重点 in the ST is used as an adverb to modify 帮助, and is expressed in the active tense in the ST. In the TT, the interpreter has rendered 应重点 in the passive tense using the LB *(priority) should be given to* because the literal translation of 重点帮助(emphatically help) is not idiomatic in the target language.

Example 12

ST: 第二 ，**应 重点 帮助** 苏丹 政府 加强 能力 建设 。

Dì èr, yīng zhòngdiǎn bāngzhù sūdān zhèngfǔ jiāqiáng nénglì jiànshè.

BT: Second, **should emphatically** help the Sudanese government strengthen capacity building.

TT: Secondly, priority **should be given to** assisting the Sudanese Government in strengthening its capacity-building.

6.1.3 Intention bundles

Table 6.6 presents the frequency of the intention bundles in the L2I (partial) and their patterns of correspondence (absolute token frequency). The equivalence pattern was

the most frequent pattern in the intention bundles at 67%, followed by shifts at 23% and addition at 10%.

Table 6.6 Correspondence pattern in intention bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
china will continue to	33	30	1	2
china stands ready to	34	31	3	0
ready to work with	28	0	0	28
China is ready to	19	18	1	0
we will continue to	13	5	8	0
will continue to play	4	4	0	0
	131	88	13	30
		67%	10%	23%

6.1.1.3.1 Equivalence in intention bundles

Most of the bundles (67%) in the intention sub-category were equivalent to both the meaning and the form in the ST. For example, 30 out of 33 occurrences of the four-word bundle *China will continue to* were the rendition of 中国将继续 (see Example 13). *China stands ready to/China is ready to* were the translations of 中国愿 (see Example 14).

Example 13

ST: **中方 将 继续** 落实 去年 9月 中 非 合作 论坛 峰会 通过 的 《 北京 行动 计划 》。

Zhōngfāng jiāng jìxù luòshí qùnián 9 yuè zhōng fēi hézuò lùntán fēnghuì tōngguò de “běijīng xíngdòng jìhuà”

BT: **China will continue to** implement the Beijing Action Plan adopted at the summit of the Forum on China-Africa Cooperation in September last year.

TT: **China will continue to** implement the Beijing action plan adopted at the Summit of the Forum on China-Africa Cooperation in September last year.

Example 14

ST: **中方 愿** 继续 为 支持 非洲 国家 实现 持久 和平 与 共同 繁荣 而 做出 自己 的 努力 。

Zhōngfāng yuàn jìxù wèi zhīchí fēizhōu guójiā shíxiàn chíjiǔ hépíng yǔ gòngtóng fánróng ér zuò chū zìjǐ de nǔlì.

BT: **China is willing to** continue to make its own efforts to support African countries in realizing lasting peace and common prosperity.

TT: **China stands ready to** continue its efforts to support African countries in achieving lasting peace and common prosperity.

6.1.1.3.2 Addition in intention bundles

The addition pattern in the intention bundles was mainly the addition of a personal reference. In other words, the interpreter tended to add the subject of *China* or *we* when

the subject was omitted in the ST, as explained with reference to Example 5. As shown in Example 15, the subject of the predicate verb (愿) was omitted in the ST because it was the same as the subject in the previous segment (中方). However, the interpreter has added the subject in the TT.

Example 15

ST： 中方 衷心 期待 也门 尽快 恢复 和平 稳定，实现 发展，愿 继续 为 此 发挥 积极 的 建设性 作用 。

Zhōngfāng zhōngxīn qídài yěmén jǐnkuài huīfù hépíng wěndìng, shíxiàn fāzhǎn, yuàn jìxù wèi cǐ fāhuī jījí de jiànshè xìng zuòyòng.

BT： China sincerely expects Yemen to restore peace, stability and achieve development as soon as possible, and **is willing to** continue to play a positive and constructive role in this regard.

TT： China sincerely hopes that the country can restore peace and stability and achieve development as soon as possible. **China stands ready to** play a constructive and positive role in that regard.

6.1.1.3.3 Shift in intention bundles

The shift pattern mainly occurred in the intention bundle *ready to work with*. The shift mainly occurred in the verb phrase *work with*, which was the rendition of 同 … 一道 in the ST; 同 is used as a conjunction in the context, which is similar to *with* in English.一 道 is both a noun and a preposition. The literal translation of 同 … 一道 should be *together with*. However, in the UNSC setting, this phrase is unanimously translated as the verb phrase *work with.* The reason for the shift in the part-of-speech

may have been due to the segmentation strategy that is often adopted by SI interpreters. Specifically, SI interpreters tend to process the ST in segments. When the first segment (我们 愿 同 国际 社会 一 道) of the ST was uttered by the speaker, the interpreter preferred to translate this segment first to avoid having to store the information in the mind, thus saving the cognitive load for the interpreting of the following segments. However, the literal translation of the first segment (我们 愿 同 国际 社会 一 道) is not grammatically correct or idiomatic in English as it does not contain a verb. Therefore, the interpreter shifted the part-of-speech 同 … 一 道 to a verb phase (*work with*) (see Example 16).

Example 16

ST: 我们 愿 **同** 国际 社会 一 **道** ，继续 为 大 湖 地区 实现 和平 稳定 与 发展 作出 贡献 。

Wǒmen yuàn tóng guójì shèhuì yīdào, jìxù wéi dàhú dìqū shíxiàn hépíng wěndìng yǔ fāzhǎn zuòchū gòngxiàn.

BT: We **are willing to work with the** international community to continue to contribute to the realization of peace, stability and development in the Great Lakes region.

TT: We stand **ready to work with** the rest of the international community in our continued effort to contribute to achieving peace, stability and development in the region.

A shift pattern only occurred twice in the bundle *China will continue to*, which was a shift in the personal reference when the interpreter rendered the pronoun (我们) as the name of the country (*China*), as shown in Example 17. The shift from a more general

reference (*we*) to a more specific subject (*China*) was relatively rare in the present corpus.

Example 17

ST: 我们 将 继续 在 力所能及 的 范围 内 为 缓解 叙利亚人道 局势 作出 自己 的 贡献 。

Wǒmen jiāng jìxù zài lìsuǒnéngjí de fànwéi nèi wèi huǎnjiě xùlìyǎ réndào júshì zuòchū zìjǐ de gòngxiàn.

BT: **We will continue to** make our own contribution to ease the humanitarian situation in Syria within our capacity.

TT: **China will continue to** contribute to alleviating the humanitarian situation in Syria to the best of its ability.

6.1.4 Action bundles

Table 6.7 displays the frequency and correspondence patterns of the L2I (partial) action bundles (absolute token frequency). The equivalence pattern represented 90% of the action bundles, followed by the addition pattern at 9% and the shift pattern at 1%.

Table 6.7 Correspondence pattern in action bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
play a constructive role	54	54	0	0
we call on the	15	7	6	2
played an important role	14	14	0	0
listened carefully to the	12	10	2	0
China calls on the	13	13	0	0

resolve their differences through	9	9	0	0
listened attentively to the	10	10	0	0
we call on all	9	4	5	0
play an active role	2	2	0	0
China has always supported	11	11	0	0
taken note of the	5	5	0	0
support the work of	14	11	3	0
strengthen communication and coordination	9	9	0	0
	177	159	16	2
		90%	9%	1%

6.1.1.4.1 Equivalence in action bundles

Most of the interpreting patterns involved equivalence. In fact, nine out of 13 action bundles were direct translations of phrases in the ST, thus implying that these LBs were prefabricated for SI interpreters. When receiving the messages from the STs, the SI interpreters immediately rendered them as fixed expressions via LBs. The four-word bundle *play(ed) a/an constructive/important/active role* was a typical example, as it is the literal translation of 发挥建设性/重要/积极作用 (see Example 18).

Example 18

ST: 中方 愿 继续 为 南 苏丹 实现 和平 、 稳定 和 发展 发挥 建设性 作用 。

Zhōngfāng yuàn jìxù wèi nán sūdān shíxiàn hépíng, wěndìng hé fāzhǎn fāhuī jiànshè xìng zuòyòng.

BT: China is willing to continue to **play a constructive role** in realizing peace, stability and development in South Sudan.

TT: China is ready to continue to **play a constructive role** in achieving peace, stability and development in South Sudan.

6.1.1.4.2 Addition in action bundles

The addition of the bundles *we call on the* and *we call on all* represented the addition of personal references by adding the subject *we*, as illustrated with reference to Example 5 (see Example 19).

Example 19

ST: 中方 对 以色列 决定 关闭 联合国 近东 巴勒斯坦 难民 救济 和 工程处 在 东 耶路撒冷 学校 的 行为 表示 遗憾 , **呼吁** 国际 社会 加大 对 近东 巴勒斯坦 难民 救济 和 工程处 的 资助 , 帮助 巴勒斯坦 改善 经济 状况 。

Zhōngfāng duì yǐsèliè juédìng guānbì liánhéguójìndōng bālèsītǎn nànmín jiùjì hé gōngchéng chù zài dōng yēlùsālěng xuéxiào de xíngwéi biǎoshì yíhàn, hūyù guó jì shèhuì jiā dà duì jìndōng bālèsītǎn nànmín jiùjì hé

BT: China regrets Israel's decision to close the UNRWA school in East Jerusalem and **calls on the** international community to increase funding for the Near East Relief and Works Agency to help Palestine improve its economic situation.

TT: China regrets Israel's decision to close schools run by the United Nations Relief and Works Agency for Palestine Refugees in the Near East in East Jerusalem, and **we call on the** international community to increase funding to the United Nations Relief and Works Agency for Palestine Refugees in the Near East and help Palestine improve its economic circumstances.

The addition to the bundle *listen carefully to the* was the addition of the modifier *carefully*, as shown in Example 20. The ST did not include the meaning of 认真 ('carefully'), but the interpreter has added 'carefully' in the TT. Chinese delegates at the UNSC usually say 认真听取 ('listened carefully to') in most of their speeches. Thus, even when a Chinese delegate occasionally omits 认真, SI interpreters will still resort to the habitual expression of 'listened carefully to' in their renditions. It should be pointed out that SI interpreters are confined to expressions in the ST, and the addition of modifiers was rarely encountered in the SI output in the present corpus (only two occurrences in this case).

Example 20

ST: **我 也 听取 了 齐塔库 女士 的 发言** 。

Wǒ yě tīngqǔle qí tǎ kù nǚshì de fāyán.

BT: I also listened to the statement of Ms.Citaku.

TT: **I also listened carefully** to the statement of Ms. Citaku.

There were also three occurrences of the addition pattern in the bundle *support the work of.* The noun phrase *the work of* was added by the SI interpreter, as shown in Example

21, possibly due to the strategies of anticipation[9] and stalling[10]. When the interpreter heard 中方支持 ('China support') in the ST, they may have anticipated that the following information would be the object of the predicate 支持, but the message had not yet been uttered. Therefore, the interpreter may have adopted the strategy of stalling by adding the padding material *the work of* to fill the gap without adding any new information while waiting for more information to be delivered. The meaning of the noun phrase *the work of* tends to be general, which will not affect the expression in the sentence.

Example 21

ST: 中方 **支持** 格里菲斯 特使 推进 也门 政治 进程。

Zhōngfāng zhīchí gé lǐ fēi sī tèshǐ tuījìn yěmén zhèngzhì jìnchéng

BT: China **supports** Special Envoy Griffith to advance the political process in Yemen.

TT: China **supports the work of** Special Envoy Griffiths in advancing the political process in Yemen.

6.1.1.4.3 Shift in action bundles

The shift pattern only occurred four times in the action bundles. Two occurrences in the bundle *we call on the* involved a shift in personal reference by shifting 中方 ('China') to the plural pronoun *we*, which is identical to the situation in Example 7.

[9] According to Seeber (2011, p. 195), anticipation refers to "the interpreter's ability to predict a part of the original discourses before it has been uttered by the speaker, and has received considerable attention in the literature".

[10] According to Seeber (2011, p. 193), stalling is intended to buy time for interpreters "during which the interpreter may receive more input before the integration and encoding stage".

6.1.5 Epistemic stance bundles

Table 6.8 displays the frequency and correspondence patterns of the L2I (partial) epistemic stance bundles (absolute token frequency). All the bundles in this type were equivalent to the expressions in the ST, as shown in Examples 22-23.

Table 6.8 Correspondence pattern in epistemic stance bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
we believe that the	4	4	0	0
china believes that the	6	6	0	0
has always believed that	3	3	0	0
	13	13	0	0
		100%	0%	0%

Example 22

ST: **我们 认为** ，国际 社会 当前 应 重点 做好 以下 三 方面 工作 。

BT: **We believe that the** international community should focus on the following three aspects.

TT: **We believe that the** international community should focus on the following three priorities.

Example 23

ST:**中方 始终 认为** ，利比亚 问题 必须 通过 政治 途径 加以 解决。

Zhōngfāng shǐzhōng rènwéi, lìbǐyǎ wèntí bìxū tōngguò zhèngzhì tújìng jiāyǐ jiějué

BT: **China always believes** that the Libyan issue must be resolved through political means.

TT: **China has always believed** that the Libyan issue must be settled through political means.

6.2 Discourse organisers

As shown in Figure 6.2, the addition pattern occurred most frequently in the introduction/translation signal bundles, but the equivalence pattern was more prevalent in the elaboration/clarification bundles. The shift pattern only occurred seven times.

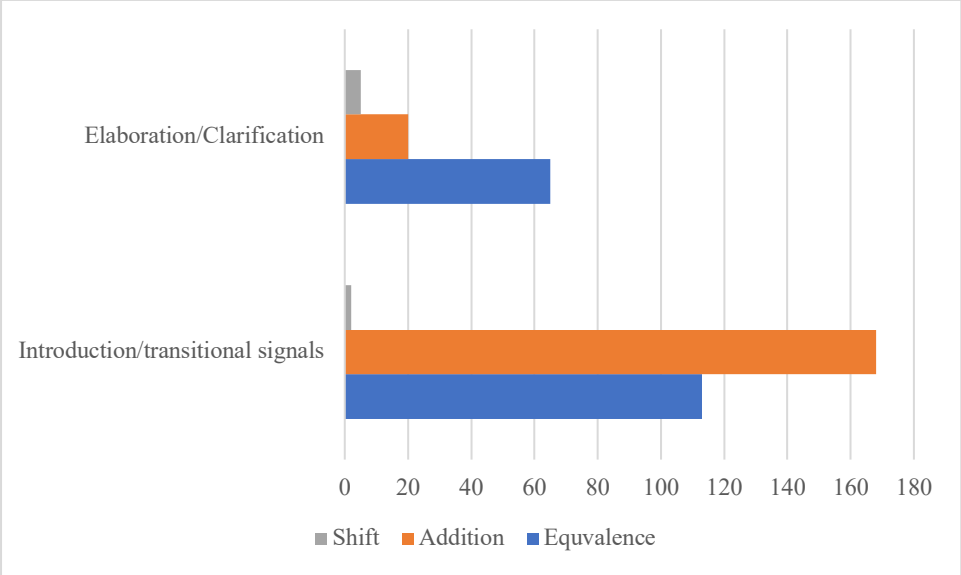

Figure 6.2 The relationship between correspondence pattern and discourse organising LBs.

6.2.1 Introduction/focus bundles

Table 6.9 presents the frequency of the desire bundles in the L2I (partial) and their correspondence patterns (absolute token frequency). The addition pattern accounted for 59%, followed by equivalence at 40% and shift at 1%. In the introduction focus bundle

group, these four-word bundles generally demonstrated either the pattern of equivalence or of addition. In other words, some introduction/focus bundles only showed the pattern of equivalence and other bundles in this type only demonstrated the pattern of addition.

Table 6.9 Correspondence pattern in introduction/focus bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
with a view to	58	0	58	0
at the same time	55	52	3	0
it is important to	21	0	21	0
it is necessary to	22	0	22	0
it is imperative to	24	0	24	0
with regard to the	18	18	0	0
on the issue of	22	21	1	0
is the only way	7	5	0	2
there is a need	11	0	11	0
as the main channel	6	6	0	0
that is acceptable to	8	8	0	0
so as to achieve	7	0	7	0
so that they can	11	0	11	0
on the question of	3	3	0	0
so that it can	10	0	10	0
	283	113	168	2
		40%	59%	1%

6.2.1.1 Equivalence in introduction/focus bundles

The vast majority of occurrences of the five four-word introduction/focus bundles demonstrated the pattern of equivalence, such as *at the same time, with regard to the, on the issue of, as the main channel* and *on the question of*. It appeared that they were immediate activations triggered by expressions in the ST, thus implying that they were prefabricated in the interpreters' minds and could be used as soon as the corresponding STs were produced. Fifty-three of the 55 occurrences of *at the same time* were the translation of 同时 in the ST, as shown in Example 24.

Example 24

ST:首先 我 想 祝愿 总统 阁下 以及 各位 同事 新春 快乐 ，**同时** 也 祝愿 世界 各 国 人民 和平 、 安宁。

Shǒuxiān wǒ xiǎng zhùyuàn zǒngtǒng géxià yǐjí gèwèi tóngshì xīnchūn kuàilè, tóngshí yě zhùyuàn shìjiè gèguó rénmín hépíng, ānníng.

BT: First of all, I would like to wish His Excellency the President and colleagues a happy Chinese New Year, and **at the same time** wish the people of all countries in the world peace and tranquility.

TT: First of all, I would like to convey my best new year greetings to you, Sir. **At the same time**, I wish peace and tranquillity to the people of the world.

The four-word bundle *as the main channel* was the translation of 主渠道 in the ST in all occurrences, as shown in Example 25.

Example 25

ST: 联合国 应 继续 加大 努力 ，发挥 斡旋 主 渠道 作用 。

Liánhéguó yīng jìxù jiā dà nǔlì, fāhuī wòxuán zhǔ qúdào zuòyòng.

BT: The United Nations should continue to intensify its efforts to play the role of **the main channel** of mediation.

The United Nations should continue to intensify its efforts and act **as the main channel** of mediation.

6.2.1.2 Addition in introduction/focus bundles

In contrast to the overall dominant pattern of equivalence, introduction/focus bundles were more often additions of expressions in the ST. Eight of these types of bundles unanimously demonstrated the addition pattern. One of them was the prepositional phrase *with a view to,* which means having the aim of doing something in the future (*Cambridge Dictionary*). In the STC, the logical relationship between the two segments is implicit. The last segment of the sentence (最终 解除 对 苏丹 的 制裁) is the effect of the previous segments based on the context. In this case, the interpreter tended to make the logical relationship explicit by employing an LB in order to conform to the typical usage in the TT, as shown in Example 26. Similarly, the SI interpreters used the four-word bundles *so that they can, so as to achieve* and *as that it can* to explicitly express the cause-effect relationship within a sentence when the logical relationship was implicit in the context but was not explicitly expressed via the use of particular words or phrases in the ST (as shown in Example 27).

