

*Article*

## **Polyseme Transfer in the Chinese to English Machine Translation Output and Chinese Students' English Writing**

**Zhenyan Ye**

University of Hong Kong, China

### **Abstract**

Machine translation (MT) systems such as Google Translate, Bing or Youdao are increasingly present in everyday life. Anecdotal evidence suggests that language students might use them to produce written work in the target language (TL) and thus possibly get around a potentially difficult writing task. The crucial question to ask would be whether it is possible to differentiate the output of MT from learner language. This paper seeks to address this question by comparing the lexical features of these two types of discourse in the Chinese context. In particular, it examines the use of English translation equivalents of polysemous Chinese words in two parallel corpora: A Chinese webpage corpus translated into English using Bing and Youdao on the one hand and a Chinese learner writing corpus on the other. While the comparison yields similar error rates, it also establishes that human learners and translation engines have difficulties with different sets of words. Word frequency also plays a significant role in differentiating between the two sets of output. The paper concludes with the finding that MT output is sufficiently different from learner language in terms of lexis. The findings could be used to create an algorithm for the detection of ethics code violation through the use of MT engines in written assignments.

### **Keywords**

Lexical transfer, polysemy, machine translation, writing

## **1 Introduction**

Vocabulary is a crucial building element of any language due to its indispensable role in general language acquisition and communication (Schmitt et al, 2001; Agustín-Llach, 2020). However, despite its importance, the research into lexical errors remains relatively scarce. In cross-linguistic activities, lexical errors are quite likely to happen because of the similarities and differences between the source language and the target language. These differences can cause language transfer according to Odlin (1989) and also influence the use of vocabulary either positively or negatively. Specifically, at the lexical level, it will lead to positive transfer if the traits of lexical units from source language and target language happen to be in accordance with each other. On the other hand, negative transfer will occur in the opposite case.

Elston-Guttler and Williams (2005) point out that one of the most prominent types of negative lexical transfer is polysemy transfer. At the lexical level, polysemy is the capacity of a word to have multiple meanings that are usually related by contiguity of meaning within a semantic field (Fillmore & Atkins, 2000). Based on this definition, as the mapping between words and concepts differs across languages, the transfer of polysemous words can easily lead to lexical errors (Elston-Guttler and Williams, 2005). Dodigovic, Liang and Yue (2015) recognize two fields where this problem can largely impact cross linguistic activities – second language (L2) learners' writing and machine translation (MT) output. This gives rise to the thought that there might be some overlap between polysemy transfer happening in L2 learners' writing and MT output.

In the field of Second Language Acquisition (SLA), many researchers (Ellis and Heimbach, 1997; Jiang, 2000; Ellis, 2008) believe that when L2 words are learnt, they are mapped onto L1 concepts and are connected to the already existing L1 schemata, which is likely to trigger the transfer of L1 lexical knowledge and lead to errors. More specifically, according to Ijaz's (1986) "semantic equivalence hypothesis", learners tend to interpret the new language symbols based on an acquired language by seeking word equivalents.

In translation of any sort, lexical transfer is also likely to play a role. With regard to MT, Melcuk and Wanner (2001) defined lexical translation as a main problem, given this obstacle has been widely discussed for several decades but never fully overcome. For Melcuk and Wanner (2001), the essential problem here is that it is difficult for MT systems to choose the optimal one from the possible equivalents in a given context when the source lexical item semantically corresponds to more than one lexical item in the target language. This inability to make the right lexical choice based on the use of a bilingual dictionary is also found in less confident L2 learners (Dodigovic, 2013).

Based on this, Dodigovic, Liang and Yue (2015) endeavored to develop a post-editing application that can amend the negative lexical transfer occurring either in the English writing of Chinese-English bilinguals or in the output automated translation from Chinese into English. However, despite such attempts to address negative lexical transfer in these two fields, the research into lexical transfer in MT fields remains scarce, which is especially true in China. Therefore, it is necessary to conduct relevant research to assess the feasibility of software design mentioned above.

This study aims to identify and analyze cases of negative lexical transfer from Chinese to English caused by polysemes in MT output and compare their pattern to those in Chinese English learners' L2 writing. The research questions of the current study are presented as follows: (1) As for the Chinese to English machine translation engine of Youdao and Bing, which MT engine performs better in coping with Chinese polysemous words? (2) To what extent is the output of Chinese-English machine translation similar to learner language in terms of lexical transfer due to polysemy?

## **2 Literature Review**

### **2.1 Vocabulary and lexical transfer**

Although the phenomenon of L1 transfer has been widely acknowledged (Wang, 2011), the debate over the accurate definition of "transfer" is still ongoing nowadays (Ma, 2015). Against this background, a highly appreciated one proposed by Odlin (1989), which is referred to for the theoretical foundation of this current study, emphasizes that transfer means the influence caused by similarities and differences between the source language and the target language.

Based on this point, there are two scenarios when language transfer happens on the lexical level: if the traits of two lexical units from SL and TL happen to be in accordance with each other, the transfer will be positive and thereby lead to native-like expression. On the other hand, negative transfer will result in non-native-like production.

In order to avoid lexical errors, both language learners and translators need adequate knowledge of words, and that is no small task. Basically, a word stands for far more than a combination of linguistic form and its literal meaning. Rather, to know a word means to know the contexts it usually occurs in, other words that often appear along with it, the way in which it is expressed idiomatically, the connotation it evokes, the syntactic structure needed to accommodate this word, and its possible function in an utterance (Dodigovic, 2005). Therefore, for L2 learners, the lexical development is not just the learning of isolated words but of words and their associations (Ellis, 2008; Nation, 2013). Given this complex nature of vocabulary and the demanding learning condition it requires, including the context-based process mentioned by Jiang (2000), many L2 learners are usually not able to achieve a comprehensive mastery of L2 words, which leads to L1 transfer errors in their L2 production (Ellis, 2008; Harutyunyan, 2020).

## 2.2 Lexical transfer caused by polysemy

### 2.2.1 Polysemy and its transfer

Traditionally, polysemy refers to a word being able to carry multiple meanings which are more or less related (Hsieh, 2011). Although this trait of words is common in most languages (Morimoto & Loewen, 2007), it can pose significant difficulties in cross-linguistic activities because words of roughly the same meaning in two languages within one context are not necessarily interchangeable in all contexts of use.

