



A Corpus-Based Analysis of Two Chinese Loanwords in English Newspapers: Evidence for Linguistic Innovation and Propagation

Wenfeng Li

In the past decades, work on different English varieties, the more well-established or emerging ones alike, has increased exponentially and investigations at various linguistic levels have produced a wealth of evidence pertaining to claims about the status of these varieties. This paper will focus on the lexical level, and offers a preliminary corpus-based analysis of two Chinese loanwords used in contemporary English media. The goal is to discover patterns of use of these two loanwords in different contexts and at different time periods.

Keywords: *Corpus linguistics; Sociolinguistics; Language contact; Linguistic change*

Introduction

In the past decades, work on different English varieties, the more well-established or emerging ones alike, has increased exponentially and investigations at various linguistic levels have produced a wealth of evidence pertaining to claims about the status of these varieties. This paper will focus on the lexical level, and offers a preliminary corpus-based analysis of two Chinese loanwords used in contemporary English media. The goal is to discover patterns of use of these two loanwords in different contexts and at different time periods.

Literature review

English varieties and lexis

Lexis or lexical borrowing lies at the heart of the definition of one variety of English. As far as Inner Circle countries are concerned, the use of distinctive lexical items to express unique social, cultural, political notions, events, and phenomena is a landmark of one particular English variety. Back in the beginning of the 20th century, Mencken (1919) argued fervently for the existence and “legitimacy” of an American variety of English, basing much of his claim on the observation of America’s innovation of a wide range of lexical items denoting objects, products and phenomena in a variety of domains (biology, geopolitics, society, culture, literature, technology, etc.) that were found

in the U.S. Görlach (1998) made a statement along the same lines in his discussion of how the settlers in the new land adapted their linguistic system to the new situations and he claimed that innovation in the lexicon constituted a major process in the formation of a new variety of English.

Lexis plays an equally, if not more important role in the definition and/or identification of English varieties in the Outer Circle and Expanding Circle. According to Kachru (1983), “the linguistic characteristics of Indian English are transparent in the Indian English sound system (phonology), sentence construction (syntax), vocabulary (lexis), and meaning (semantics)” (p. 66). Deterding (2007) listed several (borrowed) lexical items used by all races in Singapore as part of the defining features of an emerging distinct brand of Singaporean English. Commenting on the controversy on the existence of a Hong Kong variety of English, Bolton & Kwok (1990) suggested that “(i)f one can establish that (in addition to identifiable local accent) there are clusters of shared *lexical* and grammatical items which contribute to a distinctive body of shared linguistic features then this may well legitimatise recognition of Hong Kong English as a localised variety.” (p. 163, cited in Benson, 2000, p. 373; italics mine)

In the Chinese context, much importance is attached to lexis (with Chinese characteristics) in the research on a Chinese variety of English, although controversy remains on the existence, label and definition of such a variety. In the earliest research on the English language used in China, Ge (1980) pointed out the necessity of distinguishing “China English” from “Chinglish” or “Chinese English”—the two labels for “some English expressions in books or magazines published in China that didn’t conform to the norms observed by the people from the English-speaking countries” given by “foreigners who read about such usages/expressions, considering them to be completely unacceptable” (Ge, 1980, p. 2). Ge (1980) acknowledged the importance of abiding by such norms, but insisted that there was a need to express things that were unique about China using “China English”. The examples of such “China English” he gave are English translations of some Chinese concepts and phenomena (e.g. “four modernizations”), which could be understood by the English-speaking people along with some explanation (Ge, 1980, p. 2). It can be inferred that vocabulary was one, if not the only one important defining feature of “China English” in Ge (1980). Although Ge’s view on China English is considered to be incomprehensive in later research (e.g. Jin, 2003), it points to the important role lexis plays in defining a Chinese variety of English. The relevance of lexis or vocabulary, along with other dimensions of English usage, in defining a Chinese variety of English was more explicitly expressed in Li (1993), who posited that “China English is a composite of lexis, syntax, and discourse with Chinese characteristics” (p. 19). This view was echoed by He & Li (2009), Jiang & Xiang (1997) and Wang (1991) and other scholars who argued for the existence of a Chinese variety of English.

Lexical borrowing: innovation vs. propagation; synchronic vs. diachronic

With vocabulary, or lexis being placed at the center of the definition of an English variety—which in this case, refers to a Chinese variety of English, the question naturally arises as to what characterizes a lexicon of a particular English variety and how it is developed to delineate the boundaries between the variety in question and other varieties. Put in a different way, the question can be broken down into three sub-questions. First, what are the mechanisms whereby language changes at the lexical level to produce a unique lexicon for a particular variety? The second sub-question is about how such adaptation affects or contributes to the general structure of an emerging language variety. Third, what are the results of this kind of lexical adaptation that mark the distinctiveness of that variety, which are synonymous with what Benson (2000) referred to as “vocabulary of regional varieties of English” (p. 374)? This paper will address the first two sub-questions in the analysis of two new borrowings, with a view to gleaning sufficient knowledge for proposing a tentative answer for the third one.

Previous research on different varieties of English has provided an answer to the first sub-question above: one major mechanism of linguistic adaptation at the lexical level is lexical borrowing,

along with coinage or invention of new lexemes. For instance, Mencken (1919) asserted that the “Americanisms” or the linguistic innovations manifest in the lexical system of the then-emerging American variety were realized by the invention of new lexical items and borrowing of words from outside resources (p. 90). Görlach (1998) also listed borrowing as one of the three principal methods of lexical adaptation, along with coinage and semantic shift of existing words. The current investigation will focus on borrowing, in particular one type of borrowing, and how it relates to the development and demarcation of a Chinese variety of English.

