

APTARA	WENG	weng12134	Dispatch: March 28, 2015	CE: N/A
	Journal	MSP No.	No. of pages: 1	PE: Jessie

World Englishes, 2015
doi: 10.1111/weng.12134

0883-2919

The grammatical features of English in a Chinese Internet discussion forum

Q1

HAIYANG AI* AND XIAOYE YOU**

ABSTRACT: Recent research in world Englishes shows that the lexis-grammar interface, or lexicogrammar, constitutes an important area for investigating structural nativisation of local varieties of English (Schilk 2011). While most studies on lexicogrammar have focused on Inner Circle and Outer Circle varieties, the study presented in this paper focuses on an Expanding Circle variety, China English, and explores some of its patterns of structural nativisation. Utilising large-scale corpus data collected from an online discussion forum, this study examines several locally emergent linguistic patterns in China English, including new ditransitive verbs, verb-complementation, and collocation. The results suggest that there exist certain associations between specific lexical items and grammatical constructions in this local variety. It is argued in this paper that these lexicogrammatical features can be considered as concrete instantiations of structural nativisation in China English. Methodologically, this study demonstrates that, in addition to the International Corpus of English (ICE) and online newspaper articles, new media sites can also be successfully utilised for research on localised varieties of English.

INTRODUCTION

Studies on the nativisation of the English language into diverse local English varieties have received frequent attention in the field of world Englishes (see Pandharipande 1987; Schneider 2004, 2007). An important area of research on nativisation processes has been focused on structural nativisation, which is generally understood as the ‘emergence of locally characteristic linguistic patterns’ (Schneider 2007: 5–6). Recent studies on structural nativisation—particularly those studies using corpus linguistics methodologies—show that structural nativisation of English varieties can be productively explored through what is known as lexicogrammar (Mukherjee 2009; Schilk 2011; Hundt & Gut 2012). Recent studies using large corpora (e.g. the International Corpus of English, or ICE) have also investigated structural nativisation, focusing on numerous lexicogrammatical structures in a wide variety of Englishes. Many of these recent studies on nativised English varieties also include less-studied Inner Circle varieties such as New Zealand English (NZE), Irish English (IrE), Australian English (AusE), and Canadian English (CanE); as well as new Englishes, such as Bahamian English (BahE), Jamaican English (JamE), Fijian English (FijiE), Trinidad and Tobago English (TTE), among others (see Hundt & Gut 2012).

The lexis-grammar interface has so far been most extensively explored for evidence of structural nativisation in Indian English. For example, Mukherjee (2009) studied structural nativisation of present-day English in India by focusing on new locally emerging

*Department of Applied Linguistics, Pennsylvania State University, 304 Sparks Building, University Park, PA 16802, USA. E-mail: hua126@psu.edu

**Faculty of English Language and Culture, Guangdong University of Foreign Studies, Guangzhou, P. R. China; Department of English, Pennsylvania State University, 118 Burrowes Building, University Park, PA, 16802, USA. E-mail: xuy10@psu.edu

forms, including collocations, new prepositional verbs, new ditransitive verbs, and verb-complementational patterns. Schilk (2011) exploited ICE-India, ICE-GB, and the Times of India Corpus to analyse structural nativisation by focusing on collocation and verb-complementational profiles of three focal ditransitive verbs (i.e. *give*, *send* and *offer*). Similar studies on lexicogrammar in Indian English at the level of verb-complementation have been conducted by Mukherjee and Hoffmann (2006) and Schilk et al. (2012). Collectively, these studies highlight the centrality of the lexis-grammar interface in exploring structural nativisation of nativised English varieties.

While much has been learned about structural nativisation in Indian and other Inner Circle and Outer Circle varieties of English, relatively little is known about locally emergent characteristic linguistic patterns in Expanding Circle varieties, such as China English. As English is increasingly used in China, Chinese speakers are also appropriating the language and using it according to their needs and customs. To date, a large Chinese population is becoming English-literate, and English has gained a certain institutional status thanks to the nearly universal requirement of teaching and learning of English at almost every level of the Chinese education system (You 2011). While a good deal of research has been conducted on various aspects of China English (e.g. Bolton 2003; You 2008, 2010, 2011; Xu 2010), few studies have approached China English from the perspective of structural nativisation. This study appears to be one of the first attempts to investigate structural nativisation in China English at the level of lexicogrammar. Specifically, this study explores characteristic lexicogrammatical patterns—namely new ditransitive verbs, verb-complementation, and collocations—that have emerged in China English by examining large-scale corpus data. The results show that there exist certain associations between specific lexical items and grammatical constructions. These associations, or co-selections between lexis and grammar, it can be argued, constitute concrete instantiations of structural nativisation in China English. In terms of methodology, this study also demonstrates that online discussion forums can also be utilised for fruitful corpus analyses when studying local varieties of English.

THE INTERFACE BETWEEN LEXIS AND GRAMMAR

Traditionally, lexis and grammar have been treated along separate lines both in linguistic theory and language pedagogy. The link between lexis and grammar was first considered by British linguist Firth (1957) from the perspective of syntagmatic relations and paradigmatic choices. The lexis-grammar interface was later further developed by Halliday (1966) from the theoretical perspective of Systemic Functional Linguistics. Recent studies in corpus linguistics and similar, related fields have accumulated a substantial amount of empirical evidence for the understanding that lexis and grammar are in fact ‘intimately intertwined’ (Biber 2012: 10). Lexicogrammar, according to Biber et al. (1998: 84), refers to the association ‘between words and their grammatical environments, or between grammatical structures and their lexical environments.’ A good deal of corpus linguistics research has explored corpus resources to describe systematic associations between particular lexical items and target grammatical constructions. For example, Kennedy (1998: 123) showed that some of the most common verbs in English appear most often in the simple present tense (e.g. *think*, *know*, *want*, *see*, *mean*), whereas others occur more frequently in the simple past tense (e.g. *said*, *came*, *took*). Biber et al. (1998: 102) found that the nearly equivalent grammatical constructions *that*-clauses and *to*-clauses can be distinguished between their

differing lexical associations. For example, in everyday conversation, the most common verbs controlling a *that*-clause are *think*, *say*, and *know* while the most common verbs controlling a *to*-clause are *want* and *try*. In a more recent study, Römer (2005) provided a detailed description of the associations between verbs and the progressive aspect in English. Taken together, these studies illustrate that in actual language use, there are definite and specific associations between lexical items and grammatical constructions.

