

A Study on the Effect of L1 to L2 Transfer on the Production of Idiomatic Expressions in L2 Among Mandarin-speaking Intermediate Learners of English

Xinran Chen¹, Yiran Du^{2*}, Menghan Qu³, Shimei Gao⁴

¹The University of Chicago, Chicago, Illinois 60637, United States

²School of the Arts, University of Liverpool, Liverpool L69 3BX, United Kingdom

³Hongwen School Qingdao Laoshan Campus, Qingdao, Shandong 266000, China

⁴Cooperative School of International Education, Tianjin University of Commerce, Tianjin 300134, China

* Corresponding author. Email: hsydu14@liverpool.ac.uk

ABSTRACT

Acquiring idiomatic expressions for second-language learners is a big challenge. This is particularly true for native Mandarin Chinese speakers to learn English idiomatic expressions. The problem could be explained by culture differences and “language transfer”. Previous research did not pay much attention to the influence of L1 on an intermediate second-language learner’s idiomatic repertoire. Our study focuses on the extent to which knowledge of Mandarin Chinese idioms affects the accuracy with which these speakers produce the correct translation of such idioms into English. Inspired by Li’s study, we designed the English idiom translation-focused surveys, recruited 7 and 34 subjects in the pilot study and the actual experiment respectively, and analyzed the data by using the chi-square test. Due to statistical insignificance, we failed to prove our hypotheses, which may be attributed largely to deficiency in the number of the subjects. Some improvement could be done in the future including increasing the number of task takers, assessing the actual English proficiency of the participants, and improving the question set up and the sampling method. A possible topic of future study is the extent of the participants demonstrating traces of idiomatic transfer may be related to two factors: (1) the discrepancies between their environment of acquiring L2 languages; (2) the difference between the education systems.

Keywords: Idioms, idiomatic expressions, linguistic production, crosslinguistic transfer, Mandarin Chinese, English

1. INTRODUCTION

Idiomatic expressions are expressions in which mere semantic aggregation of their components does not lead to the intended interpretation [1, 2]. While they are an indispensable component of daily communication, idiomatic expressions are known to be difficult for second-language learners, whether in comprehension, retention, or production [3]. In recognition of this general idea, both the Common European Framework of Reference for Languages (CEFR) and China’s Standards of English Language Ability (CSE) refer to the command of idiomatic expressions in their proficiency-level descriptors [4]. Particularly, language learners who are in the CEFR B1, B2, and C1 levels are partly differentiated by their ability to comprehend and produce idiomatic expressions effectively [5], including by the

number of idioms that they can recognize [6].

Further, for second language (L2) learners who speak Mandarin Chinese as their native language (L1), idiomatic expressions in English may pose a special problem, as the general cultural difference between native Mandarin Chinese speakers and native English speakers can give rise to a disparity in lexical usage, including connotations [7]. Thus, apart from imported expressions that are literally translated (e.g., “Don’t put all your eggs in one basket” and “the last straw”), idiomatic expressions in the English language can remain largely elusive unless with prior knowledge (e.g., “kick the bucket”).

One factor that seems to affect an L2 learner’s grasp of English-language idioms is “language transfer” [8, 9] from the L1 to the L2. Previous research suggests that

when learners deal with idiomatic expressions cross-linguistically, L1-to-L2 transfer is attested and ranges from positive to negative depending on both the proficiency of the learners and the similarity between the corresponding idiomatic expressions in the two languages [10-14]. Based on the findings from these studies, we suspect that the L1 exerts significant influence on a second-language learner’s idiomatic repertoire, especially at the intermediate stage where the learner has not yet acquired enough proficiency to minimize “obvious searching for expressions” [5].

We are therefore interested in how native speakers of Mandarin with intermediate proficiency in English may exhibit influence from their native language (Mandarin) in their spontaneous production of English as a second language, with a focus on idiomatic expressions. Specifically, in this study we focus on the extent to which knowledge of Mandarin Chinese idioms affects the accuracy with which these speakers produce the correct translation of such idioms into English.