Example 26

ST: 安理会 应 及时 审查 对 苏丹 采取 的 制裁 措施 ，根据 最新 形势 作出 调整 ，∅ 最终 解除 对 苏丹 的 制裁 。

Ānlǐhuì yīng jíshí shěnchá duì sūdān cǎiqǔ de zhìcái cuòshī, gēnjù zuìxīn xíngshì zuòchū tiáozhěng, zuìzhōng jiěchú duì sūdān de zhìcái.

BT: The Security Council should review the sanctions against Sudan in a timely manner, make adjustments based on the latest situation, and ∅ finally lift the sanctions against Sudan.

TT: The Council should review the sanctions on the Sudan in a timely manner and make adjustments in the light of the latest developments, **with a view to** the eventual lifting of the sanctions.

Example 27

ST: 中方 高度 重视 安理会 改进 工作 方法，提高 权威 和 效率，∅更好 地 履行 《 联合国 宪章 》 赋予 的 职责 。

Zhōngfāng gāodù zhòngshì ānlǐhuì gǎijìn gōngzuò fāngfǎ, tígāo quánwēi hé xiàolǜ, gèng hǎo de lǚxíng “liánhéguó xiànzhāng” fùyǔ de zhízé.

BT: China attaches great importance to improving the Security Council's working methods, enhancing its authority and efficiency, ∅ better fulfilling the responsibilities entrusted by the "United Nations Charter".

TT: China is very committed to improving the Council's working methods and enhancing its authority and effectiveness **so that it can** better discharge the duties mandated for it by the Charter of the United Nations,

Moreover, three bundles with the structure of *it is* + *adj*+ *to* (*it is important/necessary/imperative to*) were added to LBs in the TT. There was no subject in the STC, as shown in Example 28, and the TT began with predicate verbs. However, when rendering it in English, the SI interpreter needed a subject to start their utterance. Therefore, the structure of *it is* +*adj* +*to* is often adopted by SI interpreters to give them time to wait for the subsequent information and to form a complete sentence. As the

adjectives used in this case (*important/necessary/imperative*) have general meanings, they would not alter or affect the meaning of the ST. Similarly, SI interpreters also resort to the LB *there is a need* to manage STs that lack subjects, as shown in Example 29.

Example 28

ST: 同时 ∅ 加大 对 大 湖 地区 国家 医疗 、 教育 、 互联 互通 等 基础 设施 建设 领域 的 援助 和 投资 ，有效 提高 非洲 年轻人 就业 ，改善 当地 人民 的 生活 水平 。

Tóngshí jiā dà duì dàhú dìqū guójiā yīliáo, jiàoyù, hùlián hùtōng děng jīchǔ shèshī jiànshè lǐngyù de yuánzhù hé tóuzī, yǒuxiào tígāo fēizhōu niánqīng rén jiùyè, gǎishàn dāngdì rénmín de shēnghuó shuǐpíng.

BT: At the same time, ∅ increase aid and investment in the fields of infrastructure construction such as medical care, education, and interconnection in the countries in the Great Lakes region, so as to effectively increase the employment of young Africans and improve the living standards of local people.

TT: In the meantime, **it is important to** increase assistance and investment in health, education, interconnectivity and other infrastructure areas for the countries in the Great Lakes region, effectively improve youth employment in Africa and raise the living standards of the local populations.

Example 29

ST: 第三 ，∅ 应 帮助 刚果（ 金 ）政府 尽快 控制 埃博拉 疫情 。

Dì sān, yīng bāngzhù gāngguǒ (jīn) zhèngfǔ jǐnkuài kòngzhì āi bó lā yìqíng.

BT: Third, ∞ should help the Congo (Kinshasa) government to control the Ebola epidemic as soon as possible.

TT: Thirdly, **there is a need** to help the Government of the Democratic Republic of the Congo bring the Ebola outbreaks under control as soon as possible.

6.2.1.3 Shift in introduction/focus bundles

The following example demonstrated the shift of part-of-speech.

Example 30

ST: **只 有 如此**，有关 解决 方案 才 是 公正 、 持久 的，才 能 被 国际 社会 广泛 接受 。

Zhǐyǒu rúcǐ, yǒuguān jiějué fāng'àn cái shì gōngzhèng, chíjiǔ de, cáinéng bèi guójì shèhuì guǎngfàn jiēshòu.

BT: **Only in this way** can the relevant solution be just, lasting and widely accepted by the international community.

TT: That **is the only way** to ensure a fair and durable solution that is broadly acceptable to the international community.

6.2.2 Elaboration/clarification bundles

Table 6.10 displays the frequency and correspondence patterns of desire bundles in the L2I (partial) (absolute token frequency). The equivalence pattern accounted for 72%, while the addition pattern accounted for 22% and the shift pattern accounted for 6%.

Table 6.10 Correspondence pattern in elaboration/clarification bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
on the basis of	35	34	0	1
as well as the	24	18	6	0
the rest of the	10	0	10	0

a political settlement to	6	6	0	0
within the framework of	7	7	0	0
organizations such as the	4	0	4	0
for the maintenance of	4	0	0	4
	90	65	20	5
		72%	22%	6%

6.2.2.1 Equivalence in elaboration/clarification bundles

The pattern of equivalence was observed in almost all the instances of four-word introduction/focus bundles, such as *on the basis of* (with the exception of addition), *a political settlement to* and *within the framework of.* In addition, the ST and the TT corresponded in terms of the use of LBs. Specifically, the bundle *within the framework of* corresponded to 在…框架内 in the ST (see Example 31).

Example 31

ST: 我们 主张 ，在 联合国 安理会 相关 决议 框架 内 ，由 当事 方 通过 对话 达成 彼此 均 可 接受 的 方案, 是 解决 科索沃 问题 的 最佳 途径 。

Wǒmen zhǔzhāng, zài liánhéguó ānlǐhuì xiāngguān juéyì kuàngjià nèi, yóu dāng shì fāng tōngguò duìhuà dáchéng bǐcǐ jūn kě jiēshòu de fāng'àn, shì jiějué kēsuǒwò wèntí de zuì jiā tújìng.

BT: We maintain that within the framework of the relevant UN Security Council resolutions, the parties concerned to reach a mutually acceptable solution through dialogue is the best way to resolve the Kosovo issue.

TT: We believe that **within the framework of** relevant Security Council resolutions, the parties concerned should reach a mutually acceptable solution through dialogue. We see that as the best way to resolve the Kosovo issue

6.2.2.2 Addition in elaboration/clarification bundles

The four-word bundle *the rest of the* was often incorporated in a longer bundle, *China and the rest of the international community*, which is the translation of 我们愿同国际社会一道 (see Example 32). Both the expressions occurred frequently in the ST and in the TT and tended to be fixed, thus implying that they were examples of diplomatic language that is often used by delegates in the UNSC setting. The addition of *the rest of the* implies that China and the international community are not the opposite of each other, and that China is part of the international community. It is reasonable to postulate that the addition in this case was due to the interpreting norm in the UNSC setting; in other words, SI interpreters are required or trained to render the expression by adding *the rest of the*.

Example 32

ST: 我们 愿 同∅国际 社会 一 道 ，继续 为 大 湖 地区 实现 和平 稳定 与 发展 作出 贡献 。

Wǒmen yuàn tóng guójì shèhuì yīdào, jìxù wéi dàhú dìqū shíxiàn hépíng wěndìng yǔ fāzhǎn zuòchū gòngxiàn.

BT: We are willing to work with the ∅ international community to continue to contribute to the realization of peace, stability and development in the Great Lakes region.

TT: We stand ready to work with **the rest of the** international community in our continued effort to contribute to achieving peace, stability and development in the region.

6.2.2.3 Shift in elaboration/clarification bundles

The shift pattern mainly occurred in the bundle *for the maintenance of*, which was a shift in the part-of-speech when the verb phrases in the ST were shifted to become prepositional phrases, as shown in Example 33. Although the ST used a verb phrase (维护), it functioned as a modifier. Therefore, the interpreter shifted the part-of-speech to construct their interpreting output.

Example 33

ST: 安理会 作为 **维护** 国际 和平 与 安全 的 首要 机构 ，应 释放 积极 信息。

Ānlǐhuì zuòwéi wéihù guójì hépíng yǔ ānquán de shǒuyào jīgòu, yīng shìfàng jījí xìnxī

BT: The Security Council, as the primary body for **maintaining** international peace and security, should release positive messages

TT: As the primary body **for the maintenance of** international peace and security, the Security Council should send positive messages, adjust its South Sudanese-related measures in a timely manner

6.3 Referential expressions

As demonstrated in Figure 6.3, the equivalence pattern occurs most frequently in referential bundles among all three sub-categories.

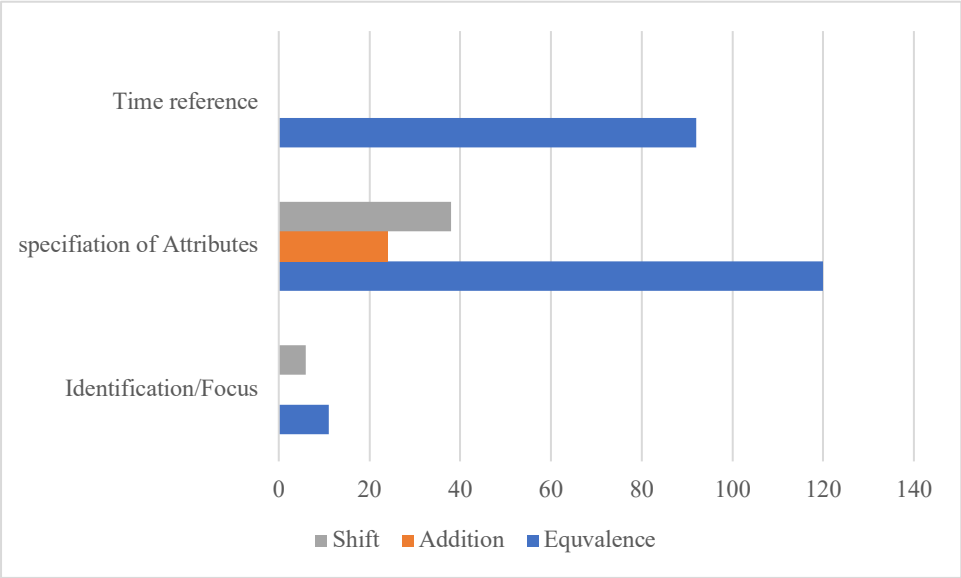

Figure 6.3 The relationship between correspondence pattern and referential LBs.

6.3.1 Identification/focus bundles

Only three bundles were included in the identification/focus bundle group, and they only showed the patterns of equivalence and shift. No pattern of addition was found in this bundle group.

Table 6.11 Correspondence pattern in identification/focus bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
between the two sides	4	2	0	2
both the symptoms and	6	6	0	0
cooperation between the two	7	3	0	4
	17	11	0	6
		65%	0%	35%

6.3.1.1 Equivalence in identification/focus bundles

The bundle *both the symptoms and* was incorporated into a longer bundle, *both the symptoms and root causes* (Example 34). Similar to Example 32, the expressions in the ST (标本兼治) and in the TT (*both the symptoms and root causes*) are both fixed expressions, and correspond to each other in the UNSC setting.

Example 34

ST: 第四，安理会 应 重视 解决 产生 冲突 的 根源性 问题，实现 标本兼治。

Dì sì, ānlǐhuì yīng zhòngshì jiějué chǎnshēng chōngtú de gēnyuán xìng wèntí, shíxiàn biāoběn jiānzhì

BT: Fourth, the Security Council should pay attention to solving the root causes of conflicts, and achieve a solution to **both the symptoms and** root causes

TT: Fourthly, the Security Council should attach great importance to tackling the underlying causes of conflicts by addressing **both the symptoms and** root causes.

6.3.1.2 Shift in identification/focus bundles

The occurrences of the shift pattern in the identification/focus bundle mainly referred to a shift from a more specific idea to a more general idea. In Example 35, the literal translation of the ST is *between China and Africa*, but the interpreter used the LB *between the two sides*. This may have been because the interpreter attempted to avoid using the same expression in their interpreting output, as the previous segment used *between China and Africa*; thus, the interpreter used *between the two sides* instead.

Example 35

ST: 日前，中国 国务 委员 兼 外交 部长 王毅 成功 访问 了 非洲，进一步 巩固 了 中 非 友好 关系，深化 了 中 非 务实 合作 。

Rìqián, zhōngguó guówù wěiyuán jiān wàijiāo bùzhǎng wáng yì chénggōng fǎngwènle fēizhōu, jìnyībù gǒnggùle zhōng fēi yǒuhǎo guānxì, shēnhuàle zhōng fēi wùshí hézuò.

BT: A few days ago, Chinese State Councilor and Foreign Minister Wang Yi successfully visited Africa, which further consolidated the friendly relations between China and Africa and deepened the practical cooperation **between China and Africa**.

TT: The successful visit by China's State Councilor and Minister for Foreign Affairs, Mr. Wang Yi, earlier this month served to further consolidate the friendly relations between China and Africa and deepen the results-oriented cooperation **between the two sides**.

6.3.2 Specification of attributes bundles

Table 6.12 Correspondence pattern in specification of attributes bundles

Lexical bundle	Freq. in L2I (partial)	Equivalence	Addition	Shift
situation on the ground	23	2	0	21
efforts should be made (to)	16	0	16	0
situation in the country	12	4	0	8
constructive role in the	10	10	0	0
a shared future for	9	9	0	0
solution to the issue	10	2	8	0

constructive role in achieving	10	10	0	0
community with a shared future	7	7	0	0
root causes of conflict	16	16	0	0
rights and interests of	10	10	0	0
interests of the country	14	14	0	0
stability in the country	9	9	0	0
concerns of all parties	13	13	0	0
progress has been made	9	0	0	9
joint force of the	9	9	0	0
a positive role in	5	5	0	0
	182	120	24	38
		66%	13%	21%

6.3.2.1 Equivalence in specification of attributes bundles

Example 36 below presents the LBs for the specification of attributes (*root causes of conflicts*) in a typical example of equivalence to the ST (冲突根源).

Example 36

ST: 第三 ，帮助 解决 **冲突 根源** 问题 。

Dì sān, bāngzhù jiějué **chōngtú gēnyuán** wèntí.

BT: Third, help resolve the **root causes of conflicts**.

TT: Thirdly, it is crucial to address the **root causes of conflicts**.

6.3.2.2 Addition in specification of attributes bundles

The addition to the bundle ‘efforts that should be made’ is the addition of a subject by adding the noun *efforts* (see Example 37).

Example 37

ST: **应 加快** 部署 联合国 驻 也门 特派团 ，协助 和 监督 协议 执行 。

Yīng jiākuài bùshǔ liánhéguó zhù yěmén tèpài tuán, xiézhù hé jiāndū xiéyì zhíxíng.

BT: Should accelerate the deployment of the United Nations Mission in Yemen, assist and monitor the implementation of the agreement.

TT: **Efforts should be made** to accelerate the deployment of the United Nations mission in Yemen in order to assist and monitor the implementation of the Agreement.

6.3.2.3 Shift in specification of attributes bundles

Example 38 shows a shift in the LBs by using a more general expression to render a more specific expression.利比亚局势 in the ST has been rendered as ‘situation on the ground’, which is a frequent LB in interpreted texts. The reason for choosing the more general expression might have been a strategy for risk management. By using the general expression *on the ground*, the interpreter avoided uttering the name of a specific country. In addition, the interpreter avoided saying ‘Libya’ repeatedly in the same sentence by using the phrase *on the ground*.

Example 38

ST: 中方 支持 任何 有助于 稳定 **利比亚 局势** 、 推动 利比亚 问题 政治 解决 进程 的 努力。

Zhōngfāng zhīchí rènhé yǒu zhù yú wěndìng lìbǐyǎ júshì, tuīdòng lìbǐyǎ wèntí zhèngzhì jiějué jìnchéng de nǔlì.

BT: China supports any efforts to stabilize **the situation in Libya** and promote the political settlement process of the Libyan issue.

TT: China supports any and all efforts to help stabilize the **situation on the ground** and promote a political settlement process for the Libyan question.

6.3.3 Time reference bundles

Table 6.12 Correspondence pattern in time reference bundles

Lexical bundle	Freq. in parallel corpus	Equivalence	Addition	Shift
as soon as possible	84	84	0	0
at an early date	8	8	0	0
	92	92	0	0
		100%	0%	0%

6.3.3.1 Equivalence in time reference bundles

All the time reference bundles demonstrated the equivalence pattern. The two bundles in this sub-category had similar meanings, as they were the translation of 尽早/早日/尽快 in the ST.

6.5 Discussion of results

This chapter aimed to describe the translation relationships of four-word LBs to their STs by employing the categories of equivalence, addition and shift. The discussion to follow is focused on answering the third main research question proposed in Chapter One: RQ3: What is the impact of STs on L2 SIs with regard to the use of LBs in interpreted texts? Two sub-questions are also addressed, namely RQ 3.1 What translation relationships do the target language LBs have to the corresponding parts of the STs? and RQ 3.2: How are these three types of translation relationships distributed in the dataset?

As shown in Table 6.3, equivalence was the most common (63%) strategy, and can be regarded as word-for-word renditions. It is reasonable to infer that the use of LBs in this category in the TT was automatic, as these LBs may require the lowest level of processing capacity on the part of L2 interpreters when interpreting at conferences. The examples of equivalence appeared to indicate a close relationship between the LBs in the ST and the TT, as generally described in form-based interpreting. However, the additive situations appeared to be more sense-oriented or communicative in nature. The majority of stance bundles and referential bundles are equivalent to the ST. Among the stance bundle group,

The addition pattern accounted for 29%, and included both grammatical supplementation and lexical specification. More than half of the discourse-organising bundles demonstrated the addition pattern. With regard to the grammatical supplementation type, the four-word bundles *it is important/imperative/necessary to*

and *there is a need* were used in the TT as the first part of the interpreting output when rendering ST sentences without a subject.

In addition, discourse-organising bundles such as *with a view to, so as to achieve* and *so that it can* were used to establish logical links in the renditions as a strategy to be more explicit in terms of meaning.

Moreover, the interpreters had a tendency to add personal references (*I would like to, we call on the, we hope that the*) as grammatical supplements when there was no subject in the source text and the equivalent English rendition required a subject in order to be grammatically correct.

The shift pattern accounted for only 7% of the total occurrences, which mainly occurred at the lexical level. There are several types of LB shifts, including a shift in personal references, a shift in tense, a shift in words (selecting a more specific expression or rendering a more general ST expression) and a shift in the part-of-speech.

A typical example of LB shift was when the noun *China* was shifted to *we.* The interpreter's use of more general language was probably due to the risk management strategy. In addition, the processing of the subsequent sentences becomes simpler when a noun in its plural form serves as the subject, particularly for interpreters whose first language is not English.

