International Journal of Linguistics, Literature and Translation
ISSN: 2617-0299 (Online); ISSN: 2708-0099 (Print)
DOI: 10.32996/ijllt
Journal Homepage: www.al-kindipublisher.com/index.php/ijllt

IJLLT
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Measuring the Success of English Borrowings in Chinese

Yunhan Jia
School of Arts, Languages and Cultures, The University of Manchester, Manchester, The United Kingdom
Corresponding Author: Yunhan Jia, E-mail: yunhan.jia@postgrad.manchester.ac.uk

| ABSTRACT
After entering the recipient language, many borrowings are in competition with their native equivalents, serving as an alternative lexical choice to express the designated concept. Concerning this phenomenon, questions arise as to whether some borrowings are used more frequently than others and what factors contribute to their success if we consider borrowings with a higher frequency more successful. To address these issues, this study takes an onomasiological approach, considers both borrowings and their native equivalents, and measures the relative frequency of borrowings in comparison with their native equivalents to assess their success. Raw frequencies are avoided in the study because they are subject to the popularity of topics in which a concept might be mentioned, meaning that some borrowings are frequently used possibly just because the relevant topics are popular, which thus cannot effectively reflect how the linguistic features of the borrowings affect their usage and success. Based on a dataset with data collected from Weibo, this study conducts a multiple regression analysis in Rbrul to investigate how the factors affect the success of English borrowings in Chinese, including concept frequency, concept novelty, lexical field, age of borrowing, number of competitors, and relative length of borrowings. The results of the statistical model indicate that borrowings are significantly more successful than others if they denote a new or low-frequency concept and if they fall into the rapidly-evolving fields rich in new inventions. In contrast, age of borrowing and relative length cannot significantly determine the success of borrowings. By measuring the success of English borrowings in Chinese, this study not only fills the gap caused by the rarity of studies taking an onomasiological approach on the contact-induced outcome in Chinese but also sheds light on the major motivations for using borrowings in Chinese.

| KEYWORDS

success, English borrowings, multiple regression model, onomasiological approach, cultural borrowings, core borrowings.

| ARTICLE INFORMATION

ACCEPTED: 25 October 2022 **PUBLISHED:** 23 November 2022 **DOI:** 10.32996/ijllt.2022.5.11.18

1. Introduction

Linguistic elements, such as words, phrases, or patterns, may be replicated from one language (i.e., the donor language or DL) to another language (i.e., the recipient language or RL) in contact situations which feature the interaction between speakers of different languages. This phenomenon is known as borrowing and the replicated elements are famously labelled "borrowings" (Myers-Scotton, 2006: 209; Matras, 2020: 158). Studies on borrowings have established that words of different types vary in their levels of borrowability, which could depend on their word class and lexical field. Regarding word class, it is universally found that nouns have the highest borrowability and that content words are more borrowable than function words mostly because of their referential values and semantic transparency (e.g., Moravcsik, 1978; Field, 2002; Poplack, 2018). Besides, the lexical field of borrowings also affects their receptivity in RL in that basic vocabularies, such as the words of motions, body parts, and kinship terms, are resistant to borrowing since they already exist in nearly all languages (Tadmor, 2009). In these studies, borrowability is often measured based on either the number of word types or the number of word tokens. However, the number of types cannot reflect the actual usage of borrowings considering that some words, despite being borrowed, are used at a very low frequency in RL. Although this limitation can be bridged by taking the number of tokens as a measure, previous studies usually count the raw frequency of the borrowings, which is problematic since the frequency of borrowings is constrained by the frequency of the topics of conversation and the related concepts designated by the borrowings (Calude et al., 2020: 30). In other words, what is really

measured by raw frequencies is the probability of occurrence of the concepts rather than the borrowability of different words, and the high frequency of some borrowings may just arise from the popularity of relevant topics instead of any linguistic attributes of the borrowings that facilitate their occurrence. Thus, to measure the potential use of borrowings, Calude et al. (2020) argue that we need to consider how likely it is to use a borrowing when people need to express a certain concept. This points to a "concept-based onomasiological approach to borrowing" which takes into account both borrowings and their native equivalents that designate the same concept (Zenner & Kristiansen, 2014: 1). Since borrowings and native equivalents are the lexicalizations of the same concept, the likelihood of using the borrowing to express the concept could be measured by the relative frequency of the borrowing compared with its native equivalents. This likelihood is also referred to as the success rate of borrowings by some linguists (e.g., Zenner et al., 2012; Winter-Froemel et al., 2014; Calude et al., 2020), and borrowings with a higher relative frequency than others are considered more successful. By contrast, the term *borrowability* is not favoured in this study because calculating the relative frequency based on token counts reveals the likelihood of using foreign items after they have entered the RL while *borrowability* can also mean the likelihood at the time of borrowing.

The onomasiological approach to measuring the success of borrowings has been operated in the studies on borrowings in Dutch (Zenner et al., 2012), German (Winter-Froemel et al., 2014), and New Zealand English (Calude et al., 2020), and these studies find that the success of borrowings hinges on multiple factors, such as word length, age of borrowing, and lexical field. However, there are few studies that have investigated the success of English borrowings in Chinese. Thus, this study aims to calculate the success rates of English borrowings in Chinese and examine how the success of these borrowings varies with different influencing factors (also known as constraints). To be specific, this study aims to answer the following questions: 1) What factors play a significant role in the success of English borrowings in Chinese? 2) How do these factors affect the success of English borrowings? To assess the combined effect of multiple factors and determine their significance, the study takes a variationist approach and conducts quantitative research through multivariate statistical models. Therefore, unlike previous studies that only focus on word class or lexical field, this study assesses multiple factors and can provide a more comprehensive picture of the borrowing scenario. Moreover, by figuring out the significant factors, this study can give insights into the motivations for using English borrowings in Chinese. Although qualitative studies can also examine influencing factors and shed light on the motivations for using borrowings, they cannot effectively explain whether the factors significantly constrain the success of borrowings (Zenner et al., 2012: 756).