Haugen’s (1950) work on linguistic borrowing laid the foundation for later research. For the purpose of this paper, I will adopt his definition of borrowing. According to Haugen, borrowing is “the attempted reproduction in one language of patterns previously found in another” (p. 212). Although borrowing was defined as a process involving reproduction, the term “borrowing” was used to refer to the results or instances of borrowing in Haugen (1950). Following Haugen, I would use the term to refer to both the process and products/instance of borrowing.

Apart from providing a definition of borrowing, Haugen (1950) also discussed the independence of the act or process of borrowing from the result of such an action and offered a detailed analysis of various kinds of interference that loans might undergo in various socio-cultural contexts, which would lead to the alteration or change in such loans. He underscored the importance of isolating the “initial leap of pattern from one language to another” and stated that “every loan now current must at some time have appeared as an innovation”, although it is difficult to “catch a speaker in the actual process of making an original borrowing” (p. 212).

This point made in Haugen (1950) has close bearing on the distinction between the two key concepts in Croft’s (2000) analysis of language change (p. 4-5): “innovation”, defined as altered replication or “the creation of new forms in the language” (p. 4), and “propagation” or the spread of the change or the new forms in the language (p. 4-5). These two constructs would serve as a good starting point for examining borrowings: in effect, the act of borrowing could be seen as a process of innovation, since new features are added to the linguistic pool, though they are not equivalent in the strict sense. Furthermore, an examination of the propagation process of borrowings has the potential of answering the question about how lexical adaptation contributes to the structure of an (emerging) English variety. In this sense, Croft’s model would be useful in the analysis of data on lexical borrowing. If Croft’s model were to be adopted for the current study, both synchronic and diachronic analyses would be necessary, since “(i)nnovation is a synchronic phenomenon,.....it occurs in speaker action at a given point in time” whereas (p)ropagation is a diachronic phenomenon” occurring “sometimes over a very long period of time, even centuries” (Croft , 2000, p. 5). The former kind of analysis can be conducted by investigating the motivations and/or functions for certain innovations, the characteristics of such innovations, the contextualization (contextual features) of these innovations, etc. , while the latter will encompass a wider range of research, including work on the use and/or frequency of use of some innovations in different domains, in different regions and countries, how the usage might have changed within and/or across regions/countries over time and whether they have been or are in the process of being accepted into general international English.

It should be pointed out that Croft’s (2000) theory dealt mainly with language change, in particular changes within a language or language family. Although there was a relatively brief discussion about borrowing, Croft approached this issue from a global perspective, focusing his analysis on contact situations, language maintenance & resistance to contact-induced borrowing. The two notions emphasized throughout the book didn’t seem to have much applicability or relevance to the discussion of borrowing in Croft (2000). In this paper, I would draw on his model of language change and point out where adaptations or modifications are needed to adequately account for the data.

Lexical borrowing: categorization

In the definition and categorization of borrowing, this paper will follow Haugen (1950), who put forward three major categories of borrowing, including loanwords, loanblends and loanshifts. According to Haugen, loanwords “show morphemic importations without substitution” (p. 214), with a varying degree of phonemic substitution; loanblends show both morphemic substitution and importation whereas only morphemic substitution is involved in loanshifts, which include both loan translations and semantic loans. In his definitions of these terms, inflectional modifications are not included in the term “morpheme”.

Of particular relevance to this paper is the first category of borrowing, namely, loanwords involving direct morphemic importation. These three categories are investigated in various studies on Chinese borrowings (e.g. Cannon, 1987, 1988; Gao, 2001; Yang, 2005, 2009). Research on China English has used the term of “transliteration” (or its Chinese equivalent, 音译, yinyi) in place of the more general “loanword” (He & Li, 2009; Jin, 2003; Li, 1993; Wan, 2005); transliteration refers to Chinese borrowings transcribed in a Roman alphabet based on their pronunciation. According to Moody (1996), the majority of the Chinese loanwords listed in Cannon (1987, 1988) were borrowed into English at a time when “there was no standard or even reliable system for transcribing Chinese characters in a Roman alphabet” (Moody, 1996, p. 412); in addition, many of the words were transcribed according to their pronunciation in different Chinese dialects (e.g. Cantonese, Amoy) from which they were borrowed. In the past three decades, however, Pinyin¹ has been gaining more recognition and influence in the transliteration of Chinese loanwords and Pinyin-based loanwords have been on the increase (see Yang, 2009 for a review). A comparison of “transliteration” and the term “loanword” as defined in Haugen (1950) reveals the superiority of the former in the study of Chinese borrowing: although no phonemic change occurs to Chinese words introduced into English in this way, such borrowings do not quite qualify as “morphemic importations without substitution”, given the orthographic changes and the loss of tonal information, which I will elaborate in the discussion. However, since the term “loanword”, along with “loan translation” has been commonly recognized in the field (Yang, 2009, p. 91), it will be used interchangeably with “transliteration” in this paper.

Chinese borrowings in English

Chinese borrowings in English have received abundant research attention and have been examined from different perspectives for different purposes: a number of studies have approached this topic from a sociolinguistic perspective, with a focus on China’s social and cultural influence on general English vocabulary and processes and forces underlying such influence (Bian et al., 2007; Du, 1999; Li, 2001; Li, 2009; Moody, 1996; Moody & Ma, 2008; Wang & Chang, 2001; Zhou, 2011). Some studies focus on existing Chinese borrowings in English, and the mechanisms whereby such borrowings came into being (e.g. transliteration, loan translation, semantic shift) (Chen & Zhou, 2012; Tao, 2011; Xiong, 1996; Zhang, 2008). For instance, Chen & Zhou (2012) reviewed the list of Chinese transliterations in Moody (2008) and in Liang (2006, cited in Chen & Zhou, 2012, p. 140-141), in addition to some recent loanwords they observed. This kind of studies provides a relatively comprehensive or updated list of Chinese borrowings but is not satisfactorily informative as far as the status, usage and spread of such borrowings are concerned.