While the lexis-grammar interface can be examined for a range of linguistic structures (e.g. Hundt & Gut 2012), scholars have approached the lexis-grammar interface by mainly focusing on new ditransitive verbs, verb-complementation, and collocation in finding evidence for structural nativisation in local varieties of English (e.g. Mukherjee & Schilk 2008; Mukherjee 2009; Schilk 2011; Schilk et al. 2012). Briefly, new ditransitive verbs refer to those verbs that cannot be complemented with two object noun phrases in Inner Circle varieties (e.g. British English, American English), but can nevertheless be used in this basic ditransitive pattern in other varieties of English (cf. Mukherjee & Hoffmann 2006; Mukherjee 2009; see also Ersson & Shaw 2003). For instance, the verb *furnish* can be used as a ditransitive in Indian English: 'Can you furnish me Dr. Shastri's address?' Such innovative use of these verbs is suggested to be based on logical and plausible analogies between the meanings and complementation patterns of existing typical templates (e.g. *give*) and those of the new forms (e.g. *furnish*) (Mukherjee 2009: 125–126). Because most of these new ditransitive verbs are relatively rare, Mukherjee (2009) recommends using larger corpora.

Verb-complementation is another area reflective of the interdependence of lexical choices and grammatical patterns. The complementation patterns of a verb not only specify the number of argument roles a verb invokes but also stipulate how its various constituents are related. A major area of study for verb-complementation patterns is the ditransitive construction (e.g. *give somebody something*) and its dative alternation (e.g. *give something to somebody*). Depending on the configuration of various constituents, structurally related patterns can be derived to account for such variations as inversion of object, passive voice, relative clauses, and so on. Thus, a productive area of study in verb-complementation is the differing preference among interlocutors between ditransitive verbs and their various complementational patterns. Typically, this line of research tends to focus on semantically related groups of words (e.g. *convey*, *submit*, *supply*) that describes a 'transfer' event in the transfer-caused-motion (TCM) construction (cf. Goldberg 1995). In this present study, the focus is on the verb-complementational patterns of the prototypical ditransitive verb *give* in our China English corpus.

Collocation has also been recognised as a lexicogrammatical feature in structural nativisation of local varieties of English (Mukherjee 2009; Schilk 2011). Collocation refers to 'the relationship of habitual co-occurrence between words' (Stubbs 1995: 23). The notion of co-occurrence is also explained in Sinclair's (1991: 112) idiom principle, as 'many uses of words and phrases show a tendency to co-occur with certain grammatical choices.' These two definitions of collocation suggest that association or co-selection occurs not only between words, but also between words and grammatical constructions. In the present study, these definitions are used for exploring the collocation behaviors of the lexical verb *give* in relation to its various grammatical environments, including verb-complementation patterns and object-slot collocates.

RESEARCH ON LEXICOGRAMMAR IN WORLD ENGLISHES

While lexis and grammar remain frequent objects of study in world Englishes, studies have tended to focus either on lexis (e.g. Granger & Tyson 1996; Ho & Wong 2001; Sand 2004; De Klerk 2005) or on grammar (Bautista 2004, 2008; Wong 2004; Collins 2007). It is only recently that some studies have begun focusing on the link between lexis and grammar, particularly from the perspective of structural nativisation. Mukherjee and Hoffmann (2006), for example, examined the frequency and distribution of ditransitive verbs and their complementational patterns in Indian English. Taking a structural nativisation perspective, Schilk (2011) analyzed collocation and verb-complementational profiles of ditransitive verbs (i.e. *give*, *send*, and *offer*) in Indian English based on large-scale corpora. In a subsequent and similar study, Schilk et al. (2012) compared verb-complementational patterns of three focal ditransitive verbs (i.e. *convey*, *submit*, and *supply*) between Indian English and Sri Lankan English. In addition, Mukherjee (2009) documented and discussed locally emergent forms—collocations, new prepositional verbs, new ditransitive verbs, and verb-complementational patterns—in Indian English using authentic corpus data. Balasubramanian (2009), taking a register variation perspective, examined a range of lexicogrammatical features (e.g. position of *also*, future *will* for simple present tense) in Indian English based on the Corpus of Contemporary Indian English. Also using corpus linguistics methodology, Bautista (2008) examined lexico-grammatical features in Philippine English, and reported that speakers of Philippine English show a tendency for: (1) using singular nouns in one of the structures; (2) omission of articles (e.g. \emptyset majority; such + \emptyset singular noun); (3) omission of indirect object for the verb *assure*; and (4) using the relative pronoun *wherein* as an all-purpose connector. Jung and Min (1999) examined the lexis-grammar interface (e.g. modals for volition and prediction, prepositions for spatial relations) in Korean English based on a corpus of newspapers articles. Notwithstanding the different foci and perspectives, these studies suggest that the lexis-grammar interface—especially verb-complementational and collocation behaviors—can be productively explored for specific instantiations of structural nativisation of local varieties of English.

