Article

Effects of EFL Learning on L1 Chinese Lexis

Wen Kong [1], Quan-Jiang Guo [2,*], Yin-Yan Dong [2] and Xuesong (Andy) Gao [3]

[1] Department of English, Changzhou University, Changzhou 213100, China; kongwen1970@sina.com
[2] College of Science and Technology, Ningbo University, Ningbo 315211, China; dongyinyan2013@163.com
[3] School of Education, University of New South Wales, Sydney, NSW 2052, Australia;
    xuesong.gao@unsw.edu.au
* Correspondence: guoquanjiang@nbu.edu.cn

Abstract: The advocates of multi-competence theory argue that the L2 learners' language system is unique because of the crosslinguistic influences of both languages. However, the influence of a foreign language on the learner's L1 has not been extensively investigated. In order to address the gap, the present study sought to investigate the effects of EFL learning on written L1 Chinese at the lexical level. Two studies were conducted on 200 abstracts of MA theses written in Chinese, half on English literature written by Chinese-L1 English majors (EMs), and half on Chinese literature written by Chinese-L1 Chinese majors (CMs). The first study investigated the differences between the two groups in terms of the frequencies of nouns, verbs, adjectives, adverbs, pronouns, prepositions, and conjunctions in the abstracts. The second study examined the differences in the lexical complexity and diversity between the two groups. The results reveal 12 significant differences in 27 investigated word classes and subclasses, as well as significant differences in lexical complexity, but no significant difference in lexical diversity. The identified differences are discussed from a multi-competence perspective.

Keywords: crosslinguistic influence; multi-competence; backward transfer; lexical complexity and diversity

1. Introduction

Since Ulrich Weinreich first introduced the concept of "interference" as "those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language" [1] in 1953, there has been extensive research into two fields: the forward transfer from a first language (L1) to a second language (L2), and the backward or reverse transfer from L2 to L1. The former, the more obvious and practical of the two interferences, has been studied for longer in the field of second language acquisition (SLA) [2–6]. Backward transfer, from L2 to L1, came into focus much later and is usually explored from two dimensions: (1) The attrition or loss of L1 in a migration setting or L2-dominant context; and (2) The effect of L2 on L1 in an L1-dominant context, where L2 is learned as a foreign language (FL) and used mainly for academic purposes. The attrition or loss of L1 can be found at the phonological, morphological, lexical, semantic, pragmatic, rhetorical, and conceptual levels. The latest research results were gathered in the *Oxford Handbook of Language Attrition*, edited by Schmid and Kopke, the first book dedicated to language attrition [7]. However, few studies have discussed the effects of L2 in an L1-dominant context, wherein L2 users have scarcely any exposure to L2 culture. In order to address this gap, the present study sought to investigate whether sustained and intensive English as a foreign language (EFL) learning can affect learners' written Chinese (L1) in an exclusively Chinese context.

2. Multi-Competence Theory

Multi-competence theory, first proposed by Cook [8], represents a novel theoretical take on backward transfer. Specifically, it questions the stability of the mother tongue and

argues for the case of L2 users, whose unique language is different from that of monolingual speakers of either language.

Research on SLA has been seen through a monolingual prism for decades; however, some researchers, including Meara, Grosjean, and Cook, disapprove of the monolingual view because it evaluates those learning to be bilingual as if they were L1 speakers in both languages [9–16]. Meara argues that, "there is no reason why a person who speaks both English and Spanish should behave in the same way as a monolingual speaker of either language" [9]. Grosjean suggests that the bilingual is "a specific and fully competent speaker/hearer who has developed a communicative competence that is equal, but different in nature, to that of the monolingual" [11]. Cook reflects on the biased monolingual view and proposes "multi-competence", taking L2 learners not as deficient monolinguals, but as people in their own right [12–16]. He first defined "multi-competence" in L2 learning as "the compound state of a mind with two grammars" [12], but later developed this into "the overall system of a mind or a community that uses more than one language". He thereby extends the concept beyond the psychological construct of the mind of the individual to the sociological construct of the multi-competence of the community, treating the diverse languages of the community as a coherent whole rather than separately [16].

Cook proposed an integration continuum to explain three possible relationships between L1 and L2 in the mind of a bilingual user: total separation, in which the two languages are independent; interconnection, in which they are connected to a greater or lesser degree; and total integration, in which they form a single system [13,14]. However, Cook suggests that neither total separation nor total integration can be absolutely true: total separation is impossible, since both languages are in the same mind, and total integration is also impossible, since L2 users can keep the languages apart. In between the two extremes, there are many different degrees and types of interconnection. This study examines one type of interconnection: the possibility that L2 learners' unique competence affects differences in their native monolingual competence.