The phenomena of borrowing from Chinese into English have also been investigated by researchers fighting to establish the legitimacy of a Chinese variety of English, be it called Chinese English or China English (see Chen, 2007; Du & Jiang, 2001; Ge, 1980; Jiang & Xiang, 1997; Jin, 2002, 2003; Li, 1993; Li, 2004; Wan, 2005; Wang, 1991; Wang, 2009; Wang, 2011; Wu, 2009; Yang, 2010; Yuan & Lu, 2003; Zhu & Liu, 2006, etc.). A similar list of Chinese borrowings can be found in these studies, including loanwords and loan translations. For example, Ge (1980) listed various (then-current) loan translations that were associated with Chinese society, culture and politics (e.g. four

¹ Pinyin is the official system for transcribing Mandarin Chinese.

modernizations, Two Hundred Policies, Four Books) as well as loanwords for that matter (e.g. *baihuawen*, *xiucaí*). Expanding on Ge's list (1980), Li (1993) provided more examples of Chinese borrowings in English, by means of transliteration (e.g. *kowtow*, *yamen*) and loan translation (e.g. Great Leap Forward, Gang of Four). Jiang & Xiang (1997) included in their list of examples both loan translation (e.g. spiritual civilization) and semantic shift. Jin (2003) also listed some borrowings to illustrate the different features of a Chinese variety of English. In his analysis, "a shoe shiner" (Hong Kong English) is an example of Sinicized English, and "one country, two systems" illustrates a kind of Chinese-flavored English, while borrowings like "kung fu" belong to expanded/enriched English, which has contributed to the international Englishes.

To sum up, Chinese borrowings have been discussed or mentioned by scholars aiming to establish China English or Chinese English—whatever the label they give it, as a legitimate variety of English or otherwise. Nonetheless, without a systematic analysis looking into the actual usage, status and acceptance of such borrowings in different contexts and situations, these studies fall short of their target. The argument that a Chinese variety of English exists or is coming into existence needs to be built on linguistic evidence such as phonological, lexical, syntactic and pragmatic innovations that are uniquely Chinese. I would argue that a list of borrowings do not measure up to evidence of lexical innovations to corroborate the claim about an English variety, unless it is demonstrated that the phenomenon of borrowing is robust and has been or is in the process of being recognized on a regional, national or even global scale. Relevant to this point is Haugen's (1950) analysis of different stages of borrowing. It can be implied from his study that if some borrowings fail to find their way into the linguistic system or become part of the linguistic repertoire of the (majority of the) speakers, with or without some modification (morphemic, phonemic, or semantic), it is not likely that they will contribute to the marking of that particular variety of language. In this sense, the results of borrowing, or the acceptance of borrowed items into the borrowing language needs to be proved, necessitating some kind of corpus- or data-based analysis. In fact, this kind of analysis has been conducted by different scholars, though the studies are not without their own problems. I will come back to this later in this section. In addition, if researchers are to use instances of borrowing to argue for the legitimacy of a variety of English, it will be of enormous significance to investigate if borrowing, a form of nativization, is a continuous process in that variety. In other words, apart from the more established borrowings, research needs to include in their list new items that are undergoing the process of borrowing, bearing in mind the notions of innovation and propagations in their investigation of such items. If the robustness of the borrowing phenomenon or processes is borne out in such empirical studies, it will give much validity to claims about the emergence and/or existence of the variety under investigation, from the point of view of vocabulary. The availability of large corpus of English media, newspapers and other forms of English communication provides a unique opportunity for researchers who are willing to undertake a project on the current usage of new borrowings and their status in their original and other contexts.

As mentioned previously, corpus- and data-based studies on this topic are not lacking in the field. Work on the acceptance of Chinese borrowings in general English can be traced back to the 1980's: The influential work by Cannon (1987, 1988) revealed the presence of 196 Chinese borrowings, mostly loanwords and loan translations, in desk English dictionaries. Although Cannon made no argument for or against the existence of a Chinese variety of English, his work provided valuable evidence of the continuing nativization of English in China and of their acceptance in general English. His research method has also inspired later research on the status of Chinese borrowings in English. For instance, Yang (2009) conducted a rigorous examination of the Chinese borrowings in English listed in Cannon (1988), comparing them to the corpus from eight up-to-date desk dictionaries of English. The resulting list was composed of contemporary Chinese borrowings accepted into general English; the change in the status of different Chinese languages (e.g. Cantonese, Amoy and Mandarin) and different orthographic systems (e.g. the Yale or Wade-Giles System; Pinyin) in the process of borrowing were also examined in the study. Of special relevance to this paper is his

data-based observation of an upward tendency for Pinyin-based loanwords (or transliterations) in the past three decades. This has important implications for future research on Chinese loanwords. If a comparable degree of acceptance can be observed for new Chinese loanwords and these loanwords are found in various semantic fields, as Yang (2009) suggested, it will lend further support to the claims made by China English scholars.