The present study's identification of translation relationships of lexical bundles between source texts and target texts is aligned with the findings of Xu and Li' (2021) study, which reported 65% equivalence cases, 19% addition cases and 16% shift cases in their corpus of Cantonese and English simultaneous interpreting. Xu and Li argue

that these translation relationship types reflect the strategic choices of interpreters in using lexical bundles. In a comparative study of professional and trainee Chinese-English consecutive interpreters, Tang and Jiang (2022) found that interpreters at different levels of proficiency use equivalence strategy more common, followed by addition and shift, which is also consistent with the current study' findings. However, professional interpreters were found to use a wider range of four-word bundles than trainee interpreters. While both studies analysed the relationship between the strategy (the translation relationships in the current study) and syntactic structures of lexical bundles, the current study focused on the corelation between the translation relationships and discoursal functional of lexical bundles, which may have practical implications for interpreter training programmes, which should focus on developing interpreters' ability to use lexical bundles strategically and appropriately in different interpreting contexts.

Chapter 7 Conclusion

7.1 Summary of major findings

This project aims to investigate the use of lexical bundles by L2 interpreters in comparison with L1 speakers and their correspondent source texts. To enable this investigation, an interpreting comparable-parallel corpus based on the United Nations Security Council Meetings (UNSCCP corpus) was designed and built. Two English components of L2 interpreted texts (referred to as L2I) and L1 original speech were included in the comparable corpus (abbreviated as L1O). One Chinese component, the source texts (abbreviated as STC), and one English component, the target texts, made up the parallel corpus (abbreviated as L2I partial). The current study focuses on four-word bundles as they contain relatively complete syntactic structures and have been widely investigated in previous studies compared with shorter or longer bundles.

7.1.1 Summary of major findings from the comparable corpus

Chapter four answers questions RQ1 and RQ2: What are the general distribution patterns of the use of LBs in L2I and L1O in terms of their frequency? What are the general distribution patterns of the use of LBs between L2 interpreting and L1 original speech in terms of their syntactic structures and discoursal functions?

Specifically, when comparing the use of four-word bundles between L2 interpreters and L1 speakers, the focus is on the general frequency of LBs (RQ 1.1-1.4) and how they are distributed in different syntactic structural groups (RQ 2.1) and discoursal functional groups (RQ 2.2). In terms of raw frequency, manually screened bundles, and the list of bundles excluding context-dependent types, the data demonstrated that four-

word bundles occurred significantly more frequently in L2I than in L1O. In the political discourse of the United Nations, the results indicate that L2 interpreters tend to rely more heavily on lexical bundles than L1 English speakers.

The prevalent use of formulaic expressions of four-word lexical bundles by L2 interpreters in the construction of their interpreting outputs is evidence of the normalisation translation pattern, as the use of LBs is considered a typical feature in L1 speech. In addition, the excessive use of formulaic expressions of lexical bundles by L2 interpreters may be attributable to interpreters' strategies for coping with the heavy cognitive load during simultaneous interpreting. The significantly greater number of context-dependent bundles utilised by L2 interpreters may be the interpreters' strategy for presenting familiar information to buy themselves time to process interpreting outputs.

Moreover, only a small number of LBs are shared between the two groups of speakers, with a small subset of bundles employed by both L2 interpreters and L1 speakers, as observed. It is safe to conclude, based on the log-likelihood results, that L2 interpreters tend to use different bundles compared to L1 speakers.

In terms of the structural taxonomy, L2I and L1O demonstrated different distribution patterns. Specifically, L2I contained the most NP/PP phrases, while L1O contained the most VP tokens. The fact that L1 speakers rely on VP more than L2 interpreters in their speech output suggests that L2I features more written language, reflecting the formal speech style of political discourse, whereas L1O is closer to the spoken language. This distinction between L2I and L1O may be explained by Shlesinger's (1989) oral-literate continuum suggesting that simultaneous interpreters tend to interpret oral texts with more literate features (Baker, 1996).

L2 interpreters used a greater number of different four-word bundle types across the three main functional categories than L1 speakers, according to the type counts of these bundles. In each subcategory, the results suggest that apart from ability bundles in the stance bundle group and introduction/transitional bundles in the discourse-organising bundle group, the type counts of L2I were generally higher than those of L1O.

Regarding frequency counts, both L2I and L1O showed a similar pattern in the distribution of functional bundles. Most stance bundles used by L2 interpreters and L1 speakers are used to express desire, obligation, intention, and epistemic stance. Therefore, it makes sense to surmise that lexical bundles of high frequency are to some extent prefabricated. Even though lexical bundles are not usually idiomatic, the fact that they are always useful suggests that they are retained in memory unaltered and used for textual or interpersonal discourse functions.

7.1.2 Summary of major findings from the parallel corpus

Chapter five aims to answer RQ3: What is the impact of source texts on L2 SI regarding the use of LBs in interpreted texts? To discover whether and to what extent source texts influence the use of lexical bundles (LBs) by interpreters, this part of the analysis examines the three types of translation relationships in which LBs are used in accordance with their distribution of discourse functions. The results showed that equivalence pattern accounts for 63% of all occurrences, followed by addition (29%) and shift (7%). In the three major functional groups, stance bundles and referential bundles are primarily used as equivalent cases. In contrast, most discourse-organizing bundles are used as addition compared to ST. The additional cases are mostly grammatical supplements when rendering Chinese ST into English TT. It is reasonable

to infer that most bundles used by L2 interpreters correspond with the expressions in the ST. The addition and shift of LBs are used by interpreters to cope with the grammatical distinctions between Chinese and English. When LBs are used as additions, L2 interpreters tend to resort to general or value expressions to avoid interfering with the meanings of the ST. They are sometimes used as a stalling strategy to allow to interpreters to wait for more information to be processed.

7.2 Innovations and significance of the study

Firstly, despite their prevalence in speech production, the study of lexical bundles is a relatively new topic in corpus-based interpreting studies. Lexical bundles are assumed to be processed in a holistic manner and stored in the speaker's memory so they are conducive to fluent speech output and mitigate interpreters' cognitive load. It is thought that this project is the first systematic comparison of lexical bundles used between English L1 speakers and L2 Chinese interpreters.

Secondly, the construction of the comparable-parallel corpus consisting of source texts, interpreted texts, and L1 speech is one of the innovations of the current project. This corpus can be used to investigate a variety of topics, including interpreting universals, SI strategies of professional interpreters, and interpreting norms. The combination of the comparable and parallel corpus in one study is also relatively limited in previous studies, and is beneficial in outlining the features of the use of LBs in a more comprehensive manner.

Thirdly, the current project combined quantitative and qualitative analysis. Both descriptive statistics, in the form of frequencies of the identified lexical bundles and

inferential statistics (i.e., log-likelihood test) are computed. The author also analysed the contextual use of LBs by examining their concordance.

Fourthly, the emphasis on the L2 interpreted texts contributes to the understanding of this underexplored field when most studies in corpus-based interpreting studies have focused on an into-L1 interpreting direction.

The current research adds to interpretese in presenting supportive evidence of normalisation as well as contributing knowledge to corpus-based interpreting studies. It also enhances the knowledge of the SI output by illustrating how L2 interpreters use LBs to form their speech.

Methodologically, the analytical steps of investigating LBs in both SI output and L1 speech may provide reference for future studies. Pedagogically, the prevalence of LBs by L2 interpreters suggest that they are essential for fluent SI output. The retrieved LBs list is helpful to interpreting training in political discourse.

7.3 Limitations and future directions

One of the limitations of the current study is that the identification and classification of four-word LBs are inevitably somewhat arbitrary, as these processes rely on manual filtration and selection. Future studies in this field could further refine the selection criteria to eliminate human factors. Moreover, bundles occur in different corpora that are worth further investigation into the different functions used by different speakers regarding the same bundle.

Although the current study supported the holistic nature of LBs, it should be noted that additional experimental studies are required to establish whether L1 speakers or interpreters actually comprehend and produce lexical bundles as unanalysed prefabricated sequences. Future studies may extend the research on how to introduce LBs in interpreting training and quality assessment.

Appendix 1 Four-word lexical bundles retrieved from L2I in Wordsmith

WordSmith Tools 8.0 Word list (WL_index.tokens)			
N	**Word**	**Freq.**	**Texts**
1	AS SOON AS POSSIBLE	163	122
2	COUNTRIES OF THE REGION	151	78
3	DEMOCRATIC REPUBLIC OF THE	149	23
4	REPUBLIC OF THE CONGO	148	23
5	I WOULD LIKE TO	130	101
6	PEACE AND STABILITY IN	121	87
7	PLAY A CONSTRUCTIVE ROLE	108	99
8	REGIONAL AND SUBREGIONAL ORGANIZATIONS	107	75
9	INTERNATIONAL PEACE AND SECURITY	107	76
10	UNITED NATIONS AND THE	98	70
11	A CONSTRUCTIVE ROLE IN	98	90
12	WITH A VIEW TO	97	82
13	CHARTER OF THE UNITED	91	72
14	INTERNATIONAL COMMUNITY SHOULD CONTINUE	90	74
15	WE HOPE THAT THE	88	78
16	COMMUNITY SHOULD CONTINUE TO	85	72
17	AT THE SAME TIME	84	74
18	CHINA WILL CONTINUE TO	76	68
19	SPECIAL REPRESENTATIVE OF THE	73	55
20	CHINA STANDS READY TO	73	73
21	ON THE BASIS OF	71	61
22	ECONOMIC AND SOCIAL DEVELOPMENT	70	55
23	UNITY AND TERRITORIAL INTEGRITY	69	69
24	STABILITY AND DEVELOPMENT IN	69	64
25	REPRESENTATIVE OF THE SECRETARY	69	55
26	WOULD LIKE TO THANK	66	66
27	PURPOSES AND PRINCIPLES OF	60	48
28	HUMANITARIAN SITUATION IN SYRIA	58	31
29	POLITICAL SETTLEMENT OF THE	57	44

30	WITH THE INTERNATIONAL COMMUNITY	55	55
31	READY TO WORK WITH	55	55
32	PRINCIPLES OF THE CHARTER	54	46
33	A POLITICAL SOLUTION TO	54	44
34	CHINA WOULD LIKE TO	51	48
35	UNITED NATIONS PEACEKEEPING OPERATIONS	50	26
36	IT IS IMPORTANT TO	50	38
37	POLITICAL SOLUTION TO THE	49	39
38	IT IS NECESSARY TO	44	33
39	CONTINUE TO PLAY A	42	40
40	AS WELL AS THE	42	39
41	AN IMPORTANT ROLE IN	42	40
42	THROUGH DIALOGUE AND CONSULTATION	40	39
43	RELEVANT SECURITY COUNCIL RESOLUTIONS	40	38
44	WITH THE UNITED NATIONS	39	32
45	SHOULD CONTINUE TO SUPPORT	39	37
46	PEACE AND SECURITY IN	39	28
47	MINISTER FOR FOREIGN AFFAIRS	39	37
48	COUNTRIES IN THE REGION	39	27
49	WORK WITH THE INTERNATIONAL	38	38
50	RELEVANT UNITED NATIONS RESOLUTIONS	38	23
51	CHINA IS READY TO	38	37
52	A POLITICAL SETTLEMENT OF	38	35
53	CONTINUE TO SUPPORT THE	37	34
54	SITUATION ON THE GROUND	36	28
55	MAINTAINING INTERNATIONAL PEACE AND	36	31
56	SPECIAL ENVOY OF THE	35	31
57	REGIONAL PEACE AND STABILITY	35	31
58	IT IS IMPERATIVE TO	35	31
59	COMMUNICATION AND COORDINATION WITH	35	33
60	AS THE AFRICAN UNION	35	34
61	IMPROVE THE HUMANITARIAN SITUATION	34	31
62	THAT THE INTERNATIONAL COMMUNITY	33	29
63	SUCH AS THE AFRICAN	33	32
64	SOVEREIGNTY AND TERRITORIAL INTEGRITY	33	33
65	ENVOY OF THE SECRETARY	33	29
66	CHINA HOPES THAT THE	33	30

67	SAFETY AND SECURITY OF	32	27
68	WITH REGARD TO THE	31	30
69	USE OF CHEMICAL WEAPONS	31	12
70	IMPLEMENTATION OF THE AGREEMENT	31	19
71	STANDS READY TO WORK	30	30
72	INTERNATIONAL COMMUNITY SHOULD SUPPORT	30	29
73	INTERNATIONAL COMMUNITY SHOULD STRENGTHEN	30	27
74	WE CALL ON THE	29	28
75	READY TO CONTINUE TO	29	27
76	ON THE ISSUE OF	29	25
77	LIKE TO THANK SPECIAL	29	29
78	EFFORTS OF THE UNITED	29	28
79	BELT AND ROAD INITIATIVE	29	19
80	NATIONS AND THE INTERNATIONAL	28	24
81	MIDDLE EAST PEACE PROCESS	28	17
82	MAINTAINING PEACE AND STABILITY	28	26
83	INTERNATIONAL COMMUNITY SHOULD PROVIDE	28	27
84	DIFFERENCES THROUGH DIALOGUE AND	28	28
85	SHOULD CONTINUE TO PROVIDE	27	24
86	SHARED FUTURE FOR HUMANKIND	27	27
87	SECURITY COUNCIL AND THE	27	27
88	ROLE OF THE UNITED	27	27
89	IS THE ONLY WAY	27	26
90	INDEPENDENCE AND TERRITORIAL INTEGRITY	27	26
91	BY THE SECURITY COUNCIL	27	26
92	WITH THE SYRIAN GOVERNMENT	26	25
93	PEACE AND STABILITY AND	26	24
94	MAINTENANCE OF INTERNATIONAL PEACE	26	23
95	BETWEEN THE UNITED NATIONS	26	10
96	AGENDA FOR SUSTAINABLE DEVELOPMENT	26	24
97	AFRICAN UNION AND THE	26	22
98	WE SHOULD CONTINUE TO	25	19
99	SUPPORTS THE UNITED NATIONS	25	22
100	SOON AS POSSIBLE AND	25	24
101	PLAYED AN IMPORTANT ROLE	25	25
102	ON THE INTERNATIONAL COMMUNITY	25	24

103	COMMITTEE ESTABLISHED PURSUANT TO	25	19
104	WOULD LIKE TO MAKE	24	24
105	WE WILL CONTINUE TO	24	24
106	WAY TO RESOLVE THE	24	24
107	SHOULD BE MADE TO	24	21
108	SETTLEMENT OF THE SYRIAN	24	17
109	PEACE AND SECURITY AND	24	24
110	MEMBERS OF THE COUNCIL	24	18
111	GOVERNMENT AND THE PEOPLE	24	22
112	FOR FOREIGN AFFAIRS OF	24	23
113	ESTABLISHED PURSUANT TO RESOLUTION	24	19
114	EFFORTS SHOULD BE MADE	24	22
115	ACCORDANCE WITH THE RELEVANT	24	24
116	UNITED NATIONS PEACE AND	23	20
117	SITUATION IN THE COUNTRY	23	23
118	REST OF THE INTERNATIONAL	23	23
119	PEACE AND RECONCILIATION PROCESS	23	17
120	OTHER REGIONAL AND SUBREGIONAL	23	21
121	LISTENED CAREFULLY TO THE	23	23
122	LIKE TO MAKE THE	23	23
123	HUMANITARIAN SITUATION IN THE	23	22
124	GOVERNMENT OF THE DEMOCRATIC	23	11
125	CONTINUE TO WORK WITH	23	23
126	AT AN EARLY DATE	23	22
127	UNITED NATIONS AS THE	22	22
128	THERE IS A NEED	22	11
129	STANDS READY TO CONTINUE	22	22
130	PEACE AND DEVELOPMENT IN	22	21
131	LASTING PEACE AND STABILITY	22	21
132	COUNCIL AND THE INTERNATIONAL	22	22
133	CONSTRUCTIVE ROLE IN THE	22	21
134	A SHARED FUTURE FOR	22	22
135	WORK OF THE UNITED	21	21
136	WITH THE REST OF	21	21
137	WILL CONTINUE TO PROVIDE	21	21
138	SOLUTION TO THE ISSUE	21	20
139	POLITICAL PROCESS IN SYRIA	21	16

140	NATIONS PEACE AND DEVELOPMENT	21	20
141	IS A NEED TO	21	11
142	INTERNATIONAL COMMUNITY SHOULD RESPECT	21	20
143	GROUP OF FIVE FOR	21	19
144	FIVE FOR THE SAHEL	21	19
145	COORDINATION WITH THE SYRIAN	21	20
146	CONSTRUCTIVE ROLE IN ACHIEVING	21	21
147	CHINA CALLS ON THE	21	19
148	BASIS OF RESPECT FOR	21	21
149	AS THE MAIN CHANNEL	21	20
150	WITH EAST JERUSALEM AS	20	20
151	WILL CONTINUE TO WORK	20	20
152	WE HOPE THAT ALL	20	20
153	THROUGH DIALOGUE AND CONSULTATIONS	20	18
154	THAT IS ACCEPTABLE TO	20	18
155	SOLUTION TO THE SYRIAN	20	17
156	RESOLVE THEIR DIFFERENCES THROUGH	20	20
157	PRINCIPLE OF LAND FOR	20	18
158	ON THE GROUND AND	20	19
159	MEMBERS OF THE SECURITY	20	19
160	LISTENED ATTENTIVELY TO THE	20	20
161	ALL THE PARTIES CONCERNED	20	17
162	A POLITICAL SETTLEMENT TO	20	18
163	WITH A SHARED FUTURE	19	18
164	WILL CONTINUE TO SUPPORT	19	18
165	WE CALL ON ALL	19	17
166	UNITED NATIONS RELIEF AND	19	18
167	THROUGH DIALOGUE AND NEGOTIATIONS	19	18
168	THAT THE PARTIES CONCERNED	19	17
169	SUPPORT THE UNITED NATIONS	19	18
170	STABILIZATION MISSION IN THE	19	18
171	SPECIAL ENVOY DE MISTURA	19	13
172	SHOULD PLAY A CONSTRUCTIVE	19	19
173	ROOT CAUSES OF CONFLICT	19	17
174	RIGHTS AND INTERESTS OF	19	16
175	RELIEF AND WORKS AGENCY	19	18
176	PRESIDENCY OF THE SECURITY	19	18