In the field of SLA, for instance, Ijaz (1986) claims that even advanced adult ESL learners are heavily influenced by native language transfer according to his “semantic equivalence hypothesis”, in which it is indicated that learners tend to interpret the new language symbols based on an acquired language by seeking word equivalents. For example, the English word “question” can be transferred into Chinese as “问题”, so it is likely for Chinese English learners to understand the concept of “question” based on the word “问题”, which is already acquired in their mind. Nevertheless, this mental process of L2 learners is quite unreliable since finding word equivalents is intricate, as Elston-Guttler and Williams (2005) pointed out that language specificity in the mapping between words and concepts varies in terms of cross-linguistic differences in lexicalization pattern. In other words, it is likely that a word from a source language has two concepts that are represented by two different words in the target language, which will possibly lead to negative transfer. For instance, the above-mentioned word “问题”(X) in Chinese signifies two non-linguistic concepts “something that you say or write in order to ask a person about something”(a) and “a situation that is unsatisfactory and causes difficulties for people”(b) simultaneously, whereas these two concepts are expressed through two different words “question”(Y) and “problem”(Z) in English. In this way, in Chinese English learners’ writing where “问题” serves as a source lexical unit to be expressed in English, it will be problematic to choose from “question” and “problem”. It is worth noting that Elston-Guttler and Williams (2005) claim that the various senses of a polysemous word must be highly or moderately related, which distinguishes polysemy from homonymy which means a group of words that share the same pronunciation or spelling but have unrelated meanings (e.g. bank as institution and bank as area next to river).

### 2.2.2 Polysemy transfer in L2 learners’ writing

A considerable number of studies have paid attention to the way in which L2 learners deal with L1 polysemous words when producing L2. Ijaz (1986) found that L2 learners tend to transfer literal meaning instead of figurative meaning from L1 to L2. This always causes errors if the L1 word is polysemous and has more than one equivalent in L2. Specifically, the various meanings of an L1 word can be distributed across several L2 words that are mutually synonymous. L2 learners are likely to choose the equivalent

that they are more familiar with, even though it might not be the target word (Hemchua & Schmitt, 2006). This problem becomes even more compelling as it was revealed that L2 learners tend to make more errors with the increase of L2 proficiency (Martin, 1984). Hemchua and Schmitt (2006) claim that polysemy transfer is exactly the issue causing advanced learners at university level to have considerable difficulties when choosing the correct word in L2. Similarly, Lennon (1996) also points out that L2 learners, even those of high proficiency, are frequently affected by negative transfer when using words which are polysemous in their native language.

In the Chinese context, many studies show that for Chinese English learners, negative polysemy transfer is also a significant lexical error frequently made in their writing. For example, Wu (2013) indicates that there is a significant difference between Chinese and English in terms of their language specificity, which means that there are few words in Chinese whose lexical domain can entirely overlap with one of English words. Thus, according to Elston-Guttler and Williams's (2005) model of polysemy transfer, it seems that Chinese English learners are more likely to face difficulty in choosing the correct equivalents in English. Zhang and Wang (2005) support this hypothesis, stating that in Chinese students' English writing, such errors occur frequently "where the choice is an equivalent to a Chinese word or expression on the literal level, but does not convey the intended meaning in English". In their research, typical Chinese to English negative polysemy transfer that originates from Chinese students' writing is examined. For example, they provide the following sentence:

We should raise our spoken English level. (improve)

In this case, the Chinese verb “提高” is mistakenly expressed by the word “raise”, as “improve” is preferred in this context, even though both meanings are equivalents of this Chinese verb.

A recent piece of research in the field lays the foundation for the current study. Dodigovic, Ma and Jing (2017) and Ma (2015) conducted research in which they collected negative polysemy transfer in two different stacks of writing assignments (50 pieces for each) that were written by Year 1 and Year 4 English major students of Xi'an Jiaotong Liverpool University respectively. As a result, there are 186 items presented in a table, in which the Chinese equivalent and the correct English expression for each item are also shown. There are mainly two reasons why this research is advanced and significant. Firstly, compared with previous research (e.g. Zhang & Wang, 2005) which examines the negative polysemy transfer, the one conducted by Dodigovic, Ma and Jing (2017) delve deeper into this category and collect a large number of authentic examples. Moreover, such examples are retrieved from the writing of Year 1 and 4 English students of a Sino-British university, and therefore the English proficiency of target learners is higher than other research (Wu, 2013) where the level of students is no more than senior high school. In this way, according to Martin's (1984) claim, these examples produced by more advanced learners will be more suitable for analysis and significant in terms of their wider range of lexicon and higher frequency of polysemy transfer.

### *2.2.3 Machine translation (MT) and its transfer*

In translation of any sort, lexical transfer is also likely to play a role, especially when translating from one's native language into L2. Machine Translation (MT) is likely prone to the same issues as human translation. With regard to MT, Melcuk and Wanner (2001) defined lexical translation as a main problem. This obstacle has been widely discussed for several decades but never fully overcome. Although the medium of learners and MT in processing language is quite different – human brain and computer respectively, they are essentially faced with the same problem in dealing with lexical items cross-linguistically. It is therefore possible that the outcomes of lexical choices might be similar. For Melcuk and Wanner (2001), the essential problem here is that it is difficult for an MT system to choose the optimal one from possible options in a given context when the source lexical item semantically corresponds to more than one target lexical item. This model actually demonstrates the same selection

pattern in processing polysemes as a language learner's, according to Elston-Guttler and Williams (2005). This inability to make the right lexical choice based on the use of a bilingual dictionary is also found in less confident L2 learners (Dodigovic, 2013).

Nevertheless, although it has been revealed that L2 learners and machine are confronted with similar problems when processing lexical transfer, it is worth noting that the recent development in MT model and computer science has more or less changed the landscape. Melcuk and Wanner (2001) proposed a concept called “linguistic clue”, which refers to any piece of information regarding the meaning of source lexical unit under consideration and its potential target equivalents. Under this paradigm, the rendering of high accuracy in dealing with lexical items can be guaranteed as long as the machine is able to retrieve linguistic clues in context. There is evidence showing that several up-to-date Chinese-English MT engines are able to identify linguistic clues in simple sentences with the assistance of example-based corpora (Zhu, 2012). For instance, it is found that the Chinese expression “提高” in the sample sentence provided (Zhang & Wang, 2005) above can be correctly translated as “improve” through Youdao and Bing Chinese to English Translation as the key linguistic clue “水平 (level)” is recognized to help MT select the optimal equivalent.