Despite the surge of studies on the lexis-grammar interface in Inner Circle and Outer Circle varieties of English, little work has been done to understand the interface in Expanding Circle varieties, including China English. Previous research on China English, has tended to focus on: (1) the concept of China English as a local variety of English (e.g. Rong 1991; Li 1993; Gu & Xiang 1997; Du & Jiang 2001; Jiang 2003; Wei & Fei 2003); (2) Chinese speakers' perceptions and attitudes towards the variety (e.g. Hu 2004; Chen & Hu 2006; He & Li 2009; He & Miller 2011); (3) the role of English in China's language curriculum (e.g. Bianco et al. 2009); (4) historical and sociolinguistic descriptions of English in China (e.g. Bolton 2003); (5) rhetorical strategies and literary and multilingual creativity (e.g. Bolton 2002; Kirkpatrick & Xu 2002; You 2008, 2010, 2011; Zhang 2002); and (6) linguistic features of China English (e.g. Xu 2010; Yang 2005). On the whole, while these studies contribute to our understanding of China English from diverse theoretical perspectives, they sometimes lack in-depth analyses of the lexis-grammar interface, which is the focus of the present study.

Among these studies, Xu's (2010) book-length treatment of China English's linguistic features comes closest to the present study in that it has touched on some linguistic structural issues. Using a dataset combining interviews, newspaper articles, short stories,

and questionnaires, Xu examined lexical and syntactic features in China English. At the lexical level, he classified the emerging lexis into Chinese loanwords in English (e.g. *chow mein*, *fengshui*), nativised English words (e.g. *save face*, *migrant workers*), and English words shared with other varieties. At the syntactic level, Xu considered such factors as regional preference, innovation (e.g. simplification, generalisation, complexification), and language transfer in determining characteristic syntactic features in China English. He documented a wide array of syntactic structures uniquely Chinese, including adjacent default tense, null-subject utterances, co-occurrence of connective pairs, subject pronoun copying, yes-no response to tag questions, among others. Overall, Xu has offered a comprehensive analysis of linguistic features that emerge in China English. However, while Xu examined both lexis and syntax in great detail, he has not examined the lexis-grammar interface. In addition, given current trend of exploiting corpus resources at the level of mega-million words, his dataset is relatively small: the written component of the dataset only consists of 20 newspaper articles and 12 short stories. The present study intends to fill this gap by documenting and describing characteristic linguistic patterns that have emerged in China English at the lexis-grammar interface. For the purpose of this study, the focus is on three major areas: verb-complementation, collocation and new ditransitive verbs. We selected these areas mainly because first, verb-complementational and collocation patterns carry constructional meanings and constitute essential components at the structural level of a language (cf. Goldberg 2006); second, their significance has been recognised in studies investigating lexicogrammatical features of local varieties of English in other locations (Mukherjee & Hoffmann 2006; Mukherjee 2009; Schilk 2011; Schilk et al. 2012); and third, they have also emerged clearly as local forms of structural nativisation in the China English dataset presented in this paper. While other linguistic features (e.g. new prepositional verbs) might also be indicative of the associations between lexis and grammar, these three areas were chosen for this study in order to present a succinct analysis of structural nativisation of China English.

With respect to data sources, previous corpus-based studies on local varieties of English have largely relied on the ICE series of corpora (e.g. Bautista 2008; Mukherjee & Hoffmann 2006; Mukherjee 2009; Schilk 2011; Schilk et al. 2012). Some studies have also incorporated corpora resources that were compiled from online newspaper articles (e.g. Jung and Min 1999; Mukherjee & Hoffmann 2006; Balasubramanian 2009; Schilk et al. 2012). Few studies, thus far, have examined English in the new media, including online discussion forums, social networking websites such as Facebook, and microblogging websites such as Twitter. These types of new media data, in our view, are qualitatively different from online newspaper articles because they are produced by numerous general users of English rather than a few journalists, which means that they are less likely to be shaped by editorial interference and that they represent a wider variety of communicative purposes and levels of formality. In the present study, we explore how data collected from an online discussion forum can be used in studying lexicogrammar in world Englishes.

DATA COLLECTION

The data analyzed in this study were collected from an online discussion forum entitled *The 21st Century Community* in November 2011.¹ Focusing on English learning, the forum is offered by 21st Century Newspapers, a popular English learning newspaper group targeting students of all proficiency levels in China. The group has attracted a large number of high

Table 1. The corpus description

Sub-forum name	English Corner
Total number of threads	2,354
Total number of running words (tokens)	7, 157, 364
Total number of different types of words (types)	144, 625

school and college students to its online space, mainly because of the popularity of its print newspapers. Participation in the discussion forum is on a voluntary basis with no access restriction; however, only registered members can post messages. At the time of data collection, this online community had 13 sub-forums for different participants (e.g. teachers, college students, and high school students) with varying interests (e.g. language pedagogy, language skills, test preparation, and opinions, among others).

For the purpose of this article, the focus is on the *English Corner* sub-forum. As the most popular sub-forum in the entire online community, this virtual space, is frequented by high school and college students, graduates, as well as white-collar workers. Some of the most popular topics discussed in this sub-forum include issues related to English language learning, as well as other daily concerns of the forum participants (e.g. work, learning, family issues, making friends, etc.). Given the nature of online discussion forums, noises (e.g. advertisements) were mixed with the threads. To ensure that the threads that were examined were highly relevant to forum participants, it was decided to focus on those with more than twenty follow-up posts. These threads were automatically downloaded using computer programs that the authors have developed.² The corpus building procedure includes: (1) determining the total number of the webpages within the sub-forum; (2) retrieving and parsing each webpage (using python modules *urllib2* and *BeautifulSoup*); (3) extracting and saving the results of each thread as a plain text file; and (4) removing irrelevant elements (e.g. HTML tags, symbols, images, block quotes). In total, the corpus consists of 2,354 threads, totalling over 7 million words (Table 1).

To conduct analyses of the collocation and verb-complementational profiles in China English, we decided to focus on the prototypical ditransitive verb *give*. Using WordSmith Tool 4.0 (Scott 2004), 500 instances of the use of *give* in our online forum corpus were sampled. Each instance of a concordance line containing *give* was manually analyzed and coded according to the classification scheme of complementation patterns of *give* following the framework of Mukherjee and Hoffmann (2006). Instances of *give* being used as part of a phrasal verb (e.g. *give up*, *give in*, *give away*) were excluded, because their semantic meanings are different from that of the basic ditransitive meaning of *give* (see also Schilk 2011).