2. METHODOLOGY

2.1. Theoretical foundations

In terms of experimental approaches, we mainly took inspiration from Li’s study [13]. It includes, crucially, a translation task in the experiment as well as a pilot study before the actual experiment. In the translation task, participants were asked to translate from Chinese idioms to English or vice versa. Meanwhile, the pilot study, which tested the extent to which the target population was familiar with the idioms to be used in the actual experiment, helped the researcher eliminate a familiar

idiom pair and ascertain the propriety of the rest of the experimental corpus.

What was also important in Li’s study was a key insight concerning the nature of the translation task: that translation from Chinese to English is a “productive task from L1 to L2” while translation from English to Chinese is a “receptive task from L2 to L1” (p. 168) [13]. Since our research focus is mainly on L2 production, we decided to use Chinese text as stimuli and include a similar kind of translation task in our main experiment.

2.2. Participants

We recruited a total of 41 participants for all parts of our study. All the participants in our study were native Mandarin Chinese speakers whose English proficiency was roughly at the intermediate level (CEFR B1-C1). Seven participants were recruited in the pilot study; 34 were recruited in the experiment.

The 34 participants in the experiment belong to two populations differentiated by their prior experience with “internationalized education”, which we define specifically to signify education conducted in an institution in which English is one of the main working languages or is the default working language. Participants in the first population (n = 17) come from a traditional Chinese education background with less than 1 year of internationalized education, while participants in the second population (n = 17) have had at least one year of internationalized education. For reference, we have also similarly categorized the 7 participants in the pilot study. Summary statistics pertaining to all the participants in our study can be found in Table 1 below.

Table 1. Summary statistics of the participants in the pilot study and the experiment

		n	Gender			Age (mean)*	Age (std. dev.)
			F	M	Other		
Pilot study	Internationalized	1	1	0	0	20	0
	Traditional	6	2	2	2	21.33	2.981
Experiment	Internationalized	17	11	6	0	20.47	2.329
	Traditional	17	7	10	0	19.88	0.471

* Mean ages not significantly different, according to an Analysis of Variance (ANOVA) test.

2.3. Experimental design

We basically adopted the idea of the Chinese-stimulus translation task in Li’s study, albeit with some changes [13]. First, in terms of question type, in order to ensure standardization during grading, instead of free-response (open) questions, we opted for multiple-choice questions. In addition, in order to make it difficult for the participants to realize the true purpose of our study, we

decided that there should be a large number of “dummy questions” (i.e., fillers) randomly dispersed between our target questions (i.e., our questions of interest). These “dummy questions” should resemble those typically found in English proficiency tests/quizzes in China and, with their sheer frequency of occurrence in the exam, contribute to giving our participants the general impression (though not entirely false) that the exam somehow tests their English proficiency.

2.3.1. The target questions

Each target question had a textual stimulus in the form of a Chinese idiom as well as 3 choices (also referred to as “items”) which served as candidates for an English translation of the idiom. For ease of reference, we defined the following terms:

Thought-image: A word-level fragment that serves a similar function as an “imagery,” using the Merriam-Webster definition of “imagery” as in “mental images [...] especially the products of imagination” [15].

Isomorphic item: an English expression that is identical or almost identical to the Chinese idiom in terms of thought-images; alternatively, a literal (or quasi-literal) translation of the Chinese idiom.

Intruder item: an English expression that is similar to (but not identical with) the Chinese idiom in terms of thought-images.

Heteromorphic item: an English expression that is significantly different from the Chinese idiom in terms of thought-images.

Convergent Chinese–English idiom pair: a pair that consists of a Chinese idiom and an isomorphic item, wherein the two expressions have identical or nearly identical meanings.

Divergent Chinese–English idiom pair: a pair that consists of a Chinese idiom and a heteromorphic item, wherein the two expressions have identical or nearly identical meanings.

Based on the above definitions, we further defined the following two types of target questions:

Convergent question: a target question in which the Chinese idiom in a convergent Chinese-English idiom pair serves as the translation stimulus while the 3 choices are the English idiom in the pair, an intruder item, and a heteromorphic item, respectively. The latter two items should either (1) not have idiomatic meanings in English or (2) have meanings that deviate from that of the Chinese idiom in the pair.