3. Research on the Effects of FL on L1 in an L1-Dominant Environment

A review of the literature reveals a paucity of research on the effects of FL in an L1-dominant environment. The study by Van Hell and Dijkstra found that foreign language knowledge can influence native language performance in exclusively native language contexts and that backward transfer becomes noticeable in highly proficient L2 users [17]. Tokumara found that monolingual Japanese speakers associated the English loanword 'bosu' with 'gangsters' or 'monkeys', while monolingual English speakers associated 'boss' with 'office' or 'work' [18]. However, as Japanese learners acquired proficiency in English, they associated 'bosu' less with 'gangsters' or 'monkeys', and more with 'office' or 'work.' Sun, investigating the reverse transfer effects of EFL learning on L1 Chinese [19], did not find a significant difference in Chinese word diversity or complexity among three groups of EFL learners with different English proficiencies, although they differed significantly in their use of synonyms. The EFL learners who had studied English for longer and had acquired higher proficiency in English used more synonyms in their written Chinese, which she ascribed to the knowledge of a foreign language opening up new lexical routes to the same concept in the Common Underlying Conceptual Base [19,20]. Liu and Ni demonstrated the influence of English (L2) tag question knowledge on Chinese (L1) tag questions in advanced Chinese EFL learners, and a semantic extension from English to Chinese proved that intensive L2 learning in an L1-dominant environment can trigger the effects of FL on a strongly built L1 [21].

This review of the relevant literature shows that backward transfer from FL to L1 can occur in an exclusively L1 context. However, there is still room left to explore more nuanced questions. For instance, the participants in Sun's research were high school students, undergraduate non-English majors, and English majors. If participants were advanced EFL learners, such as postgraduates majoring in English, what impact would intensive EFL learning have on their L1 written language? In her study, Sun focused on

the differences in the frequency of using nouns, verbs, prepositions, and conjunctions without exploring the subclasses of each word class in depth. Is it possible that while no significant differences exist in the total frequency of one word class, they may still exist in the subclasses of the same word class? For instance, there are seven subclasses in the Chinese verb category; thus, just because no significant difference in the total number of verbs between independent groups was found, this does not mean that no significant difference in the subclasses of verbs exists. Besides this, the writing task in her study was conducted as part of a controlled experiment, yet there is a key difference between a timed writing task and an authentic writing task. When the writing product is acquired in a natural and authentic environment, participants can obtain feedback from their classmates, teachers, or supervisors; make use of the suggested references or sources; and have enough time to think about and repeatedly revise their writing structure and language. Would a change in the time constraint lead to a difference in the writing performance, especially regarding lexical and syntactic features? With these questions in mind, we attempted to pursue more specific instances of multi-competence caused by EFL's effect on the written Chinese vocabulary. The reason for choosing to focus only on the Chinese vocabulary is that it is an "open class" system, allowing for greater adaptability in both bilingual and monolingual environments because new items are relatively easy to add [22]. Therefore, if EFL has any effects on L1, they may be more notable in the vocabulary than in other aspects of language.

4. Research Design

This study is driven by two research questions:

(1) Is there any difference between advanced EFL learners and advanced Chinese users regarding the frequency of using nouns, verbs, adjectives, adverbs, pronouns, prepositions, and conjunctions in L1 Chinese writing?

(2) Is there any difference between advanced EFL learners and advanced Chinese users regarding lexical complexity and diversity in their L1 Chinese writing?

4.1. Subjects and Materials

For the purpose of the study, 100 Chinese abstracts of MA theses on English literature written by English majors (EMs), and 100 Chinese abstracts of MA theses on Chinese literature written by Chinese majors (CMs) were selected randomly from the subdatabase of Master's and Doctoral Theses of Wanfang Data (http://g.wanfangdata.com.cn, accessed on 30 July 2021), an online database providing various academic writings published in China. We confined the materials to the field of literature to keep the language style and discourse features in the 200 Chinese abstracts as similar as possible.

As the literature review showed that backward transfer occurs with increasing L2 proficiency, postgraduate EMs' English proficiency was good enough for us to detect whether EFL had an influence on L1 Chinese. Although postgraduate CMs are not pure monolinguals, it is undeniable that most of them did not learn English as intensively as the EMs. Generally, EMs represented advanced EFL learners and CMs represented advanced Chinese users.

The total number of Chinese characters in the EMs' abstracts was 75,541, and that in the CMs' abstracts was 71,538. The descriptive statistics and independent samples of a *t*-test revealed no significant difference in length between the EMs' and CMs' abstracts ($p = 0.195$).

4.2. Instruments and Procedure

To investigate Research Question 1, the following three steps were followed:

Step 1: The segmentation of Chinese words and the tagging of each word with its part of speech were automatically conducted using the Natural Language Processing and Information Retrieval Sharing Platform (NLPIR), a platform developed by Zhang Huaping of the Web Search, Mining and Security Lab, Beijing Institute of Technology, China. Because

the total numbers of Chinese characters for the EMs and CMs were not equal, we used the following formula to standardize the data to enable a comparison between the two independent groups:

$$\frac{\text{the total number of each word class in each abstract}}{\text{the total number of Chinese charaters in each abstrat}} \times 1000$$

Step 2: After standardization, the frequency of each word class and its subclasses was calculated with Microsoft Excel 2010 (Microsoft Corporation, Redmond, WA, USA) and AntConc 3.4.0, a software package developed by Laurence Anthony of the Faculty of Science and Engineering, Waseda University, Japan.

Step 3: A Mann–Whitney test was conducted in SPSS 19.0 to check the differences in the frequency of each of the word classes and its subclasses between the two groups.

To investigate Research Question 2, the following steps were followed:

(1)   Lexical complexity

Step 1: The lexical complexity of Chinese words in the present study is defined as the percentage of infrequent vocabulary in an abstract. The *Syllabus of Graded Words and Characters for Chinese Proficiency* (SGWCCP), compiled by the Testing Center of the National Chinese Proficiency Test Committee, was adopted as a reference to divide Chinese words into four levels on the basis of their frequency of use, which ranged from the most frequent (Level 1) to the least frequent (Level 4). There are 8822 Chinese words contained in the SGWCCP, covering more than 95% of Chinese words [23].