2. Possible Constraints on the Success of Borrowings

The constraints that might have an impact on the success of borrowings can be divided into three types: word-related constraints, concept-related constraints, and stylistic and social constraints. Concerning some of the constraints, different studies do not always agree on whether they are significant. Since an onomasiological and multivariate approach to studying borrowings has been a recent innovation conducted by few studies, this section mainly reports the results in the studies of Zenner et al. (2012, 2014), Winter-Froemel et al. (2014), and Calude et al. (2020).

2.1 Word-related Constraints

Besides word class mentioned in **Introduction**, word-related constraints generally involve the length and age of borrowings. Word length could be a major concern in speakers' selection of borrowings, as a shorter form can achieve speech economy and trigger less processing load (Chesley & Baayen, 2010: 1352). The significance of word length is reported in Calude et al.'s study (2020) on Māori borrowings in New Zealand English, which confirms that shorter borrowings tend to be more successful. Similarly, the study conducted by Winter-Froemel et al. (2014) on English borrowings in German demonstrates that word length turns out to be a significant factor in a model that compares borrowings with each of their German equivalents. However, after a closer look into the English borrowings that are shorter than their native equivalents but are less frequent, they hypothesize that one syllable difference might not significantly contribute to the economy of expression, indicating that word length matters only up to a certain point. Likewise, although the quantitative research in Schaefer's study (2018) on English borrowings in German media discovers that most of the borrowings are shorter than their near equivalents, the qualitative analysis based on the interview data from the journalists suggests that the significance of word length could be limited as the respondents state that word length is not the primary concern. By contrast, Zenner et al. (2012) report no significance of word length in the success of English borrowings in Dutch. Nonetheless, when considering the interaction between word length and concept frequency, they reveal that when borrowings are the shortest lexicalization of their concepts, concept frequency no longer determines their success. This implies that length still plays a role in people's selection of borrowings, although its importance tends to be conditioned by other factors. Regarding the age of borrowing, it is found in both Zenner et al.'s (2012) and Winter-Froemel et al.'s (2014) studies that older borrowings are significantly more successful than the younger ones since the acceptance and diffusion of newcomers in the RL speech community takes time. Besides, other potential factors such as sense pattern (i.e., being monosemous or polysemous) and the level of integration into RL have also been tested in some studies, but they are either non-significant or only significant in a restricted context (e.g., Chesley & Baayen, 2010; Winter-Froemel et al., 2014; Calude et al., 2020).

2.2 Concept-related Constraints

As the use of borrowings pertains to the occurrence of the designated concepts, concept-related features may significantly govern the success of borrowings, including concept novelty, the number of lexicalizations of the concept, lexical field, and concept frequency. Concept novelty concerns whether a borrowing denotes a concept new to the RL speech community (Zenner et al., 2014: 91). The borrowings designating a new concept tend to be borrowed for filling lexical gaps, for which native equivalents may be introduced at a later stage, while the borrowings with an already lexicalized concept have to compete with existing native equivalents at the first place. The former type is often called cultural borrowings or necessary borrowings, whereas the latter is labelled core borrowings or luxury borrowings (e.g., Myers-Scotton, 2006; Haspelmath, 2009; Onysko & Winter-Foremel, 2011). According to Zenner et al. (2012: 775), necessary borrowings are preferred to luxury borrowings because the former are more likely to be absorbed by the RL, given the lack of native alternatives, at least at the time of borrowing. However, the difference is only borderline significant, and the effect of concept novelty could be neutralized by word length and age of borrowing (Zenner et al., 2012; Zenner et al., 2014). Concept novelty also makes a difference in the use of Māori words in New Zealand English, as it is found that speakers regardless of their ethnicities prefer cultural borrowings to core borrowings (Calude et al., 2020). Considering that the existence of native equivalents impacts the success of borrowings, Zenner et al. (2014: 88) hypothesize that the number of native competitors also comes into play in the success of luxury borrowings in that borrowings with fewer competitors at the time of borrowing face less competition and could be more successful, which, however, does not display statistical significance in their model. Another factor tied to concept novelty is the lexical field of borrowings. The fields of IT technology, media, and sports are found to generate a significantly higher success rate for English borrowings not only because these fields are "strongly influenced by the Anglo-American world" (Zenner et al., 2012: 766) but also because the rapidly-evolving feature of these fields makes them more likely to embrace borrowings that designate new concepts (Winter-Froemel et al., 2014: 123). Lastly, Zenner et al. (2012: 774) discover that the concept frequency of borrowings is significantly correlated with their success in that borrowings whose concepts have a higher frequency tend to be less successful. They argue that concepts with a higher frequency are more connected with the core vocabulary (or basic vocabulary) which is highly conventionalized in the RL speech community and resistant to borrowing.

2.3 Stylistic and Social Constraints

Regarding stylistic and social constraints, Zenner et al. (2012) assess the register variation between quality and popular newspapers and the regional variation between Belgian Dutch and Netherlandic Dutch, but neither of the factors can significantly predict the success of borrowings. Besides, probing into the social constraints of gender and ethnicity, Calude et al.'s study (2020) reveals that Māori borrowings are favoured by Māori females, not males, and that polysemous Māori words are disfavoured by speakers of non-Māori ethnicity.

3. Methodology

This study follows an onomasiological and multivariate approach to studying borrowings. Based on the onomasiological approach, this study distinguishes borrowings without native equivalents from those with native equivalents, determines the potential native equivalents for the latter, and calculates their success rates relative to their native equivalents. Then, under the multivariate approach, potential constraints are coded, and a multivariate statistical model is employed to investigate the impact of the constraints on the success of borrowings.

3.1 Identifying and Selecting English Borrowings

Instead of opting for a large corpus with hundreds of millions of tokens, this study relies on a relatively small dataset for quantitative research. This is mainly because in large corpora it is hard to disambiguate the meaning of homonymous or polysemous words in each context, especially for the high-frequency ones, which could interfere with the identification of the designated concept and the calculation of relative frequency and thus skew the results. The dataset in this study is formed by collecting data from hot topics on Weibo, a major online social platform in China. Weibo is chosen because the informal register of social media can encourage the use of English borrowings which tends to be resisted in formal settings partly due to the strict regulation on code-mixing by the Chinese government (Zhou et al., 2021: 61). Hot topics, introduced with hashtags, are created anonymously and made publicly available for Weibo users to comment on a certain topic, such as #浙江疫情# 'the epidemic in Zhejiang'. The study retrieves hot topics from the website https://weibo.zhaoyizhe.com/ and collects 4,000 hot topics with the most hits from each month throughout the year 2021. Hence, the dataset is composed of 48,000 hot topics in total, with around 514,000 tokens. Each word type roughly ranges between 1 to 500 occurrences in the dataset, making it possible for the study to disambiguate the meaning of homonymous or polysemous words and count the lexicalizations of a concept accurately.