Divergent question: a target question in which the Chinese idiom in a divergent Chinese-English idiom pair serves as the translation stimulus while the 3 choices are the English idiom in the pair, an intruder item, and an isomorphic item. The latter two items should either (1) not have idiomatic meanings in English or (2) have meanings that deviate from that of the Chinese idiom in the pair.

2.3.2. The dummy questions

As stated before, the dummy questions should not only resemble questions typically found in English proficiency exams in China, but also look sufficiently

similar to the target questions.

To achieve the first goal, we referred to our experience as students to design questions that test on some frequently tested grammatical and lexical (vocabulary) knowledge. The grammatical knowledge featured in our dummy questions includes dative alternation [16], adjective order, past subjunctive, and special verb argument structure transformations including null instantiation and the resultative [16]. The lexical knowledge featured includes spelling, synonym disambiguation, and idiomatic expressions (though with no overlap with the target questions).

To achieve the second goal, i.e., resemblance to the target questions in the formal aspect, we used in the dummy questions a similar Chinese-to-English translation prompt and similarly offered 3 choices per Chinese stimulus.

2.4. Time limit

Then it comes the question of time limit. We needed to make sure that participants would not be overwhelmed by fatigue or loss of focus anywhere during the experiment, while leaving enough time for them to react to the stimuli and answer the questions (and not, for example, guess randomly due to lack of time). It seems that approximately the first 15 minutes into a series of tasks is a period in which attention may be kept (personal experience; personal communication), even though it is disputed how long exactly this time limit should be [17]. Moreover, data from Hurtado’s experiment illustrates that when an action was done by a user to an item (e.g., a user clicks on a product on an ecommerce site), the users spent a total of 195,745 seconds interacting with 6,041 items – i.e., the average attention span spent on an item per user was 32.5 seconds [18]. Further, results from Huang et al.’s experiment on primary school students, middle school students and undergraduates demonstrate that attention span was positively correlated with age [19].

Therefore, we decided that 12 minutes was a good upper limit such that the participants in our study – the majority of whose age would correspond to the “undergraduates” group in the Huang et al.’s study – would be able to maintain their attention during the entire exam [19].

2.5. Idiom selection

The Chinese idioms used in this study were drawn from various sources, including, importantly, a Chinese-English bilingual idiom dictionary and the corpus of idiom pairs used in Li’s study [13, 20]. Although in general we picked different idioms for the pilot study and the experiment respectively, we controlled their usage frequency within a narrow range (900-1100), as

indicated by the BLCU corpus used by us [21], which with its “up to ten billion characters” of textual data mainly in Chinese (p. 93), is what we believe to be a suitable reference for the actual usage frequencies of the

Chinese idioms. The list of idioms that we used for both the pilot study and the actual experiment is given below in Table 2:

Table 2. List of idioms used in the study

Idioms*	Literal translation	Idiomatic translation**	Frequency***	Notet
baishouqijia	build a house with bare hands	'start from scratch' (1)	1077	P
kailvdeng	turn on the green light	'give the green light' (2)	1022	P/E
zhishangtanbing	talk about warfare on paper	'fight only on paper' (1) 'theory without practice' (4)	947	E
huolinghuoxian	vivid as if alive	'come to life' (1)	902	E
duanzhangquyi	take a meaning out of its context	'garble a statement' (1)	1006	E
anranshise	lose color and become dark/dim	'fall into the shade' (1)	992	P/E
yimingjingren	once (the bird) cries, it surprises everyone	'make a great coup' (1)	995	E
renyunyiyun	say what others say	'follow the herd' (4)	1050	E
guzhuyizhi	bid the last bit that you have	'send the helve after the hatchet' (3)	1056	P/E

*Chinese idioms here are spelt in Pinyin.

**Other translations also possible.

Idiomatic translation sources: (1) = [20]; (2) = [13]; (3) = [22]; (4) = the authors of this paper.