Step 2: A program developed by the authors was used to count the total number of Chinese words at each frequency level.

Step 3: The following formula was adopted to standardize the data to allow a comparison between the two groups:

$$\frac{\text{the total number of Chinese words at each level in each abstract}}{\text{the total number of Chinese words in each abstrat}} \times 1000$$

Step 4: The differences in the frequency of words at the four different levels between the groups were tested with the Mann–Whitney test in SPSS 19.0.

(2)   Lexical diversity

Step 1: Lexical diversity, in the present study, is defined as the ratio of the number of different Chinese words to the total number of Chinese words. On the basis of the automatic segmentation and tagging of each word class by NLPIR, AntConc3.5.0 was used to count the total number of Chinese words and the number of different Chinese words in each text.

Step 2: The following formula was adopted to standardize the data to facilitate a comparison between the two groups:

$$\frac{\text{the total number of different Chinese word in each abstract}}{\text{the total number of Chinese words in each abstrat}} \times 1000$$

Step 3: The differences in lexical diversity between the two independent groups were tested with the independent samples *t*-test in SPSS 19.0.

5. Results and Discussion

5.1. Research Question 1: Frequency of Word Class

(1)   Nouns

NLPIR segmented and tagged four classes of proper nouns, and noun phrases or idioms. The focus in this study was on proper nouns, which include the names of people, the names of places, the names of an institutes or organizations, and other proper nouns. The results of the Mann–Whitney test are shown in Table 1.

Table 1. Ranks and Test Statistics [a] (nouns).

	Group	N	Mean Rank	Sum of Ranks	Mann–Whitney's U	Wilcoxon's W	Z	Asymp. Sig. (2-Tailed)	r
Nouns	CM	100	90.76	9075.50	4025.500	9075.500	−2.381	0.017	0.168
	EM	100	110.25	11,024.50					
	Total	200							
Names of people	CM	100	76.07	7606.50	2556.500	7606.500	−5.972	0.000	0.422
	EM	100	124.94	12,493.50					
	Total	200							
Names of places	CM	100	86.49	8649.00	3599.000	8649.000	−3.439	0.001	0.243
	EM	100	114.51	11,451.00					
	Total	200							
Names of institutes or organizations	CM	100	102.00	10,200.00	4850.000	9900.000	−1.741	0.082	0.123
	EM	100	99.00	9900.00					
	Total	200							
Other proper nouns	CM	100	100.81	10081.00	4969.000	10,019.000	−0.106	0.915	0.007
	EM	100	100.19	10,019.00					
	Total	200							

[a] Grouping variable: group.

The results in Table 1 show that the frequencies of nouns, names of people, and names of places used by EMs were significantly higher than the frequencies used by CMs. This difference may be because English majors must translate many English names of people and places into Chinese. The high frequency of these two kinds of nouns may lead to the total number of nouns in EMs' abstracts being significantly higher than in the CMs'. However, despite this explanation, there are other reasons that need to be explored. Some Chinese scholars have noticed that, compared with Chinese, nouns are preferred in English to express ideas or describe a person or thing [24,25]. Some linguists claim that English is primarily a "nominalizing language" [26,27]. In the process of the nominalization of English, three types of nouns are generally used to denote an action: an action noun derived from a verb with the addition of a derivational suffix (e.g., preparation, adjustment); a gerund derived from a verb with the addition of *-ing* (e.g., setting, building); and the same word used as a noun without any morphological change (e.g., control, search). English majors are more likely to keep the same nominal class when they translate English into Chinese or write in Chinese. For example:

第二章对弗吉尼亚·伍尔夫独特的女性主义视角进行了**探讨** . . . 第三章聚焦在对文本《一间自己的房间》中的**分析**。(EM No. 19)

(In Chapter two, **discussions** about Virginia Woolf's unique feminist perspective will be made . . . Chapter Three will focus on an **analysis** of the text of *A Room of One's Own*.)

探讨 (discussion) and 分析 (analysis) (in italics) are verbs in traditional Chinese; however, both of them are used as nouns in this EM's written Chinese. This also occurs in CMs' abstracts because CMs are not real monolinguals and every university student in China has studied English for at least 10 years. Their Chinese is thus likely to have also been affected by English, although the change in their Chinese is less noticeable.

Another reason is that the serial verb construction, also known as verb serialization or verb stacking, a syntactic phenomenon in which two or more verbs or verb phrases are strung together in a single clause [28], is frequently used in traditional Chinese, but EMs are inclined to transfer the English nominal structure into Chinese. For example:

牛仔少年们是有血有肉，有生有死，能力有限却也努力逐梦的人。(EM No.34)

(The cowboys are flesh-and-blood mortal men who insist on the pursuit of their dreams in spite of limited ability.)

A CM may express the same meaning with a serial verb construction without the linking verb, 是, and the noun, 人, which makes Chinese more concise and idiomatic:

牛仔少年们有血有肉，有生有死，能力有限却也努力逐梦。

(The cowboys (are) flesh-and-blood mortal (men), pursue their dreams in spite of limited ability.)

Although the EM's longer structure is acceptable in Chinese, it sounds verbose and unidiomatic to Chinese native speakers.