The aforementioned dictionary-based analyses can effectively test the membership of Chinese borrowings in general English; however, for loanwords and/or loan translations that are not as well-established in this respect, the use of corpus will be a more powerful diagnostic tool of their current status in general English. It also has more potential of bringing to light the initial stage of borrowing (innovation) as well as the later stage, i.e. the propagation of new features or innovations and making discoveries which are otherwise not possible. A recent corpus-based study was carried out by Gao (2001) who investigated the acculturation of English at the lexical level in his analysis of data from *Beijing Review*, an English magazine in China and from *China Daily*, a widely-read China-based English newspaper. Gao presented a number of instances of loan translation and semantic shift, used in a wide range of domains (politics, economy, customs and arts, etc.), and in doing so, he aligned himself with scholars arguing for the legitimacy of a Chinese variety of English (p. 85). Two things were especially noteworthy in Gao's paper: loanwords were excluded from his analysis and no reasons were offered for this decision. One possibility is that he didn't find any loanwords other than the well-established ones in his data, as Yang (2005) speculated. However, this was still surprising, according to Yang (2005), since he found more loanwords than loan translations in his data from similar sources. The second point I would like to make concerns the nature of Gao's study: it is essentially a short-period synchronic analysis of the innovations as defined by Croft (2000), since the paper focused on the lexical items as they were used at certain points of time over a short period and provided only one instance of each lexical innovation, with no reference to the frequency, distribution and diffusion of such innovations, except for a brief mention of such expressions and phrases not being part of the linguistic repertoire of native speakers (Gao, 2001, p. 75) and an equally short comment on how some expressions belonging to different certain fields were used with varying degrees of frequency at different certain historical periods, which was not supported by any statistics. In this respect, his study fell short of its potential. Yang's (2005) study based on *China Daily* and *Beijing Weekend* was more carefully designed and provided more details about the frequency and use of the borrowings identified. He also offered some tentative explanations about the context in which borrowing might occur in the form of loanwords instead of loan translations or vice versa. I will come back to this in the discussion of my own data. In summary, Yang's (2005) study fills a gap in the research on Chinese borrowings, but still fails to dig into the process of borrowing due to the type of data he used—the publications of two China-based news media at a short period of time.

Taken together, the dictionary- or data-based studies mentioned in this section have given us new understandings on lexical borrowing that is of interest in this paper. Nevertheless, the dictionary-based studies in a whole haven't been able to shed much light on how lexical items—especially the newer ones, have been working their way into the language. The aforementioned corpus-based studies failed to provide a clear picture of where the borrowings examined stand in a broader context, with respect to the more established counterparts. The dictionary-based and corpus-based research should complement each other so that a fuller picture and a better understanding of the status and significance of certain borrowings can be obtained. This paper will make such an endeavor by scrutinizing the use of two loanwords in different contexts from the perspective of innovation and propagation; a large online corpus of English used in global media, along with online up-to-date dictionaries will be used for data collection and analysis. The hopes are that patterns and evidence of linguistic penetration and/or acceptance will be found for these items at the regional or global scale.

Data collection and analysis

The corpora

In order to investigate the innovation and propagation of some new features in operation at the level of lexis in an arguably emerging variety of English, I decided to focus on analyzing a couple of new borrowings that haven't established themselves as members of the general English; therefore, I discarded all the items listed in Cannon (1987, 1988), Gao (2001) and Yang (2009) and most of the other studies mentioned in the previous sections. Inspired by Yang (2005), I adopted a corpus-based approach to this topic and based my selection of target items, collection and analysis of the target items on available databases. I used several sources of data in this paper: Factiva, the official website of China Daily, China Daily in print/in PDF, and six (6) online dictionaries.

Factiva is a large Internet-based database of news, covering more than 31,000 sources of news from more than 200 countries in 26 languages, "including nearly 600 continuously updated newswires (including Dow Jones, Reuters and the Associated Press). More than 2,300 sources are available on or before the date of publication." (Factiva User's Guide). The data—or news articles in which the target items appeared were obtained from this database. I will elaborate on the process of data collection in the corresponding section.

In addition to Factiva, I also used the official website of China Daily for an informal survey prior to the collection of data, with a view of choosing appropriate items for this paper. China Daily is one of the most widely-read English newspapers in China, and offers free PDF download of a large proportion of the newspaper in print. This would be useful when entries of the news articles I obtained from Factiva need to be compared with newspapers in print/PDF. I will discuss this in the section on data analysis. The third type of corpus used in this paper was the online version of 6 English dictionaries, listed in Appendix 1. The choice of dictionaries was made on the basis of logistical and financial considerations. Furthermore, the websites of the six dictionaries supposedly represent their newer if not the latest version of their desk counterparts, as stated in the introduction in one of the websites. Thus, they should provide accurate information about the acceptance of newer loanwords in general English. The idea to include three American dictionaries and three British ones was adapted from Cannon (1988) and Yang (2009).

Selection of target items

The selection of the target items has been based on the following considerations: first, the items should have been in use with a sufficient degree of frequency over a relatively long period of time in written communication, i.e. newspapers in this case, to serve the purpose of this paper as described above. Secondly, they shouldn't have been identified by most of the studies (e.g. Cannon, 1988; Moody, 1996; Yang, 2009) as well-established "naturalized" borrowings. Ideally, these items should also reflect mainstream cultural and social situations, events and/or phenomena in contemporary China. To determine which target items could satisfy the prerequisites, an informal survey was carried out on the loanwords that have been found with relatively high frequency in China Daily. The survey was conducted by means of browsing the newspapers available in China Daily's official website and asking frequent readers of the newspaper questions about such loanwords. The borrowings selected for analysis in this paper were two loanwords *hukou* and *guanxi*. The choice of loanwords was the result of a number of factors—linguistic and non-linguistic. A later search in the newspaper's website corroborated the results of the survey, showing that these two terms have been used with high frequency over a long span of time. The definitions of these two items are given in the following section.

***Hukou* and *guanxi*: Denotations and connotations**

The definition of the Chinese word for "hukou" (户口) given by Online Xinhua Dictionary, one of the most authoritative Chinese dictionaries, is short and uninformative: "Hukou has two senses:

(1) one household and (2) the brochure on which the civil administration of the local government keeps record the registration of the residents or the ID of local residents. A more detailed explanation is found on the Wikipedia page for Hukou system². Although a similar household registration system exists in some other Asian countries, and there might be some parallel the system in China and that in other Western countries, the hukou system has a unique history, distinct features and a range of social, economic and political implications in contemporary China. It is tied up with education, health care, employment, and housing in China. The dichotomy between rural and urban household registrations and the superiority of hukou of certain big cities have posed a major obstacle to upward mobility in modern China. The hukou system has long been and will still be a topic of much debate and controversy. It is thus not surprising that the word has been borrowed into English by means of transliteration, instead of direct translation, though it has often found to be accompanied by some explanation or definition in English newspapers. I will come back to this point in the analysis of data.