177	PEACEFUL SETTLEMENT OF DISPUTES	19	13
178	PARTIES TO THE CONFLICT	19	16
179	NATIONS RELIEF AND WORKS	19	18
180	JERUSALEM AS ITS CAPITAL	19	19
181	INTERESTS OF THE COUNTRY	19	18
182	I WISH TO THANK	19	19
183	EAST JERUSALEM AS ITS	19	19
184	COMMUNITY WITH A SHARED	19	18
185	COMMUNITY SHOULD SUPPORT THE	19	18
186	CHINA SUPPORTS THE UNITED	19	18
187	WORKS AGENCY FOR PALESTINE	18	17
188	WITHIN THE FRAMEWORK OF	18	18
189	UNITED NATIONS MULTIDIMENSIONAL INTEGRATED	18	18
190	UNITED NATIONS DEVELOPMENT PROGRAMME	18	16
191	UNITED NATIONS ASSISTANCE MISSION	18	18
192	THEIR DIFFERENCES THROUGH DIALOGUE	18	18
193	THAT THE SECURITY COUNCIL	18	18
194	STATE IN IRAQ AND	18	16
195	STABILITY IN THE COUNTRY	18	15
196	SO AS TO ACHIEVE	18	18
197	REFUGEES IN THE NEAR	18	17
198	POLITICAL SETTLEMENT TO THE	18	16
199	PLAY AN ACTIVE ROLE	18	17
200	PALESTINE REFUGEES IN THE	18	17
201	NATIONS MULTIDIMENSIONAL INTEGRATED STABILIZATION	18	18
202	MULTIDIMENSIONAL INTEGRATED STABILIZATION MISSION	18	18
203	LEAGUE OF ARAB STATES	18	16
204	LAND FOR PEACE AND	18	18
205	ISLAMIC STATE IN IRAQ	18	16
206	INTERNATIONAL COMMUNITY SHOULD FULLY	18	18
207	INTEGRATED STABILIZATION MISSION IN	18	18
208	FOR PEACE AND THE	18	18
209	FOR PALESTINE REFUGEES IN	18	17
210	CONCERNS OF ALL PARTIES	18	18

211	CHINA HAS ALWAYS SUPPORTED	18	17
212	BY THE UNITED NATIONS	18	18
213	AGENCY FOR PALESTINE REFUGEES	18	17
214	WITH REGIONAL AND SUBREGIONAL	17	16
215	WEAPONS OF MASS DESTRUCTION	17	8
216	TAKEN NOTE OF THE	17	14
217	SUPPORT THE WORK OF	17	17
218	STRENGTHEN COMMUNICATION AND COORDINATION	17	16
219	SO THAT THEY CAN	17	17
220	SETTLEMENT ACTIVITIES IN THE	17	16
221	SECURITY SITUATION IN THE	17	15
222	POLITICAL AND DIPLOMATIC MEANS	17	11
223	PEACE PROCESS IN THE	17	16
224	PEACE AND RECONCILIATION IN	17	16
225	ON THE QUESTION OF	17	14
226	LEGITIMATE RIGHTS AND INTERESTS	17	14
227	IS ACCEPTABLE TO ALL	17	15
228	IRAQ AND THE LEVANT	17	15
229	IMPORTANT ROLE IN MAINTAINING	17	17
230	IMPLEMENTATION OF THE PEACE	17	12
231	ECONOMIC COMMUNITY OF WEST	17	17
232	COOPERATIVE AND SUSTAINABLE SECURITY	17	17
233	COMMUNITY OF WEST AFRICAN	17	17
234	ACTIVITIES IN THE OCCUPIED	17	16
235	WOULD ALSO LIKE TO	16	16
236	WILL CONTINUE TO PLAY	16	16
237	WE MUST CONTINUE TO	16	13
238	THANK SPECIAL REPRESENTATIVE OF	16	16
239	SUMMIT OF THE FORUM	16	16
240	SO THAT IT CAN	16	16
241	SITUATION IN SYRIA AND	16	13
242	SHOULD BE GIVEN TO	16	13
243	PROMOTE THE POLITICAL PROCESS	16	16
244	PROGRESS HAS BEEN MADE	16	16
245	ORGANIZATIONS SUCH AS THE	16	16
246	NATIONS AS THE MAIN	16	16

247	MAINTENANCE OF PEACE AND	16	14
248	MAINTAIN PEACE AND STABILITY	16	16
249	JOINT FORCE OF THE	16	16
250	IS READY TO WORK	16	16
251	INTERNATIONAL COMMUNITY TO CONTINUE	16	16
252	INTERGOVERNMENTAL AUTHORITY ON DEVELOPMENT	16	16
253	I WOULD ALSO LIKE	16	16
254	GULF COOPERATION COUNCIL INITIATIVE	16	16
255	FULLY RESPECT THE SOVEREIGNTY	16	16
256	FOR THE MAINTENANCE OF	16	15
257	FOR A POLITICAL SETTLEMENT	16	15
258	COUNCIL SHOULD CONTINUE TO	16	15
259	AN ACTIVE ROLE IN	16	15
260	ALLEVIATE THE HUMANITARIAN SITUATION	16	13
261	AFRICAN SOLUTIONS TO AFRICAN	16	16
262	A POSITIVE ROLE IN	16	16
263	WILL CONTINUE TO STRENGTHEN	15	15
264	WEST AFRICA AND THE	15	8
265	VOTED IN FAVOUR OF	15	15
266	UNITED NATIONS AND OTHER	15	15
267	SOLUTION IS THE ONLY	15	15
268	SO AS TO CREATE	15	15
269	ON ALL PARTIES TO	15	14
270	IT IS ESSENTIAL TO	15	13
271	INTERNATIONAL COMMUNITY SHOULD INCREASE	15	15
272	HUMANITARIAN SITUATION IN YEMEN	15	9
273	FOR HIS BRIEFING AND	15	15
274	COUNTRY AND ITS PEOPLE	15	15
275	COMMUNITY TO CONTINUE TO	15	15
276	CALL ON ALL PARTIES	15	15
277	BY THE INTERNATIONAL COMMUNITY	15	15
278	BEIJING SUMMIT OF THE	15	15
279	ALL THE RELEVANT PARTIES	15	13
280	ALL SETTLEMENT ACTIVITIES IN	15	15
281	ADDRESS THE ROOT CAUSES	15	14

282	A VIEW TO ACHIEVING	15	14
283	WITH THE GOVERNMENT OF	14	14
284	WITH THE AIM OF	14	12
285	TERRITORIAL INTEGRITY OF SYRIA	14	14
286	SUPPORT THE EFFORTS OF	14	13
287	STABILITY AND DEVELOPMENT OF	14	13
288	SPECIAL COORDINATOR MLADENOV FOR	14	14
289	SOLUTIONS TO AFRICAN PROBLEMS	14	13
290	SOLUTION TO THE QUESTION	14	10
291	SITUATION IN THE MIDDLE	14	10
292	SHOULD FULLY RESPECT THE	14	14
293	SETTLEMENT OF THE PALESTINIAN	14	10
294	SECURITY COUNCIL RESOLUTIONS AND	14	14
295	ROLE IN ACHIEVING PEACE	14	14
296	REPORT OF THE SECRETARY	14	12
297	REGIONAL PEACE AND SECURITY	14	12
298	PROCESS IN SOUTH SUDAN	14	8
299	PREVENT VIOLENCE AGAINST CIVILIANS	14	14
300	PEACE AND THE ARAB	14	14
301	ON THE PART OF	14	12
302	NORMS GOVERNING INTERNATIONAL RELATIONS	14	14
303	MLADENOV FOR HIS BRIEFING	14	14
304	LIKE TO THANK UNDER	14	14
305	INTERNATIONAL CONFERENCE ON THE	14	14
306	INTERNATIONAL COMMUNITY SHOULD HELP	14	13
307	HOPE THAT THE INTERNATIONAL	14	13
308	HAS PLAYED AN IMPORTANT	14	14
309	GOOD OFFICES OF THE	14	12
310	FORCE OF THE GROUP	14	14
311	FOR A POLITICAL SOLUTION	14	12
312	EFFORTS TO IMPROVE THE	14	14
313	COORDINATOR MLADENOV FOR HIS	14	14
314	CONTINUE TO PROVIDE SUPPORT	14	12
315	CLOSE ATTENTION TO THE	14	13
316	CHINA CALLS ON ALL	14	14
317	BY THE PURPOSES AND	14	14
318	BETWEEN PALESTINE AND ISRAEL	14	10

319	ATTACHES GREAT IMPORTANCE TO	14	13
320	ADVANCE THE POLITICAL PROCESS	14	13
321	ACHIEVING LASTING PEACE AND	14	13
322	WOULD LIKE TO EMPHASIZE	13	13
323	WE BELIEVE THAT THE	13	11
324	UNITED NATIONS MISSION IN	13	11
325	UNITED NATIONS HIGH COMMISSIONER	13	9
326	THANK SPECIAL COORDINATOR MLADENOV	13	13
327	SO AS TO ENSURE	13	12
328	SHOULD WORK TOGETHER TO	13	13
329	SECURITY AND STABILITY IN	13	13
330	ROLE IN THAT REGARD	13	12
331	ROLE IN MAINTAINING PEACE	13	13
332	RESPECT FOR THE SOVEREIGNTY	13	13
333	QUESTION OF PALESTINE IS	13	12
334	PROVIDE CONSTRUCTIVE ASSISTANCE TO	13	13
335	PRIMARY RESPONSIBILITY FOR MAINTAINING	13	13
336	POLITICAL SOLUTION IS THE	13	13
337	POLITICAL SETTLEMENT IS THE	13	13
338	PARTIES IN SOUTH SUDAN	13	11
339	ORGANIZATION FOR THE PROHIBITION	13	13
340	ON THE GREAT LAKES	13	13
341	NATIONS HIGH COMMISSIONER FOR	13	9
342	MISSION IN THE DEMOCRATIC	13	13
343	IS READY TO CONTINUE	13	13
344	INTERNATIONAL COMMUNITY AND THE	13	12
345	INTERESTS OF ALL PARTIES	13	13
346	IMPROVING THE HUMANITARIAN SITUATION	13	11
347	HOPE THAT THE PARTIES	13	13
348	FUTURE OF THE COUNTRY	13	13
349	FOR THE SAHEL G	13	13
350	FOR MAINTAINING INTERNATIONAL PEACE	13	12
351	COUNCIL SHOULD REMAIN UNITED	13	13
352	CONSTRUCTIVE ROLE IN PROMOTING	13	13
353	CONFERENCE ON THE GREAT	13	13
354	COMMUNITY SHOULD FULLY RESPECT	13	13
355	CHINA IS DEEPLY CONCERNED	13	13

356	CHINA BELIEVES THAT THE	13	12
357	CALL ON THE INTERNATIONAL	13	13
358	AN OBJECTIVE AND IMPARTIAL	13	13
359	ALSO LISTENED CAREFULLY TO	13	13
360	AFRICA PEACE AND SECURITY	13	12
361	A POLITICAL SETTLEMENT IS	13	13
362	YORK TO PRESIDE OVER	12	12
363	WITH THE PURPOSES AND	12	12
364	THAT THE UNITED NATIONS	12	11
365	TERRITORIAL INTEGRITY OF THE	12	12
366	SUFFERING OF THE SYRIAN	12	9
367	STABILIZATION MISSION IN MALI	12	12
368	STABILITY IN THE REGION	12	12
369	STABILITY IN SOUTH SUDAN	12	9
370	RESPECT THE LEADERSHIP OF	12	12
371	REPUBLIC OF SOUTH SUDAN	12	11
372	PROHIBITION OF CHEMICAL WEAPONS	12	12
373	PRINCIPLE OF A SYRIAN	12	12
374	PEACE AND DEVELOPMENT TRUST	12	11
375	OTHER COUNTRIES IN THE	12	11
376	ON THE IMPLEMENTATION OF	12	11
377	ON PEACE AND RECONCILIATION	12	12
378	OBJECTIVE AND IMPARTIAL POSITION	12	12
379	NEW YORK TO PRESIDE	12	12
380	MEMORANDUM OF UNDERSTANDING ON	12	11
381	MAINTAIN INTERNATIONAL PEACE AND	12	11
382	MAIN CHANNEL OF MEDIATION	12	11
383	ITS ECONOMIC AND SOCIAL	12	12
384	IS AT THE CORE	12	12
385	INTERNATIONAL COMMUNITY SHOULD ACTIVELY	12	12
386	INITIATIVE AND ITS IMPLEMENTATION	12	12
387	HIGH COMMISSIONER FOR REFUGEES	12	8
388	FOR THE RESUMPTION OF	12	12
389	FOR THE PROHIBITION OF	12	12
390	EFFECTIVE IMPLEMENTATION OF THE	12	11
391	ECONOMIC AND HUMANITARIAN SITUATION	12	10

392	EASE THE HUMANITARIAN SITUATION	12	12
393	COUNCIL INITIATIVE AND ITS	12	12
394	COUNCIL FOR THIS MONTH	12	12
395	COOPERATION COUNCIL INITIATIVE AND	12	12
396	COMMUNITY SHOULD RESPECT THE	12	12
397	COMMENDS THE UNITED NATIONS	12	12
398	CHINESE PRESIDENT XI JINPING	12	12
399	CHINA HAS ALWAYS BEEN	12	12
400	CHINA ATTACHES GREAT IMPORTANCE	12	11
401	CALLS ON THE INTERNATIONAL	12	12
402	BETWEEN THE TWO SIDES	12	9
403	ALSO LIKE TO THANK	12	12
404	ALL THE PARTIES TO	12	10
405	AGREEMENT ON PEACE AND	12	12
406	ABIDE BY THE PURPOSES	12	12
407	A SOLUTION TO THE	12	10
408	A POLITICAL SOLUTION IS	12	12
409	A COMMUNITY WITH A	12	11
410	WORK WITH ALL PARTIES	11	11
411	WITH THE PARTIES CONCERNED	11	11
412	WE STAND READY TO	11	11
413	UNITED NATIONS ORGANIZATION STABILIZATION	11	11
414	THREAT TO INTERNATIONAL PEACE	11	10
415	THAT ALL THE PARTIES	11	11
416	TAKE THIS OPPORTUNITY TO	11	11
417	SUBREGIONAL ORGANIZATIONS SUCH AS	11	11
418	STABILITY IN THE MIDDLE	11	9
419	SOVEREIGNTY OF THE COUNTRIES	11	11
420	SOLUTION THROUGH DIALOGUE AND	11	11
421	SOLUTION THAT IS ACCEPTABLE	11	9
422	SHOULD RESPECT THE SOVEREIGNTY	11	11
423	SHOULD CONTINUE TO HELP	11	9
424	SHOULD ABIDE BY THE	11	11
425	SETTLEMENT IS THE ONLY	11	11
426	REVITALIZED AGREEMENT ON THE	11	11
427	RETURN OF REFUGEES AND	11	11

428	RESPONSIBILITY FOR MAINTAINING INTERNATIONAL	11	11
429	RESPECT THE SOVEREIGNTY AND	11	11
430	RESOLUTIONS AND INTERNATIONAL LAW	11	11
431	POSITIVE AND CONSTRUCTIVE ROLE	11	10
432	POLITICAL PROCESS IN YEMEN	11	9
433	PLAY AN IMPORTANT ROLE	11	11
434	PEACE PROCESS IN SOUTH	11	8
435	PEACE IN THE MIDDLE	11	11
436	PARTIES IN SYRIA TO	11	10
437	ORGANIZATION STABILIZATION MISSION IN	11	11
438	ON THE RESOLUTION OF	11	11
439	OBJECTIVE AND IMPARTIAL INVESTIGATION	11	9
440	NATIONS ORGANIZATION STABILIZATION MISSION	11	11
441	NATIONS ASSISTANCE MISSION IN	11	11
442	MAKE THE FOLLOWING POINTS	11	11
443	LOWCOCK FOR HIS BRIEFING	11	11
444	IS AN IMPORTANT COUNTRY	11	11
445	INTERNATIONAL COMMUNITY SHOULD WORK	11	9
446	FROM THE INTERNATIONAL COMMUNITY	11	11
447	EFFORTS TO PROMOTE PEACE	11	10
448	CREATE FAVOURABLE CONDITIONS FOR	11	11
449	CONTINUE TO CONTRIBUTE TO	11	11
450	CONFLICT IN THE REPUBLIC	11	11
451	CHINA VOTED IN FAVOUR	11	11
452	CHINA EXPRESSES ITS APPRECIATION	11	10
453	CHINA APPRECIATES THE EFFORTS	11	11
454	CAREFULLY TO THE STATEMENT	11	11
455	CALLS ON ALL PARTIES	11	10
456	BOTH THE SYMPTOMS AND	11	11
457	BASIC NORMS GOVERNING INTERNATIONAL	11	11
458	ASSISTANCE AND ECONOMIC SUPPORT	11	10
459	ALSO LISTENED ATTENTIVELY TO	11	11
460	AGREEMENT ON THE RESOLUTION	11	11
461	ADVANCING THE POLITICAL PROCESS	11	10
462	ACHIEVE PEACE AND STABILITY	11	11

463	ACCEPTABLE TO ALL PARTIES	11	10
464	WHEN IT COMES TO	10	9
465	WE WOULD LIKE TO	10	10
466	WE ARE READY TO	10	10
467	UNITED NATIONS MISSION FOR	10	10
468	UNITED NATIONS AND ITS	10	10
469	THROUGH DIALOGUE AND NEGOTIATION	10	9
470	THROUGH BILATERAL AND MULTILATERAL	10	10
471	TERRITORIAL INTEGRITY OF ALL	10	10
472	TALKS AS SOON AS	10	10
473	TAKES NOTE OF THE	10	9
474	SYMPTOMS AND ROOT CAUSES	10	10
475	SUPPORT THE SYRIAN GOVERNMENT	10	8
476	SUPPORT THE ROLE OF	10	10
477	SUPPORT THE MEDIATION EFFORTS	10	9
478	STATE SOLUTION IS THE	10	10
479	SPEAK WITH ONE VOICE	10	10
480	SITUATION IN THE DEMOCRATIC	10	9
481	SHOULD FOCUS ON THE	10	10
482	SECURITY COUNCIL SHOULD REMAIN	10	10
483	SECURITY COUNCIL COMMITTEE ESTABLISHED	10	8
484	ROOT CAUSES OF THE	10	10
485	RESPECT THE SOVEREIGNTY OF	10	10
486	RESOLUTION OF THE CONFLICT	10	10
487	RELEVANT UNITED NATIONS AGENCIES	10	10
488	RELEVANT COUNCIL RESOLUTIONS AND	10	10
489	REFUGEES AND DIScontentD PERSONS	10	8
490	PROPER SETTLEMENT OF THE	10	8
491	PRESIDENCY OF THE COUNCIL	10	10
492	PLAY A POSITIVE ROLE	10	10
493	PEACE AND DEVELOPMENT FUND	10	10
494	OTHER MEMBERS OF THE	10	10
495	ON THE PARTIES CONCERNED	10	9
496	ON THE COUNCIL'S AGENDA	10	9
497	ON ASSUMING THE PRESIDENCY	10	10
498	ON AN EQUAL FOOTING	10	10
499	OFFICES OF THE UNITED	10	8