RESULTS

New ditransitive verbs

Transitivity is a linguistic term that refers to the number of objects a verb requires or takes in a given instance. Traditional grammar makes a binary distinction between transitive verbs and intransitive verbs. Transitive verbs are further differentiated between monotransitive verbs (with only one object) and ditransitive verbs (with two objects). Recent research on structural nativisation of local varieties of English has revealed a new category, that

is new ditransitive verbs, which refers to those verbs that cannot take two noun phrases as objects in Inner Circle varieties of English (e.g. British English, American English), but can nevertheless be used in the basic ditransitive patterns in other varieties of English (Mukherjee 2009; Mukherjee & Hoffmann 2006; see also Ersson & Shaw 2003). In the China English corpus, eight new ditransitive verbs were observed, each of which is exemplified as follows:

- (1) Would it be more helpful if you could **supply** me some oral or listening English website on business? (t555496.txt)³
- (2) I'd appreciate it if you **present** me those websites. (t542153.txt)
- (3) As Sandy **advised** me this forum, thanks for Sandy to let me know you. (t575265.txt)
- (4) I'm afraid I may not be able to come up with the correct answer. I tell her to **inform** me the topic in advance. (t548050.txt)
- (5) Fortunately, we now have a very good math teacher. He **explained** us the complex theories in plain words which make them understandable. (t555496.txt)
- (6) If you love that girl, try your best to protect her, to give her happiness and **provide** her a safe shelter. (t504299.txt)
- (7) We didn't know who would **put** him this delicate question when my friend offered to bell the cat. (t533883.txt)
- (8) This word **reminds** me a saying 'Unity is strength'. (t559814.txt)

Requesting information was found to be a very common pattern in the *English Corner* forum. In (1), the speaker uses the verb *supply* to request information about websites that provide English listening materials for business purposes. In doing so, the speaker uses the verb *supply* in a ditransitive manner (i.e. *supply somebody something*). Similar use has been observed for verbs such as *present*, *advise*, *inform*, *explain*, *provide*, *put*, and *remind* for various purposes, including seeking information, as in (2) and (4), as well as making statements, as in (3) and (8). In all these instances of new ditransitive use, the direct object of the verb is either first or third person pronoun (e.g. *me*, *us*, *her*, *him*), indicating a high level of personal involvement in the activities described by these statements.

One might also take a language-acquisition perspective which would, thus, beg the question as to whether such creative uses are in fact instances of erroneous use, that is one simply leaves out prepositions (e.g. *supply somebody with something*; *advise somebody of something*). Such a view would necessarily essentialise the Chinese variety of English as, in Mukherjee's words (2009: 131), a 'deficient learner variety.' However, it is argued here that this pattern is to be understood in a less prescriptive manner, and, instead from a more descriptive perspective. As Mukherjee (2009: 126) posits, these new ditransitive verbs do not emerge randomly in the discourse, but are in fact based on 'logical and plausible analogies' that speakers draw between the meanings and complementation patterns of established ditransitive verbs and those of the new forms. For example, the new ditransitive verb *provide* resembles the prototypical ditransitive verb *give* in semantic meaning, that is both verbs denote a 'transfer' event with the typical ditransitive meaning of 'X causes Y to receive Z' (cf. Goldberg 1995, 2006). The semantic closeness between these two verbs thus leads to a grammatical resemblance for the verb *supply* to be used in a similar ditransitive fashion (i.e. *give somebody something* → *provide somebody something*). As Mukherjee (2009: 126–127) argues, there exist a 'nativized semantico-structural analogy' between the semantic meaning and syntactic pattern of *give* and other established ditransitive verbs and semantically closely related verbs (e.g. *provide*, *supply*). Such uses of verbs in this

Table 2. New ditransitive verbs observed in the corpus

New ditransitive verb	Instances
Advise	1
Inform	2
Explain	2
Present	4
Provide	30
Put	2
Remind	42
Supply	2

ditransitive manner, as exemplified in (1) to (9), were easily comprehensible to the forum participants in terms of their semantic meanings and syntactic structures, and did not hinder the realisation of the participants' communicative purposes. Note that the examples in (1) to (3) show that the speakers demonstrated a rather high level of pragmatic awareness by using polite forms when requesting information (*would it be more helpful . . . ; I'd appreciate it if . . .*). In our view, the innovative uses of the new ditransitive verbs in our corpus are best viewed not as 'learner errors', but as locally emerged forms in China English that are capable of achieving the writers' communicative goals in the online discussion forum. Further insights can be gleaned by examining the frequency and distribution of the verbs used innovatively in ditransitive structures. Table 2 provides the frequencies of the verbs identified in our corpus.

As can be seen, with the exception of the verb *provide* and *remind*, most of these new ditransitive verbs are relatively infrequently used, despite the fact that the size of our corpus exceeds 7 million words. The new ditransitive verbs found in our China English corpus differ from those found in Indian English as reported in Mukherjee (2009) and Mukherjee and Hoffmann (2006). First, a number of verbs used in ditransitive patterns in Indian English were not used in such a way in our China English corpus. These verbs include *brief*, *confer*, *dispatch/despach*, *father*, *furnish*, *gift*, *impart*, *intimate*, *notify*, *print*, *rob*, *submit*, and *threaten*. There are, however, some new ditransitive verbs that have occurred in both varieties. For instance, the verb *provide* is among the most frequent in both varieties: it has occurred 30 times in our 7+ million corpus, compared to 217 times in a 31 million corpus (i.e. *The Statesman Archive* corpus) in Indian English. This means that the ditransitive uses of *provide* in Indian English are almost twice as frequent as those in China English (7.0 vs. 4.2 times per million), reflecting a trend that this new innovative use of *provide* appears to be more entrenched in Indian English than in China English.