Similar to “hukou”, the word *guanxi* (关系) is also unique in terms of its cultural connotation. It should be pointed out that the original Chinese word-关系 has multiple senses, many of which are often translated into English using their (near)equivalents (e.g. relationship, significance) and that the loanword *guanxi* (transliterated from the original word) expresses only one of the original senses. This sense is often translated as “connection” or “relationship”. According to Wikipedia, “(*g*)*uanxi* describes the rudimentary dynamic in personalized networks of influence (which can be best described as the relationships individuals cultivate with other individuals) and is a central idea in Chinese society. In Western media, the pinyin romanization of this Chinese word is becoming more widely used instead of the two common translations—‘connections’ and ‘relationships’—as neither of those terms sufficiently reflects the wide cultural implications that *guanxi* describes”³. The explanation by Wikipedia can partially account for why this (sense of the) word has been borrowed into English via transliteration instead of via loan translation or semantic shift. The claims made on the Wikipedia page will be examined against empirical data in this paper.

Data collection

The data was collected differently from the three sources. The examination of the two target items in the 6 dictionaries was quick and straightforward. They were keyed into the search column in each online dictionary and the entries (or a lack thereof) were recorded separately. The results are shown in Chart 1 and Chart 2. The collection of data from Factiva database involved many more steps and trials, which would be described in greater detail.

In order to locate the target items in different news articles in Factiva, each of the two items, i.e. *hukou* and *guanxi*, was used as the key word in the initial search⁴, with the option of searching for “All dates”. This produced a large number of entries/articles for each word (5,740 for “hukou” and 5,746 for “guanxi”), but many turned out to be irrelevant. Such result was not totally unexpected, as the spelling of the loanwords was identical to that of many other Chinese characters transcribed in a Roman Alphabet. Since such Chinese transliterations have lost their tonal marking, the one feature distinguishing Chinese words written in Pinyin, it is difficult, if not impossible for the search engine to separate the expected results from other irrelevant items, with only one key word provided as the input.

Different combinations of key words were used in an attempt to locate the target items and eventually, I chose a relatively effective way of getting only the right entries, albeit with its own drawbacks. For example, I would use the target item “hukou” as well as several key words that might

² According to Wikipedia, (“Hukou”, 2018), hukou refers to “a record in the system of household registration in mainland China and Taiwan.....A hukou can also refer to a family register in many contexts since the household registration record..... is issued per family, and usually includes the births, deaths, marriages, divorces, and moves, of all members in the family”.

³ The definition and explanation of *guanxi* in Wikipedia (“Guanxi”, 2018) was retrieved on March 31, 2018.

⁴ The search was performed in December, 2013.

be found in the same articles, made somewhat ambiguous and flexible and connected by “or” (e.g. (hukou) and (residen* or regist* or ID). This gave me 4,188 entries. It is still possible that the results included articles that didn’t match the topic, but a quick look at the entries confirmed the superiority of this set of search command over other methods deployed. Another potential problem that might be caused by this kind of search command is that some articles might have been left out that satisfied the requirements, simply because a different key word was used in the same article as “hukou” was. Therefore, a comprehensive search and manual elimination of the mismatch in the results are needed in future research. The same method was applied for “guanxi”, using the most relevant words in combination with the target word (e.g. (guanxi) and (relationship or connection or network*), with a total of 2,611 entries listed. The following section discusses the data with respect to the two notions crucial to the analysis: innovation and propagation.

Data analysis

Examining innovation

Haugen (1950) stated that a clear understanding of the process of borrowing could only be obtained by isolating the initial leap of a pattern from one language to another (p. 212). The database available for this paper has provided a unique chance for looking into this initial stage at which a borrowing appears as an innovation. The question is: how can one identify an instance of borrowing as an innovation? There are different ways to approach this problem. First, Croft’s (2000) definition of innovation (i.e. altered replication or the creation of new forms) indicates that it is a process involving some change in feature. Drawing on Croft (2000), I maintain that the change could occur to different linguistic aspects of the target items, although phonology was the focus in Croft’s (2000) discussion. It naturally follows that these items (called “replicates” by Croft, 2000, p. 3) become distinguishable from the original. Not confusing Croft’s definition of innovation with Haugen’s conceptualization of the same term, I would treat an innovation as actualization of the process in a linguistic unit. The process such items undergo is innovation, and what results from the process are innovations, which might in turn participate in the propagation process until they become established in the new language/variety.

In the case of the current investigation, the change affects the phonetics, and orthography-phonology and orthography-semantic interface of the lexical items borrowed from Chinese to English. There are also other contextual features the borrowed items don’t share with the original forms in the donor language and such features contribute to the identification of such items as innovations in the borrowing language. Another dimension of innovation that should also be addressed in the analysis is the functions and/or motivations for such process. In short, innovation as a process can be investigated from multiple angles and this will be done for the two loanwords in their respective contexts, using the corpora from the above-mentioned sources.

To start with, both loanwords have been affected at the representation level to an equal degree: the diacritics marking the tones of each of the constituent characters in *guanxi* and *hukou*, both written in Pinyin, were abandoned in the borrowing process of these words from Chinese into English. The orthographic change led to the loss of different kinds of linguistic information, including phonetic and phonological information needed for the processing of the sounds, semantic information the retrieval of which depends largely on the reading of characters and to a lesser extent, Pinyin transliterations with tonal diacritics. This type of change means that the two loanwords are distinct from the original forms and are mostly likely to be treated as erroneous or foreignized representations of some Chinese characters if they stand alone instead of appearing in an English text.