500	OFFICE OF THE UNITED	10	8
501	NATIONS AND THE SECURITY	10	9
502	MISSION IN MALI MINUSMA	10	10
503	MEET EACH OTHER HALFWAY	10	10
504	LIKE TO EMPHASIZE THE	10	10
505	JOINT COMPREHENSIVE PLAN OF	10	8
506	ISSUES THROUGH DIALOGUE AND	10	9
507	INTERNATIONAL LAW AND THE	10	10
508	INTERNATIONAL COMMUNITY SHOULD UPHOLD	10	10
509	INTERNATIONAL COMMUNITY SHOULD STEP	10	10
510	INTERNATIONAL COMMUNITY SHOULD FOCUS	10	10
511	INTERNATIONAL COMMITTEE OF THE	10	9
512	INTERFERENCE IN THE INTERNAL	10	9
513	HUMANITARIAN ASSISTANCE AND ECONOMIC	10	9
514	HOPE THAT ALL THE	10	10
515	HAVE PLAYED AN IMPORTANT	10	10
516	HAS TAKEN NOTE OF	10	10
517	HAS BEEN MADE IN	10	10
518	HAS ACTIVELY PARTICIPATED IN	10	10
519	GIVE FULL PLAY TO	10	9
520	FOR THE UNITED NATIONS	10	10
521	FOR THE MONTH OF	10	8
522	FOR THE IMPLEMENTATION OF	10	8
523	FOR FOREIGN AFFAIRS AND	10	10
524	FOR CONVENING TODAY'S MEETING	10	10
525	FOCUS ON THE FOLLOWING	10	10
526	FIND A POLITICAL SOLUTION	10	9
527	ECONOMIC COMMUNITY OF CENTRAL	10	10
528	DEVELOPMENT AS SOON AS	10	10
529	DEEPLY CONCERNED ABOUT THE	10	10
530	CREATE CONDITIONS CONDUCIVE TO	10	10
531	COUNCIL RESOLUTIONS AND INTERNATIONAL	10	10
532	COUNCIL COMMITTEE ESTABLISHED PURSUANT	10	8
533	COOPERATION WITH THE UNITED	10	9
534	COOPERATION WITH REGIONAL AND	10	10
535	COOPERATION BETWEEN THE TWO	10	8
536	CONSTRUCTIVE ROLE IN THAT	10	10

537	COMPREHENSIVE PLAN OF ACTION	10	8
538	COMMUNITY SHOULD STEP UP	10	10
539	COMMUNITY SHOULD INCREASE ITS	10	10
540	COMMUNITY OF CENTRAL AFRICAN	10	10
541	COMMITTEE OF THE RED	10	9
542	COMMENDS THE EFFORTS OF	10	10
543	CHINA'S POSITION ON THE	10	10
544	CHINA WELCOMES THOSE DEVELOPMENTS	10	10
545	CHINA TAKES NOTE OF	10	9
546	CHINA COMMENDS THE EFFORTS	10	10
547	CHAIRPERSON OF THE AFRICAN	10	10
548	AT THE CORE OF	10	10
549	ASSISTANCE TO THE COUNTRY	10	10
550	AS CHAIR OF THE	10	10
551	ALL PARTIES IN SYRIA	10	9
552	ADHERE TO THE PURPOSES	10	10
553	ACHIEVE LASTING PEACE AND	10	10
554	ACCORDANCE WITH ITS MANDATE	10	9
555	A PROPER SOLUTION TO	10	8
556	A POSITIVE AND CONSTRUCTIVE	10	9
557	WOULD FIRST LIKE TO	9	9
558	WORK IN ACCORDANCE WITH	9	9
559	WITH THE RELEVANT SECURITY	9	9
560	WITH THE COUNTRIES OF	9	9
561	UNITED NATIONS RESOLUTIONS AND	9	9
562	UNITED NATIONS OFFICE ON	9	8
563	UNITED NATIONS OFFICE FOR	9	9
564	UNITED NATIONS HYBRID OPERATION	9	9
565	UNITED NATIONS COUNTRY TEAM	9	9
566	UNDER THE LEADERSHIP OF	9	8
567	TOGETHER WITH THE REST	9	9
568	THAT THE MISSION WILL	9	9
569	TERRORISM AND VIOLENT EXTREMISM	9	8
570	TERRITORIAL INTEGRITY OF YEMEN	9	9
571	SUPPORTS THE EFFORTS OF	9	9
572	SUPPORT AND ASSISTANCE TO	9	9
573	SO AS TO REACH	9	9

574	SHOULD REMAIN UNITED AND	9	9
575	SETTLEMENT OF THE QUESTION	9	8
576	SECURITY COUNCIL SHOULD CONTINUE	9	9
577	SECURITY COUNCIL FOR THIS	9	9
578	ROOT CAUSES OF CONFLICTS	9	9
579	ROLE AS THE MAIN	9	9
580	RETURN OF SYRIAN REFUGEES	9	8
581	RESOLVE DIFFERENCES THROUGH DIALOGUE	9	9
582	PROVISIONS OF INTERNATIONAL LAW	9	8
583	PROMOTE THE IMPLEMENTATION OF	9	8
584	PROMOTE A POLITICAL SETTLEMENT	9	9
585	PRIME MINISTER AND MINISTER	9	9
586	PRIMARY RESPONSIBILITY FOR THE	9	8
587	PRESIDENT XI JINPING ANNOUNCED	9	9
588	POSITION ON THE ISSUE	9	9
589	PLAYING A CONSTRUCTIVE ROLE	9	9
590	PEACE IN THE REGION	9	9
591	PEACE AND SUSTAINABLE DEVELOPMENT	9	8
592	PAY CLOSE ATTENTION TO	9	9
593	PARTS OF THE COUNTRY	9	9
594	PARTIES IN THE COUNTRY	9	9
595	PALESTINE IS AT THE	9	9
596	ON DRUGS AND CRIME	9	8
597	OFFICE ON DRUGS AND	9	8
598	NATIONS OFFICE ON DRUGS	9	8
599	NATIONS HYBRID OPERATION IN	9	9
600	MINISTER AND MINISTER FOR	9	9
601	MIDDLE EAST AND THE	9	8
602	MEASURES TAKEN BY THE	9	9
603	MAINTAINING PEACE AND SECURITY	9	9
604	LIKE TO THANK MR	9	9
605	LIKE TO BEGIN BY	9	9
606	ITS SUPPORT FOR THE	9	9
607	ITS ROLE AS THE	9	9
608	IS DEEPLY CONCERNED ABOUT	9	9
609	INTEGRITY OF ALL COUNTRIES	9	9
610	HYBRID OPERATION IN DARFUR	9	9

611	HOPES THAT THE PARTIES	9	9
612	HOPES THAT ALL PARTIES	9	9
613	HOPE THAT ALL PARTIES	9	9
614	HAS ALWAYS SUPPORTED THE	9	8
615	GREAT IMPORTANCE TO THE	9	8
616	FOR THEIR RESPECTIVE BRIEFINGS	9	9
617	EXPLOSIVE REMNANTS OF WAR	9	9
618	EFFORTS TO FIND A	9	9
619	EFFORTS OF ALL PARTIES	9	9
620	DIALOGUE AND CONSULTATION AND	9	9
621	DEVELOPMENTS ON THE GROUND	9	9
622	CONTINUE TO PROVIDE ASSISTANCE	9	9
623	COMMUNITY OF A SHARED	9	9
624	CHINA WELCOMES THE EFFORTS	9	9
625	CHINA SUPPORTS THE EFFORTS	9	8
626	CHINA IS WILLING TO	9	9
627	CHINA HOPES THAT ALL	9	9
628	CHINA HAS NOTED THAT	9	9
629	CHINA HAS BEEN CLOSELY	9	9
630	CHEMICAL WEAPONS BY ANY	9	9
631	CALL ON THE PARTIES	9	9
632	BUILD A COMMUNITY OF	9	9
633	BILATERAL AND MULTILATERAL CHANNELS	9	9
634	BEST OF ITS ABILITY	9	9
635	BEEN CLOSELY FOLLOWING THE	9	9
636	BEARS THE PRIMARY RESPONSIBILITY	9	9
637	AUTHORITY ON DEVELOPMENT IGAD	9	9
638	AT THE HEART OF	9	9
639	AT A CRITICAL JUNCTURE	9	9
640	ASSUMING THE PRESIDENCY OF	9	9
641	ASSISTANCE MISSION IN AFGHANISTAN	9	9
642	AS ONE OF THE	9	9
643	ADHERE TO THE PRINCIPLE	9	8
644	ACT IN ACCORDANCE WITH	9	9
645	ABSTAINED IN THE VOTING	9	8
646	A JOINT EFFORT TO	9	9
647	A COMMUNITY OF SHARED	9	9

648	WOULD LIKE TO EXPRESS	8	8
649	WOULD LIKE TO BEGIN	8	8
650	WITH THE RELEVANT PARTIES	8	8
651	WITH THE RELEVANT COUNCIL	8	8
652	WITH THE COUNTRIES CONCERNED	8	8
653	WITH ALL THE PARTIES	8	8
654	WITH ALL PARTIES TO	8	8
655	WITH ALL PARTIES IN	8	8
656	WILL CONTINUE TO TAKE	8	8
657	WILL CONTINUE TO ACTIVELY	8	8
658	WE LOOK FORWARD TO	8	8
659	UNITED NATIONS MISSION TO	8	8
660	UNION AND THE LEAGUE	8	8
661	UNANIMOUS ADOPTION OF RESOLUTION	8	8
662	THAT THE SITUATION IN	8	8
663	TERRORIST FIGHTERS AND THEIR	8	8
664	STRENGTHEN ITS SECURITY CAPACITY	8	8
665	STRENGTHEN COORDINATION WITH THE	8	8
666	STABILITY AND SUSTAINABLE DEVELOPMENT	8	8
667	STABILITY AND PROSPERITY IN	8	8
668	SOLUTION TO THE PALESTINIAN	8	8
669	SHOULD STRENGTHEN COORDINATION AND	8	8
670	SHOULD PROVIDE CONSTRUCTIVE ASSISTANCE	8	8
671	SHOULD CONTINUE TO STRENGTHEN	8	8
672	SHOULD ADHERE TO THE	8	8
673	SETTLEMENT TO THE ISSUE	8	8
674	SETTLEMENT OF HOTSPOT ISSUES	8	8
675	RESPONSIBILITY OF THE INTERNATIONAL	8	8
676	RESPONSIBILITY FOR THE MAINTENANCE	8	8
677	REMAIN SEIZED OF THE	8	8
678	REMAIN COMMITTED TO THE	8	8
679	PROVIDE SUPPORT AND ASSISTANCE	8	8
680	PROCESS IN THE COUNTRY	8	8
681	PRESIDING OVER TODAY'S MEETING	8	8
682	POLITICAL AND SECURITY SITUATION	8	8
683	PLAY A POSITIVE AND	8	8
684	ONLY WAY TO RESOLVE	8	8

685	ON THE ONE HAND	8	8
686	ON THE MIDDLE EAST	8	8
687	ON THE FOLLOWING THREE	8	8
688	NATIONS PEACEKEEPING OPERATIONS IN	8	8
689	NATIONS MISSION TO SUPPORT	8	8
690	NATIONS AND THE BASIC	8	8
691	NATIONAL SECURITY AND STABILITY	8	8
692	MISSION TO SUPPORT THE	8	8
693	MISSION IN AFGHANISTAN UNAMA	8	8
694	MEDIATION AND GOOD OFFICES	8	8
695	MEASURES TO PREVENT VIOLENCE	8	8
696	LOWCOCK FOR THEIR BRIEFINGS	8	8
697	LIKE TO THANK YOU	8	8
698	LEVERAGE THE ROLE OF	8	8
699	LEGITIMATE CONCERNS OF ALL	8	8
700	LAST BUT NOT LEAST	8	8
701	JOINT EFFORTS OF THE	8	8
702	INTERNATIONAL COMMUNITY SHOULD TAKE	8	8
703	INTERNATIONAL COMMUNITY SHOULD BE	8	8
704	INTEGRITY OF THE COUNTRIES	8	8
705	I WOULD FIRST LIKE	8	8
706	HUMANITARIAN SITUATION OF THE	8	8
707	HUMANITARIAN AND ECONOMIC ASSISTANCE	8	8
708	HOPE THAT THE MISSION	8	8
709	HAS BEEN CLOSELY FOLLOWING	8	8
710	HAS ALWAYS BELIEVED THAT	8	8
711	HAS ALWAYS BEEN COMMITTED	8	8
712	GUTERRES FOR HIS BRIEFING	8	8
713	GENERAL LOWCOCK FOR HIS	8	8
714	FULLY RESPECT THE LEADERSHIP	8	8
715	FULL PLAY TO THE	8	8
716	FOREIGN TERRORIST FIGHTERS AND	8	8
717	FOR THE SYRIAN PEOPLE	8	8
718	FOR PEACE AND SECURITY	8	8
719	ESTABLISHMENT OF THE CONSTITUTIONAL	8	8
720	EMPHASIZE THE FOLLOWING POINTS	8	8
721	EFFORTS TO PROMOTE THE	8	8

722	EFFORTS TO IMPLEMENT THE	8	8
723	EFFORTS OF THE INTERNATIONAL	8	8
724	EFFORTS OF THE GOVERNMENT	8	8
725	EFFORTS MADE BY THE	8	8
726	DIALOGUE AND CONSULTATIONS AND	8	8
727	DEPUTY PRIME MINISTER AND	8	8
728	COORDINATION AND COOPERATION WITH	8	8
729	COOPERATION IN ORDER TO	8	8
730	CONTINUE TO HELP THE	8	8
731	CONCERTED EFFORTS OF THE	8	8
732	COMMUNITY SHOULD FOCUS ON	8	8
733	COMMITTED TO A POLITICAL	8	8
734	CHINA WISHES TO THANK	8	8
735	CHINA THANKS SPECIAL REPRESENTATIVE	8	8
736	CHINA HAS TAKEN NOTE	8	8
737	CHINA HAS ALWAYS BELIEVED	8	8
738	CHINA ABSTAINED IN THE	8	8
739	CEASE ALL SETTLEMENT ACTIVITIES	8	8
740	CAREFULLY TO THE STATEMENTS	8	8
741	BY CHINESE PRESIDENT XI	8	8
742	BETWEEN THE COUNCIL AND	8	8
743	BEST OF OUR ABILITY	8	8
744	ASSUMPTION OF THE PRESIDENCY	8	8
745	AS TO REACH A	8	8
746	AS PRESIDENT OF THE	8	8
747	AN IMPORTANT COUNTRY IN	8	8
748	ALWAYS BEEN COMMITTED TO	8	8
749	ALL THE PARTIES IN	8	8
750	ALL PARTIES CONCERNED TO	8	8
751	ADVANCE THE POLITICAL SETTLEMENT	8	8
752	ADDRESSING THE ROOT CAUSES	8	8
753	ACHIEVING PEACE AND STABILITY	8	8
754	A LARGE NUMBER OF	8	8
755	A COMMUNITY OF A	8	8

Appendix 2 Four-word lexical bundles retrieved from L1O in Wordsmith

WordSmith Tools 8.0 Word list (WL_index.tokens)			
N	**Word**	**Freq.**	**Texts**
1	I WOULD LIKE TO	235	131
2	REPRESENTATIVE OF THE SECRETARY	86	56
3	SPECIAL REPRESENTATIVE OF THE	84	54
4	REPUBLIC OF THE CONGO	78	19
5	DEMOCRATIC REPUBLIC OF THE	78	19
6	INTERNATIONAL PEACE AND SECURITY	55	44
7	WOULD ALSO LIKE TO	54	50
8	I WOULD ALSO LIKE	51	47
9	USE OF CHEMICAL WEAPONS	46	19
10	UNITED KINGDOM WELCOMES THE	40	34
11	WOULD LIKE TO THANK	37	34
12	MEMBERS OF THE COUNCIL	35	33
13	WITH REGARD TO THE	31	28
14	UNITED NATIONS AND THE	31	29
15	WITH THE UNITED NATIONS	30	22
16	TAKE THIS OPPORTUNITY TO	29	28
17	SEXUAL VIOLENCE IN CONFLICT	29	15
18	IT IS IMPORTANT THAT	28	25
19	WILL BE ABLE TO	27	21
20	IT IS VITAL THAT	26	23
21	INTERNATIONAL HUMANITARIAN LAW AND	25	20
22	AS WELL AS THE	25	23
23	UNITED KINGDOM WILL CONTINUE	24	23
24	SITUATION ON THE GROUND	24	22
25	ROLE TO PLAY IN	24	24
26	FOR THE PEOPLE OF	24	21
27	FOR THE COORDINATION OF	24	23
28	CHARTER OF THE UNITED	24	19
29	UNITED NATIONS HIGH COMMISSIONER	23	20

30	PROHIBITION OF CHEMICAL WEAPONS	23	18
31	ORGANIZATION FOR THE PROHIBITION	23	18
32	OFFICE FOR THE COORDINATION	23	22
33	NATIONS HIGH COMMISSIONER FOR	23	20
34	KINGDOM WILL CONTINUE TO	23	22
35	FOR THE PROHIBITION OF	23	18
36	COORDINATION OF HUMANITARIAN AFFAIRS	23	22
37	ALSO LIKE TO THANK	22	21
38	MEMBERS OF THE SECURITY	21	19
39	AT THE SAME TIME	21	20
40	WOMEN AND PEACE AND	20	17
41	REPORT OF THE SECRETARY	20	20
42	LIKE TO THANK THE	20	19
43	CONTINUE TO SUPPORT THE	20	20
44	BY THE UNITED NATIONS	20	20
45	UNITED NATIONS MULTIDIMENSIONAL INTEGRATED	18	16
46	PER CENT OF THE	18	16
47	ON BEHALF OF THE	18	18
48	NATIONS MULTIDIMENSIONAL INTEGRATED STABILIZATION	18	16
49	MULTIDIMENSIONAL INTEGRATED STABILIZATION MISSION	18	16
50	IT IS CLEAR THAT	18	15
51	INTEGRATED STABILIZATION MISSION IN	18	16
52	HIGH COMMISSIONER FOR HUMAN	18	16
53	COMMISSIONER FOR HUMAN RIGHTS	18	16
54	WEAPONS OF MASS DESTRUCTION	17	11
55	PAY TRIBUTE TO THE	17	16
56	OFFICE OF THE UNITED	17	16
57	MEMBERS OF THE UNITED	17	17
58	LOOK FORWARD TO THE	17	15
59	IT IS VERY GOOD	17	14
60	I THINK THAT WE	17	14
61	I THINK IT IS	17	15
62	WOULD LIKE TO SAY	16	15
63	VERY MUCH WELCOME THE	16	13