#### 2.2.4 Word Frequency and polysemes

Schmitt (2000) points out that words having high frequency are more likely to be polysemous and polysemes having high frequency are more likely to have more related meanings. For Chinese polysemes, Wang (2009) proves the positive correlation that Chinese polysemes with higher frequency tend to have more related meanings by examining the frequency of 10,632 Chinese polysemous words in the “Universal Corpus of Modern Chinese”. Based on the data collected by Wang (2009), the correlation between Chinese polysemes' frequency and the number of meanings they have can be presented as follows:

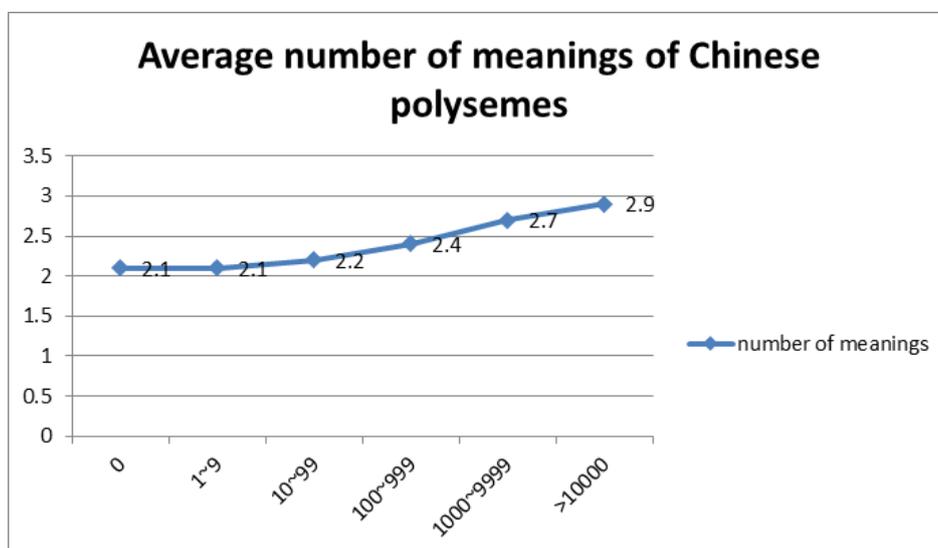


Figure 1. Correlation between Chinese polysemes frequency and number of meanings

As can be seen from Figure 1, the number of meanings of Chinese polyseme steadily rises with the increase of their frequency and finally peaks at 2.9 for words that appear more than 10,000 times in the corpus.

As was mentioned before, for L2 learners, the difficulty in transferring polysemes is to choose the correct equivalent from the target language (Ijaz, 1986). Therefore, it can be expected that this difficulty will become more pronounced in transferring polysemes with high frequency because these words have

more meanings and therefore learners have to select a correct equivalent in the target language from more given meanings. This inference can be attested by the research conducted by Ma (2015), in which it is found that Chinese students tend to make more mistakes in their English writing when transferring the Chinese polysemes with higher frequency, such as “看”, “地方”, and “认识”.

As for machine translation engine, since it is confronted with a similar process in transferring polysemes as L2 learners (Melcuk & Wanner, 2001), it is possible that the positive correlation between polyseme frequency and error rate in transferring that occurs in L2 learners' writing is also common to MT engines. However, this viewpoint has not yet been proved as previous studies rarely paid attention to this issue, especially in terms of Chinese to English MT engines.

### 2.3 Machine translation and lexis

Overall, negative polysemy transfer is a major problem for both L2 learners and machine translation. Carter (1998) suggests that lexical selection mistakes can be less generously tolerated than syntactic mistakes, as they can directly cause the misunderstanding of the information and add to the interpreting load of the audience. To date, software for dealing with students' writing and MT output respectively, while helpful, still cannot cover a crucial aspect of natural language processing, namely its native-like quality (Dodigovic, Liang & Yue, 2015). For instance, automated translation applications, such as Bing and Youdao, can produce a comprehensible Chinese to English translation, but it is as a rule not idiomatic enough. On the other hand, some applications (like YiGai and Grammarly) are still unable to identify and correct most of the problems that appear in texts generated by MT and non-native speakers.

Based on the previous research (Melcuk & Wanner, 2001; Zhang & Wang, 2005; Zhu, 2012), it can be expected that there are some differences between L2 learners' writing and the output of MT in terms of polysemy transfer due to the progress in MT engine advances. However, despite the complexity of the way in which negative polysemy transfer happens in these two fields, there are rarely studies that compile an authentic Chinese to English MT output corpus or delve into comparison between them. For this reason, this study aims to collect polysemy transfer from Chinese to English MT output and then examine the similarities and differences between the two kinds of linguistic performance in terms of lexical transfer by polysemy.

## 3 Methodology

### 3.1 Data and instruments

#### 3.1.1 The webpage corpus used

This corpus-based study investigates the lexical transfer by polysemy in the output of machine translation. The webpage corpus used in the present study is comprised of 30 articles/reports extracted mainly from Chinese websites of the government, universities, and news agencies. There are two reasons for building the corpus based on these three types of website. Firstly, the quality and timeliness of information provided in these types of website are highly recognized in China, and therefore the context in which polysemous words appear is more authentic and representative. Secondly, the formality of language is a conspicuous feature of these websites, which resembles students' final year dissertation that was chosen for the current research. In this way, the similarity in formality will render these two corpora more comparable and thereby validate the results.

#### 3.1.2 MT engines

In this study, Chinese to English Translation using Youdao and Bing online software is employed to

translate 30 texts extracted from websites for the collection of transfer examples. Youdao Translation, developed by Netease, is a program based on an Internet corpus and a search engine. It is also one of the most satisfactory applications in its category. Bing translator, developed by Microsoft, uses similar core MT technology as the one of Youdao.

### *3.1.3 Contrastive data – learners' transfer*

Dodigovic, Ma and Jing (2017) as well as Ma (2015) collected Chinese lexical transfer based on the learner corpus that consists of 100 pieces of writing (541,482 English words in total) by 50 Year 1 students and 50 Year 4 students at Xi'an Jiaotong Liverpool University (XJTLU), a Sino-foreign university in China jointly run by the University of Liverpool and Xi'an Jiaotong University. The students studying in XJTLU can be regarded as typical L2 learners of English, since all the modules provided in XJTLU are instructed in English (except for those related to Chinese language and culture).