A second type of change that distinguishes the two loanwords from their original forms is lexico-grammatical in nature. These two words have been used, to a varying degree, as regular English nouns in the data, preceded by an English article (“a”) or marked for plurality (by the suffix “-s”). This is not surprising given the morphological differences between Chinese and English in the

marking of plurality and the use of quantifiers. It should be noted, however, the two loanwords exhibited some difference in the frequency in which they were marked for plurality. For instance, *hukou* is marked for plurality in 69 search entries or news articles (from both Chinese and Western media) found in Factiva. Nonetheless, there is only 1 entry from a UK-based trade magazine in which *guanxi* bears the plural marker (-s). The gap between the two loanwords can be explained by the difference in the denotations for the two loanwords. The original form for *hukou* with the sense of residence ID or permit is a countable noun and can be immediately preceded by a sequence of a number + a classifier (e.g. 两个户口 or two Classifier + *hukou*), while the original form for *guanxi* with the sense of personal networking is usually considered an uncountable noun and cannot be modified by a specific number in Chinese, though it can sometimes take a modifying quantifier. The use of *hukou* as a regular countable noun in English is evidenced by the search results showing that the word is modified by “a” that is used to mark a countable noun in the singular form.

It has been showed that the two loanwords exhibit different patterns from their original forms in the donor language, primarily due to the typological differences between the donor language and the borrowing language. What is more worth noting about these two loanwords is that they are separated from other items in the borrowing language in many ways and stand out as innovations in the borrowing language, working to make their way into the linguistic repertoire of the speakers of the “new” language (variety). The following section looks at how the loanwords behave or are treated by speakers of that language (variety) differently.

One way the two loanwords differ from other English words in news articles concerns their “appearance”, for lack of a better word. This includes the font style for the loanwords and the quotation marks that accompany them. The quotation marks were noticed when I was browsing some of the news articles. A systematic investigation of this approach to the loanwords indifferent newspapers and of the consistency of such an approach in the same newspaper(s) across time wasn’t conducted given the scope of this paper. Instead, the investigation was restricted to a couple of newspapers published over a short period of time: I first identified the newspapers with articles in which the loanwords were used within quotations marks (e.g. Xinhua News) and then looked at five of the latest entries and five of the oldest entries containing the target loanwords located by the kind of search command discussed in the section on data collection. The findings showed that the kind of treatment given to the loanwords by each of these newspapers was fairly consistent across time. The discovery of a different font used for these loanwords was made possible by the availability of the PDF version of China Daily newspapers. The decision to restrict my search to a small number of newspapers led to the examination of the news articles in the form of PDF that corresponded to the articles I found by searching on the website of China Daily. It turned out that the loanwords appeared in italics in all the articles in PDF format I was able to download from China Daily. In contrast, in the online issues or versions of newspaper or magazine articles, there were no differences in font style between the two loanwords and other words in the same articles. It remains to be found out whether the same strategy was adopted in the articles in print for other editions of China Daily and in other newspapers. Apart from these features that separate the loanwords from the other words in the media texts, some other common strategies were found in different newspapers in presenting these loanwords in their articles. For example, in a number of newspapers, Chinese (e.g. South China Morning Post) or otherwise (e.g. The Wall Street Journal, The Toronto Star), the first mention of one of the two loanwords was almost always accompanied by a translation, explanation or definition of the term, usually immediately following the introduction of the loanword, regardless of the way the loanword per se was presented, i.e. with or without quotation marks. The data was examined in terms of the consistency of this kind of strategy and the results mirrored those for the check on the use of quotation marks.

Taken together, the evidence supports the claim about the distinctiveness of the two loanwords as compared to their original forms in the donor language and to other words in the borrowing language. These features help define the loanwords as “innovations”. On the surface, the loanwords seem a misfit on both sides, and the justifications seem lacking for the use of such

innovations, since alternatives exist to express the meanings they convey. A legitimate question to ask in this situation is what are the motivation and/or functions of this kind of innovation. A tentative answer was offered in Croft (2000), as summarized by Von Rooy (2010). Innovation “is a linguistic-functional process.....driven by forces such as expressiveness, economy and an attempt to avoid misunderstanding” (Von Rooy, 2010, p. 9). However, this kind of theory wasn’t completely borne out by the findings made hitherto. There is no denying that the use of loanwords might have been based on the intention of avoiding misunderstanding, since the rich social and cultural connotations these loanwords are unlikely to be effectively conveyed by a translation or (near) equivalent in the language. This might also be considered expressive (creative); however, the use of the loanwords along with some form of definition or explanation didn’t seem to conform to the economy principle as discussed in Croft (2000, p. 75), since extra effort was needed to avoid misunderstanding by providing any definition or explanation to the loanwords used in the text. There must be another layer of consideration for the creation of a new form, in addition to the ones proposed in Croft’s (2000) model of language change. In fact, in his mention of borrowing of content words, Croft cited the hypothesis made by Thomason & Kaufman (1988) about the motivation for such borrowing. Non-basic words are borrowed is “for cultural and functional...reasons” (Thomason & Kaufman, 1988, cited in Croft, 2000, p. 205). This complements Croft’s original model and provides a more satisfactory account for the innovation investigated in this paper. Without going into too much detail about the social dimension of the motivation, I would argue that the writers of the news articles who used each of the two loanwords and provided explanations/definitions for them have done so for social reasons (e.g. asserting their social identity or showing alignment to their readers). This means that, different from other kinds of innovation discussed in Croft, this kind of innovation concerning the two loanwords overlaps with the propagation process in terms of motivation, contrary to the claim made by Croft about the separateness of motivations for innovation and propagation⁵. Propagation of the loanwords is examined in the next section.