These new ditransitive verbs emerging in our China English corpus cannot be explained by the influences of the Inner Circle varieties of English (e.g. British English, American English), since generally these verbs are not used in such a way in these varieties. Nor can it be said to be inspired by varieties of English in Outer Circle nations (e.g. Indian English) as China English speakers generally have little exposure to these varieties or speech communities. Thus, the new ditransitive uses documented in this study can be viewed as output of 'a genuinely creative process' (Mukherjee 2009: 127) on the part of Chinese users of English. In our view, these locally emerged ditransitive verbs can be considered as concrete instantiations of structural nativisation of the Chinese variety of English at the level of lexicogrammar. Since the use of new ditransitive verbs are primarily

influenced by more established ditransitive verbs such as *give*, this verb is next examined for its verb-complementational and collocation behavior.

Verb-complementation

The focus in this paper is on the prototypical ditransitive verb *give* and this study examines the verb-complementational and collocation profiles in China English. Before the data analysis is presented, a short review of the argument structure of *give* is discussed. From the perspective of Construction Grammar (cf. Goldberg 1995), the verb *give* invokes three argument roles (i.e. the 'agent' = X, the 'recipient' = Y, and the 'patient' = Z), which are involved in a transfer event with the ditransitive meaning of 'X causes Y to receive Z'. The verb *give* is generally considered as a prototypical ditransitive verb (i.e. *give somebody something*) on which other semantically related verbs can map its complementation patterns. Broadly speaking, the verb complementation profile of *give* can be categorized into five basic types, as described in (9) to (13) with examples from our online forum corpus. Note that O_i :NP denotes noun phrase serving as indirect object, and O_d :NP denotes noun phrase serving as direct object.

Type I (S) GIVE [O_i :NP] [O_d :NP]

(9) Then my teacher *gave* me a forced smile. (t555496.txt)

Type II (S) GIVE [O_d :NP] [O_i :PP_{to}]

(10) I want to *give* my hand to a man before 28. (t537523.txt)

Type III (S) GIVE [O_d :NP] O_i

(11) Everyone who have a certain knowledge can *give* the answer instantly. (t509689.txt)

Type IV (S) GIVE O_i O_d

(12) You know it is said that it is much better to *give* than to receive. (t555496.txt)

Type V (S) GIVE [O_i :NP] O_d

(13) So *give* him as a present! (t539567.txt)

A Type I pattern takes both indirect and direct objects, and is considered as the most basic type of ditransitive complementation (cf. Schilk 2011). In a Type II pattern, the indirect object is realised as a *to*-phrase and positioned after the direct object. Type III, IV, and V patterns all involve the omission of certain element(s), such as indirect object (Type III), direct object (Type V), or both (Type IV). It is important to note that not all arguments need to be made explicit in all contexts at all times, as long as they can be recovered from the context or inferred from world knowledge (cf. Jackson 1990; Matthews 1981; Newman 1996; Biber et al. 1999). Structurally related patterns can be derived from the five basic patterns to accommodate factors such as elements in fronted position, relative clause structures, participle constructions, and passive constructions. For example,

Table 3. Complementations of GIVE in China English—the most frequent patterns in the data

Type	Pattern	China English		Indian English		British English	
		F	%	F	%	F	%
I	(S) GIVE [O_i :NP] [O_d :NP]	355	71	407	22.6	404	38
IP	[S < O_i active] BE <i>given</i> [O_d :NP] (by-agent)	0	0	130	7.2	84	7.9
II	(S) GIVE [O_d :NP] [O_i :PP _{to}]	63	12.6	310	17.3	123	11.6
IIP	[S < O_d active] BE <i>given</i> [O_i :NP _{to}] (by-agent)	2	0.4	70	3.9	23	2.2
III	(S) GIVE O_i [O_d :NP]	47	9.4	528	29.4	247	23.2
IIIP	[S < O_d active] BE <i>given</i> O_i (by-agent)	1	0.2	123	6.8	38	3.6
IIIPb	[antecedent] _{co} (S < O_d) _{co} (BE) <i>given</i> O_i (by-agent)	4	0.8	49	2.7	28	2.6
Other		28	5.6	180	10.1	117	10.9
Sum		500	100	1797	100	1064	100

a passive construction involves the occurrence of copula BE, followed by past participle(s), and optionally followed by an agent or agents in a sentence. To mark the various patterns of GIVE, we have followed the notation system of Mukherjee and Hoffmann (2006). For instance, IIP denotes a passive construction variation of a Type II pattern, while IIIPb denotes a passive construction variation of a Type III pattern with an antecedent in front of a sentence. See Mukherjee and Hoffmann (2006: 172) for a complete list of major and derivative complementation patterns of GIVE.

In order to examine the verb-complementational profile of the verb *give* in our corpus, we randomly sampled 500 instances and coded their complementational profiles according to the classification scheme as outlined in Mukherjee and Hoffmann (2006: 172). Based on our analyses, eight patterns with the most frequent occurrences have emerged (See column 1 and 2 in Table 3). For illustration, several of these patterns are exemplified and described below.