Regarding English borrowings, this study settles on alphabetic words that have a foreign origin and are also used in native English. Alphabetic words are the items that still retain their alphabetic writing without undergoing orthographic adaptation to the logographic writing of Chinese, which makes them orthographically marked and easily recognizable as foreign items for Chinese speakers. In this case, we can make sure that their foreign status and non-Chinese character play a role in Chinese speakers'

selection of them, as suggested by Geeraerts and Grondelaers (2000: 56). Thus, measurement units are not included because they are usually written and perceived as symbols (e.g.,℃ for *degree Celsius*) rather than linguistic borrowings. It should also be noted that the foreign origin is not restricted to an English origin, given that the etymology of borrowings is generally opaque to the RL speakers. For example, despite initially originating from Latin, the English word *versus* is still treated as an English borrowing. Besides, although borrowings may consist of morphs, words, phrases, and even constructions (Matras, 2020: 158), only lexical borrowings are considered in this study, which, in the narrow sense, are also called "loanwords" and defined as the "unanalysable units in the recipient language" (Haspelmath, 2009: 37). Lexical borrowings are chosen because unlike other types of borrowings, their concepts can be easily identified, and they usually do not differ from the Chinese equivalents syntactically. Additionally, proper names are excluded from the analysis because they cannot be used productively by RL speakers (Onysko, 2007: 106).

When identifying borrowings, it is worth mentioning a closely related concept, namely, codeswitching, which refers to the alternation of different languages used within a conversation (Matras, 2020: 107). The status of a foreign lexical item may be ambiguous since it can be either a single-word codeswitch or a lexical borrowing. Although some linguists advocate using structural integration to distinguish the two phenomena based on the criterion that, as opposed to borrowings, codeswitches are not integrated into the recipient language phonologically, morphologically, or syntactically (e.g., MacSwan, 2016; Poplack, 2018), it is found that foreign items may have different levels of integration, indicating that borrowing and codeswitching are on a continuum without a clear boundary (e.g., Winford, 2010; Matras, 2020). Following the hypothesis of the codeswitching-borrowing continuum, this study does not unambiguously differentiate borrowings from codeswitches. Moreover, since the study collects data from hot topics on Weibo, it can be assumed that the English items in the topics have already been widespread and accepted in the Weibo community, approaching the borrowing end of the continuum. Otherwise, the topics cannot invite numerous comments and become popular. However, the English words that only occur once in someone's reported speech or in the title of a song, TV program, or other proper names, are excluded because they are more likely to arise from someone's idiosyncratic use of foreign words without following the constraints shared in the online community.

After all the steps of exclusion, 121 word types are selected for an onomasiological analysis.

3.2 Selecting Native Equivalents and Calculating Success Rates

Since an onomasiological approach requires taking native equivalents of borrowings into account, English borrowings in the dataset are checked for the existence of native equivalents. Drawing on the terminologies proposed by Onysko and Winter-Froemel (2011), this study refers to borrowings without native equivalents as catachrestic borrowings and those with native equivalents as non-catachrestic borrowings, considering that the traditional terminologies of cultural and core borrowings (or necessary and luxury borrowings) are prescriptive and misleading (for discussion, see Onysko & Winter-Froemel, 2011: 1551-1552). Catachrestic borrowings would be excluded because as the only lexicalization of a concept, they are exempt from the potential constraints on the success of borrowings. Without native equivalents, their relative frequency will always be 100%.

To identify non-catachrestic borrowings and select native equivalents, firstly, the concept of each borrowing is determined by referring to the contexts where they occur and widely consulting numerous dictionaries, encyclopaedia, and online sources, such as *A Dictionary of Loan Words and Hybrid Words in Chinese*, *A Dictionary of New Chinese Words in the Last 100 Years*, Baidu Baike (the Chinese counterpart of Wikipedia), Wikipedia, *The Language Situation in China*. At the same time, the meaning of homonymous and polysemous words is disambiguated, and two concepts would be identified for the English form with two senses. It should be noted that for these words, only the senses used in the dataset are considered. Secondly, these sources are consulted again to find out all the native expressions of each concept. However, not all the gathered expressions would be viewed as valid native equivalents. The expressions that fit the following criteria are excluded: 1) the expressions are not lexicalized, which instead take the form of descriptive phrases or sentences; 2) the expressions are not in use or only used as translations following the English words; 3) the expressions are the hypernyms of the English words. The exception to the first criterion is the phrases that have been standardised in Chinese, which are determined by consulting multiple Chinese dictionaries and the Standardized Chinese Translations of Foreign Words recommended by the Ministry of Education of the People's Republic of China (accessible from http://www.moe.gov.cn/s78/A18/A18_ztzl/ztzl_wyfygf/). The second criterion is included in line with a usage-based perspective, considering that some words, although recorded in the sources, are not picked up by Chinese speakers or conventionalized in their mental lexicon. Thus, for the speakers, these expressions are still unavailable for use. To apply this criterion, the gathered expressions are checked for their usage on the platform of Weibo (https://weibo.com) and in several corpora, such as the Weibo subcorpus of the BCC Corpus (http://bcc.blcu.edu.cn/) and the corpus of Chinese Web 2017 on Sketch Engine (https://app.sketchengine.eu/). The third criterion is adopted because hypernyms cannot express the exact denotation of the borrowings that assume a more specific meaning. After the exclusion, the remaining expressions are treated as the native equivalents of English borrowings. The borrowings that do not have a native equivalent are classified as catachrestic borrowings, while those with native equivalents are labelled non-catachrestic borrowings and then calculated for their success rates. Altogether, there are 81 non-catachrestic borrowings (with two polysemous words) and 83 concepts.