The writing included in the learner corpus was all written by students who are native speakers of Chinese from the department of English Culture and Communication. It is worth noting that the writing of XJTLU students is supposed to be highly formal and academic, as XJTLU students must enroll in EAP (English for Academic Purpose) course in the year 1 and year 2, which aims to develop students' English skills comprehensively and acquaint them with academic English writing style for their future research.

### *3.1.4 Chinese words frequency list*

As was mentioned in the previous section, polysemes having high frequency are more likely to have more related meanings (Schmitt, 2000; Wang, 2009), which poses more difficulty for students and cause more errors in transferring Chinese polysemes to English (Ma, 2015). In order to establish whether the Chinese polysemous words with higher frequency are more likely to cause negative transfer in MT output as in students' writing, the study utilizes data found in "Chinese Word and Character Frequencies Based on Film Subtitles" (SUBTLEX-CH-WF, Cai and Brysbaert, 2010), which is available from <http://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexch/subtlexchwf.zip>. The frequency of Chinese polysemous words/characters that belong to negative transfer in MT output is checked in the frequency list, which is provided in the appendix.

## **3.2 Procedure**

After the webpage corpus was built, the errors caused by MT engine in translating those texts were to be identified. Following the definition of polysemy and Chinese polysemy transfer proposed by Elston-Guttler and Williams (2005) and Dodigovic (2015) respectively, the researcher firstly read through all the 30 webpage texts carefully in order to discover and identify typical Chinese polysemy that is likely to be mistranslated by MT engines. These lexical items were all highlighted by placing “#” next to them with serial numbers for the sake of data retrieval afterwards. It is worth noting that these “#” marks and serial numbers were not included when the texts were input into MT software in case of interference on the translation process. Subsequently, each text was translated by both MT engines of Youdao and Bing, and the English output of every text from Bing was saved as another 30 texts, while for Youdao, only the translation of highlighted polysemous words was recorded to serve as comparison data. With the help of serial number, each highlighted Chinese polysemous word and its translation in documents can be conveniently traced and compared to attribute the examples of transfer to either positive or negative. For the sake of contrasting with learners' transfer, this data was then organized and input into an Excel document, which includes not only the Chinese polysemes, their translation, and their correctness, but also the Chinese pinyin, literal meaning of each examples, as well as the Chinese sentence in which the

polyseme appears. Afterwards, the Excel document of both negative transfer by students and transfer (both positive and negative) were examined thoroughly to identify all the identical items (the polysemous words that appear in both documents). Then, pairs of identical items were compared and the results were classified into three types: (1) the lexical translation from both MT engines belong to negative transfer, which was marked as “2”; (2) the lexical translation from either one of MT engines are negative transfer, which was signed as “1”; (3) the lexical translation from both MT engines attribute to positive transfer, which is labeled as “0”. Since this approach is mainly data-driven, the emerging patterns are further analyzed and conclusions from those drawn.

## 4 Results

There were totally 185 typical Chinese polysemes found in the 30 articles in the webpage corpus that was scrutinized in this study. While Youdao MT caused 55 negative transfer instances and 13 omissions (i.e. the case where MT failed to provide any equivalent) in its rendering, Bing translator yielded 58 negative transfer instances and 6 omissions in its output. To answer the first research question, the number of negative transfer and omission caused by these two MT engines was counted. The difference between the performances of two MT engines, while not conspicuous, does exist in terms of their accuracy and omission rate. Compared with the slight difference between the number of negative transfer instances caused by Youdao and Bing (30% versus 31%), the proportion of positive transfer and omission in their output differs more strongly, which is 63% versus 66% and 7% versus 3% respectively.

The second research question could be answered from two perspectives. Firstly, all the words that appear in both learners’ transfer data and MT output data were juxtaposed and then compared to show the similarity between the transfer pattern of students and MT engines. Based on the contrastive data (students’ negative transfer), for identical lexical items, if it is negatively transferred by both MT engines as what was done by students, it will be regarded as completely similar. If it is negatively transferred by only one MT engine, it will be considered as partly similar. If it is positively transferred by both MT engines as opposed to students’ result, it will be marked as dissimilar. As a result, it was found that for 37 identical lexical items, a considerable number of them fell into the “dissimilar” category. In other words, both MT engines were able to positively transfer more than half of the identical Chinese polysemes that were not correctly transferred by students.

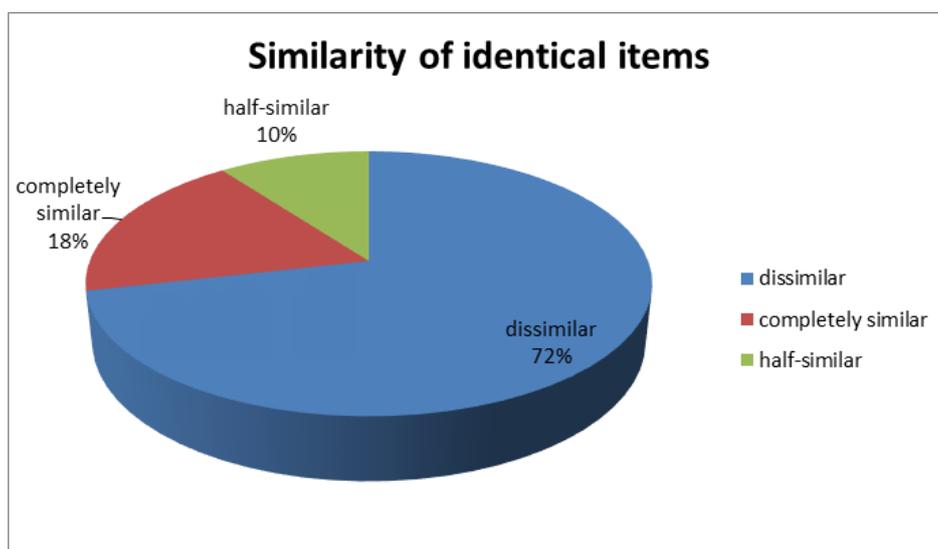


Figure 2. Three types of similarity in identical items

*Note:* In Figure 2, “dissimilar”, “partly similar”, and “completely similar” are represented by number “0”, “1”, and “2” respectively. The negative transfer instances in students’ expression are bolded. The instances of transfer in MT output are bolded, and if negative, it is further italicized.