- (14) I think that your wife definitely **gives** you a good answer. (t636933.txt)
 (15) I appreciate what 21st Century **gives** me. (t515334.txt)
 (16) When one encounters love, when he/she **gives** all his/her heart, it is not so easy to forget the person who was loved for a long time. (t532560.txt)
 (17) Finally, thank you very much the corrections **given**. (t559814.txt)

As a moderator in the discussion forum, the speaker in (14) makes recommendations of cities that foreigners should visit in China. In addition to pointing out the necessity of considering Chinese traditions and local cultures, he also highlights the importance of consulting the wife's opinion before making such a decision. In doing so, he uses the Type I pattern of *give*, as in *gives you a good answer*. The speaker in (15) shares his story of discovering the 21st Century newspaper 'by accident' at a newsstand, and how he liked the paper and wanted to be a reader forever. In his statement, he uses a derivative Type Ib

pattern by fronting the direct object with a *wh*-clause (see Table 3). In (16), the speaker participates in a thread on the topic of how time can change everything. He comments that if one *gives all his/her heart* and loves someone for long, it is truly difficult to forget. In (17), the speaker was expressing her gratitude of the corrections that she has received from other participants on the forum. She acknowledges that whenever she receives corrections, she would print out her own writings alongside the corrections, and then mark the mistakes on paper. In this process, she produces a Type IIIPb pattern of *give*, with the indirect object omitted (i.e. *me*) and direct object fronted in a passive voice (i.e. *corrections that was given to me*).

Table 3 summarizes the distribution of different complementation patterns of *give* in the corpus data. Compared with the patterns in Indian English and British English as reported in Mukherjee and Hoffmann (2006), interesting observations can be made regarding the results in the corpus data presented in this current paper. For example, the complementation patterns of *give* in China English are far from evenly distributed. In fact, the Type I pattern alone accounts for over three-quarters of all the 500 instances of the use of *give* that were coded and examined. Conversely, several other complementation patterns of *give* either occur sporadically (i.e. Type IIP, Type IIIP, and Type IIIPb) or virtually do not occur (i.e. Type IP). These findings are in sharp contrast with those reported in Indian English and British English. Specifically, the occurrences of the Type I pattern in China English are nearly two times as frequent as those in British English, and almost three and half times as frequent as those in Indian English. By contrast, the uses of other patterns (i.e. Type IP, Type IIP, Type IIIP, Type IIIPb and Other) in China English are considerably less frequent than those in British or Indian English.

The preference of certain verb-complementation patterns can be interpreted from the angle of the interlocutors' perception of verb transitivity. It is generally recognised that the number of arguments is indicative of the degree of transitivity of the verb. For example, the verb *arrive* invokes only one argument, that is the subject (e.g. *He arrived*). The verb *eat*, however, can invoke more than one argument, that is the subject and an optional object (e.g. *He didn't eat*. vs. *He ate an apple*). In other words, *eat* can function both as a transitive and as an intransitive verb. Thus, the more arguments a verb can take, the greater its transitivity. With respect to the verb *give*, the Type I and Type II patterns, by definition, require the most arguments: the subject, the direct object and the indirect object. If we group patterns with the most arguments together, that is Type I, Type II and all their derivative patterns, the combined cluster would account for the vast majority (84 per cent) of the instances of *give* analyzed in our corpus. This strong preference for the Type I and Type II patterns, together with their derivative forms, suggests that the verb *give* has been perceived by China English speakers to be *highly transitive*. Such a preference, it can be argued, can be considered a distinctive feature of the structural nativisation of China English at the level of verb-complementation.

Another interesting aspect revealed from the corpus analysis was China English speakers' multilingual creativity. For example, various constituents of the basic Type I pattern are often substituted by phrases written in Chinese characters, as exemplified in (18) to (24).

(18) 华尔街英语 [Wall Street English] **gave** us some discount last year, I and many [of] my colleagues were seduced. (t526399.txt)

(19) Y'd better **give** the 音标 [phonetic script] to us. (t538345.txt)

1 222 Haiyang Ai and Xiaoye You

- 2 (20) I logined the website that you **gave** to 西雅 [Xiya] (t572844.txt)
 3 (21) What they **gave** me is a 俄罗斯破解版 [Russian-cracked version], an old one.
 4 (t540843.txt)
 5 (22) Every day my parents would **give** me 1角8分 [1 dime and 8 cents] so I could buy two
 6 of them. (t549678.txt)
 7 (23) Stop that or I will **give** you a color see see! (t553282.txt)
 8 (24) You **Give** Me Stop! 你给我站住! [Stop! Freeze!] (t553282.txt)
 9

10
 11 In (18), the Chinese name of an English language training company called Wall Street
 12 English is used as the subject of the *give* structure. Similarly, phrases written in Chinese
 13 characters are used in indirect object slot, as in (20) ‘Xiya’, a person’s name, or in direct
 14 object slot, as in (19) ‘phonetics’, (21) ‘the Russian-cracked version’, and (22) ‘a dime
 15 and eight cents’. These instances suggest that the speakers appear to have a sufficient
 16 understanding of complementational requirements of *give* and are capable of replacing
 17 its various constituents with a noun phrase in Chinese characters. In (23) and (24), the
 18 speakers creatively meshed the verb *give* with a similar Chinese word 给 *gěi* in literal
 19 translations of Chinese sayings. While (23) is a literal translation of a threat ‘给你点颜
 20 色看看’ (I will give you a color [to] see see), meaning ‘I’ll teach you a lesson’, (24) is
 21 a direct translation of the command 你给我站住! (You give me stop!), meaning ‘Stop!
 22 Freeze!’ Note that in both instances the speakers have made reference to the Chinese word
 23 *gěi* either explicitly or implicitly. Thus, the Chinese word *gěi* is linked to the English word
 24 *give* in the two instances in an interesting way. Semantically, the word *gěi* can function as
 25 a verb signifying ‘to supply, to provide’, a meaning close to the verb *give*. Grammatically,
 26 *gěi* is typically followed by personal pronouns, resulting in a similar pattern to *give* being
 27 followed by a direct object in English (i.e. *give me* vs. *gěi wo*). Phonetically, the word
 28 *gěi* sounds somewhat like *give* in English. Thus, we would argue that while the direct
 29 translation of the Chinese word *gěi* as *give* does not fully conform to the Type I pattern
 30 of the prototypical verb *give* in English, they nevertheless share considerable similarities,
 31 semantically, grammatically, and even phonetically. Such use often creates a playful and
 32 humorous tone and is not uncommon in online discourse.