After the native equivalents are clustered with the borrowings under the same concept, the number of tokens is counted for each word based on the search results in the dataset. Then, the relative frequency of each borrowing is calculated vis-à-vis all the native equivalents that designate the same concept, which is used as the success rate. The meaning of the homonymous or polysemous native equivalents is also disambiguated in each context, and only the occurrences of the equivalents that denote the same concept as the borrowings are counted. Besides, when the native equivalents appear in compounds or fixed collocations, they are not counted. The calculation of the success rates is illustrated in Table 1: *rap* occurs 17 times while 说唱 *shuochang* occurs 5 times, so the success rate of *rap* is 77%; by contrast, the figure for *ending* is 4%.

Table 1 : The Calculation of the Success Rates of *rap* and *ending*

Concept	English borrowing (the number of tokens)	Native equivalent (the number of tokens)	Concept frequency	Relative frequency/success rate
RAP	rap (17)	说唱 (5)	22	17/22=77%
ENDING	ending (2)	结尾 (2) 结局 (49)	53	2/53=4%

3.3 Coding Constraints and Building the Statistical Model

In accord with previous studies, six constraints are coded for each borrowing, which might predict the success of the borrowings, as described below:

1) **Relative length** of the borrowings. This is measured by the difference in the number of syllables between the borrowing and its native equivalent. If a borrowing has more than one native equivalent, the comparison is made with its shortest native equivalent. For example, rap has one syllable while its equivalent 说唱 shuochang has two syllables, so rap is coded as -1 for this constraint. It could be hypothesized that the shorter a borrowing is, the more successful it will be.
2) **Concept novelty**. Borrowings that designate a new concept at the time of borrowing are coded as "cultural borrowings", while borrowings whose concept has been lexicalized by native equivalents at the time of borrowing are coded as "core borrowings". For clarification, the terminologies of cultural and core borrowings in this study concern the availability of native equivalents at the time of borrowing, whereas catachrestic and non-catachrestic borrowings are classified based on the existence of native equivalents at the time of research. Since cultural borrowings "have an easier time establishing themselves as a successful lexicalization for a concept" (Zenner et al., 2012: 775), they are assumed to be more successful than core borrowings.
3) **Age of borrowing**. For cultural borrowings, the study uses the time when the new concept or product was first introduced into China to gauge their age. By contrast, for core borrowings, the study uses the time when the borrowings first went viral in China to measure their age. It is hard to accurately trace the history of borrowings and determine their exact age. Thus, the study uses several age groups to code the borrowings instead of exact values, including 1-5 (years old), 5-10, 10-20, 20-30, and more than 30. The age group of 1-10 is further divided into 1-5 and 5-10 because it is easier to detect the recent introduction of a borrowing, which is not the case for the borrowings more than 10 years old. It could be hypothesized that older borrowings tend to be more successful.
4) **Concept frequency**. Concept frequency is calculated by adding the token counts of borrowings and native equivalents under the same concept. A negative correlation between concept frequency and the success rate of borrowings could be expected, since a more frequent concept is more likely to be lexicalized by core vocabularies that resist borrowing.
5) **Number of competitors**. Based on the number of competitors (i.e., native equivalents) at the time of research, borrowings are divided into three types: borrowings with 1-2 competitors, borrowings with 3-4 competitors, and borrowings with more than 4 competitors. Although lexical competition at the time of borrowing does not significantly determine the success of borrowings, as indicated by Zenner et al.'s study (2014), it is likely that the number of competitors at the time of research could account for the relative frequency of borrowings in that borrowings with fewer competitors might be more successful.
6) **Lexical field**. Four lexical fields are used to code the borrowings, namely, economy, business, and trade (e.g., *GDP*), entertainment, sports, and video games (e.g., *vlog*), technology, biology, and health (e.g., APP), and the general field (e.g., *pick*). Borrowings coded as "general" are not domain-specific and can be used in various fields. It could be hypothesized that borrowings in the fields that are rich in new inventions tend to be more successful.

Concerning the variables of concept novelty, age of borrowing, and number of competitors, numerous sources are consulted to gain information on the borrowings, the same as those mentioned in the last section. As for the borrowings whose age is not recorded anywhere, three corpora are used to gauge their age: the LCMC corpus (built around 1991), the CCL corpus (built around 2003), and the Corpus of Chinese Web 2017 (built around 2017). For example, if a word is only recorded in the last corpus, it would be gauged to be 5-10 years old. Although word class may also be correlated with success rates, this potential constraint is not included in the study, considering that many English borrowings and their native equivalents have more than one word class as a

result of category shift. For example, *battle* has both a verbal and nominal usage, and so does its native equivalent. Distinguishing their word class in different contexts would be too time-consuming.

After the coding procedure, the study conducts a multiple regression analysis in Rbrul (Version 4.1.3). In the statistical model, success rates are used as a continuous dependent variable, while the constraints are used as predictors (i.e., independent variables). The predictors of relative length and concept frequency are set as continuous, whereas the other predictors are binary. Following Zenner et al. (2012: 769), the values of concept frequency are logged to neutralize extremes. Lastly, the one-level modelling of Rbrul is used to predict the success of borrowings.

4. Results and Discussion

The results of the multiple regression model are presented in Table 2, which can explain 71% of the attested variation (R2 =0.714). Since a P-value < 0.05 indicates statistical significance, the results show that concept frequency, concept novelty, lexical field, and the number of competitors significantly contribute to the variation in the success of borrowings, while age of borrowing and relative length are non-significant variables.