Figure 2 shows the proportion of three similarity type found in the comparison of identical lexical items. Specifically, 72% of identical lexical items are categorized as “dissimilar”, and this indicates that approximately 3/4 of these Chinese polysemes were successfully processed by both MT engines. Meanwhile, there are 18% of identical items that belong to “completely similar” type, which means nearly 1/5 of them were negatively transferred by both the MT engines and students. This is followed by 10% partly similar cases, where the negative polysemy transfer was caused in the MT output of either Youdao or Bing. Several typical examples from these three categories are presented as follows:

Table 1

*Typical Examples of Three Types of Identical Polysemes*

Chinese equivalent	Chinese Pinyin	Expression found in students’ writing	Expression found in MT output (Youdao/Bing)	Similarity
差距	cha ju	difference	gap	0
认识	ren shi	have his own recognition	talk about their knowledge	2
		ability to realize western goods		
提升	ti sheng	growing the electronics business	improve/raise ourselves	1

Subsequently, the second perspective of showing the similarity between the polysemy transfer instances in students’ writings and MT output is analyzing the correlation between their accuracy in transferring Chinese polysemes and the word frequency of these Chinese polysemes. In other words, it will be shown whether negative transfer was more likely to be caused in more frequent Chinese polysemous words in MT output than in students’ writing. The results show that while both MT engines have a high error rate in dealing with those most frequent Chinese polysemes, their performance differs in translating the words with medium and low frequency. The results show that while both MT engines have a high error rate in dealing with those most frequent Chinese polysemes, their performance differs in translating the words with medium and low frequency.

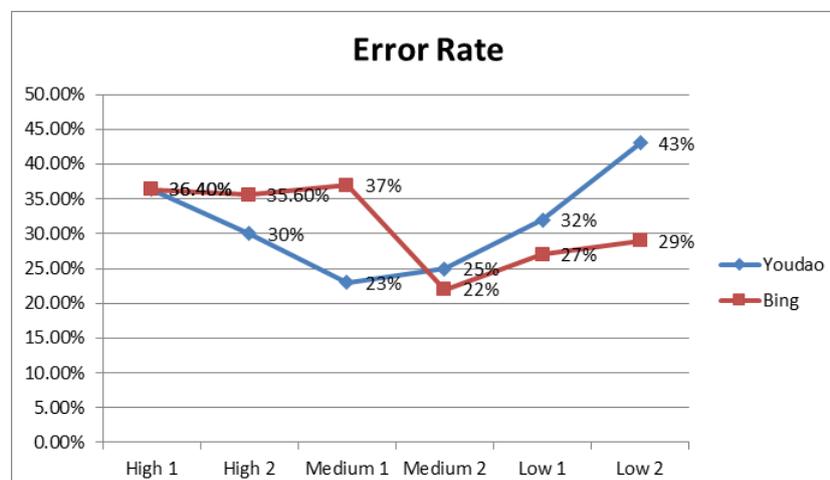


Figure 3. Error rate in transferring Chinese polysemes

Figure 3 shows the error rate of two MT engines in transferring words from six different ranges (from high to low), which were divided in terms of how many times the words appear in the subtitle corpus chosen by Cai and Brysbaert (2010). The six ranges of words frequency were defined in terms of the following standard (Wang, 2009):

Table 2

*The Standard of Defining Six Frequency Ranges*

	Time of appearance/one million words	Frequency ranking range
High 1	More than 500	1-234
High 2	50-500	235-1626
Medium 1	10-50	1627-2757
Medium 2	5-10	2758-8373
Low 1	1-5	8374-20945
Low 2	Less than 1	>20945

It can be seen from Figure 3 that two MT engines have the same error rate (36.4%) in translating words from High 1, which is the second highest rate among six ranges for both of them. However, a significant difference is that whereas Youdao's error rate bottoms at 23% in Medium 1, the one of Bing peaks at 37% in this range. Apart from this, although the error rate of both MT engines rises gradually from Medium 2, this tendency of Youdao is much more dramatic and reaches its summit (43%) at Low 2 range.

## 5 Discussion

### 5.1 The performance of Youdao and Bing

The first part of the statistical results shows that both Youdao and Bing MT engines positively transferred nearly 2/3 of the total Chinese polysemes identified in the webpage corpus. This finding supports Zhu's (2012) claim that recent Internet corpus-based MT engines have developed significantly in word sense disambiguation which makes the production of target word far from a random selection from possible equivalents. In other words, given the feature of polysemes that they have at least two related meanings, it is impossible for MT engines to have the accuracy more than 1/2 in transferring them if the process is based on random selection. Following examples might be helpful to explain the way in which MT engines managed to transfer Chinese polysemes positively:

- (1) 需要参加相关学位考试

Youdao: (1) Need to attend relevant degree examination

Bing: Need to take the degree examination

- (2) 在此一并表示感谢

Youdao: In here show appreciation

Bing: Also show appreciation

These two examples represent the typical positive polyseme transfer done by both MT engines (the polysemes being transferred and the target lexical items being produced are underlined in the above sentences). As was mentioned earlier, Melcuk and Wanner (2001) point out that identifying "linguistic clues" in the context where a word occurs is becoming a crucial approach for MT engines to choose a correct equivalent from the target language. It seems that the above examples are supportive of Zhu's

(2012) claim that several up-to-date Chinese to English MT applications are able to utilize this strategy. In example (1), while the Chinese word “参加” has at least three possible literal equivalents in English, namely “attend”, “join”, “participate”, Youdao MT engine succeeded in selecting the most appropriate one (“attend”). In addition, in example (2), where the Chinese word “表示” can be translated into English as “present”, “indicate”, “demonstrate”, and “show”, both Youdao and Bing selected “show” for the context, which it is the optimal result. The most tenable account for such positive polyseme transfer is that the crucial “linguistic clues” (bolded in the above sentences) are close to the source lexical items and therefore the MT engines are able to identify them and thereby produce the correct results.

Nevertheless, despite the considerable proportion of positive transfer accomplished by both MT engines, it is still noticeable that nearly 1/3 of the total Chinese polysemes found in the website corpus were negatively transferred. Some typical instances are presented as follows (the Chinese polyseme being transferred and the negative transfer of both MT engines are underlined, and the correct version is provided in the parentheses):

- (3) 精彩地演绎了马头琴独特的演奏技艺 (Perform)

Youdao: To demonstrate the Mongol stringed instrument unique art.

Bing: Beautifully interpret the matouqin's unique skills.