The third important reason is that null-subject sentences are common in Chinese when a subject can be inferred from the sentence or context. Usually, it is a noun that functions as the subject. For example, the following is an excerpt from a CM's abstract without the subject, 本文, in parentheses:

最后，(***本文***)
探索了《檀香刑》民间叙事在现当代文学史的意义，梳理五四到90年代民间叙事的种种变迁，着眼于当前中国传统文化复苏的大背景 . . . (CM No. 8)

(Finally, (***this paper***) explored the significance of *Sandalwood Penalty* to contemporary literature, reviewed the various development of folk narrative from the May Fourth Movement to the 1990s, focused on the recovery of traditional Chinese culture . . .)

However, EMs will usually insert the typical English syntactic feature of the subject into their sentences. For example:

本文通过对吉姆·凯尔高动物小说主题的探讨，揭示了其作品的独特魅力 . . . (EM No.6)

(Through the discussion of Jim Kjelgaard's animal fiction, ***this paper*** reveals the unique charm of his work . . .)

The common phenomenon of EMs adding subjects to each sentence may lead to the increased number of nouns in their written Chinese.

(2)   Verbs

NLPIR segmented and tagged seven classes of verbs: adverbial verbs (verbs functioning as an adverbial, such as 翻 in 翻读); nominal verbs (verbs functioning as a noun, such as 影响 in 政治影响); 是 verbs (verbs similar to the function of link verbs in English); 有 verbs (verbs expressing the meaning of 'possess' or 'have/has'); directional verbs (verbs expressing the direction of an action, such as 上); formal verbs (verbs without notional content, but requiring another verb to be its object, such as 给以 in 给以帮助); and intransitive verbs. The results of the Mann–Whitney Test are shown in Table 2.

The results in Table 2 indicate that EMs use the different classes of verbs less frequently than CMs, except for adverbial verbs, and the differences in the total number of verbs, nominal verbs, and directional verbs is great enough to be significant. Nominal verbs in Chinese are similar to action nouns and gerunds in English. In fact, since the May Fourth Movement in 1919, through which Western culture was introduced into China through the translation of foreign works during that period, which had a great impact on Chinese lexicon and syntax, nominal verbs have continued to grow in Chinese under the influence of foreign languages [25]. The results in Table 2 are in agreement with the results in Table 1. Compared with their Chinese counterparts, EMs use more nouns and fewer verbs in their written Chinese, seemingly due to prolonged and intensive English learning.

Table 2. Ranks and Test Statistics [a] (verbs).

	Group	N	Mean Rank	Sum of Ranks	Mann–Whitney's U	Wilcoxon's W	Z	Asymp. Sig. (2-Tailed)	r
Verbs	CM	100	121.04	12,104.00					
	EM	100	79.96	7996.00	2946.000	7996.000	−5.019	0.000	0.355
	Total	200							
Adverbial verbs	CM	100	95.10	9510.00					
	EM	100	105.90	10,590.00	4460.000	9510.000	−1.656	0.098	0.117
	Total	200							
Nominal verbs	CM	100	122.20	12,220.00					
	EM	100	78.80	7880.00	2830.000	7880.000	−5.302	0.000	0.375
	Total	200							
是 verbs	CM	100	107.32	10,732.00					
	EM	100	93.68	9368.00	4318.000	9368.000	−1.666	0.096	0.118
	Total	200							
有 verbs	CM	100	105.75	10,575.00					
	EM	100	95.25	9525.00	4475.000	9525.000	−1.368	0.171	0.097
	Total	200							
Directional verbs	CM	100	112.00	11,199.50					
	EM	100	89.01	8900.50	3850.500	8900.500	−2.813	0.005	0.199
	Total	200							
Formal verbs	CM	100	106.66	10,666.00					
	EM	100	94.34	9434.00	4384.000	9434.000	−1.548	0.122	0.109
	Total	200							
Intransitive verbs	CM	100	107.41	10,740.50					
	EM	100	93.60	9359.50	4309.500	9359.500	−1.687	0.092	0.119
	Total	200							

[a] Grouping variable: group.

(3)   Adjectives and adverbs

NLPIR segmented and tagged three classes of adjectives: adverbial adjectives (adjectives functioning as an adverb, such as 轻 in 轻放); nominal adjectives (adjectives functioning as a noun, such as 骄傲 in 骄傲使人落后); and distinguishing adjectives (adjectives functioning to classify different things or people, or to designate the nature of things, such as 男 or 重型). NLPIR does not further classify adverbs. Therefore, adjectives and adverbs are discussed together in this section to save space. The results of the Mann–Whitney test are shown in Table 3.

Table 3. Ranks and Test Statistics [a] (adjectives and adverbs).

	Group	N	Mean Rank	Sum of Ranks	Mann–Whitney's U	Wilcoxon's W	Z	Asymp. Sig. (2-Tailed)	r
Adjectives	CM	100	106.95	10,695.00					
	EM	100	94.05	9405.00	4355.000	9405.000	−1.576	0.115	0.111
	Total	200							
Adverbial adjectives	CM	100	94.22	9421.50					
	EM	100	106.79	10,678.50	4371.500	9421.500	−1.543	0.123	0.109
	Total	200							
Nominal adjectives	CM	100	93.62	9362.00					
	EM	100	107.38	10,738.00	4312.000	9362.000	−1.718	0.086	0.121
	Total	200							
Distinguishing adjectives	CM	100	90.68	9067.50					
	EM	100	110.33	11,032.50	4017.500	9067.500	−2.401	0.016	0.170
	Total	200							
Adverbs	CM	100	104.15	10,414.50					
	EM	100	96.86	9685.50	4635.500	9685.500	−0.891	0.373	0.063
	Total	200							

[a] Grouping variable: group.