Table 2 : Constraints on the Success of Borrowings

	Coefficient	**N**	**P-value**
Log (concept frequency)	-0.291	83	1.04e-0.8***
Concept novelty			0.000211***
Cultural borrowings	0.146	38	
Core borrowings	-0.146	45	
Lexical field			0.00928**
Technology, biology, and health	0.086	14	
Entertainment, sports, and video games	0.075	26	
Economy, business, and trade	0.038	8	
General	-0.199	35	
Number of competitors			0.0316*
>4	0.133	7	
3-4	-0.029	21	
1-2	-0.104	55	
Age of borrowing			0.578
10-20	0.048	28	
1-5	0.046	16	
20-30	0.014	11	
5-10	-0.049	23	
>30	-0.060	5	
Relative length	0.005	83	0.84

4.1 Significant Factors

4.1.1 Concept Frequency, Concept Novelty, and Lexical Field

Concept frequency significantly impacts the success of borrowings, and the negative value of the coefficient (i.e., -0.291) suggests that borrowings designating a more frequent concept tend to be less successful. This supports the hypothesis that a more frequent concept is more likely to be lexicalized by core vocabularies in the RL that have been conventionalized, thus incurring more resistance to the use of foreign words.

As expected, concept novelty also significantly contributes to the success of borrowings, and the values of the coefficients suggest that cultural borrowings are more successful than core borrowings. For one thing, this is because at the time of borrowing, cultural borrowings which denote a new concept lack native equivalents and have an easier time establishing themselves in the RL as the only choice to express a concept; on the contrary, core borrowings confront lexical competitions with their native equivalents at the time of borrowing. For another, the advantage of cultural borrowings over core borrowings may have a lasting effect even after their native equivalents are introduced at a later stage. Since cultural borrowings have been used for a longer time, they are more entrenched and conventionalized than their native equivalents, which makes cultural borrowings the preferred choice. To put it in another way, "people have accepted the borrowed term and gotten used to it" (Alzahrani, 2010: 34).

Concerning the lexical field, the results show that English borrowings are significantly disfavoured in the general field. By contrast, other fields which are more likely to absorb new inventions and more exposed to the Anglo-American world result in higher success rates of borrowings. As new inventions tend to be designated by cultural borrowings, a correlation between concept novelty and the lexical field can be expected. As table 3 displays, other fields contain a significantly higher proportion of cultural borrowings, whereas the general field is dominated by core borrowings which are usually less successful. Besides concept novelty, concept frequency is also associated with the lexical field. The general field is composed of borrowings that can be used in various domains, and so can their native equivalents. Therefore, the concepts designated by these borrowings will be more frequent than those designated by the borrowings in the other fields with their domain-specific usage and occurrences in limited topics. As illustrated by Figure 1, overall, the concept frequency of borrowings in the general field is higher than that in the other fields. Hence, given the association between lexical fields, concept novelty, and concept frequency, two generalizations can be made: 1) borrowings in the general field tend to be core borrowings whose concepts have a higher frequency which inhibits the use of borrowings; 2) borrowings in the other fields tend to be cultural borrowings whose concept frequency is relatively low because they are used in limited fields and topics, thus being more successful.

Table 3: The Proportion of Cultural and Core Borrowings in Different Lexical Fields

	Technology, biology, and health	Entertainment, sports, and video games	Economy, business, and trade	General
Cultural borrowings	93%	65%	88%	3%
Core borrowings	7%	35%	12%	97%

Despite the association with concept frequency and concept novelty, lexical field still exerts its own influence on the success of borrowings. To illustrate, the field of entertainment, sports, and video games (abbreviated into ESV hereafter) contains a much higher proportion of core borrowings than the field of economy, business, and trade (abbreviated into EBT hereafter), and the concept frequency of borrowings in the ESV field is also slightly higher than that in the EBT field. Nonetheless, the former witnesses a higher success rate than the latter, which indicates that the individual effect of lexical field is also manifested in the success of borrowings. This also draws attention to the comparison between the ESV field and the general field as the majority of core borrowings fall into the two fields (at 96%).

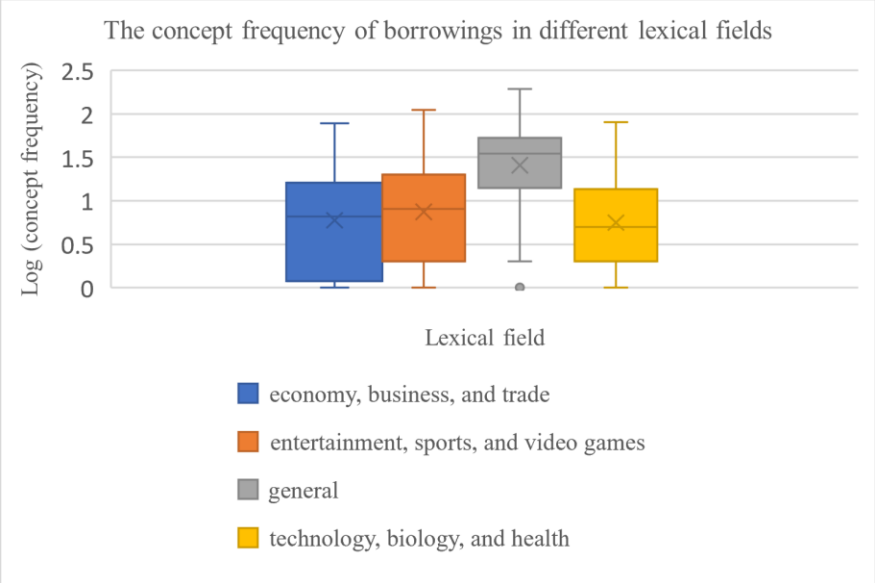

Figure 1 : The Logged Concept Frequency of Borrowings in Different Lexical Fields

To do a more nuanced examination of the effect of lexical field, the study builds a multiple regression model by only including core borrowings in the ESV and general fields. Core borrowings in the other two fields are excluded because both only contain one core borrowing, which can skew the results. As Table 4 shows, core borrowings in the ESV field are significantly more successful than their counterparts in the general field. This pattern could be well-explained by referring to the pragmatics of borrowings. On the one hand, since English borrowings, especially the unintegrated alphabetic ones in this study, are usually restricted to informal

settings, the association with informality can endow these expressions with the indexicality of relaxation and entertaining effects, as Chen (2007: 185) notes in the codeswitching between Taiwanese and Mandarin; by contrast, the native equivalents are often assigned the value of seriousness. Thus, in the ESV field which features a relaxing and entertaining atmosphere, English borrowings are preferrable to their native equivalents thanks to their closer connection with relaxation and entertainment. On the other hand, borrowings in the general field are used at a much lower frequency than their native equivalents, which thus function as a marked choice to convey "pragmatic markedness" (Onysko & Winter-Froemel, 2011: 1555). In other words, these borrowings are mainly used for pragmatic functions, and this usage entails their marked status and their lower frequency, whereas the native equivalents serve as a default term without carrying special pragmatic effects. For example, in the dataset, only constituting 4% of the occurrences of the designated concept, baby is used in a marked way to add special pragmatic meanings. As in (1a), *baby*, modified by 老 *lao* 'old', refers to the senior citizens who pursue romanticism like the youth. The use of the English borrowing here performs the pragmatic function of projecting the youth-like image of the old generation. By contrast, the Chinese equivalent 婴儿 *yinger* 'baby' is used as a default choice with a higher frequency to provide a stereotypical interpretation of the concept, as in (1b).