64	LIKE TO TAKE THIS	16	15
65	I ALSO WANT TO	16	15
66	AS WE HAVE HEARD	16	14
67	ABOUT THE IMPORTANCE OF	16	15
68	WHY THE UNITED KINGDOM	15	14
69	UNITED NATIONS ASSISTANCE MISSION	15	14
70	THERE CAN BE NO	15	14
71	SUPPORT THE UNITED NATIONS	15	13
72	SET OUT IN THE	15	13
73	PARTIES TO THE CONFLICT	15	14
74	ON ALL PARTIES TO	15	14
75	IT IS IMPORTANT TO	15	14
76	FOR THE UNITED NATIONS	15	14
77	FOR THE UNITED KINGDOM	15	15
78	ESTABLISHED PURSUANT TO RESOLUTION	15	14
79	DEEPLY CONCERNED ABOUT THE	15	13
80	CALL ON ALL PARTIES	15	15
81	ALL MEMBERS OF THE	15	14
82	UNITED NATIONS HYBRID OPERATION	14	14
83	UNITED KINGDOM REMAINS COMMITTED	14	14
84	UNITED KINGDOM BELIEVES THAT	14	12
85	STABILIZATION MISSION IN MALI	14	13
86	PEACE AND SECURITY IN	14	12
87	ON THE IMPORTANCE OF	14	12
88	ON THE GROUND IN	14	14
89	NEED OF HUMANITARIAN ASSISTANCE	14	14
90	NATIONS HYBRID OPERATION IN	14	14
91	LET ME BEGIN BY	14	13
92	KINGDOM REMAINS COMMITTED TO	14	14
93	IS WHY THE UNITED	14	13
94	IMPLEMENTATION OF THE PEACE	14	11
95	HYBRID OPERATION IN DARFUR	14	14
96	FOR HIS BRIEFING AND	14	14
97	COMMITTEE ESTABLISHED PURSUANT TO	14	14
98	WILL CONTINUE TO SUPPORT	13	13
99	WHAT THE SPECIAL REPRESENTATIVE	13	8
100	STATES MEMBERS OF THE	13	13

101	PROGRESS THAT HAS BEEN	13	12
102	PREVENTING SEXUAL VIOLENCE IN	13	9
103	PEACE AND RECONCILIATION IN	13	13
104	ENVOY OF THE SECRETARY	13	13
105	ARE IN NEED OF	13	13
106	URGE ALL PARTIES TO	12	11
107	UNITED KINGDOM FULLY SUPPORTS	12	10
108	UNITED KINGDOM CONTINUES TO	12	11
109	THINK THAT IT IS	12	10
110	SOLUTION TO THE CONFLICT	12	11
111	ON THE IMPLEMENTATION OF	12	12
112	IT IS VERY IMPORTANT	12	12
113	IT IS ESSENTIAL THAT	12	11
114	IS IMPORTANT THAT WE	12	11
115	IMPLEMENTATION OF THE AGREEMENT	12	8
116	I THINK WE ALL	12	11
117	I THINK THAT IT	12	10
118	HUMAN RIGHTS VIOLATIONS AND	12	11
119	HAVE A RESPONSIBILITY TO	12	10
120	FREE AND FAIR ELECTIONS	12	9
121	BY THE GOVERNMENT OF	12	10
122	AS THE UNITED KINGDOM	12	10
123	AS SET OUT IN	12	12
124	AS A RESULT OF	12	10
125	AFRICAN UNION AND THE	12	11
126	WOULD LIKE TO REITERATE	11	10
127	WOULD LIKE TO BEGIN	11	11
128	WELCOME THE EFFORTS OF	11	11
129	VIOLATIONS OF INTERNATIONAL HUMANITARIAN	11	8
130	UNITED KINGDOM'S FULL SUPPORT	11	11
131	UNITED KINGDOM IS COMMITTED	11	10
132	STABILIZATION MISSION IN THE	11	10
133	PEOPLE OF THE SUDAN	11	8
134	ON THE SITUATION IN	11	11
135	LIKE TO BEGIN BY	11	11
136	KINGDOM IS COMMITTED TO	11	10

137	IS VITAL THAT THE	11	10
138	IS VERY IMPORTANT THAT	11	11
139	IMPORTANT ROLE TO PLAY	11	11
140	IF WE ARE TO	11	10
141	I THINK THAT THE	11	8
142	FOR THE COUNCIL TO	11	11
143	BY THE SYRIAN REGIME	11	9
144	BELIEVE THAT IT IS	11	11
145	ALL MEMBER STATES TO	11	10
146	A NUMBER OF SPEAKERS	11	9
147	WOULD THEREFORE LIKE TO	10	10
148	WOULD LIKE TO TAKE	10	10
149	WOULD LIKE TO START	10	10
150	WORK OF THE UNITED	10	10
151	WITH THE AFRICAN UNION	10	8
152	WILL CONTINUE TO WORK	10	9
153	VOTE IN FAVOUR OF	10	8
154	VIABLE AND SOVEREIGN PALESTINIAN	10	8
155	SECURITY COUNCIL AND THE	10	10
156	RULE OF LAW AND	10	10
157	OUR SUPPORT FOR THE	10	9
158	OPERATION IN DARFUR UNAMID	10	10
159	ON THE GROUND AND	10	10
160	ON THE COUNCIL'S AGENDA	10	10
161	NATIONS ASSISTANCE MISSION IN	10	10
162	LOOK FORWARD TO WORKING	10	10
163	LIKE TO THANK OUR	10	10
164	LIKE TO START BY	10	10
165	LET ME CONCLUDE BY	10	10
166	IS ONE OF THE	10	10
167	I WILL BE BRIEF	10	10
168	I AM SURE THAT	10	9
169	HIGH COMMISSIONER FOR REFUGEES	10	9
170	HAD TO SAY ABOUT	10	10
171	GROUP OF FIVE FOR	10	9
172	FROM THE UNITED NATIONS	10	9
173	FROM PEACEKEEPING TO PEACEBUILDING	10	8

174	FOR THE SECURITY COUNCIL	10	10
175	FOR THE SAKE OF	10	9
176	FOR THE IMPLEMENTATION OF	10	10
177	FIVE FOR THE SAHEL	10	9
178	BY THE SECURITY COUNCIL	10	10
179	AN IMPORTANT ROLE TO	10	10
180	ALL THE PARTIES TO	10	9
181	A VITAL ROLE IN	10	10
182	A VIABLE AND SOVEREIGN	10	8
183	WOULD LIKE TO WELCOME	9	9
184	WORKS AGENCY FOR PALESTINE	9	9
185	WITH THE SUPPORT OF	9	9
186	WITH INTERNATIONAL HUMANITARIAN LAW	9	8
187	UNITED KINGDOM'S SUPPORT FOR	9	9
188	UNITED KINGDOM IS PROUD	9	9
189	THERE NEEDS TO BE	9	8
190	SPECIAL ENVOY OF THE	9	9
191	SEXUAL EXPLOITATION AND ABUSE	9	8
192	RIGHTS VIOLATIONS AND ABUSES	9	8
193	RELIEF AND WORKS AGENCY	9	9
194	REFUGEES IN THE NEAR	9	9
195	PALESTINE REFUGEES IN THE	9	9
196	OTHER MEMBERS OF THE	9	9
197	ME BEGIN BY THANKING	9	9
198	LET ME START BY	9	9
199	KINGDOM'S FULL SUPPORT FOR	9	9
200	JOINT COMPREHENSIVE PLAN OF	9	9
201	JOIN OTHERS IN THANKING	9	9
202	IT IS RIGHT THAT	9	9
203	IS VERY GOOD TO	9	9
204	IS THE ONLY WAY	9	9
205	INTERNATIONAL COMMITTEE OF THE	9	9
206	I WOULD THEREFORE LIKE	9	9
207	I HOPE THAT WE	9	9
208	I ALSO THANK THE	9	9
209	HOPE THAT WE CAN	9	9
210	HAVE BEEN ABLE TO	9	9

211	HAS BEEN ABLE TO	9	9
212	HAS AN IMPORTANT ROLE	9	9
213	FULL IMPLEMENTATION OF THE	9	9
214	FOR THE GOVERNMENT OF	9	9
215	FOR PEACE AND RECONCILIATION	9	8
216	FOR PALESTINE REFUGEES IN	9	9
217	ECONOMIC COMMUNITY OF WEST	9	9
218	COMPREHENSIVE PLAN OF ACTION	9	9
219	COMMUNITY OF WEST AFRICAN	9	9
220	COMMITTEE OF THE RED	9	9
221	CAUSES OF THE CONFLICT	9	9
222	BEHALF OF THE UNITED	9	9
223	AS WE HEARD FROM	9	8
224	ARE HELD TO ACCOUNT	9	8
225	ALL PARTIES TO THE	9	9
226	AGENCY FOR PALESTINE REFUGEES	9	9
227	WOULD LIKE TO HIGHLIGHT	8	8
228	WITH THE SPECIAL ENVOY	8	8
229	WEST AFRICA AND THE	8	8
230	UNITED NATIONS RELIEF AND	8	8
231	UNITED NATIONS OFFICE FOR	8	8
232	UNITED NATIONS JOINT INVESTIGATIVE	8	8
233	UNITED KINGDOM IS DEEPLY	8	8
234	UNITED KINGDOM CALLS ON	8	8
235	UNION MISSION IN SOMALIA	8	8
236	THIS OPPORTUNITY TO THANK	8	8
237	THIS IS THE FIRST	8	8
238	THANK SPECIAL REPRESENTATIVE OF	8	8
239	THANK HIM FOR HIS	8	8
240	SUPPORT OF THE UNITED	8	8
241	STATE IN IRAQ AND	8	8
242	SECURITY AND COOPERATION IN	8	8
243	ROOT CAUSES OF THE	8	8
244	REVITALIZED AGREEMENT ON THE	8	8
245	RESOLUTION OF THE CONFLICT	8	8
246	PEACE AND STABILITY IN	8	8
247	PARTICIPATION OF WOMEN IN	8	8

248	OVER THE PAST FEW	8	8
249	ORGANIZATION FOR SECURITY AND	8	8
250	ON WOMEN AND PEACE	8	8
251	ON THE BASIS OF	8	8
252	NATIONS RELIEF AND WORKS	8	8
253	NATIONS JOINT INVESTIGATIVE MECHANISM	8	8
254	MISSION IN MALI MINUSMA	8	8
255	LOOK FORWARD TO HEARING	8	8
256	LIKE TO WELCOME THE	8	8
257	LET ME SAY THAT	8	8
258	KINGDOM IS DEEPLY CONCERNED	8	8
259	IT IS CRUCIAL THAT	8	8
260	IT IS ALSO IMPORTANT	8	8
261	ISLAMIC STATE IN IRAQ	8	8
262	IS DEEPLY CONCERNED ABOUT	8	8
263	I SHOULD LIKE TO	8	8
264	I JOIN OTHERS IN	8	8
265	FORWARD TO WORKING WITH	8	8
266	FOR THEIR BRIEFINGS TODAY	8	8
267	FOR SECURITY AND COOPERATION	8	8
268	EFFORTS TO ENSURE THAT	8	8
269	DOING ON THE GROUND	8	8
270	CONTINUE TO WORK WITH	8	8
271	CALL ON THE GOVERNMENT	8	8
272	BUT IT IS ALSO	8	8
273	AT THE HEART OF	8	8
274	ASSISTANCE MISSION IN AFGHANISTAN	8	8
275	AS WE HAVE SAID	8	8
276	AS THIS IS THE	8	8
277	ALL OF US TO	8	8
278	AGREEMENT ON THE RESOLUTION	8	8
279	AFRICAN UNION MISSION IN	8	8
280	AFRICA AND THE SAHEL	8	8
281	A MEMBER OF THE	8	8

Appendix 3 LBs list analysed in L2I (partial)

N	Lexical bundle	Freq. in Comparable corpus	Freq. in parallel corpus
1	AS SOON AS POSSIBLE	163	84
2	I WOULD LIKE TO	130	61
3	PLAY A CONSTRUCTIVE ROLE	108	54
4	WITH A VIEW TO	97	58
5	WE HOPE THAT THE	88	48
6	AT THE SAME TIME	84	55
7	CHINA WILL CONTINUE TO	76	33
8	CHINA STANDS READY TO	73	34
9	ON THE BASIS OF	71	35
10	READY TO WORK WITH	55	28
11	CHINA WOULD LIKE TO	51	26
12	IT IS IMPORTANT TO	50	21
13	IT IS NECESSARY TO	44	22
14	AS WELL AS THE	42	24
15	SHOULD CONTINUE TO SUPPORT	39	19
16	CHINA IS READY TO	38	19
17	SITUATION ON THE GROUND	36	23
18	IT IS IMPERATIVE TO	35	24
19	CHINA HOPES THAT THE	33	10
20	WITH REGARD TO THE	31	18
21	WE CALL ON THE	29	15
22	ON THE ISSUE OF	29	22
23	IS THE ONLY WAY	27	7
24	WE SHOULD CONTINUE TO	25	22
25	PLAYED AN IMPORTANT ROLE	25	14
26	WE WILL CONTINUE TO	24	13
27	EFFORTS SHOULD BE MADE	24	16
28	SITUATION IN THE COUNTRY	23	12
29	REST OF THE INTERNATIONAL	23	10
30	LISTENED CAREFULLY TO THE	23	12
31	AT AN EARLY DATE	23	8
32	THERE IS A NEED	22	11

33	CONSTRUCTIVE ROLE IN THE	22	10
34	A SHARED FUTURE FOR	22	9
35	SOLUTION TO THE ISSUE	21	10
36	CONSTRUCTIVE ROLE IN ACHIEVING	21	10
37	CHINA CALLS ON THE	21	13
38	AS THE MAIN CHANNEL	21	6
39	WE HOPE THAT ALL	20	10
40	THAT IS ACCEPTABLE TO	20	8
41	RESOLVE THEIR DIFFERENCES THROUGH	20	9
42	LISTENED ATTENTIVELY TO THE	20	10
43	A POLITICAL SETTLEMENT TO	20	6
44	WITH A SHARED FUTURE	19	7
45	WE CALL ON ALL	19	9
46	ROOT CAUSES OF CONFLICT	19	16
47	INTERESTS OF THE COUNTRY	19	10
48	I WISH TO THANK	19	14
49	WITHIN THE FRAMEWORK OF	18	6
50	STABILITY IN THE COUNTRY	18	7
51	SO AS TO ACHIEVE	18	9
52	PLAY AN ACTIVE ROLE	18	7
53	CONCERNS OF ALL PARTIES	18	2
54	CHINA HAS ALWAYS SUPPORTED	18	13
55	TAKEN NOTE OF THE	17	11
56	SUPPORT THE WORK OF	17	5
57	STRENGTHEN COMMUNICATION AND COORDINATION	17	14
58	SO THAT THEY CAN	17	9
59	ON THE QUESTION OF	17	11
60	WOULD ALSO LIKE TO	16	3
61	WILL CONTINUE TO PLAY	16	9
62	WE MUST CONTINUE TO	16	4
63	SO THAT IT CAN	16	12
64	SHOULD BE GIVEN TO	16	10
65	PROMOTE THE POLITICAL PROCESS	16	7
66	PROGRESS HAS BEEN MADE	16	9
67	ORGANIZATIONS SUCH AS THE	16	4

68	JOINT FORCE OF THE	16	9
69	FOR THE MAINTENANCE OF	16	4
70	FOR A POLITICAL SETTLEMENT	16	8
71	COUNCIL SHOULD CONTINUE TO	16	5
72	BETWEEN THE TWO SIDES	12	4
73	BOTH THE SYMPTOMS AND	11	6
74	COOPERATION BETWEEN THE TWO	10	7
75	WE BELIEVE THAT THE	13	4
76	CHINA BELIEVES THAT THE	13	6
77	HAS ALWAYS BELIEVED THAT	8	3

References

Aarts, J. M., & Meijs, W. (1984). *Corpus linguistics: Recent developments in the use of computer corpora in English language research*: Rodopi.

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for specific purposes (New York, N.Y.), 31*(2), 81-92. doi:10.1016/j.esp.2011.08.004

Adolphs, S. (2006). *Introducing electronic text analysis: A practical guide for language and literary studies.* New York: Routledge.

Aijmer, K., Altenberg, B., & Johansson, M. (1996). *Languages in contrast: Papers from a symposium on text-based cross-linguistic studies, Lund 4-5 March 1994* (Vol. 88): Lund studies in English.

Albl-Mikasa, M., & Tiselius, E. (2021). *The Routledge handbook of conference interpreting.* New York: Routledge.

Altenberg, B. (1993). *Recurrent word combinations in spoken English.* Paper presented at the Proceedings of the fifth Nordic association for English studies conference.

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and application* (pp. 101-122). Oxford: Clarendon Press.

Altenberg, B., & Granger, S. (2002). Recent trends in cross-linguistic lexical studies. *Lexis in contrast: Corpus-based approaches*, 3-48.

Alves, F., & Gonçalves, J. L. V. (2003). A relevance theory approach to the investigation of inferential processes in translation. *Triangulating translation: perspectives in process oriented research*, 3-24.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of memory and language, 62*(1), 67-82. doi:10.1016/j.jml.2009.09.005

Aston, G. (2018). Acquiring the language of interpreters: A Corpus-based Approach. In *Making way in corpus-based interpreting studies* (pp. 83-96). Singapore: Springer.

Baayen, R. H. (2001). *Word frequency distributions* (Vol. 18): Springer Science & Business Media.

Baigorri-Jalón, J., & Travieso-Rodríguez, C. (2017). Interpreting at the United Nations: the impact of external variables. The Interpreters' View. *CLINA: An Interdisciplinary Journal of Translation, Interpreting and Intercultural Communication3 (2)*, 53-72.

Baker, M. (1993). Corpus Linguistics and Translation Studies: Implications and Applications. In M. Baker, G. Francis, E. Tognini-Bonelli, & J. M. Sinclair (Eds.), *Text and technology: In honour of John Sinclair* (pp. 233-250). Amsterdam: John Benjamins

Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target. International journal of translation studies, 7*(2), 223-243.

Baker, M. (1996). Corpus-based Translation Studies: The Challenges that Lie Ahead. In H. Somers (Ed.), *Terminology, LSP, and Translation: Studies in Language Engineering in Houour of Juan C. Sager* (pp. 175-186). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Baker, M. (2004). A corpus-based view of similarity and difference in translation. *International journal of corpus linguistics, 9*(2), 167-193. doi:10.1075/ijcl.9.2.02bak

Baker, M. (2007). Patterns of Idiomaticity in Translated vs: Non-Translated Text. *Belgian journal of linguistics, 21*(1).

Baker, P. (2010). *Sociolinguistics and corpus linguistics*: Edinburgh University Press.

Bao, A. (2020, June 08). 联合国译员训练班：新中国第一代同传译员的诞生 (United Nations Interpreter Training Course: The Birth of the First Generation of Simultaneous Interpreters in New China). *中国新闻周刊 (China Newsweek).*

Bartłomiejczyk, M. (2006). Strategies of simultaneous interpreting and directionality. *INTERPRETING, 8*(2), 149-174.

Bartłomiejczyk, M. (2015). Directionality. *Routledge Encyclopedia of Interpreting Studies, London/New York: Routledge*, 108-110.

Bartłomiejczyk, M., & Stachowiak-Szymczak, K. (2021). *Modes of conference interpreting: Simultaneous and consecutive.* Paper presented at the The Routledge Handbook of Conference Interpreting.

Becker, J. D. (1975). *The phrasal lexicon.* Paper presented at the Theoretical issues in natural language processing, Cambridge.

Bendazzoli, C. (2010). The European Parliament as a source of material for research into simultaneous interpreting: advantages and limitations. In N. L. Zybatow (Ed.), *Translationswissenschaft – Stand und Perspektiven. Innsbrucker Ringvorlesungenzur Translationswissenschaft VI (Forum Translationswissenschaft, Bd. 12)* (Vol. 12, pp. 51-68). Frankfurt am Main: Peter Lang.