Propagation: the other side of the coin

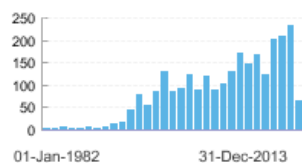
Within the scope of this paper, only the crude findings of a preliminary analysis of the data were presented here. The preliminary analysis approached the question about the propagation of the loanwords by examining three dimensions: their regional distributions, the yearly distributions and the acceptance of the loanwords into general English. The two sets of data collected from Factiva, with 4,188 entries for hukou and 2,611 entries for guanxi were used to address the first two dimensions. The six online dictionaries were used to investigate the third dimension. The findings concerning each of the three dimensions were presented below.

With respect to the regional distributions, the data provided by Factiva showed that the loanword guanxi was found most often in China, followed by Asia (referring to other Asian countries), Emerging Marketing Countries, Eastern Asia, the United States, Hong Kong, Beijing, Singapore, UK, Australia. For the word loanword, hukou, the top ten regions included China, Beijing, Guangdong, Emerging Market Countries, Asia, Eastern Asia, Hong Kong, Shenzhen, Shanghai, the U.S. For both loanwords, the U.S. and Hong Kong were ranked high in the list probably because China Daily has a U.S. edition and Hong Kong edition. However, the data did provide some evidence of the spread of the words at the national and broader Asian level.

As regards the yearly distributions of the loanwords in all media and all regions, two charts were provided by Factiva, as shown below:

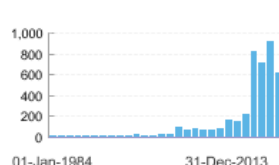
⁵ According to Croft (2000), while functional considerations such as expressiveness, economy, and avoidance of misunderstanding motivate the process of innovation, the process of propagation is motivated by a set of social reasons (e.g. accommodation, acts of identity) (p79).

Chart 1 Guanxi



Distribution: Yearly

Chart 2: Hukou



Distribution: Yearly

It is clear from the charts that two loanwords didn't enter the English media until relatively recently. Guanxi appeared slightly earlier and the use of this loanword showed a steady increase over time and only dropped in the year of 2013. In sharp contrast, hukou had a negligible presence in the 1980's and 1990's and started to gain some more recognition towards the end of the last century—the period from 2009 to 2013 in particular, saw a spike in the use of this loanword. Although the frequency of use of the loanwords could have been significantly affected by the social, economic and political situations in the regional, national and international contexts, the patterns observed in the charts could find a good explanation grounded on the sociolinguistics and/or social realities of the two loanwords. The concept *guanxi* has always been part of the social landscape in China and the discussion of this notion bears on various aspects of Chinese people's social and economic life. The adoption of this innovation in China-based media was a natural process, and its spread is also predictable on the basis of the increasing level of China's influence in the globe. By comparison, it was only recently that the issues related to hukou have attracted public attention, because the hukou system had remained constant in ancient China, until the new hukou policies were implemented in New China, and the consequences of the policies were more clearly felt by the younger and older generation whose life has been most affected by the process of urbanization.

The third issue addressed in this paper concerns the status of the two loanwords in general English. This was examined by checking their presence or the lack thereof in six online dictionaries⁶. The results of the search were as follows. On the British side, there was 0 entry for both loanwords in Longman Dictionary of Contemporary English; there was 0 entry for *hukou*, and 1 entry for *guanxi* in Oxford English Dictionary (3 ed). In Collins English Dictionary, *guanxi* had 1 entry while *hukou* has none. Among the American dictionaries, Dictionary.com has 1 entry for *guanxi* but not for *hukou* while the other two dictionaries had 0 entry for both *guanxi* and *hukou*. Apparently, *guanxi* has gained some international recognition (found in three of the six dictionaries) whereas *hukou* still has a long way to go (found in none of the six dictionaries). It is hypothesized that if the debate about China's household registration system continues, *hukou* will have a good chance of emulating *guanxi*, although it is not clear whether a loanword related to a temporary social phenomenon can compete with one that is part of the underlying social structure.

Conclusion

Research on lexical borrowing has given us much insight into how a language or language variety develops in the context of international contact. In effect, many languages around the world, English in particular, have benefited from the process of borrowing and have been enriched by borrowings from other languages.

With a view of investigating the innovation and propagation processes of two loanwords, this paper has focused on a synchronic analysis of corpora from different sources, and has only attempted a preliminary diachronic analysis and future research is necessary to address such issues by conducting a more systematic and careful diachronic analysis, as well as a synchronic one, so that

⁶ The search of these two loanwords in the six dictionaries was performed twice, once in December, 2013 and the second time in March, 2018.

more insight can be gained into how new vocabulary items enter and consequently contribute to the establishment of a new language variety.