Table 4: The Effect of Lexical Field on the Success of Core Borrowings in the ESV and General Fields

	Coefficient	N	P-value
Lexical field			1.32e-05***
ESV	0.231	9	
general	-0.231	34	

(1) a. 老 **baby** 的七夕太上头了
'The senior citizens celebrated the Chinese Valentine's Day in an intoxicating way.'
b. 阿富汗**婴儿**被从铁丝网上递进机场
'An Afghan baby was handed into the airport over barbed wire.'

4.1.2 Number of Competitors

Contrary to the hypothesis that less lexical competition leads to higher success rates, Table 2 shows that borrowings with the largest number of competitors are significantly more successful. Besides *MV* (shorter for music video) whose success may primarily arise from its status as a cultural borrowing in the ESV field, the other borrowings with more than 4 competitors are core borrowings in the general field, and a possible explanation for their high success rates is that they are more versatile than their native equivalents with a broader range of referents. To illustrate, the English borrowing *logo* (with 12 tokens) can refer to all kinds of specific signs of an organization, whereas its native equivalents have a more specialized denotation: 会徽 *huihui* (with 1 token) stands for a logo for a sports meeting, 队标 *duibiao* (with 2 tokens) refers to a logo for a team, 团徽 *tuanhui* (with 2 tokens) refers to a logo for a league, etc. Likewise, *emo* (with 18 tokens), which might be short for emotional or emotion, has undergone semantic shift after borrowing and now refers to various negative emotions designated by different native equivalents, such as 丧 *sang* 'depressed' (with 6 tokens), 悲伤 *beishang* 'sad' (with 1 token), and 难过 *nanguo* 'upset' (with 5 tokens). Therefore, the broader meaning of the borrowings relative to their native equivalents can make them occur in more contexts and contribute to their success. Although one may argue that these borrowings resemble polysemous words whose different senses should be disambiguated, this is not done in the study considering that their different meanings are highly correlated and can be clustered under a shared label.

Although number of competitors appears to be relevant to the success of borrowings, the result should be interpreted with caution not only because its significance is relatively weak (p-value: 0.0316, slightly lower than 0.05) but also because there are only seven borrowings with more than four competitors, which indicates that the explanation provided by the study might not be generalizable.

4.2 Non-significant factors

4.2.1 Age of Borrowing

The P-value shows that the success of borrowings does not significantly depend on the age of borrowing, but the different values of the coefficient of different age groups suggest that age of borrowing may slightly affect the success of borrowings. Borrowings in the age groups of 5-10 (a younger one) and more than 30 (the oldest one) are less successful than those in other age groups, reflecting two opposing patterns. This may result from the different behaviours of cultural and core borrowings, given that the majority of the borrowings of 5-10 years old are constituted by core borrowings, which is reversed in the borrowings older than 30, as illustrated in Table 5. According to Zenner et al. (2012: 778), cultural borrowings are negatively correlated with the age of borrowing since they "can gradually lose ground to the upcoming alternatives", whereas the success rates of core borrowings may increase with age because they "need time to become a worthy competitor" of their conventionalized native equivalents. Hence,

borrowings in the age group of 5-10 are less successful because most of them are younger core borrowings; conversely, borrowings in the age group of more than 30 are less successful due to the dominance of older cultural borrowings. However, it should be noted that this explanation is debatable, considering that the oldest age group only contains five borrowings in this study. Interestingly, the age group of 1-5, which has a higher success rate despite the preponderance of the youngest core borrowings, behaves contrary to the hypothesis that a higher success rate can be found in older core borrowings. Concerning this result, the study proposes two explanations. Firstly, core borrowings may be favoured for their indexicality of trendiness and innovation, but these values might fade away over time. For example, although *style* was a buzzword about ten years ago (Zhang, 2015), in this dataset it only occurs once (relative frequency: 10%). Therefore, the youngest core borrowings have the strongest association with the favourable values, which contributes to their success. Secondly, since the study uses the time when core borrowings went viral to measure their age of borrowing, it is not surprising that the borrowings that gain popularity in recent years are used quite frequently. All in all, the importance of the age of borrowing should not be overstated, and a larger dataset is needed in future studies to test the diverging hypotheses related to cultural and core borrowings.

Table 5: The Proportion of Cultural and Core Borrowings in Three Age Groups

	1-5 years old	5-10 years old	>30 years old
Cultural borrowings	6%	30%	80%
Core borrowings	94%	70%	20%