*** This is the sum of the frequencies of all variants of each Chinese idiom that we could think of. E.g., for “kailvdeng” (‘give the green light’) there are at least four variants possible: “kailvdeng”, “kai...lvdeng”, “lianglvdeng”, and “liang...lvdeng”. It needs to be noted that for the specific case of “kailvdeng” (‘give the green light’) (and its variants), we excluded occurrences in which the literal meaning (i.e., the traffic light) was evidently the intended interpretation.

† P = included in the pilot study; E = included in the experiment.

2.6. Pilot study

In order to know how well our initial experimental designs would work, we conducted a pilot study. The aforementioned 7 participants were recruited to finish the pilot exam, which we deliberately named “English-Level Quiz”, on the online survey platform Wenjuanxing. They were requested to fill in some background information, finish the exam section in 720 seconds and take part in a post-exam interview.

The pilot exam section that we designed took center stage. This section included 28 questions. 4 questions were target questions, including 2 divergent questions

and 2 convergent questions. The other 24 questions were dummy questions, including 8 grammar questions, 15 vocabulary questions and 1 question on the interface between grammar and vocabulary (see Appendix A for details). The order of all questions and the order of all the choices were randomized for each participant.

In the post-exam interview, we asked each participant 5 pre-written open questions via WeChat to get a sense of how the participants thought of our prototype exam. Below are the questions that we asked, originally in Mandarin Chinese, but here rendered into their English translations for the reader’s convenience:

How do you feel about the volume of questions?

- Do you think the time was tight?
- How do you feel about the difficulty of the questions?
- Any other suggestions for our exam?
- What are the 3 questions that impressed you the most?

The responses from the participants indicated that the timing of the exam was clearly in excess, while the exam questions were generally of moderate difficulty, and the dummy questions mostly distracted their attention away from the target questions.

2.7. The experiment

2.7.1. General setup

Based on the participants' performance and feedback in the pilot study, we revised the setup of the experiment. An overview of the differences between the pilot study and the experiment is summarized in Table 3 below:

Table 3. Comparison between the pilot study and the experiment

		Pilot study	Experiment
Time limit		720 seconds for all questions	10 seconds per question*
Are participants allowed to return to previous questions?		Yes	No*
Total number of questions		28	32
Number of target questions		4	8
in which	Divergent questions	2	4
	Convergent questions	2	4
Number of dummy questions		24	24**
Randomization of order of all questions and choices		Yes	Yes
Interview/Survey format		Interview; all open Q's	Survey; some MCQ's†

* In the experiment, the participants had to complete each question in 10 seconds. They could not skip a question unless they end up making no choice within the 10-second window, by the end of which the system automatically skips to the next question.

** The individual dummy questions in the experiment were identical to those in the pilot study.

† MCQ's = Multiple-choice questions.

2.7.2. Hypotheses

If the accuracy of Chinese-to-English idiom translations by native speakers of Mandarin Chinese (with English at around intermediate level) is indeed affected by their knowledge of Mandarin Chinese idioms, then the mean frequency at which the speakers choose the isomorphic items in the target questions will be significantly larger than the null value (e.g., one third) that assumes random guessing.

A difference in educational experience (specifically, less than one year vs. at least one year of internationalized education) will also be manifested in a difference in performance for the two populations studied.

2.7.3. Procedure

We used the Wenjuanxing platform again, changing the name of the exam slightly to "English Ability Quiz". All the participants were asked to fill in the background information, answer the questions in the exam section (see Appendix B for details), and complete a short survey attached to the end of the exam section, instead of a separate interview as used in the pilot study.

Below are the survey questions in the experiment, which were originally asked in Mandarin Chinese, but have been rendered into their English translations for the reader's convenience. The first three questions were multiple-choice.

What do you think of the number of questions? (Too few / okay / too many)

How do you feel about the time setting of this exam?
(Too short / okay / too long)

What do you think of the difficulty of the questions?
(Too easy / okay / too hard)

Which three questions were the most memorable?

Did you use exam tricks? Please give examples.

What suggestions do you have for our question design?