The results in Table 3 show that the frequency of distinguishing adjectives in EMs was significantly higher than that in CMs. Distinguishing adjectives cannot be modified by the

adverb, 很 (very), in Chinese, as other adjectives can. In English, adjectives that are usually used as attributives, such as 'earthen', 'atomic', and 'medical', as well as some prefixes, such as super-, mini-, multi-, and non-, have been increasingly translated into Chinese and have become an important part of Chinese distinguishing adjectives. Sustained EFL learning may encourage EMs to use these sorts of adjectives more frequently in their Chinese writing.

The results of the Mann–Whitney Test demonstrate that there was no significant difference in adverbs between EMs and CMs, suggesting that adverbial use is not affected by EFL learning. Since the May Fourth Movement in 1919, the number of Chinese adverbs has been increasing under the influence of foreign languages; however, the increase in Chinese adverbs does not occur through the adoption or creation of new words, but rather through adding an auxiliary word, 地 (similar to the English suffix "-ly"), after Chinese verbs, nouns, or adjectives, such as in the adverbs 讽刺地, 科学地, or 冷酷地 [29]. Perhaps the special word formation of Chinese adverbs caused there to be no significant difference found in the frequency of adverbs between the two groups. Another reason may be that the grammatical features and semantic categories of Chinese adverbs roughly correspond to those of English adverbs. Thus, no notable backward transfer from English was found for Chinese adverbs.

(4)   Pronouns

NLPIR segmented and tagged three kinds of pronouns: personal pronouns, demonstrative pronouns, and interrogative pronouns. The results of the Mann–Whitney test are shown in Table 4.

Table 4. Ranks and Test Statistics [a] (pronoun).

	Group	N	Mean Rank	Sum of Ranks	Mann–Whitney's U	Wilcoxon's W	Z	Asymp. Sig. (2-Tailed)	r
Pronouns	CM	100	82.55	8255.00	3205.000	8255.000	−4.386	0.000	0.310
	EM	100	118.45	11,845.00					
	Total	200							
Personal pronouns	CM	100	85.16	8516.00	3466.000	8516.000	−3.749	0.000	0.265
	EM	100	115.84	11,584.00					
	Total	200							
Demonstrative pronouns	CM	100	93.73	9373.00	4323.000	9373.000	−1.654	0.098	0.117
	EM	100	107.27	10,727.00					
	Total	200							
Interrogative pronouns	CM	100	96.04	9603.50	4553.500	9603.500	−1.405	0.160	0.099
	EM	100	104.97	10,496.50					
	Total	200							

[a] Grouping variable: group.

The results in Table 4 indicate that EMs used three classes of pronouns more frequently than CMs did, and there were significant differences in the total number of pronouns and personal pronouns. In English, anaphoric personal and demonstrative pronouns are important to maintain intersentential cohesion and discourse coherence. However, in written Chinese, if a reader knows or can infer the subject from the context, a zero pronoun is preferred and a formal pronoun is considered unnecessary. For example:

凯特·格伦维尔是一位杰出的澳大利亚女性作家，**她**因《神秘的河流》这本小说而闻名。 (EM No. 5)

(Kate Grenville is a prominent Australian female writer. *She* is famous for her fiction *The Secret River*.)

丰子恺生活在传统向现代过渡的特定历史时期, (**他**)受到传统与现代的双重影响。 (CM No.3)

(Because Feng Zikai lived in a particular transitional period in China from traditional ages to modern times, (*he*) was influenced by traditional culture as well as modem culture.)

In the EM's Chinese, 她 (she) is used to refer to Kate Grenville, but in the CM's Chinese, 他 (he) is deleted because it is unnecessary in written Chinese.

It is plausible to claim that EMs use pronouns more frequently than CMs do because of the influence of prolonged English learning.

(5)   Prepositions

NLPIR segmented and tagged two special classes of Chinese prepositions: 把 prepositions (used for constructing longer sentences that focus on the result or influence of an action by putting the object closer to the front of the sentence and preceding it with a 把), and 被 prepositions (used to form passive constructions). The results of the Mann–Whitney test are shown in Table 5.

Table 5. Ranks and Test Statistics [a] (prepositions).

	Group	N	Mean Rank	Sum of Ranks	Mann–Whitney's U	Wilcoxon's W	Z	Asymp. Sig. (2-Tailed)	r
Prepositions	CM	100	102.76	10,276.00	4774.000	9824.000	−0.552	0.581	0.039
	EM	100	98.24	9824.00					
	Total	200							
把 prepositions	CM	100	104.81	10,480.50	4569.500	9619.500	−1.693	0.090	0.120
	EM	100	96.20	9619.50					
	Total	200							
被 prepositions	CM	100	88.18	8818.00	3768.000	8818.000	−3.424	0.001	0.242
	EM	100	112.82	11,282.00					
	Total	200							

[a] Grouping Variable: group.