4.2.2 Relative Length

Another non-significant factor is the relative length of borrowings in comparison with their native equivalents. Although word length may not significantly determine the success of borrowings, this does not deny the importance of speech economy in people's language use. For one thing, among the 83 concepts, only 6 concepts (7%) are designated by borrowings that are longer than their native equivalents, and all of them are only one syllable longer. This may indicate that as long as the borrowing is no longer than its native equivalents, the increase in the relative length does not significantly impact its success. Besides, far from being a coincidence, the marginal proportion of the longer borrowings appears to result from speakers' deliberate selection or manipulation. Among the 77 concepts designated by borrowings no longer than their native equivalents, around 50% (38 concepts) are designated by borrowings in the form of abbreviations. Some of them have been abbreviated at the time of borrowing, which are also used in native English, such as *CEO* and *TV*, while some have undergone abbreviation after they entered Chinese, such as *cosplay* shortened into *cos*. Besides, except that cos alternates with the longer form cosplay, in all other cases, only the abbreviation is used in the dataset. Therefore, it seems that the factor of length is taken into account at the time of borrowing so that the shortest form of a word is more likely to be selected, and after borrowing, for the economy of expression, speakers opt for shortening the borrowings instead of employing another shorter word, which suggests that word length, although it is not a paramount factor, still influences speakers' language use. For another, word length may affect the success of cultural and core borrowings differently. Concerning cultural borrowings, Figure 2 shows that the shorter the cultural borrowings are in relation to their native equivalents, the more successful they are; by contrast, the success of core borrowings is nearly unaffected by length difference, as reflected by the approximately horizontal trendline in Figure 3. For cultural borrowings, the shorter length can, to some degree, add to their competitiveness, complying with the hypothesis on speech economy. Regarding core borrowings, the limited effect of word length may relate to people's motivations for using them. To elaborate, as core borrowings are always in competition with their native alternatives, they are primarily used for special pragmatic effects, which thus diminishes the effect of word length since this factor is not closely connected with additional pragmatic meanings. Given that relative length does not reach statistical significance, its potential effects should be considered cautiously, and future research can further distinguish the different patterns of cultural and core borrowings based on a larger dataset.

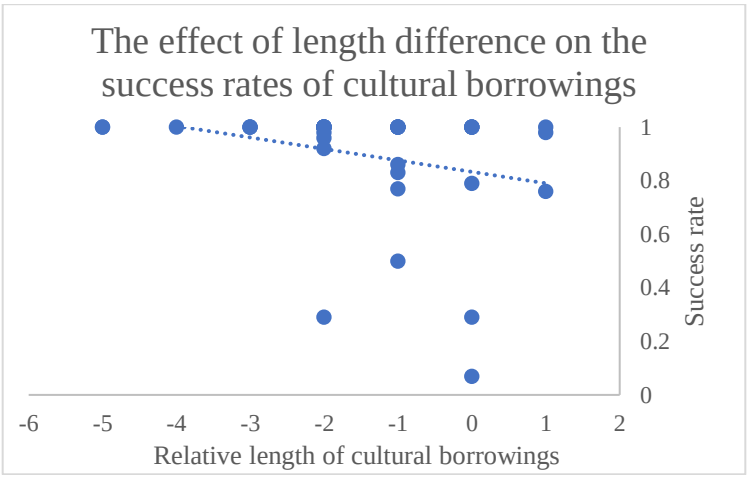

Figure 2 : The Effect of Length Difference on the Success Rates of Cultural Borrowings

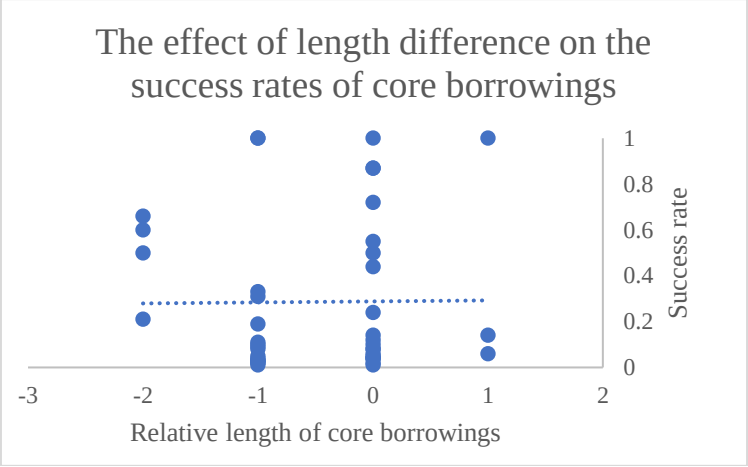

Figure 3 : The Effect of Length Difference on the Success Rates of Core Borrowings

5. Conclusion

To determine what factors significantly affect the success of English borrowings in Chinese, this study built a multiple regression model in Rbrul by including the factors of concept frequency, concept novelty, lexical field, the number of competitors, age of borrowing, and relative length of borrowings. Through the analysis of the statistical results, it can be concluded that borrowings designating a new or low-frequency concept and borrowings falling into the lexical fields rich in new inventions are significantly more successful than others. Although borrowings with more than 4 competitors at the time of research also turn out to be significantly more successful, this may not count as a robust pattern due to the relatively weak significance of the number of competitors and the limited data of the borrowings of this type. In contrast, age of borrowing and relative length do not reach statistical significance in the study, having a weak explanatory power for the variation in the success of borrowings. However, the result that relative length exerts limited influence on the success of borrowings does not necessarily contradict the importance of speech economy in people's language use. As reflected by the high proportion of abbreviations and borrowings no longer than their native equivalents, instead of being deterred from using a borrowing by its length, speakers opt for abbreviating the borrowing and selecting the shorter form for speech economy. Besides, the correlation between concept frequency, concept novelty, and lexical field points to the prominent distinction between cultural borrowings and core borrowings: cultural borrowings which designate new and low-frequency concepts and mostly fall into the rapidly-evolving fields are more successful than core borrowings which designate existing and high-frequency concepts and mostly fall into the general field. The significance of the factors also gives an insight into speakers' motivations for using borrowings. To elaborate, cultural borrowings are employed mainly because speakers have been used to them as they are more entrenched than their native equivalents, whereas core borrowings as a marked choice with a lower frequency are primarily used to create special pragmatic effects. Thus, it is not only important to distinguish borrowings with native equivalents from those without native equivalents at the time of research but also important to do so at the time of borrowing, both highlighting the significant value of an onomasiological approach to borrowing. Since cultural borrowings and core borrowings may have different developmental patterns after they enter the RL, future studies can do a more nuanced analysis of how the influencing factors affect the success of the two types of borrowings differently.