3. RESULTS AND DISCUSSION

3.1. Survey results

The responses from the survey suggest that our final experimental design was generally acceptable for both the Internationalized and Traditional participants, and that the exam was appropriate in terms of time limit, number of questions, and difficulty. Importantly, the responses to the recall question (No. 4 in the list above) show that the target questions could not be easily recalled.

This indicates that the dummy questions were still relatively successful, despite the change in question proportion resulting from the increase in the number of target questions, compared to that in the pilot study. Responses to the question on exam trick usage (No. 5 in the list above) will be discussed in §4 “Conclusion” below.

3.2. Data analysis

Below are some summary statistics (Table 4 and Table 5) for the participants’ performance in the experiment. In contrast to our processing of the survey results (all survey questions were answered by every participant and counted), for the exam section we excluded invalid responses, i.e., those from participants (Traditional $n = 2$; Internationalized $n = 2$) who did not finish all the eight target questions. We assume that this incompleteness might have been owing to the time constraint, the participants’ limited vocabulary range, or their indecisions between idiomatic expressions with comparable thought-images.

Table 4. Summary statistics for the participants’ performance in all target questions*

	Mean % isomorphic	Mean % intruder	Mean % heteromorphic
Traditional (n = 15)	25.00%	25.00%	50.00%
Internationalized (n = 15)	20.83%	36.67%	42.50%
Pooled population (n = 30)	22.92%	30.83%	46.25%

*Invalid responses not counted.

Table 5. Summary statistics for the participants’ performance, sorted by question type*

	Mean % isomorphic	Mean % intruder	Mean % heteromorphic
Traditional in C	31.67%	28.33%	40.00%
Internationalized in C	28.33%	46.67%	25.00%
Traditional in D	18.33%	21.67%	60.00%
Internationalized in D	13.33%	26.67%	60.00%

*Invalid responses not counted.

The data suggest that the participants generally had the inclination to choose the heteromorphic items over either the isomorphic or the intruder items. Meanwhile, both two groups demonstrated a stronger inclination to choose the isomorphic or intruder item when faced with convergent questions.

As for our first and initial hypothesis, we applied the ordinary two-tailed one-proportion test (based on a normal sampling distribution) at the significance level of 0.05. Using $p_0 = 22.92\%$ (from the pooled population), we arrived at a p-value of $0.2266 > 0.05$. We concluded

that there is no evidence substantial enough to support the claim that the mean frequency at which the speakers choose the isomorphic items in the target questions is significantly different from random guesswork.

The second hypothesis, regarding differences of educational experiences causing differences in performance among the two populations, was also evaluated by using the chi-square test of homogeneity. The “Chi-Square Calculator” [23] returned a p-value of $0.7869 > 0.05$ ($df = 2, \chi^2 = 0.4794$), which suggests that the differences in performance were statistically insignificant.

4. CONCLUSION

Due to the statistical insignificance, we cannot prove our hypotheses. However, some improvement could be done in our study and it could give suggestions on subsequent studies. First, the aforementioned statistical insignificance may be attributed largely to a deficiency in the number of test takers. The limited sample led to a high standard error, i.e., high uncertainty, as shown by a high p-value. In addition, our calculation process regarding the assessment of the two hypotheses was somewhat limited, owing to our relative deficiency in applying more sophisticated statistical methods in the analysis.

Further, our experiment could be improved by means of assessing the actual English proficiency of the participants to minimize discrepancies in English proficiency among each group, thereby potentially arriving at more consistent and significant tendencies. This could be achieved by involving a formal pre-test (instead of a simple self-diagnosis) to ensure that the participants from each group have English language skills in accordance with our criterion of “intermediate” [14]. Sadly, at the current stage, we do not yet have adequate resources to construct a test that can accurately evaluate the participants’ English proficiency.

With respect to our question setup, we have to confess that the complexity of the target questions turns out to be not quite consistent, since our selection of intruders could contain thought-images ranging from somewhat similar to largely different compared to the correct answer (heteromorphic for divergent questions; isomorphic for convergent questions). This can be improved in terms of dividing questions into different difficulties based on the “trickiness” (distracting effect) of the intruder items. A comparison of the roles that the intruders play in each kind of questions may also be a focus of future analysis.