The results in Table 5 reveal that there is no significant difference in 把 prepositions between the two groups. The explanation for this result may lie in the fact that no equivalent 把 construction can be found in English. Therefore, English learning may not affect the use of the 把 construction in Chinese. However, a significant difference in 被 prepositions can be found between the two groups. 被 prepositions introduce an agent in a Chinese passive construction. The 被 construction in Chinese, however, has a different semantic prosody from that of the passive voice in English. The 被 construction is usually used to denote something bad or tragic that happened in the past, for example, 他被敌人杀害了 (He was killed by the enemy). The passive construction in English does not have such connotations. Since the May Fourth Movement, the semantic and temporal constraints on 被 constructions have been gradually resolved, and the 被 construction can be adopted to express something neutral or positive. For example:

除了引言和结论部分，此篇论文**被**分为三个章节。 (EM No. 3)

(Except for the part of introduction and conclusion, this thesis *is divided* into three chapters.)

菲茨杰拉德**被**誉为"迷惘的一代"的代表作家，"爵士乐时代"的桂冠诗人。 (EM No. 49)

(Scott Fitzgerald *is known* as the representative of the "Lost Generation", the poet laureate of the "Jazz Age".)

The difference between EMs and CMs lies in the frequency of the 被 construction. The statistics reveal that EMs used the 被 construction 112 times, among which it was used 54 times in a negative sense, 26 times in a neutral sense, and 32 times in a positive sense. CMs used the same construction 50 times, among which it was used 24 times in a negative sense, 22 times in a neutral sense, and 4 times in a positive sense. The results indicate that

EMs tend to transfer the English passive construction into Chinese more frequently than CMs do because of long and intensive contact with English.

(6)   Conjunctions

NLPIR only differentiated coordinating conjunctions with segments and tags. The other types of conjunction were all grouped under the tag "conjunction." The results of the Mann–Whitney test are shown in Table 6.

Table 6. Ranks and Test Statistics [a] (conjunctions).

	Group	N	Mean Rank	Sum of Ranks	Mann–Whitney's U	Wilcoxon's W	Z	Asymp. Sig. (2-Tailed)	r
Conjunction	CM	100	87.27	8727.00	3677.000	8727.000	−3.233	0.00	0.229
	EM	100	113.73	11,373.00					
	Total	200							
Coordinating conjunctions	CM	100	85.74	8574.00	3524.000	8574.000	−3.606	0.000	0.255
	EM	100	115.26	11,526.00					
	Total	200							

[a] Grouping variable: group.

The results in Table 6 show that EMs used both conjunctions and coordinating conjunctions significantly more often than CMs. Compared with English, Chinese is a parataxis language, in which phrases or clauses are usually arranged independently without coordinating conjunctions. However, coordinating conjunctions and other conjunctions are needed to maintain intersentential cohesion and discourse coherence in English. Because of prolonged and intensive EFL learning, EMs are more likely to transfer the syntactic features of English into Chinese by keeping English conjunctions. The following example is taken from an EM's abstract:

小说中的父亲或父亲般角色的缺失揭示了在维多利亚家庭中父权统治的丧失**并**证明了男性在公共领域的无能。 (EM No. 31)

(The absence of father or fatherly roles in the novels reveals the loss of the patriarchal control in the Victorian home ***and*** demonstrates men's incompetence in the public spheres.)

The coordinating conjunction, 并 (and), is used to link two verbs; however, it is not necessary in traditional Chinese.

5.2. Research Question 2: Lexical Complexity and Diversity

(1)   Lexical complexity

The results of the Mann–Whitney test of four levels of words are shown in Table 7.

Table 7. Ranks and Test Statistics [a] (lexical complexity).

	Group	N	Mean Rank	Sum of Ranks	Mann–Whitney's U	Wilcoxon's W	Z	Asymp. Sig. (2-Tailed)	r
Level 1	CM	100	150.50	15,050.00	0.000	5050.000	−12.217	0.000	0.864
	EM	100	50.50	5050.00					
	Total	200							
Level 2	CM	100	50.94	5094.00	44.000	5094.000	−12.109	0.000	0.856
	EM	100	150.06	15,006.00					
	Total	200							
Level 3	CM	100	50.50	5050.00	0.000	5050.000	−12.217	0.000	0.864
	EM	100	150.50	15,050.00					
	Total	200							
Level 4	CM	100	150.48	15,048.00	2.000	5052.000	−12.212	0.000	0.864
	EM	100	50.52	5052.00					
	Total	200							

[a] Grouping variable: group.

The results in Table 7 show that there are significant differences at all four levels between the two groups. It is interesting to note that CMs used Level 1 and Level 4 words much more frequently than EMs did, but much fewer Level 2 and Level 3 words. This result does not support Sun's research, which showed no differences among the three groups of different English proficiency levels [19]. In Sun's experiment, she used the same SGWCCP to divide Chinese words into four levels, but the participants were high school students and undergraduates. The length of their EFL learning was shorter and the intensity of EFL learning was less compared with the participants of this study, who are MA graduates. Different EFL proficiencies may lead to different research results. Logically, the longer a person studies an L2, the greater the effect it is likely to have on L1, which may lead to a decline in L1 lexical complexity. The EMs' lexical complexity in the present study is indeed significantly lower than that of CMs at Level 1 and Level 4. In our opinion, this result may demonstrate CMs' ability to shuttle back and forth, from the most frequent Chinese words to the most difficult ones, to achieve their writing purposes. However, it is a rather challenging task for EMs, who may not have the ability to convey their ideas simply and concisely, nor the ability to express themselves with the most complex words when necessary. However, we would prefer to take a more positive and optimistic attitude to this result. As Laufer points out, some L1 attrition is a small price to pay for achieving multi-competence [30]. In addition to the difference in the length and intensity of EFL learning, another possible reason is that the participants in Sun's experiment were asked to write a composition entitled, "My future plan", in Chinese in 30 min. However, the materials we used are Chinese abstracts, which have been revised numerous times through the communal effort of the authors and their supervisors and/or classmates. Perhaps factors regarding writing conditions and tasks, such as time restrictions and genre, combined with EFL proficiency, play a role in the different research results.