Considering a dearth of studies on measuring the success of borrowings, this study provides new insights into the factors that inhibit or encourage speakers' selection of borrowings and shed light on the motivations for using borrowings. In line with previous studies, this study confirms the significance of concept-related constraints in determining the success of borrowings and discusses the lasting effect of concept novelty and the correlation between different factors in greater detail, which can enrich borrowing research and advance the understanding of the borrowing scenario in areas where English serves as a foreign language. Nevertheless, since this study only relies on data from a single domain, i.e., hot topics on Weibo, the data under investigation are far from representative, failing to project a comprehensive picture of the use of borrowings in Chinese. Due to the restricted data source, several borrowings reach a relative frequency of 100% in the dataset, although their variation with the native equivalents is attested somewhere else. For this reason, the results of the study only reflect the usage of borrowings in a restricted domain, which may not be generalizable to other domains. Another limitation is that this study does not include word class as a predictor in the model, considering the complexity of determining the word class of each instance of borrowings and native equivalents. However, previous studies have indicated that structural ease and lexical content boost the borrowability of nouns. Thus, it would be interesting for future studies to conduct advanced corpus searches to test whether these structural values which affect the adoption of foreign words at the time of borrowing still exert influence on the use of these words after they have been incorporated into the RL. Besides, it is worth noting that the success of borrowings in this study is measured based on the synchronic data without considering the diachronic dimension. Therefore, future studies could also investigate how the relative frequency of borrowings changes diachronically and whether the factors tested in this study are tied to the persistence of borrowings in the RL.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Alzahrani, A. (2010). *The use of English lexical borrowings among Saudis* [Master's dissertation, Southern Illinois University Carbondale]. Retrieved from https://www.proquest.com/openview/617b6bd19166688eb5017521ee424ade/1?pq-origsite=gscholar&cbl=18750
[2] Calude, A. S., Miller, S., & Pagel, M. (2020). Modelling loanword success – a sociolinguistic quantitative study of Māori loanwords in New Zealand English. *Corpus Linguistics and Linguistic Theory*, *16*(1), 29–66. https://doi.org/10.1515/cllt-2017-0010
[3] Chen, H. (2007). *Code-switching in conversation: a case study of Taiwan* [Doctoral Dissertation, The University of Manchester]. Retrieved from https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.532992
[4] Chesley, P., & Baayen, R. H. (2010). Predicting new words from newer words: Lexical borrowings in French. *Linguistics*, *48*(6), 1343-1374. https://doi.org/10.1515/ling.2010.043
[5] Field, F. W. (2002). *Linguistic borrowing in bilingual contexts*. J. Benjamins. https://doi.org/10.1075/slcs.62
[6] Geeraerts, D., & Grondelaers, S. (2000). Purism and fashion. French influence on Belgian and Netherlandic Dutch. *Belgian Journal of Linguistics*, *13*(1), 53–67. https://doi.org/10.1075/bjl.13.04gee
[7] Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages: A comparative handbook* (pp. 35–54). Walter de Gruyter. https://doi.org/10.1515/9783110218442.35
[8] MacSwan, J. (2016). Code-switching in adulthood. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 183–200). American Psychological Association. https://doi.org/10.1037/14939-011
[9] Matras, Y. (2020). *Language contact*. Cambridge University Press. https://doi.org/10.1017/9781108333955
[10] Moravcsik, E. (1978). Universals of language contact. In J. H. Greenberg (Ed.), *Universals of human language* (pp. 94–122). Stanford: Stanford University Press.
[11] Myers-Scotton, C. (2006). *Multiple voices: an introduction to bilingualism*. Blackwell Publishing.
[12] Onysko, A. (2007). *Anglicisms in German: borrowing, lexical productivity, and written codeswitching*. Walter De Gruyter. https://doi.org/10.1515/9783110912173
[13] Onysko, A., & Winter-Froemel, E. (2011). Necessary loans – luxury loans? Exploring the pragmatic dimension of borrowing. *Journal of Pragmatics*, *43*(6), 1550–1567. https://doi.org/10.1016/j.pragma.2010.12.004
[14] Poplack, S. (2018). *Borrowing: loanwords in the speech community and in the grammar*. Oxford University Press. https://doi.org/10.1093/oso/9780190256388.003.0004
[15] Schaefer, S. J. (2019). Anglicisms in German media: Exploring catachrestic and non-catachrestic innovations in radio station imaging. *Lingua*, *221*, 72–88. https://doi.org/10.1016/j.lingua.2019.01.002
[16] Tadmor, U. (2009). Loanwords in the world's languages: Findings and results. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages: a comparative handbook* (pp. 55–75). Walter de Gruyter. https://doi.org/10.1515/9783110218442.55
[17] Winford, D. (2010). Contact and borrowing. In R. Hickey (Ed.), *The handbook of language contact* (pp. 170–187). John Wiley & Sons, Inc. https://doi.org/10.1002/9781444318159.ch8
[18] Winter-Froemel, E., Onysko, A., & Calude, A. (2014). Why some non-catachrestic borrowings are more successful than others. A case study of English loans in German. In A. Koll-Stobbe & S. Knospe (Eds.), *Language contact around the globe* (pp. 119–142). Frankfurt am Main: Peter Lang.
[19] Zenner, E., & Kristiansen, G. (2014). Introduction: Onomasiological, methodological and phraseological perspectives on lexical borrowing. In E. Zenner & G. Kristiansen (Eds.), *New perspectives on lexical borrowing: Onomasiological, methodological and phraseological innovations* (pp. 1–18). Walter de Gruyter. https://doi.org/10.1515/9781614514305.1

[20] Zenner, E., Speelman, D., & Geeraerts, D. (2012). Cognitive Sociolinguistics meets loanword research: Measuring variation in the success of anglicisms in Dutch. *Cognitive Linguistics*, *23*(4), 749–792. https://doi.org/10.1515/cog-2012-0023
[21] Zenner, E., Speelman, D., & Geeraerts, D. (2014). Core vocabulary, borrowability and entrenchment. *Diachronica*, *31*(1), 74–105. https://doi.org/10.1075/dia.31.1.03zen
[22] Zhang, W. (2015). Multilingual creativity on China's Internet. *World Englishes, 34*(2), 231–246. https://doi.org/10.1111/weng.12135
[23] Zhou, Y., Zhou, J., Teng, Y., & Cheng, N. (2021). Survey on Lettered Words in Chinese Government Microblog. *Applied Linguistics*, *2*, 55–65. https://doi.org/10.16499/j.cnki.1003-5397.2021.02.005