Bendazzoli, C. (2018). Corpus-based Interpreting Studies: Past, Present and Future Developments of a (Wired) Cottage Industry. In R. M., B. C., & D. B. (Eds.), *Making Way in Corpus-based Interpreting Studies. New Frontiers in*

Translation Studies. (pp. 1-19). Singapore: Springer.

Bendazzoli, C., Sandrelli, A., & Russo, M. (2011). Disfluencies in simultaneous interpreting: A corpus-based analysis. *Corpus-based translation studies: Research and applications*, 282-306.

Bernardini, S. (2011). Monolingual comparable corpora and parallel corpora in the search for features of translated language. *SYNSPS: A journal of Professional Communication*, 26:2-13.

Bernardini, S., Ferraresi, A., & Miličević, M. (2016). From EPIC to EPTIC - Exploring simplification in interpreting and translation from an intermodal perspective. *Target. International journal of translation studies, 28*(1), 61-86. doi:10.1075/target.28.1.03ber

Bernardini, S., & Russo, M. (2017). Corpus linguistics, translation and interpreting. In *The Routledge Handbook of Translation Studies and Linguistics* (pp. 342-356). London/New York: Routledge.

Berūkštienė, D. (2017). A corpus-driven analysis of structural types of lexical bundles in court judgments in English and their translation into Lithuanian. *Linguistics, 70*(70), 7-31. doi:10.15388/Klbt.2017.11181

Bestgen, Y. (2018). Evaluating the frequency threshold for selecting lexical bundles by means of an extension of the Fisher's exact test. *Corpora, 13*(2), 205-228.

Bestgen, Y. (2020). Comparing lexical bundles across corpora of different sizes: The Zipfian problem. *Journal of quantitative linguistics, 27*(3), 272-290.

Biber, D. (1993). Representativeness in corpus design. *Literary and linguistic computing, 8*(4), 243-257.

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International journal of corpus linguistics, 14*(3), 275-311.

Biber, D. (2010). Corpus-based and corpus-driven analyses of language variation and use. In H. Narrog (Ed.), *The Oxford handbook of linguistic analysis*.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes (New York, N.Y.), 26*(3), 263-286. doi:10.1016/j.esp.2006.08.003

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied linguistics, 25*(3), 371-405. doi:10.1093/applin/25.3.371

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999a). Lexical bundles in conversation and academic prose. *Language and computers, 26*, 181-190.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999b). *Longman grammar of spoken and written English*. Essex: Longman.

Biel, Ł. (2016). Mixed corpus design for researching the Eurolect: A genre-based comparable-parallel corpus in the PL EUROLECT project.

Biel, Ł. (2018). Lexical bundles in EU law The impact of translation process on the patterning of legal language. In G.-R. Stanisław & G. Pontrandolfo (Eds.), *Phraseology in Legal and Institutional Settings: A Corpus-based interdisciplinary perspective* (pp. 11-26). Abingdon: Routledge.

Biel, Ł., Koźbiał, D., & Wasilewska, K. (2019). The formulaicity of translations across EU institutional genres: A corpus-driven analysis of lexical bundles in translated and non-translated language. *Translation spaces, 8*(1), 67-92. doi:10.1075/ts.00013.bie

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language teaching research, 10*(3), 245-261.

Breeze, R. (2013). Lexical bundles across four legal genres. *International journal of corpus linguistics, 18*(2), 229-253. doi:10.1075/ijcl.18.2.03bre

Buendía, C. T. (2010). Community interpreting: breaking with the 'norm'through normalisation. *Journal of Specialized Translation, 14*, 11-12.

Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for academic purposes, 30*, 38-52.

Cabrera Castro, T. (2017). Applications of Simultaneous-Interpreting Corpora in Terminology Research. *New Voices in Translation Studies*(17).

Cammoun, R., Davies, C., Ivanov, K., & Naimushin, B. (2009). Simultaneous interpreting with text. Is the text "friend" or "foe". *Unpublished Master Thesis, FTI, University of Geneva*.

Campoy, M. C., & Luzón, M. J. (2007). *Spoken corpora in applied linguistics* (Vol. 51): Peter Lang.

Cao, F. (2021). A comparative study of lexical bundles across paradigms and disciplines. *Corpora, 16*(1), 97-128.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English : a comprehensive guide: spoken and written English grammar and usage*. Cambridge: Cambridge University Press.

Chang, C.-c., & Schallert, D. L. (2007). The impact of directionality on Chinese/English simultaneous interpreting. *INTERPRETING, 9*(2), 137-176.

Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language learning & technology, 14*(2), 30-49.

Chen, Y.-H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied linguistics, 37*(6), 849-880.

doi:10.1093/applin/amu065

Cheung, A. K. F. (2012). *Omission in Simultaneous Interpreting: Word order differences. FORUM. Revue internationale d'interprétation et de traduction/International journal of interpretation and translation*, *10*(2), 19-33. doi:10.1075/forum.10.2.02che

Chmiel, A., Janikowski, P., & Lijewska, A. (2020). Multimodal processing in simultaneous interpreting with text: Interpreters focus more on the visual than the auditory modality. *Target. International Journal of Translation Studies, 32*(1), 37-58. doi: 10.1075/target.18157.chm

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied linguistics, 29*(1), 72-89. doi: 10.1093/applin/amm022

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual. review of applied linguistics, 32*, 45-61. doi:10.1017/S0267190512000074

Conrad, S. M., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica*.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for specific purposes, 23*(4), 397-423. doi:10.1016/j.esp.2003.12.001

Cortes, V. (2012). Lexical Bundles and Technology. *The Encyclopedia of Applied Linguistics*.

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for academic purposes, 12*(1), 33-43. doi: 10.1016/j.jeap.2012.11.002

Cortes, V. (2015). Situating lexical bundles in the formulaic language spectrum: Origins and functional analysis developments. In V. Cortes & E. Csomay

(Eds.), *Corpus-based research in applied linguistics* (pp. 197-216): John Benjamins.

Dam, H. V. (2001). On the option between form-based and meaning-based interpreting: The effect of source text difficulty on lexical target text form in simultaneous interpreting. *The interpreters' Newsletter*, 11, 27-55.

Davitti, E., & Pasquandrea, S. (2014). Enhancing research-led interpreter education: an exploratory study in Applied Conversation Analysis. *The interpreter and translator trainer, 8*(3), 374-398.

Dawrant, A. (1996). Word order in Chinese-English simultaneous interpretation: An initial exploration. *Unpublished MA thesis, Fu Jen University*.

Dayter, D. (2018). Describing lexical patterns in simultaneously interpreted discourse in a parallel aligned corpus of Russian-English interpreting (SIREN). *Forum, 16*(2), 241-264. Doi:10.1075/forum.17004.day

Dayter, D. (2020). Strategies in a corpus of simultaneous interpreting. Effects of directionality, phraseological richness, and position in speech event. *Meta: journal des traducteurs/Meta: translators' journal, 65*(3), 594-617. Doi: 10.7202/1077405a

De Cock, S., & Granger, S. (2021). Stance in press releases versus business news: a lexical bundle approach. *Text & talk, 41*(5-6), 691-713. doi: 10.1515/text-2020-0040

De Sutter, G., & Lefer, M.-A. (2020a). On the need for a new research agenda for corpus-based translation studies: a multi-methodological, multifactorial and interdisciplinary approach. *Perspectives, studies in translatology, 28*(1), 1-23. doi:10.1080/0907676X.2019.1611891

De Sutter, G., & Lefer, M.-A. (2020b). On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and

interdisciplinary approach. *Perspectives, 28*(1), 1-23. doi: 10.1080/09076X.2019.1611891

Defrancq, B. (2018). The European Parliament as a discourse community: its role in comparable analyses of data drawn from parallel interpreting corpora. *The intepreters' newsletter*, 23, 115-132.

Defrancq, B., & Collard, C. (2020). Using data from simultaneous interpreting in contrastive linguistics. In *New Approaches to Contrastive Linguistics* (pp. 159-182). Berlin: De Gruyter Mouton.

Defrancq, B., De Clerck, B., & De Sutter, G. (2015). Corpus-based translation studies: Across genres, methods and disciplines. *Across languages and cultures, 16*(2), 157-162. doi: 10.1556/084.2015.16.2.1

Defrancq, B., & Plevoets, K. (2022a). Linguistic convergence in the European Parliament: a correspondence analysis of n-grams used by members of parliament and interpreters. *Bandung (Leiden)*.

Defrancq, B., & Plevoets, K. (2022b). Ut interpres: Linguistic convergence between orators and interpreters in the European Parliament. *Mediated discourse at the European Parliament*, 1.

Delaere, I., De Sutter, G., & Plevoets, K. (2012). Is translated language more standardized than non-translated language?: Using profile-based correspondence analysis for measuring linguistic distances between language varieties. *Target. International journal of translation studies, 24*(2), 203-224. doi: 10.1075/target.24.2.01del

Diur, M. (2015). *Interpreting at the United Nations: an empirical study on the Language Competitive Examination (LCE).* Universidad Pablo de Olavide,

Donato, V. (2003). Strategies adopted by student interpreters in SI: A comparison between the English-Italian and the German-Italian language-pairs. The

Interpterters' Newsletter 12, 101-134.

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied linguistics, 38*(2), 165-193. doi:10.1093/applin.amv011

Ebeling, J., & Ebeling, S. O. (2018). Comparing n-gram-based functional categories in original versus translated texts. *Corpora, 13*(3), 347-370.

Ebeling, S. O., & Ebeling, J. (2017). A functional comparison of recurrent word-combinations in English original vs. translated texts. *ICAME journal, 41*(1), 31-52. doi:10.1515/icame-2017-0002

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual review of applied linguistics, 32*, 17-44. doi: 10.1017/S0267190512000025

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & talk, 20*(1), 29-62. doi: 10.1515/text.1.2000.20.1.29

Eyckmans, J. (2007). Taking SLA research to interpreting: does knowledge of phrases foster fluency? In *Multilingualism and Applied Comparative Linguistics, volume 1: pedagogical perspectives* (pp. 89-105): Cambridge Scholar Publishing.

Ferraresi, A., Bernardini, S., Petrović, M., & Lefer, M.-A. (2018). Simplified or not simplified? The different guises of mediated English at the European Parliament. *Meta: journal des traducteurs/Meta: Translators' Journal, 63*(3), 717-738. doi: 10.7202/1060170ar

Ferraresi, A., & Miličević, M. (2017). Phraseological patterns in interpreting and translation. Similar or different? In *Empirical Translation Studies* (Vol. 300, pp. 157-182). Berlin/Boston: De Gruyter.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic*

analysis.

Frawley, W. (1984). *Translation: literary, linguistic, and philosophical perspective*. Newark : University of Delaware Press.

Gile, D. (1995/2009). *Basic concepts and models for interpreter and translator training Revised edition* (Vol. 8). Amsterdam / Philadelphia: John Benjamins Publishing Company.

Gile, D. (2005a). Directionality in conference interpreting: A cognitive view. *Communication & cognition. Monographies, 38*(1-2), 9-26.

Gile, D. (2005b). interpreting: Methodological aspects. *Knowledge systems and translation, 7*, 149.

Granger, S. (2014). A lexical bundle approach to comparing languages: Stems in English and French. *Languages in contrast, 14*(1), 58-72. Doi: 10.1075/lic.14.1.04gra

Gumul, E. (2006). Explicitation in simultaneous interpreting: A strategy or a by-product of language mediation? *Across languages and cultures, 7*(2), 171-190.

Gumul, E. (2017). Explicitation and directionality in simultaneous interpreting. *Lisguistica Silesiana, 38*, 311-329.

Gumul, E. (2021). Explicitation and cognitive load in simultaneous interpreting: Product-and process-oriented analysis of trainee interpreters' outputs. *Interpreting, 23*(1), 45-75.

Guz, E. (2014). Formulaic sequences as fluency devices in the oral production of native speakers of Polish. *Research in language, 12*(2), 113-129.

Han, C. (2018). Mixed-methods research in interpreting studies: A methodological review (2004–2014). *Interpreting, 20*(2), 155-187.

Hansen, G. (2010). Integrative description of translation processes. *Shreve, GM & E. Angelone (eds.)*, 189-212.

Henriksen, L. (2007). The song in the booth: Formulaic interpreting and oral textualisation. *Interpreting : international journal of research and practice in interpreting, 9*(1), 1-20. doi:10.1075/intp.9.1.02hen

Hertog, E., Van Gucht, J., & de Bontridder, L. (2006). Musings on methodology. *Linguistica Antverpiensia, New Series–Themes in Translation Studies, 5*, 121-132.

Hild, A. (2004). Establishing rigour in a between-methods investigation of SI expertise. In Y. Gambier, M. Shlesinger, & R. Stolze (Eds), *Doubts and Directions in Translation Studies*, *9*(1), 99-115.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied linguistics, 19*(1), 24-44.

Hu, K., & Tao, Q. (2013). The Chinese-English conference interpreting corpus: Uses and limitations. *Meta: journal des traducteurs/Meta: Translators' Journal, 58*(3), 626-642.

Hudson, J. (1998). *Perspectives on fixedness: applied and theoretical* (Vol. 94). Lund: University Press Lund.

Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International journal of applied linguistics, 18*(1), 41-62. doi: 10.1111/j.1473-4192.2008.00178.x

Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes, 27*(1), 4-21. doi:10.1016/j.esp.2007.06.001

Hyland, K. (2012). Bundles in academic discourse. *Annual review of applied linguistics, 32*, 150-169. doi: 10.1017/S0267190512000037

Hyland, K., & Jiang, F. (2022). Bundles in advanced EAL authors' articles: How do they compare with world Englishes practices? *World Englishes*. *41*(4), 554-570. doi: 10.1111/weng.12605

Hyland, K., & Jiang, F. K. (2018). Academic lexical bundles: How are they changing? *International journal of corpus linguistics, 23*(4), 383-407. doi: 10.1075/ijcl.17080.hyl

Isham, W. P. (1994). Memory for sentence form after simultaneous interpretation: Evidence both for and against deverbalization. In S. Lambert & S. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 191-211). New York: John Benjamins Publishing.

Jablonkai, R. (2010). English in the context of European integration: A corpus-driven analysis of lexical bundles in English EU documents. *English for specific purposes, 29*(4), 253-267. doi: 10.1016/j.esp.2010.04.006

Jiang, N. A., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The modern language journal, 91*(3), 433-445. doi: 10.1111/j.1540-4781.2007.00589.x

Jiang, X., & Jiang, Y. (2022). Menzerath-Altmann law in consecutive and simultaneous interpreting: Insights into varied cognitive processes and load. *Journal of quantitative linguistics*, *29*(4), 541-559. doi: 10.1080/09296174.2022.2027657

Jones, R. (2014). *Conference interpreting explained*. New York: Routledge.

Kader, S., & Seubert, S. (2014). *Anticipation, segmentation… stalling? How to teach interpreting strategies.* Paper presented at the To know how to suggest… Approaches to teaching conference interpreting.

Kajzer-Wietrzny, M. (2012). Interpreting universals and interpreting style. *Unpublished PhD dissertation. Adam Mickiewicz University, Poznań, Poland*.

Kajzer-Wietrzny, M. (2015). Simplification in interpreting and translation. *across languages and cultures, 16*(2), 233-255. doi:https://doi.org/10.1556/084.2015.16.2.5

Kajzer-Wietrzny, M. (2018). Interpretese vs. non-native language use: The case of optional that. In *Making way in corpus-based interpreting studies* (pp. 97-113). Singapore: Springer.

Kajzer-Wietrzny, M., & Grabowski, Ł. (2021). Formulaicity in constrained communication. *MonTI. Monografías de Traducción e Interpretación*(13), 148-183.

Kajzer-Wietrzny, M., & Ivaska, I. (2020). A multivariate approach to lexical diversity in constrained language. *across languages and cultures, 21*(2), 169-194. doi: 10.1556/084.2020.00011

Kendall, T. (2011). Corpora from a sociolinguistic perspective. In (Vol. 11, pp. 361-389): SciELO Brasil.

Kenning, M.-M. (2010). What are parallel and comparable corpora and how can we use them? In *The Routledge handbook of corpus linguistics* (pp. 487-500). New York: Routledge.

Kenny, D. (2006). Corpus-based translation studies: A quantitative or qualitative development. *Journal of translation studies, 9*(1), 43-58.

Klaudy, K., & Károly, K. (2005). Implicitation in translation: Empirical evidence for operational asymmetry in translation. *Across languages and cultures, 6*(1), 13-28. doi: 10.1556/acr.6.2005.1.2

Kruger, H., & De Sutter, G. (2018). Alternations in contact and non-contact varieties: Reconceptualising that-omission in translated and non-translated English using the MuPDAR approach. *Translation, cognition & behavior, 1*(2), 251-290. doi: 10.1075/tcb.00011.kru

Lanstyák, I., & Heltai, P. (2012). Universals in language contact and translation. *Across languages and cultures, 13*(1), 99-121. doi:10.1556/Acr.13.2012.1.6

Laviosa, S. (1997). How comparable can'comparable corpora'be? *Target.*

International journal of translation studies, 9(2), 289-319.

Laviosa, S. (1998). Core patterns of lexical use in a comparable corpus of English narrative prose. *Meta, 43*(4), 557-570. doi:10.7202/003425ar

Laviosa, S. (2002). *Corpus-based translation studies: theory, findings, applications*. Amsterdam, New York: Rodopi.

Le, T. N. P., & Harrington, M. (2015). Phraseology used to comment on results in the discussion section of applied linguistics quantitative research articles. *English for specific purposes, 39*, 45-61. doi: 10.1016/j.esp.2015.03.003

Lee, C. (2013). Using lexical bundle analysis as discovery tool for corpus-based translation research. *Perspectives, studies in translatology, 21*(3), 378-395. doi:10.1080/0907676X.2012.657655

Lee, C. (2022). Looking under the hood for evidence of normalization: Multivariate exploratory analysis of lexical bundles. *INContext: Studies in translation and interculturalism, 2*(1).

Lee, Y.-S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of psycholinguistic research, 25*(5), 527-542.

Li, C. (2010). Coping strategies for fast delivery in simultaneous interpretation. *The journal of specialised translation, 13*, 19-25.

Li, J., & Li, S. (2015). New trends of Chinese political translation in the age of globalisation. *Perspectives, 23*(3), 424-439. doi: 10.1080/09076X.2014.983530

Li, X. (2015). Putting interpreting strategies in their place: Justifications for teaching strategies in interpreter training. *Babel, 61*(2), 170-192. doi: 10.1075/babel.61.2.02li

Li, Y. (2016). A corpus-based interpreting exploration into pragmatic functions of

prefabricated chunks. *Modern foreign languages, 39*(2), 246-256.

Li, Y. (2017). Corpus-based exploration into the use of lexical bundles and students' interpreting proficiency. *Foreign languages and their teaching, 296*(5), 88-96.