(2) Lexical diversity

The results of the independent sample *t*-test are shown in Tables 8 and 9.

Table 8. Group Statistics (lexical diversity).

	Group	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Lexical diversity	CM	100	500.28214672	60.790342097	6.079034210
	EM	100	503.21789452	58.210241848	5.821024185

Table 9. Independent Sample Test (lexical diversity).

		Levene's Test for Equality of Variances		**t-Test for Equality of Means**					**95% Confidence Interval of the Difference**	
		F	**Sig.**	**t**	**df**	**Sig. (2-Tailed)**	**Mean Difference**	**Std. Error Difference**	**Lower**	**Upper**
Lexical diversity	Equal variances assumed	0.575	0.449	−0.349	198	0.728	−2.935747801	8.416589540	−19.533409477	13.661913875
	Equal variances assumed			−0.349	197.629	0.728	−2.935747801	8.416589540	−19.533601208	13.662105606

The results in Tables 8 and 9 demonstrate there is no significant difference between CMs and EMs in lexical diversity (*p* = 0.728 d = −0.0494). The data in Table 7 may help explain the results. Although EMs used words at Level 1 and Level 4 less frequently than CMs, they used words at Level 2 and Level 3 more, which led to no significant differences in diversity between the two groups on average.

This result is in agreement with Sun's research, in which she used the same formula for calculating lexical diversity. This indicates that intensive and successful EFL learning does not decrease lexical diversity. This result conflicts with Laufer's research, however, which found that L1 lexical diversity declines in an L2-dominant environment when the

contact time with L2 increases [30]. This reveals that the degree of L2's influence on L1 may depend on different language environments. The subjects in this study were in an L1-dominant environment. In such a context, sustained contact with an FL may have some effect on L1, but it seems impossible for L1 to decline noticeably.

6. Conclusions

To sum up, in answer to Research Question 1, our research demonstrates that, in 27 investigated word classes and subclasses, EMs used nine categories significantly more than CMs: nouns, names of people, names of places, distinguishing adjectives, pronouns, personal pronouns, 被 prepositions, conjunctions, and coordinating conjunctions. Conversely, EMs used three items significantly less frequently than CMs: verbs, nominal verbs, and directional verbs. In answer to Research Question 2, we found a significant difference in lexical complexity at four levels, but there was no significant difference in lexical diversity between EMs and CMs.

Before we draw conclusions, we acknowledge that there may be differences between EMs' and CMs' processes of producing Chinese abstracts. CMs wrote their MA theses in Chinese, so it would be natural that they wrote their Chinese abstracts first, and then translated them into English abstracts. However, EMs are required to write their MA theses in English, so it is very possible that most wrote English abstracts first, and then translated their English abstracts into Chinese, instead of writing Chinese abstracts directly. Meanwhile, we cannot exclude the possibility that some EMs wrote Chinese abstracts first and then translated them into English. It is likely that Chinese abstracts translated from English will retain more English lexical and syntactic features than those written directly in Chinese. However, whether the EMs translated their English abstracts into Chinese or wrote them directly in Chinese, we contend that frequent and years-long reading and writing in English appears to have had an impact on their L1 Chinese, just as their English is certainly influenced by Chinese. The bidirectional transfer resulting from the interconnectedness of the two language channels shapes EMs' unique language features and multi-competence.

The following tentative conclusions may be drawn from the present research: First, even when an L1 is mature, it is still a dynamic system that may be affected by subsequent languages that are being used frequently or intensively. Second, the differences from CMs that EMs present in L1 written Chinese evince the multi-competence of the EMs' language, which is not the sum of two or more well-developed systems, but is, rather, a whole system in its own right. Both their L1 and their L2 are unique because of the bidirectional transfer across languages. This should not be considered a detriment to those L2 or FL learners, as it is not feasible to expect them to have two equally developed, native-like language systems. Third, the FL environment and L2 environment have different influences on L1. In an L2-dominant environment, L1 will be noticeably affected, whereas even intensive contact with an FL within an L1-dominant environment will have a less perceivable impact.