References

- Bian, F., Zhang, Y. & Shi, Y. 卞凤莲, 张艳莲, & 石颖. (2007). 从英汉互借词看中国的改革开放. *商场现代*, 14, 397-98.
- Benson, P. (2000). Hong Kong words: variation and context. *World Englishes*, 19(3), 373-380.
- Cannon, G. (1987). Dimensions of Chinese borrowings in English. *Journal of English Linguistics*, 20(2), 200-206.
- Cannon, G. (1988). Chinese borrowings in English. *American Speech*, 63(1), 3-33.
- Chen, X. 陈先锋. (2007). 英语的中国本土化倾向研究. *安阳工学院学报*, 3.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. Pearson Education.
- Chen, L. & Zhou, L. 陈俐, & 周烈红. (2012). 英语中的汉语音译借词研究. *湖北经济学院学报: 人文社会科学版*, 9(2), 140-142.
- Deterding, D. (2007). *Singapore English*. Edinburgh University Press.
- Du, Mi. 杜明甫. (1999). 英语中的汉语外来语探析. *河南财政税务高等专科学校学报*, 4, 49-51.
- Du, R. & Jiang, Y. 杜瑞清, & 姜亚军. (2001). 近二十年“中国英语”研究述评. *外语教学与研究*, 1.
- Eaves, M. (2011). English, Chinglish or China English?. *English Today*, 27(04), 64-70.
- Factiva User's Guide - Mason School of Business. Accessed on April, 25, 2013. Available through <https://mason.wm.edu/about/library/documents/factiva.pdf>.
- Gao, L. (2001). The lexical acculturation of English in the Chinese context. *Studies in the Linguistics Sciences*, 31, 73-88.
- Ge, C. 葛传棻. (1980). 漫谈汉译英问题 [J]. *翻译通讯*, (20).
- Görlach, M. (1998). Varieties of English world-wide: where we stand. *Links & Letters*, 5, 13-36.
- Guanxi. (March 31, 2018). In Wikipedia. Retrieved March 31, 2018, from <https://en.wikipedia.org/wiki/Guanxi>
- Haugen, E. (1950). The analysis of linguistic borrowing. *Language*, 26(2), 210-231.
- He, D., & Li, D. (2009). Language attitudes and linguistic features in the 'China English' debate1. *World Englishes*, 28(1), 70-89.
- Hukou system. (March 31, 2018). In Wikipedia. Retrieved March 31, 2018, from https://en.wikipedia.org/wiki/Hukou_system
- Jia, G. & Xiang, M. 贾冠杰, & 向明友. (1997). 为中国英语一辩. *外语与外语教学*, 5(11).
- Jin, H. 金惠康. (2003). 中国英语与中式英语讨论. *广东技术师范学院学报*, 5, 66-70.
- Kachru, B. B. (1983). *The Indianization of English: The English Language in India*. Oxford: Oxford University Press.
- Li, C. 黎昌抱. (2001). 英汉外来词对比研究. *外语教学*, 5, 92-96.
- Li, Z. 李宗奎. (2009). 从英语中的汉语借词看中国文化的输出. *山西广播电视大学学报*, 5, 029.
- Li, W. 李文中. (1993). 中国英语与中国式英语. *外语教学与研究*, 4(1).
- Li, X. 李小飞. (2004). 从英汉互借词看中外文化渗透. *山东外语教学*, 1, 47-49.
- Mencken, H. L. (1919/1980). The two streams of English. *The American Language*. 4th edition. H.L. Mencken, pp. 3-28. New York: Alfred A. Knopf.
- Moody, Andrew J. (1996) Transmission language and source languages of Chinese borrowings in English. *American Speech* 71(4), 405-20.
- Moody, J, A. & M, Y. 安德鲁, & 马睿颖. 英语中汉语借词的译介语与源出语 (下). *福建论坛: 社科教育版*, (4), 141-144.
- Richards, J. C. (1979). Rhetorical and communicative styles in the new varieties of English. *Language Learning*, 29(1), 1-25.
- Romaine, S. (1995). *Bilingualism* (2nd edn). Malden, MA: Blackwell Publishers.

- Tao, H. 陶红梅. (2011). 从英语中的汉语借词看语言的包容性. *江淮论坛*, (001), 188-191.
- Van Rooy, B. (2010). Social and linguistic perspectives on variability in world Englishes. *World Englishes*, 29(1), 3-20.
- Wan, P. 万鹏杰. (2005). 中国英语与中式英语之比较. *上海翻译*, 2, 012.
- Wang, J. 王建国. (2009). 中国英语研究存在的问题. *解放军外国语学院学报*, (006), 14-19.
- Wang, L. 王莉. (2011). 中国英语和中式英语. *当代教育理论与实践*, 3(12), 129-130.
- Wang, R. 汪榕培. (1991). 中国英语是客观存在的. *解放军外国语学院学报*, (1), 1-8, 56.
- Wang, R. & Chang, Y. 汪榕培, & 常骏跃. (2001). 英语词汇中汉语借词的来源, *四川外语学院学报*, 4, 70-73.
- Wu, Z. 吴祝霞. (2009). 从英语中汉语借词看中国英语的发展趋势. *漯河职业技术学院学报*, 8(1), 99-101.
- Xiong, W. 熊文华. (1996). 汉语和英语中的借词. *语言教学与研究*, 2(3), 126-140.
- Yang, J. (2005). Lexical innovations in China English. *World Englishes*, 24(4), 425-436.
- Yang, J. (2009). Chinese borrowings in English. *World Englishes*, 28(1), 90-106.
- Yajun, J. (2003). English as a Chinese language. *English Today*, 19(2), 3-8.
- Yuan, B. & Lu, C. 袁斌业, & 卢澄. (2003). 得体的中式英语:“送去主义”中不可或缺的部分. *广西师范学院学报(哲学社会科学版)*, 24(3).
- Zhang, J. 张婕. (2008). 英语中的汉语借词类别及借用方式. *河南工业大学学报(社会科学版)*, 4(4), 53-55.
- Zhou, Y. 周跃西. (2011). 英语向汉语传统色彩文化借词考释. *宁波大学学报(人文科学版)*, 5, 018.
- Zhu, W. & Liu, T. 朱万忠, & 刘婷. (2008). 中国英语和中式英语的对比研究. *高等建筑教育*, 17(1), 78-80.

Appendix 1: Six online dictionaries used for analysis

Country of publication/origin	Name	Website
U.S.	Fifth Edition of The American Heritage Dictionary of the English Language	http://www.ahdictionary.com/
	Webster international dictionary Merriam-Webster online	http://www.merriam-webster.com/
	Dictionary.com	http://dictionary.reference.com/
Britain	Longman Dictionary of Contemporary English	http://www.ldoceonline.com/
	Oxford English Dictionary	http://www.oed.com/
	Collins English Dictionary	http://www.collinsdictionary.com/dictionary/english