Also, we have realized that we used convenience sampling [24] in our choice of participants, for the sake of saving cost and time. However, this may easily have undermined the representativeness of our sample. For instance, all the 41 participants in our study were friends or acquaintances of one of the researchers, and the Traditional group of participants mostly came from the same region in China.

Finally, we present a specific suggestion concerning a possible topic of future studies. We have identified a potential difference between the performance of the Traditional and Internationalized participants in the convergent questions; in addition, we have noticed that the survey question on “exam tricks used” (Question No. 5) returned several responses related to language intuitions or similar concepts particularly among Internationalized participants, whereas Chinese-style exam tricks, such as choosing the longest or the shortest

item, were more often cited by the Traditional participants. Therefore, we suspect that the extent of the participants demonstrating traces of idiomatic transfer may be related to two factors: (1) the discrepancies between their environment of acquiring L2 languages (in this study, English); (2) the difference between the education systems, since specific tactics (exam tricks) have often been introduced or even recommended to learners in China for tackling with multiple-choice questions. We believe that a fuller explanation of how these two factors could exert influence requires further research using more rigorous and systematic approaches.

ACKNOWLEDGMENTS

We are grateful for the guidance, support, and suggestions from Prof. Andrew Nevins, Ms. Yujing Su, Mr. Yunchong Huang, our classmates, and all who responded to our exams.

REFERENCES

- [1] Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of memory and language*, 27(6), 668-683.
- [2] Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18(5), 523-534. DOI: [https://doi.org/10.1016/S0022-5371\(79\)90284-6](https://doi.org/10.1016/S0022-5371(79)90284-6)
- [3] Karlsson, M. (2019). *Idiomatic Mastery in a First and Second Language*. Bristol: Multilingual Matters.
- [4] Zhao, W., Wang, B., Coniam, D., & Xie, B. (2017). Calibrating the CEFR against the China Standards of English for College English vocabulary education in China. *Language Testing in Asia*, 7(1). DOI: <https://doi.org/10.1186/s40468-017-0036-1>
- [5] Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Strasbourg: Council of Europe Publishing. <https://www.coe.int/lang-cefr>
- [6] English Profile. (n.d.). *English Vocabulary Profile Online - British English*. Retrieved August 11, 2021, from <https://www.englishprofile.org/wordlists/evp>
- [7] Bi, Y. (2021). The influence of differences between Chinese and English thinking patterns on Chinese-English translation. *International Journal of Social Science and Education Research*, 4(5), 284-289. DOI: 10.6918/IJOSSER.202105_4(5).0038
- [8] Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- [9] Ye, Z. (2019). Chinese to English Lexical Transfer Errors in the Writing of Rural Senior High School Students. *Theory and Practice in Language Studies*,