Data Availability Statement: The data were collected from the available written documents accessible by the general public and they are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Weinreich, U. *Languages in Contact*; Mouton: The Hague, The Netherlands, 1953; ISBN 90-279-2689-1.
2. Gass, S.M.; Selinker, L. *Language Transfer in Language Learning*; Newbury House: Rowley, MA, USA, 1983; ISBN 0-88377-305-8.
3. Odlin, T. *Language Transfer: Cross-Linguistic Influence in Language Learning*; Cambridge University Press: Cambridge, UK, 1989; ISBN 9781139524537.
4. Gass, S.M. Fundamentals of second language acquisition. In *Handbook of Undergraduate Second Language Education*; Rosenthal, J.W., Ed.; Lawrence Erlbaum Associate, Inc., Publishers: Mahwah, NJ, USA, 2000; pp. 29–46.
5. Yuan, B.; Lin, Y. Directionality and complexity of L1 transfer in L2 acquisition: Evidence from L2 Chinese discourse. *Int. Rev. Appl. Linguist. Lang. Teach.* **2019**, *57*, 377–416. [CrossRef]
6. Al-Thubaiti, K. Pre-emption of L1 properties in the L2 acquisition of English wh-interrogatives: Effects of L2 proficiency and age of onset. *Int. Rev. Appl. Linguist. Lang. Teach.* **2020**, *58*, 443–473. [CrossRef]
7. Schmid, M.S.; Kopke, B. *Oxford Handbook of Language Attrition*; Oxford University Press: Oxford, UK, 2019; ISBN 978-0198793595.
8. Cook, V. The poverty-of-the-stimulus argument and multi-competence. *Second Lang. Res.* **1991**, *7*, 103–117. [CrossRef]
9. Meara, P. (Ed.) Introduction. In *Vocabulary in a Second Language*; CILTR: London, UK, 1983; pp. ii–iv.
10. Grosjean, F. Neurolinguistics, beware! The bilingual is not two monolinguals in one person. *Brain Lang.* **1989**, *36*, 3–15. [CrossRef]
11. Grosjean, F. Individual Bilingualism. In *The Encyclopedia of Language and Linguistics*; Asher, R., Ed.; Pergamon: Oxford, UK, 1994; pp. 1656–1660.
12. Cook, V. *Portraits of the L2 User*; Multilingual Matters Limited: Clevedon, UK, 2002; ISBN 1-85359-583-7.
13. Cook, V. Introduction: The Changing L1 in the L2 User's Mind. In *Effects of the Second Language on the First*; Cook, V., Ed.; Multilingual Matters Ltd.: Clevedon, UK, 2003; pp. 1–18.
14. Cook, V. The goals of ELT: Reproducing native-speakers or promoting multi-competence among second language users? In *Handbook on English Language Teaching*; Cummins, J., Davison, C., Eds.; Kluwer: Dordrecht, The Netherlands, 2007; pp. 237–248.
15. Cook, V. Multi-competence. In *The Encyclopedia of Applied Linguistics*; Chapelle, C.A., Ed.; Wiley-Blackwell: New York, NY, USA, 2012; pp. 3768–3774.
16. Cook, V. Premises of multi-competence. In *The Cambridge Handbook of Linguistic Multi-Competence*; Cook, V., Wei, L., Eds.; Cambridge University Press: Cambridge, UK, 2016; pp. 1–25.
17. Van Hell, J.G.; Dijkstra, T. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychon. Bull. Rev.* **2002**, *9*, 780–789. [CrossRef] [PubMed]
18. Tokumaru, Y. Cross-linguistic influences of L2 English on L1 Japanese in Japanese-English bilinguals. In *Bilingual Socialization and Bilingual Language Acquisition: Proceedings from the Second International Symposium on Bilingualism*; University of Vigo: Galicia, Spain, 2002; pp. 399–406. ISBN 84-8158-280-8.
19. Sun, D. An Empirical Study on Reverse Transfer of Chinese EFL Students. Unpublished Ph.D. Dissertation, Northeast Normal University, Nanjing, China, 2014.
20. Kecskes, I.; Papp, T. *Foreign Language and Mother Tongue*; Lawrence Erlbaum: Mahwah, NJ, USA, 2000; ISBN 0-8058-2759-5.
21. Peiyun, L.; Chuanbin, N. Effects of L2 on the L1 at semantic Level: An empirical study. *J. Lang. Teach. Res.* **2016**, *7*, 425–431. [CrossRef]
22. Schmid, M.S. *Language Attrition*; Cambridge University Press: New York, NY, USA, 2011; ISBN 9780521759939.
23. Testing Center of National Chinese Proficiency Test Committee. *Syllabus of Graded Words and Characters for Chinese Proficiency*; Economy and Science Press: Beijing, China, 2001; ISBN 9787505825307.
24. Shuneng, L. *Contrastive Studies of English and Chinese*; Higher Education Press: Beijing, China, 2010; ISBN 9787040292657.
25. Yang, H. *A Study of Europeanized Grammar in Modern Chinese*; Commercial Press: Beijing, China, 2008; ISBN 9787100060660.
26. Potter, S. *Changing English*; Andre Deutsch: London, UK, 1969; ISBN -10 0233960597.
27. LaPalombara, L.E. *Introduction to Grammar: Traditional, Structural, Transformational*; Winthrop Publishers, Inc.: Cambridge, MA, USA, 1976; ISBN 10: 0876260199.
28. Tallerman, M. *Understanding Syntax*; Arnold: London, UK, 1998; ISBN 9780340603772.
29. Department of Chinese, Beijing Normal University. *The Evolution and Change in Written Chinese since the May Fourth Movement*; Commercial Press: Beijing, China, 1959; ISBN 9017160.
30. Laufer, B. The Influence of L2 on L1 Collocational Knowledge and on L1 Lexical Diversity in Free Written Expression. In *Effects of the Second Language on the First*; Cook, V., Ed.; Multilingual Matters Ltd.: Clevedon, UK, 2003; pp. 19–31.