- (4) 其他成员也先后谈了他们的认识和体会 (Understanding)

Youdao: Other members also has talked about their knowledge and experience

Bing: other members have talked about their knowledge and understanding

It can be seen that while there is also a “linguistic clue” ( 技艺 ) in example (3), both MT engines translated the “演绎” literally into “demonstrate” and “interpret”, which is not as appropriate as the words like “perform” and “play” as this Chinese word means “the showing of a skill” in this context. This finding is to a degree against Zhu’s point stated above, which means that even the identification of “linguistic clues” is not even stable within a simple sentence. As for (4), there is no obvious “linguistic clue” in the given sentence to hint the exact meaning of “认识”. By examining the whole text, it is found that the article revolves around the control of corruption – a social issue rather than a subject. Therefore, it can be decided that “understanding” is more suitable for this context. This case is intricate since the selection of equivalent is based on the understanding of the entire context, which is undoubtedly beyond the ability of current MT engines and hence has caused several other similar instances of negative transfer in this study.

## 5.2 The comparison of polysemes found in both corpora

As was shown in the Results Chapter, all the polysemes that were found to appear in both students’ negative transfer data and MT output were juxtaposed and compared. As a result, it was discovered that only 18% of all identical lexical items were negatively transferred by both MT engines as by students, while 72% of them were positively transferred by both MT engines, with 10% of them positively transferred by either Youdao or Bing. This statistical result suggests that for most of the Chinese polysemes that the students have failed to cope with, the MT engines managed to transfer them in an appropriate way. In order to obtain a better understanding of how human beings and MT engines differ in dealing with these Chinese polysemes, some typical juxtaposed items are discussed here in detail as follows (the Chinese polysemes to be transferred and its transferred version in English are underlined. The original ideas that were to be expressed by students are given in Chinese in italics. The negative transferred version of students are marked by “\*”. The correct version of each negative transfer is given in brackets):

- (1) 实验室采用了全新的技术 (adopt)

MT: Laboratory has adopted new technology.

美国采用的避免不确定要素的方式 (adopt)

\* The way of avoiding uncertainties used by American

In this instance, the Chinese polyseme being transferred is “采用”, which can be generally translated to English as “adopt” and “use”. In the cases where the emphasis on “something is being used”, this two lexical items are interchangeable to serve as the equivalent for “采用”. However, “adopt” has the meaning of “begin to have or use something after consideration or evaluation”, which cannot be expressed by “use”. Taking the context into account, both sentences indicate something (“技术” and “方式” respectively) is newly accepted and used, and therefore “adopt” is the optimal choice here.

(2) 符合双方的利益 (interests)

MT: in line with the interests of both sides

这样会损害股东的利益 (interests)

\*This will hurt shareholder wealth

The Chinese polyseme being transferred in this example is “利益”, which have at least two equivalents in English – “benefit” and “interest”. Compared with “interest”, “benefit” is a more general word in terms of something improves personal life (result from appropriate way). When referring to “利益”, “interest” is always used in plural form and more likely to be used in formal register. As for the sentences in the example, it is clear that both of them are in the context of talking about the topic in relation to politics or business, which is rather formal. Thus, “interests” fits both sentences and is selected by the MT engine. Nonetheless, the production of students is “wealth”, which belongs to none of the possible literal translation of “利益” that can be retrieved from Chinese to English dictionary. This kind of transfer may reflect the case mentioned by Jiang (2000), in which learners transfer the word that has not even been memorized through dictionary definition and hence they will find a learnt word that is similar to the target meaning for replacement. In this example, it is likely that the students failed to recall the literal correspondence between “利益” and “benefit”/“interest” and then resorted to a seemingly relevant word based on the concept of “money”.

(3) 政府调整了工资标准 (standard)

MT: The government adjusted the salary standard.

这是它的标准定义 (standard)

\*This is its benchmark definition

The Chinese polyseme “标准” in (3) is usually translated to English as “standard” or alternatively “norm”. While “norm(al)” means something that is usual and expected by people, “standard” always refers to a level that is used to judge the quality of something else or something that is widely evaluated and accepted. In the context of two Chinese sentences above, it is clear that “standard” can fit both of them with either of its senses, yet only MT managed to choose it. It can be seen that students, again, produced a word (“benchmark”) that pertains to none of the literal equivalents of “标准” in a Chinese to English dictionary. By checking the English definition of “benchmark”, it can be found that although it may also refer to a level with which other things can be compared, it is specifically used in technical fields and has no usage in relation to academic texts or works. This type of transfer also attests the claim of Hemchua and Schmitt (2006) that except for being confused by related meanings in the same lexical field, students can sometimes overspecialize words in the target language, namely adopting a word whose range of use is far narrower than the source word being transferred. The existence of this kind of result suggests a significant difference between the polyseme transfer patterns of L2 learners and MT engines: while the transfer of MT strictly originates from the literally correspondent equivalents in bilingual dictionary, the transfer of L2 learners can be occasionally beyond this pattern due to their change to lexical range such as overspecialization.

Despite the disparity between the polyseme transfer patterns of MT and L2 learners, the results also show some similarity between them (18% of the identical lexical items were negatively transferred by both MT engines), which is worthy of discussion. The following examples present the Chinese polysemes that were negatively transferred by both students and MT engines:

(4) 通过这些方式 (in)

MT: Through these ways

Students: By these ways

The English prepositions “by” and “through” are the literal translation of Chinese polyseme “通过”, but neither of them can be used in the above expression as “in these ways” is the accepted set phrase that equals “通过这些方式”.

(5) 他写下了几个字 (character)

MT: He wrote down some words.

Students: He wrote some words down.

In this case, both MT and L2 learners ignored a unique feature of Chinese – the smallest meaningful unit is characters rather than words. This type of negative transfer reflects that the recognition of cross-language feature shift can be a crucial weakness of both MT engines and L2 learners in transferring Chinese polysemes to English.

In addition to these specific instances, another illuminating finding is that most of these negatively transferred examples (by both MT and learners) have high frequency ranking in the frequency list, namely “SUBTLEXT-CH-WF”, used in this study (e.g. “通过”=544; “重要”=306 “国家”=590). More accurately, the average frequency ranking of these words is 854 according to the calculation, which belongs to the range of “high frequency words” in terms of the classification stipulated in the previous chapter. This discovery may imply a crucial similarity between the polyseme transfer patterns between MT engines and L2 learners – both tend to cause more negative polyseme transfer when dealing with those with high word frequency. The detail and reason in relation to this conjecture are further discussed in the following section with the reference to the “error rate” line graph provided before.