33 In summary, this section has documented and discussed the verb-complementation pro-
 34 file of the prototypical ditransitive verb *give* in China English. We find that the basic Type
 35 I pattern of *give* is predominately preferred by China English speakers. This preference has
 36 been discussed in relation to the high transitivity nature of the verb *give* as perceived by
 37 China English speakers. Additionally, these results also show some intriguing evidence of
 38 China English speakers’ multilingual creativity. The participants were shown to substitute
 39 various constituents with Chinese phrases in Type I pattern, and mix *give* and the Chinese
 40 word *gěi* while still maintaining a certain level of semantic, phonetic and grammatical
 41 similarities between these two words.

42 Collocation

43
 44 In this subsection, the verb *give* is explored in terms of its collocation patterns in China
 45 English. As discussed earlier, collocation refers to the association or co-occurrence not
 46 only between words but also between lexical items and grammatical environments (Sinclair
 47 1991). In what follows, the collocation behaviors of *give* and its collocates in the direct
 48

object slot are presented. First, an examination of some typical examples that involve the verb *give* and its objects are discussed.

- (25) I want to thank her, because she give us the **chance** to do that. (t549830.txt)
- (26) Maybe I didn't do enough for her, but she could give me more **time** to contribute. (t551583.txt)
- (27) Maybe you can give me some good **idea**? (t543239.txt)
- (28) I am not sure whether she is gonna give you a **kiss** or kick your ass? who know? (t521819.txt)
- (29) Different life style will give us different **feelings** and experiences. (t555496.txt)
- (30) I prefer to give my son more **freedom** but my wife show a little bit strict with my son. (t572844.txt)
- (31) In big city, it can give you many **opportunities** and bring you so much challenges as well. (t555496.txt)
- (32) Will you give me some **advice** about the first day lessons? (t542153.txt)
- (33) If you are free, would you please have a read and give me some **suggestion** about writing? (t542153.txt)
- (34) If you want, I think I can give you my MSN **address**. (t525156.txt)
- (35) My parents refuse to give me **money** on my cellphone, they think I spend much on my cellphone. (t555496.txt)
- (36) I hope that you can give me some **corrections** about my last response. (t555496.txt)
- (37) What major? Maybe I can give him some **help**. (t542153.txt)

The speaker in (25) expresses gratitude to her English teacher for the *chance* to participate in a project involving research, debate, presentation, portfolio, and teamwork. The speaker in (26) conveys his sadness and confusion over his love affair that ended, in a thread entitled 'How can I know she really love me?' He laments over the lack of *time* for his contribution to the relationship. Other direct object collocates occur in the give complementation patterns include *kiss*, *feelings*, *address*, *help*, *etc.* Overall, these direct object slot collocates are closely related to the general topics discussed in the online forum. For example, in the corpus data it was found that the forum participants were primarily concerned with seeking or giving *advice/suggestion* (79 times), *help/hand* (16 times), or *answers/corrections* (19 times), or talking about *happiness/joy/feelings* (14 times) and *stress/pressures* (4 times).

Table 4 summarizes the top 15 direct object collocates of *give* analyzed in the corpus data, ranked according to log-likelihood score, generated by WordSmith Tool 4.0 (Scott 2004). The log-likelihood test is a statistical measure commonly used in corpus linguistics to determine whether distributional differences of words in two corpora are statistically significant by comparing the observed value, that is the actual frequencies extracted from corpora and the expected value, that is the frequency one would expect by chance (cf. McEnery et al. 2006). The data of Indian English and British English comes from Mukherjee and Hoffmann (2006), arranged here also by log-likelihood order, not by its original alphabetical order.

Overall, the collocation profiles of *give* in the direct object slot in our China English corpus differ markedly from those in British English and Indian English. The vast majority of the collocates with *give* in our top 15 list in China English do not occur in the top 15 list of British English at all. Even the three overlapped instances (*chance*, *impression*, and *time*) differ in ranking in the two lists. For instance, the most frequent collocation in British

Table 4. Top 15 direct object collocates of *give* in China English, Indian English, and British English

China English	Log-likelihood	British English	Log-likelihood	Indian English	Log-likelihood
advice	4727.15	time	365.39	importance	269.77
suggestion	2497.43	details	314.42	example	260.06
chance	2225.65	example	303.92	details	240.30
hand	1378.34	way	277.42	address	214.66
answer	1053.17	opportunity	201.69	news	211.27
money	1016.07	rise	187.39	information	209.16
love	1006.23	impression	171.46	chance	195.56
wishes	825.41	information	164.78	money	192.12
impression	802.23	indication	152.33	answer	167.52
time	738.40	prescription	149.71	idea	165.26
gift	689.00	chance	139.60	amount	85.10
try	679.90	idea	104.58	advice	80.92
life	581.96	ring	93.06	explanation	80.92
feeling	529.65	support	81.36	description	63.08
hope	493.99	sense	78.95	meaning	58.30

English is *give time*, but it is only ranked 10th in our corpus. Our most frequent collocation, by contrast, is *give advice*. Similarly, 11 out of the top 15 direct object collocates of *give* identified in the corpus do not occur at all in Indian English. That is to say, in the top 15 list, the two local varieties of English share only four common collocates (i.e. *advice*, *chance*, *answer*, and *money*), all of which are ranked higher in the list for China English than the list for Indian English.

In addition to collocation strength, that is the extent to which two or more words not co-occurring in a corpus by chance, the collocation profile of *give* can also be examined in the associations between direct object collocates and their differing preference of verb-complementation patterns. While the verb *give* can function in several structurally related patterns (e.g. relative clauses, fronted elements, passive voice, etc.), these patterns are relatively infrequent in general. Distinctive analysis of them is not likely to yield salient results (see Schilk 2011: 63). Thus, it was decided to follow Schilk's (2011) practice and these patterns were collapsed as derivative patterns Type-Ider, Type-IIder, etc. Note that for the analysis of the collocate *address*, a distinction between website address, email address, or IM address (e.g. Skype, MSN or QQ) was not made, but rather generally treated as 'address'.