- 9(6), 645-654. DOI: 10.17507/tpls.0906.06
- [10] Cáceres-Guerrero, L. D. (2017). Transfert dans l'acquisition des expressions idiomatiques en français langue étrangère [Transfer in the acquisition of idiomatic expressions in French as a Foreign Language]. *Rastros Rostros*, 19(35), 1-15. DOI: <https://doi.org/10.16925/2382-4921.2017.35.06>
- [11] Forssten, R. A. (2019). A comparison of factors affecting Estonian EFL learners' idiom comprehension. *Eesti Rakenduslingvistika Ühingu Aastaraamat* [Estonian Papers in Applied Linguistics], 15, 21-34.
- [12] Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, 20(2), 287-304. DOI: 10.2307/3586545
- [13] Li, C. (2016). Study of L1 Transfer in Idiom Learning. *Jiazhi Gongcheng* [Value Engineering], 35(28), 167-170.
- [14] Taki, S., & Namy Soghady, M. R. (2013). The Role of L1 in L2 Idiom Comprehension. *Journal of Language Teaching and Research*, 4(4), 824-833.
- [15] Merriam-Webster. (n.d.). Imagery. In *Merriam-Webster.com dictionary*. Retrieved August 11, 2021, from <https://www.merriam-webster.com/dictionary/imagery>
- [16] Hilpert, M. (2014). *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press Ltd.
- [17] Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education*, 40, 509-513. DOI: 10.1152/advan.00109.2016
- [18] Hurtado, J. F. (2016, July 30). Attention span for personalisation. Retrieved August 8, 2021, from the arXiv database: <https://arxiv.org/abs/1608.00147>
- [19] Huang, C., Lorusso, M. L., Luo, Z., & Zhao, J. (2019). Developmental differences in the relationship between visual attention span and Chinese reading fluency. *Frontiers in Psychology*, 10. DOI: 10.3389/fpsyg.2019.02450
- [20] Zhang, C. (Ed.). (2003). *Shiyong Han-Ying Xiyu Cidian* [A Practical Chinese-English Dictionary of Idioms]. Shanghai: Dongfang Publishing Center.
- [21] Xun, E., Rao, G., Xiao, X., & Zang, J. (2016). The construction of the BCC Corpus in the age of Big Data. *Corpus Linguistics*, 3(1), 93-118. Retrieved August 1, 2021, from [http://bcc.blcu.edu.cn/downloads/papers/The development of BCC corpus in the context of big data_Xun Endong.pdf](http://bcc.blcu.edu.cn/downloads/papers/The%20development%20of%20BCC%20corpus%20in%20the%20context%20of%20big%20data_Xun%20Endong.pdf)
- [22] Farlex. (2015). Send the helve after the hatchet. In *Farlex Dictionary of Idioms*. Retrieved August 15, 2021, from <https://idioms.thefreedictionary.com/send+the+helve+after+the+hatchet>
- [23] Pierce, R. (2019, Dec 22). *Chi-Square Calculator*. Math Is Fun. Retrieved 8 Aug 2021 from <http://www.mathsisfun.com/data/chi-square-calculator.html>
- [24] Crossman, A. (2020, August 27). *Convenience Samples for Research*. Retrieved from <https://www.thoughtco.com/convenience-sampling-3026726>
- [25] IELTS Partners. (n.d.). *IELTS in CEFR scale*. Retrieved August 1, 2021, from www.ielts.org: <https://www.ielts.org/about-ielts/ielts-in-cefr-scale>
- [26] Xue, T. (Ed.). (2016, November 13). *Jiaoyubu Guanyuan: Guojia Yingyu Nengli Dengji Kaoshi Jihua Zhubu Tuichu* [Ministry of Education Official: Gradual Implementation of the National English Test System]. Retrieved August 1, 2021, from Xinhua http://www.xinhuanet.com/politics/2016-11/13/c_1119900089.htm
- [27] off-color. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/off-color_26168
- [28] go out like a light. (2015). *Farlex Dictionary of Idioms*. Retrieved August 15, 2021, from <https://idioms.thefreedictionary.com/go+out+like+a+light>
- [29] take the guilt off the gingerbread. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/take-the-guilt-off-the-gingerbread_26659
- [30] hit a home run. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/hit-a-home-run_16712
- [31] throw for a loop. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/throw-for-a-loop_7495
- [32] put all eggs in one basket. (2015). *Farlex Dictionary of Idioms*. Retrieved August 15, 2021, from <https://idioms.thefreedictionary.com/put+all+eggs+in+one+basket>
- [33] bet bottom dollar. (2015). *Farlex Dictionary of Idioms*. Retrieved August 15, 2021, from <https://idioms.thefreedictionary.com/bet+bottom+dollar>
- [34] cut some slack. (2015). *Farlex Dictionary of Idioms*. Retrieved August 15, 2021, from <https://idioms.thefreedictionary.com/cut+some+slack>
- [35] pave the way. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/pave-the-way_6030
- [36] speak out of turn. (2015). *Farlex Dictionary of*

- Idioms*. Retrieved August 15, 2021, from <https://idioms.thefreedictionary.com/speak+out+of+turn>
- [37] get on someone's nerves. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/get-on-someones-nerves_12506.
- [38] perk up. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/perk-up_18373.
- [39] all bark and no bite. (n.d.). *Phrases.com*. Retrieved August 15, 2021, from https://www.phrases.com/phrase/all-bark-and-no-bite_1919