Li, Y., & Halverson, S. L. (2020). A corpus-based exploration into lexical bundles in interpreting. *across languages and cultures, 21*(1), 1-22. doi:10.1556/084.2020.00001

Li, Y., & Halverson, S. L. (2022). Lexical bundles in formulaic interpreting: A corpus-based descriptive exploration. *Translation and interpreting studies*. doi: 10.1075/tis.19037.li

Li, Y., & Zhao, Y. (2019). Frequency effects of prefabricated chunks on interpreing-An exploration based on PACCEL from the pragmatic perspective. *Foreign languages and their teaching*(4), 37-44.

Lim, H.-O. (2005). Working into the B language: The condoned taboo? *Meta: Journal des traducteurs/Meta: translators' journal, 50*(4).

Lin, Y., Lv, Q., & Liang, J. (2018). Predicting fluency with language proficiency, working memory, and directionality in simultaneous interpreting. *Frontiers in psychology, 9*, 1543.

Liu, K., & Afzaal, M. (2021a). Syntactic complexity in translated and non-translated texts: A corpus-based study of simplification. *Plos one, 16*(6), e0253454. doi:10.1371/journal.pone.0253454

Liu, K., & Afzaal, M. (2021b). Translator's Style Through Lexical Bundles: A Corpus-Driven Analysis of Two English Translations of Hongloumeng. *Frontiers in psychology, 12*. doi:10.3389/fpsyg.2021.633422

Liu, L. (1988). Conference interpratation at the United Nations. *Foreign language teaching and research* (2), 42-46.

Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of

lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *Journal of English for academic purposes, 39*, 21-36.

Lu, X., Kisselev, O., Yoon, J., & Amory, M. D. (2018). Investigating effects of criterial consistency, the diversity dimension, and threshold variation in formulaic language research Extending the methodological considerations of O'Donnell et al. . *International journal of corpus linguistics, 23*(2), 158-182. doi:10.1075/ijcl.16086.lu

Lv, Q., & Liang, J. (2019). Is consecutive interpreting easier than simultaneous interpreting?–a corpus-based study of lexical simplification in interpretation. *Perspectives, 27*(1), 91-106. doi:https://doi.org/10.1080/0907676X.2018.1498531

Mackintosh, J. (1999). Interpreters are made not born. *Interpreting, 4*(1), 67-80.

Malamatidou, S. (2017). *Corpus triangulation: Combining data and methods in corpus-based translation studies.* New York: Routledge.

Marco, J. (2009). Normalisation and the Translation of Phraseology in the COVALT Corpus. *Meta: Journal des traducteurs/Meta: translators' journal, 54*(4), 842-856.

Marco, J. (2018). Connectives as indicators of explicitation in literary translation: A study based on a comparable and parallel corpus. *Target. International journal of translation studies, 30*(1), 87-111.

Mauranen, A. (2007). Universal tendencies in translation. In M. Rogers & G. Anderman (Eds.), *Incorporating Corpora: The linguist and the translator* (pp. 32-48). Clevedon: Multilingual matters.

May, R. (1997). Sensible elocution: How translation works in & upon punctuation. *The translator, 3*(1), 1-20.

McEnery, T. (2003). Corpus linguistics. In R. Mitkov (Ed.), *Oxford handbook of*

computational linguistics (pp. 448-463). Oxford: Oxford University Presee

McEnery, T., & Gabrielatos, C. (2006). English corpus linguistics. In B. Aarts & A. McMahon (Eds.), *The handbook of English linguistics* (pp. 33-71). Oxford: Blackwell.

McEnery, T., & Xiao, R. (2007). Parallel and comparable corpora: What is happening? In *Incorporating corpora* (pp. 18-31): Multilingual Matters.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Taylor & Francis.

McEnery, T., & Xiao, Z. (2007). Parallel and comparable corpora: The state of play. *Corpus-based Perspectives in Linguistics, 6*.

Mellinger, C., & Hanson, T. (2016). *Quantitative research methods in translation and interpreting studies*. Routledge.

Milton, J., & Freeman, R. (1996). Lexical variation in the writing of Chinese learners of English. *Language and computers*(16), 121.

Monacelli, C. (2009). Self-preservation in simultaneous interpreting. John Benhamins

Monti, C., Bendazzoli, C., Sandrelli, A., & Russo, M. (2005). Studying directionality in simultaneous interpreting through an electronic corpus: EPIC (European Parliament Interpreting Corpus). *Meta: Journal des traducteurs/Meta: translators' journal, 50*(4).

Moon, R. (1997). Vocabulary connections: Multi-word items in English. *Vocabulary: Description, acquisition and pedagogy, 40*, 63.

Murillo, S. (2004). A relevance reassessment of reformulation markers. *Journal of pragmatics, 36*(11), 2059-2068.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Nesi, H., & Basturkmen, H. (2006). Lexical bundles and discourse signalling in

academic lectures. *International journal of corpus linguistics, 11*(3), 283-304.

Olohan, M. (2004). *Introducing corpora in translation studies*. Routledge.

Olohan, M., & Baker, M. (2000). Reporting that in translated English. Evidence for subconscious processes of explicitation? *Across languages and cultures, 1*(2), 141-158.

Øverås, L. (1998). In search of the third code: An investigation of norms in literary translation. *Meta: Journal des traducteurs/Meta: Translators' journal, 43*(4), 557-570.

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for academic purposes, 21*, 60-71. doi:10.1016/j.jeap.2015.11.003

Pan, F., Reppen, R., & Biber, D. (2020). Methodological issues in contrastive lexical bundle research The influence of corpus design on bundle identification. *International journal of corpus linguistics, 25*(2), 215-229. doi:10.1075/ijcl.19063.pan

Pawley, A., & Syder, F. H. (1983). *Two puzzles for linguistic theory: Nativelike selection and nativelike fluency*. New York: Longman.

Plevoets, K., & Defrancq, B. (2018a). The cognitive load of interpreters in the European Parliament A corpus-based study of predictors for the disfluency uh(m). *Interpreting : international journal of research and practice in interpreting, 20*(1), 1-28. doi:10.1075/intp.00001.ple

Plevoets, K., & Defrancq, B. (2018b). The cognitive load of interpreters in the European Parliament A corpus-based study of predictors for the disfluency uh(m). *Interpreting, 20*(1), 1-28. doi:10.1075/intp.00001.ple

Pochhacker, F. (2006). Research and methodology in healthcare interpreting.

Linguistica antverpiensia, 5, 135-159. doi:10.52034/LANSTTS.V5I.157

Raupach, M. (1984). Formulae in second language speech production. In H. Dechert, D. Mo¨hle, & M. Raupach (Eds.), *Second language productions* (pp. 114-137). Tu¨bingen: Gunter Narr.

Rayson, P., Berridge, D., & Francis, B. (2004). *Extending the Cochran rule for the comparison of word frequencies between corpora.* Paper presented at the 7th International Conference on Statistical analysis of textual data (JADT 2004).

Rayson, P., & Garside, R. (2000). *Comparing corpora using frequency profiling.* Paper presented at the The workshop on comparing corpora.

Ren, J. (2021). Variability and functions of lexical bundles in research articles of applied linguistics and pharmaceutical sciences. *Journal of English for academic purposes, 50*, 100968. doi: 10.1016/j.jeap.2021.100968

Reppen, R., & Olson, S. B. (2020). Lexical bundles across disciplines: A look at consistency and variability. In U. Romer, V. Cortes, & E. Friginal (Eds.), Advances in corpus-based research on academic writing: Effects of discipline, register, and writer expertise (pp. 170-182). John Benjamins Publishing Company. Doi: 10.1075/scl.95.07rep.

Riccardi, A. (2005). On the evolution of interpreting strategies in simultaneous interpreting. *Meta: Journal des traducteurs/Meta: Translators' journal, 50*(2), 753-767.

Rosendo, L. R., & Diur, M. (2021). Conference interpreting at the United Nations. In *The routledge handbook of conference interpreting* (pp. 115-125): Routledge.

Rosendo, L. R., & Marie, D. (2017). Admission exams in international organisations: The United Nations' Language Competitive Examination (LCE). *CLINA: an interdisciplinary journal of translation, interpreting and intercultural communication, 3*(2), 33-52.

Russo, M., Bendazzoli, C., & Sandrelli, A. (2006). *Looking for lexical patterns in a trilingual corpus of source and interpreted speeches: extended analysis of EPIC (European Parliament Interpreting Corpus). Forum, 4*(1) 221-254.

Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing : applying a corpus-based study to language teaching*. Netherlands/Philadelphia John Benjamins Publishing Company.

Saldanha, G., & O'Brien, S. (2014). *Research methodologies in translation studies*: Routledge.

Sandrelli, A., & Bendazzoli, C. (2005). *Lexical patterns in simultaneous interpreting: A preliminary investigation of EPIC (European Parliament Interpreting Corpus).* Paper presented at the Corpus Linguistics 2005.

Sandrelli, A., Bendazzoli, C., & Russo, M. (2010). European Parliament Interpreting Corpus (EPIC): methodological issues and preliminary results on lexical patterns in simultaneous interpreting.

Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing, and use* (Vol. 9): John Benjamins Publishing.

Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid. *Formulaic sequences: Acquisition, processing and use*, 127-151.

Schoenfeld, M., Eckhard, S., Patz, R., Meegdenburg, H. v., & Pires, A. (2019). *The UN Security Council Debates*. Retrieved from: https://doi.org/10.7910/DVN/KGVSYH

Schourup, L. (1999). Discourse markers. *Lingua, 107*(3-4), 227-265.

Scott, M. (2020). WordSmith Tools (Version 8.0)[Computer software]. Lexical Analysis Software.

Seeber, K. (2016). Modeling multimodal processing in Simultaneous Interpreting.

Poster presented at the 75th anniversary of FTI, Genve, September 30.

Seeber, K. G. (2011). Cognitive load in simultaneous interpreting: Existing theories—new models. *Interpreting, 13*(2), 176-204.

Seeber, K. G. (2015). Simultaneous interpreting. In *The Routledge handbook of interpreting* (pp. 91-107): Routledge.

Seleskovitch, D. (1999). The teaching of conference interpretation in the course of the last 50 years. *Interpreting, 4*(1), 55-66.

Setton, R. (1999). *Simultaneous interpretation: A cognitive-pragmatic analysis of simultaneous interpretation*. Ámsterdam: John Benjamins.

Setton, R. (2011). Corpus-based interpreting studies (CIS): Overview and prospects. In A. Kruger, K. Wallmach, & J. Munday (Eds.), *Corpus-based translation studies: Research and applications* (pp. 33-72). London: Continuum.

Setton, R., & Motta, M. (2007). Syntacrobatics: Quality and reformulation in simultaneous-with-text. *Interpreting, 9*(2), 199-230.

Shahriari, H. (2017). Comparing lexical bundles across the introduction, method and results sections of the research article. *Corpora, 12*(1), 1-22.

Shao, X. (2018a). 基于语料库汉英交替传译中专业译员的语块特征研究. *外语电化教学* (5), 69-75.

Shao, X. (2018b). 基于语料库的汉英交替传译中专业译员语块特征研究. *上海翻译* (6), 50-56.

Shin, Y. K. (2018). The construction of English lexical bundles in context by native and nonnative freshman university students. *English teaching, 73*(3), 115-139.

Shirazizadeh, M., & Amirfazlian, R. (2021). Lexical bundles in theses, articles and textbooks of applied linguistics: Investigating intradisciplinary uniformity and

variation. *Journal of English for academic purposes, 49*, 100946. doi:10.1016/j.jeap.2020.100946

Shlesinger, M. (1989). *Simultaneous interpretation as a factor in effecting shifts in the position of texts on the oral-literate continuum.* Unpublished MA Thesis, Tel Aviv University.

Shlesinger, M. (1995). Shifts in cohesion in simultaneous interpreting. *The Translator, 1*(2), 193-214.

Shlesinger, M. (1998a). Corpus-based interpreting studies as an offshoot of corpus-based translation studies. *Meta, 43*(4), 486-493. doi:10.7202/004136ar

Shlesinger, M. (1998b). Corpus-based interpreting studies as an offshoot of corpus-based translation studies. *Meta: Journal des traducteurs/Meta: translators' journal, 43*(4), 486-493.

Shlesinger, M. (2003). Effects of presentation rate on working memory in simultaneous interpreting. *The interpreters' newsletter, 12*, 37-49.

Shlesinger, M. (2009). Towards a Definition of Interpretese. An Intermodal, Corpus-Based Study. In A. Chesterman, H. Gerzymisch-Arbogast & G. Hansen (Eds.), *Efforts and Models in Interpreting and Translation Research. A tribute to Daniel Gile. Amsterdam/Philadelphia: John Benjamins*, 237-253.

Shrefler, N. (2011). Lexical bundles and German bibles. *Literary and linguistic computing, 26*(1), 89-106.

Shreve, G. M., & Angelone, E. (2010). Translation and cognition. *Amsterdam/Philadelphia: John Benjamins*

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied linguistics, 31*(4), 487-512.

Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second language research, 27*(2), 251-272.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for academic purposes, 12*(3), 214-225.

Straniero Sergio, F., & Falbo, C. (2012). *Breaking ground in corpus-based interpreting studies*.Peter Lang.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of verbal learning and verbal behavior, 18*(5), 523-534.

Tang, F., & Li, D. (2017). A corpus-based investigation of explicitation patterns between professional and student interpreters in Chinese-English consecutive interpreting. *The interpreter and translator trainer, 11*(4), 373-395.

Tang, F., & Jiang, S. (2022). Four-word lexical bundles in Chinese-English consecutive interperting-A comparative study between professional and trainees. *Frontiers in Psychology*, 13.

Teich, E. (2003). *Cross-linguistic variation in system and text. Methodology for the investigation of translation and comprable texts*. Berlin/Boston: De Gruyter Mouton.

Tichý, O. (2021). Corpus driven identification of lexical bundle obsolescence in Late Modern English. In S. Kranich & T. Breban (Eds) *Lost in Change: Causes and processes in the loss of grammatical elements and constructions* (pp.101-130). John Benjamins Publishing.

Toury, G. (1995). *Descriptive translation studies and beyond*. Amsterdam, Philadelphia: John Benjamins Publishing.

Toury, G. (2012). *Descriptive translation studies–and beyond: Revised Edition* (Vol.

100): John Benjamins Publishing.

Tsao, F.-F. (1982). English and Chinese (Mandarin). *Annual review of applied linguistics, 3*, 99-117.

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 153-172). Amsterdam: John Benjamins.

Van Lancker-Sidtis, D., & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language & communication, 24*(3), 207-240.

Van Rietvelde, S., Eyckmans, J., & Bauwens, D. (2010). As time goes by: Phraseological competence and linguistic anticipations in the interpreting performance. *Artesis VT working papers*.

Vandepitte, S. (2001). Anticipation in conference interpreting: A cognitive process. *Revista alicantina de estudios ingleses, 14*, 323-335.

Vinay, J.-P., & Darbelnet, J. (1958). *Stylistique comparée du français et de l'anglais*. Paris: Didier.

Vinay, J.-P., & Darbelnet, J. (1958/1995). *Comparative stylistics of French and English: A methodology for translation*. John Benjamins Publishing. (Original work published, 1958)

Wang, B. (2012). A descriptive study of norms in interpreting: Based on the Chinese-English consecutive interpreting corpus of Chinese premier press conferences. *Meta 57*(1), 198-212. doi.org/10.7202/1012749ar

Wang, B., & Tang, F. (2020). Corpus-based Interpreting Studies in China: Overview and Prospects. In Hu. K. & Kim. K-H. (Eds.), *Corpus-based translation and interpreting studies in Chinese contexts* (pp. 61-87). Palgrave Macmillan.

Wang, J. (2016). Effects of chunk cognition training on fluency of simul-interpreting: An empirical study. *Foreign language teaching and research, 48*(5), 765-775.

Wang, Y. (2017). Lexical bundles in spoken academic ELF: Genre and disciplinary variation. *International journal of corpus linguistics, 22*(2), 187-211. doi:10.1075/ijcl.22.2.02wan

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *Relc journal, 42*(2), 155-166.

Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian modern language review, 63*(1), 13-33.

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian journal of applied linguistics, 12*(1), 39-57.

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. Bloomsbury Publishing.

Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. London/ New York: Bloomsbury Publishing Plc.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge England : Cambridge University Press.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & communication, 20*(1), 1-28.

Wu, B., Cheung, A. K., & Xing, J. (2021). Learning Chinese political formulaic phraseology from a self-built bilingual united nations security council corpus: A pilot study. *Babel*, 67(4), 500-521. doi: 10.1075/babel.00233.wu

Wu, Y. (2021). Lexical bundles in English EU parliamentary discourse-variation across interpreted, translated, and spoken registers. *Compilation & translation*

review, 14(2), 37-86.

Wu, Y., & Liao, P. (2018). Re-conceptualising interpreting strategies for teaching interpretation into a B language. *The interpreter and translator trainer, 12*(2), 188-206. doi:10.1080/1750399X.2018.1451952

Xia, D., Ai, H., & Pae, H. K. (2022). “Please let me know”: Lexical bundles in business emails by business English learners and working professionals. *International journal of learner corpus research, 8*(1), 1-30.

Xiao, R. (2010). *Idioms, word clusters, and reformulation markers in translational Chinese: Can “translation universals” survive in Mandarin?* Paper presented at the international symposium on using corpora in contrastive and translation studies. .

Xiao, R., & Dai, G. (2014). Lexical and grammatical properties of translational Chinese: Translation universal hypotheses reevaluated from the Chinese perspective. *Corpus linguistics and linguistic theory, 10*(1), 11-55.

Xu, C., & Li, D. (2021). A corpus-based comparative study of prefabricated chunks used by professional interpreters in simultaneous interpreting. *Foreign language teaching and research, 53*(4), 582-593.

Xu, J., & Li, J. (2021). A syntactic complexity analysis of translational English across genres. *Across languages and cultures, 22*(2), 214-232. doi: 10.1556/084.2021.00015

Yang, S., Li, D., & Lei, V. L. C. (2020). The impact of source text presence on simultaneous interpreting performance in fast speeches: Will it help trainees or not? *Babel, 66*(4-5), 588-603.

Yin, X., & Li, S. (2021). Lexical bundles as an intradisciplinary and interdisciplinary mark: A corpus-based study of research articles from business, biology, and applied linguistics. *Applied corpus linguistics, 1*(1), 100006.

Yusof, B. (2021). Weaving critical discourse analysis into Brunei history: Examining shifts in Brunei-Britain political discourse. *Engaging modern Brunei: research on language, literature, and culture*, 95-109.

Zanetti, R. (1999). Relevance of anticipation and possible strategies in the simultaneous interpretation from English into Italian. *The Interpreters' Newsletter, 9*, 79-98.

Zanettin, F. (2014). Corpora in translation. In *Translation: A multidisciplinary approach* (pp. 178-199). Springer.

Zipf, G. K. (1935/1965). *The psycho-biology of language: An introd. to dynamic philology*. Cambridge, MA: MIT Press.