### 5.3 Correlation between word frequency and error rate

It has been presented in the result that the error rate of both MT engines in the High 1 range is 36.4%, which is the second highest figure among all ranges for both. As was found by Ma (2015), Chinese English learners also tend to produce more negative transfer instances in coping with Chinese polysemes with higher word frequency. Therefore, it seems that the performance of both MT engines and students is likely to decrease with the high frequency Chinese polysemes, which is in line with the hypothesis stated in this study. The occurrence of this similarity between MT engines and students can be ascribed to their resemblance in transferring polysemes. Specifically, despite the huge difference between MT engines and students, they are essentially confronted with the same process in transferring polysemes – selecting a correct equivalent from a group of possible words in target language. In this way, since the Chinese polysemes tend to have more related senses with the increase of frequency, both MT engines and students will have to choose an optimal word from more potential lexical items when facing Chinese polysemes of high frequency, which poses greater difficulty for both in the transfer process.

Another compelling result that can be seen from the line graph is that the error rate of both MT engines starts to rise again after slumping to their lowest point in Medium 1 (Youdao) and Medium 2 (Bing) respectively, though this upward trend of Youdao appears to be more drastic. This trend indicates that whereas both MT engines reach their ideal transferring performance in the Medium frequency range, the error rate somehow re-increases after that with the decline of words frequency. This negative correlation between the word frequency and MT engines’ accuracy in translating Chinese polysemes was

also found in the study conducted by Zhu (2016), in which it is shown that the error rate of MT engines increases dramatically with the word frequency becoming lower after the medium frequency range. This correlation pattern is not surprising given the operating principle of MT engines like Youdao and Bing. As was pointed out by Zhu (2012), the MT engine utilized by Youdao and Bing is Internet corpus-based one that collected enormous number of authentic human translation versions of various words occurring in different contexts, based on which the MT engine is able to select the optimal target equivalent for a given word. In other words, the reliability of MT engine in transferring a word is largely dependent on the number of well-translated lexical items available in the Internet corpus. Since the Chinese words with low frequency are rarely used in daily life communication and written texts, the authentic translation versions of these words are also more scarce in the Internet corpus. In this way, with the words frequency becoming lower, there will be less available example translation for this type of MT engines to retrieve, which renders the results of their polyseme transfer less reliable.

Based on the two features of MT engines discussed above, it will be possible to account for the reason why the error rate of both MT engines bottoms at the Medium frequency range. At first, the number of related meanings of a polyseme has been halved in the Medium frequency range according to the data of Wang (2009), which relieve the processing load exerted on MT engines by reducing possible equivalents in the target language. At the same time, as words of medium frequency have important status in various spoken texts and written texts (especially those that are formal) for the expression of relatively complex ideas, the total number of appearances in the corpus is still considerable (Wang, 2009), which ensure the number of human translated versions in the Internet corpus is enough for MT engines to obtain reliable results.

However, with the negative polyseme transfer caused by students, it was found that the error rate starts to remain unchanged at a relatively high level (over the one of MT engines) from the word frequency ranking of 1,000 (Zhu, 2016). That is to say, for words whose frequency belongs to medium range and low range, the performance of students in transferring Chinese polysemes is almost constant, regardless of how the word frequency of those polysemes varies. This uncorrelated trend between words frequency and error rate in the Medium and Low range is apparently different from the one of MT engines discussed above. According to Jiang (2000), although the connection between L1 and L2 words may become looser with the learners advancing in L2 competence, it will never disappear completely. In other words, the influence of L1 on L2 acquisition will continue to exist at every level of learners L2 competence, which might be the essential reason for the constantly noticeable error rate of year 4 students in dealing with polysemes of medium and low word frequency (more advanced words).

## **6 Conclusion**

Based on the two fields (L2 learners' writing and MT output) in which lexical transfer can exert a large influence, this current research attempted to explore the lexical transfer caused by Chinese polysemes in MT output and compare its patterns with the ones yielded by Chinese students.

Results collected from MT output show that the two MT engines performed similarly in translating Chinese polysemes into English as both of them negatively transferred approximately 30% of the total Chinese polysemes. Afterwards, by comparing the results from MT output with the contrastive data, it was found that among all the identical lexical items, nearly 3/4 of them were positively transferred by both MT engines, while only about 1/5 of them were negatively transferred by both MT engines. This compelling difference may indicate that the overall performance of MT engines is better than the one of students in dealing with Chinese polysemes, which is probably because of the development of MT engines, especially its ability to select optimal equivalents by identifying "linguistic clues" in the given sentence. By calculating the correlation between the word frequency of Chinese polysemes and the error

rate of MT engines, it was revealed that both MT engines are more likely to yield negative polyseme transfer in High frequency range and Low frequency range. By comparing this correlation with the one of students that was researched previously, it was discovered that both students and MT engines have more difficulty in coping with Chinese polysemes with high frequency. The reason for this similarity can be explained by the essence of lexical transfer, namely choosing the optimal item from a number of possible equivalents in the target language. As more frequent words tend to have more related senses, the processing load exerted on both students and MT engines becomes greater, although they are based on a totally different mechanism.

The findings obtained from this study can have several implications for the development of MT engines and other applications in coping with polysemous words. At first, the process of dealing with polysemous words is a significant problem to be considered by the developers of MT engines, as it is still a major difficulty faced by both domestic and overseas MT engines (e.g. Youdao and Bing). Specifically, attention should be paid to the polysemes of high and low frequency. For the polysemes of high frequency, it will be necessary to improve MT engines' ability of identifying linguistic clues in the context, the scope of which should be extended to the whole text but not only sentences. For those of low frequency, it will be important to increase the size and comprehensiveness of human-translated corpus in order to create enough retrievable items for MT engines. Secondly, due to the disparity between the transfer patterns of MT engines and learners in processing Chinese polysemes, it will be unlikely for the post-editing application developers to treat them as totally the same and develop the correction mode applicable for both. However, despite the overall difference, partial similarity such as their error rate in High frequency range does exist. This suggests that it is possible for developers to use similar strategy for the correction of negative transfer caused by Chinese polysemes of high frequency in Chinese learners' English writing and MT output. Finally, the findings in this study may also be used to create an algorithm for the detection of ethics code violation through the use of MT engines in written assignments. For example, further analysis of lexical errors in both MT and learner English writing could be conducted to build a list of polysemes likely to provide fruitful points for comparison. For situations where producing an essay in the target language without using MT is desirable or required, such a list could make detection of over-reliance on MT computationally feasible both in terms of narrowing down the amount of data to be processed and in terms of applicability to a wide range of essay topics

Some limitations can be recognized in this study. Although effort was made to improve the comparability of two corpora in terms of formality, it may still be insufficient as there are differences between webpage articles and student writing regarding their register and genre. In addition, the researcher assumed the possibility of mistranslation by MT engines and preselected Chinese polysemes to be analyzed in webpage corpus. This method might lead to the miss of available data since the potential of mistranslation of some source Chinese words can be hard to judge beforehand due to various contexts. Therefore, future research could consider build corpora of more similar register and analyze the whole translation produced by MT engines and pick out mistakenly transferred polysemes from lexical errors being made.

## References

- Agustín-Llach, M. P. (2020). Semantic and Conceptual Transfer in FL: Multicultural and Multilingual Competences. In M. Dodigovic & M. P. Agustín-Llach (Eds.), *Vocabulary in Curriculum Planning: Needs, Strategies and Tools* (pp. 165-180). Cham: Springer International Publishing.
- 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.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film

- subtitles. *Plos one*, 5(6), e10729.
- Carter, R. (1998). *Vocabulary: Applied linguistics perspectives* (2nd ed.). London: Routledge.
- Dodigovic, M. (2005). Vocabulary profiling with electronic corpora: A case study in computer assisted needs analysis. *Computer Assisted Language Learning*, 18(5), 443-455.
- Dodigovic, M. (2013). Vocabulary learning with electronic flashcards: Teacher design vs. student design. *Voices in Asia Journal*, 1(1), 1-33.
- Dodigovic, M., Liang, H., Yue, Y. (2015). Language technology project: an application for Chinese to English machine translation post-editing and Chinese-English bilingual writing correction. Collaboration proposal.
- Dodigovic, M., Ma, C., Jing, S. (2017). Lexical Transfer in the Writing of Chinese Learners of English, *TESOL International Journal*, 12(1), 75-90.
- Ellis, H. (1965). *The transfer of learning*. New York: Macmillan.
- Ellis, R. , & Heimbach, R. . (1997). Bugs and birds: children's acquisition of second language vocabulary through interaction. *System*, 25(2), 247-259.
- Ellis, R. (2008). *The study of second language acquisition* (2<sup>nd</sup> edition). Oxford University Press.
- Elston-Güttler, K. E., & Williams, J. N. (2008). First language polysemy affects second language meaning interpretation: Evidence for activation of first language concepts during second language reading. *Second Language Research*, 24(2), 167-187.
- Fillmore, C. J. & Atkins, B. T. S. (2000). "Describing polysemy: The case of 'crawl'". In Ravin, Y; Leacock, C. Polysemy: *Theoretical and computational approaches*. Oxford: Oxford University Press. pp. 91–110.
- Harutyunyan, N. & Dodigovic, M. (2020). Lexical Errors in the Writing of EFL Students in the Armenian Context. In M. Dodigovic & M. P. Agustín-Llach (Eds.), *Vocabulary in Curriculum Planning: Needs, Strategies and Tools* (pp. 165-180). Cham: Springer International Publishing.
- Hemchua, S. & Schmitt, N. (2006). An analysis of lexical errors in the English compositions of Thai learners. *Prospect*, 21(3), 3-25.
- Hsieh, S. K. (2011). Sense Structure in Cube: Lexical Semantic Representation in Chinese Wordnet. *International Journal of Computer Processing Of Languages*, 23(3), 243-253.
- Ijaz, I. H. (1986). Linguistic and Cognitive Determinants of Lexical Acquisition in a Second Language. *Language Learning*, 36, 401-451.
- Jiang, N. (2000). Lexical representation and development in a second language, *Applied linguistics*, 21(1), 47-77.
- Lennon, P. (1996). Getting "easy" verbs wrong at the advanced level', *IRAL-International Review of Applied Linguistics in Language Teaching*, 34(1), pp. 23-36.
- Ma, C. (2015). Lexical Transfer from Chinese to English in the Writing of XJTLU students. Project report.
- Martin, M. (1984). Advanced vocabulary teaching: The problem of synonyms. *The Modern Language Journal*, 68, 130–137.
- Melčuk, I., & Wanner, L. (2001). Towards a lexicographic approach to lexical transfer in machine translation (illustrated by the German–Russian language pair). *Machine Translation*, 16(1), 21-87.
- Melcuk, I. & Wanner, L. (2001). Towards a lexicographic approach to lexical transfer in machine translation (illustrated by the german-russian language pair). *Machine Translation*, 16(1), 21-87.
- Morimoto, S., & Loewen, S. (2007). A comparison of the effects of image-schema-based instruction and translation-based instruction on the acquisition of L2 polysemous words, *Language Teaching Research*, 11(3), 347-372.

- Nation, I. S. P. (2013). *Learning vocabulary in another language. 2<sup>nd</sup> edition*. Cambridge: Cambridge University Press.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. London: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Ernst Klett Sprachen.
- Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55-88.
- Schmitt, N., & Hemchua, S. (2006). An analysis of lexical errors in the English composition of Thai learners: 2698. *Prospect: an Australian journal of TESOL*, 21(3), 3-25.
- Wang, H. (2009). Sense, length, and frequency of words – a quantitative analysis of polysemes in Modern Chinese Dictionary. *Chinese Language*, 329(2), 120-130.
- Wang, D. (2011). 'Language Transfer and the Acquisition of English Light Verb+ Noun Collocations by Chinese Learners', *Chinese Journal of Applied Linguistics*, 34(2), pp. 107-125.
- Yule, G. (2014). *The study of language*. London: Cambridge University Press.
- Wu, S. (2013). Negative lexical transfer in Chinese English writing. *Journal of Jixi University*, 13(3), 99-100.
- Yule, G. (2014). *The study of language*. London: Cambridge University Press.
- Zhang, C., & Wang, D. (2005). Negative Lexical Transfer of Native Language in Chinese College Students' English Writing. *Sino-US English Teaching*, 2(2), 13-16.
- Zhu, L. (2012). The status quo and development of Chinese-English machine translation. *Xinshijiao*, 24, 95-96.
- Zhu, H. (2016). A Contrastive Analysis of L1's Negative Transfer in the Chinese -English Translation between Machine Translation and L2 Students in terms of MWU. Project report.

**Zhenyan Ye** obtained his MA TESOL degree at Hong Kong University. He worked at EF Education as an English teacher for three years after graduation and is now working as a freelance English teacher. His research interest is in the second language acquisition of English and Japanese by speakers of Chinese.