Table 5 summarizes the frequent complementation patterns of the verb *give* in relation to its various direct objects in the China English data. By taking the collocate *advice* as an example: out of a total of 50 instances, 44 of them, or 88 per cent, have occurred in the Type-I pattern; one instance, or 2 per cent, occurred in the Type-II pattern; and five instances, or 10 per cent, occurred in the Type-III pattern. As can be seen, collocates in the direct object slot of *give* show differing preference for the complementation patterns. Some collocates always occur (i.e. 100 per cent preference) in the Type-I complementation pattern (e.g. *chance*, *time*, *idea*, *kiss*, *feeling*, *freedom*, *opportunity*), and others show a strong preference (75–89 per cent preference, e.g. *advice*, *suggestion*, *address*, *money*, *correction*, *help*) for the Type I pattern. Together, these two groups of collocates, both preferring the Type-I complementation pattern, accounts for the majority of the cases in

Table 5. Frequent complementation patterns of *give* in relation to direct object collocates in China English

Collocates	Type I	Type Ider	Type II	Type IIder	Type III	Type IIIder	Other	N
chance	100.0%	0%	0%	0%	0%	0%	0%	14
time	100.0%	0%	0%	0%	0%	0%	0%	9
idea	100.0%	0%	0%	0%	0%	0%	0%	7
kiss	100.0%	0%	0%	0%	0%	0%	0%	7
feeling	100.0%	0%	0%	0%	0%	0%	0%	6
freedom	100.0%	0%	0%	0%	0%	0%	0%	5
opportunity	100.0%	0%	0%	0%	0%	0%	0%	5
advice	88.0%	0%	2.0%	0%	10.0%	0%	0%	50
suggestion	89.7%	0%	3.4%	0%	6.9%	0%	0%	29
address	81.8%	9.1%	9.1%	0%	0%	0%	0%	11
money	80.0%	0%	20.0%	0%	0%	0%	0%	15
correction	83.3%	0%	0%	0%	0%	16.7%	0%	6
help	75.0%	0%	25.0%	0%	0%	0%	0%	8
happiness	60.0%	0%	40.0%	0%	0%	0%	0%	5
wishes	40.0%	0%	60.0%	0%	0%	0%	0%	5
answer	53.8%	0%	0%	0%	46.2%	0%	0%	13
comment	33.3%	0%	0%	0%	66.7%	0%	0%	6
hand	37.5%	0%	25.0%	12.5%	25.0%	0%	0%	8
website	42.9%	14.3%	14.3%	0%	14.3%	14.3%	0%	7

the dataset, a pattern consistent with the discussion of the verb-complementation profile examined in the previous section. In Indian English, by contrast, the majority of direct object collocates (i.e. *address*, *money*, *answer*, *details*, *information*, *explanation*) shows a strong preference for the Type-III complementation pattern, not the Type I pattern (Schilk 2011: 88). Other collocates exhibit somewhat different preference for the complementation patterns. For example, while the collocate *wishes* is profiled more prominently in the Type-II pattern, *answer* and *comment* occurred more often in both the Type I and Type III patterns in Indian English. Still others show a tendency for more diverse patterns (e.g. *hand* and *website*).

The lexis-grammar interface can also be explored from the association between direct object collocates and verb forms. Table 6 summarizes the top 15 direct object collocates and their relationship with different tenses of the verb *give*. Collocates are analyzed here, both singular and plural, in a five-word span to the right of the node word *give*.

Taking the collocate *chance* as an example, it can be observed that *chance* occurs somewhat differently with different forms of *give*: that is *chance(s)* occurs with *give*, *gives*, *gave*, *given* 231, 13, 17, 18 times respectively. In other words, the simple present form (i.e. *give*) accounts for the majority of the cases that collocate with the noun *chance* in our dataset. As can be seen in Table 6, among the several tenses of *give*, a vast majority of the top 15 collocates show a predominant preference for the simple present over other forms. One potential explanation for this overwhelming preference of simple present tense is that this tense is learned early and could be considered very familiar to the majority of China English speakers. This preference of simple present tense can also be explained as a reflection of the everyday discourse nature of the online forum, where this tense is commonly used. In sum, this section has explored the collocation profile of the verb *give* in China English. The analysis presented here suggests that there is a strong link between

Table 6. Top 15 direct object collocates of *give* and its relationship with forms in China English

Collocates	Simple present	Third person singular	Simple past	Past participle
advice	505	0	6	6
suggestion	238	0	11	5
chance	231	13	17	18
hand	176	0	0	0
answer	138	0	13	16
money	149	8	9	8
love	147	4	15	8
wishes	131	0	0	0
impression	39	15	25	4
time	116	8	17	8
gift	46	0	5	8
try	104	0	3	0
life	91	6	7	10
feeling	61	30	8	0
hope	32	3	2	7

specific direct object collocates and the Type I complementation pattern, as well as the simple present tense of *give* in China English. This association between lexis and grammar, it could be argued, can be considered as concrete instantiation of the structural nativisation of China English in collocation profiles.

CONCLUSION

The emergence of a unique form of English in the largest Expanding Circle country necessitates investigation into this new variety (Berns 2005). Previous studies have examined various aspects of China English, including the existence of China English as a local variety, people's perceptions and attitudes toward the notion of China English, the role of English in China's language curriculum, historical and sociolinguistic descriptions of English in China, rhetorical strategies and literary and multilingual creativity, and linguistic features of China English. Taking a structural nativisation perspective, this present study has explored the lexis-grammar interface in China English. Utilising large-scale corpus data, three major areas in lexicogrammar were examined, that is new ditransitive verbs, verb-complementation, and collocation.

This study reveals that a number of new ditransitive verbs have emerged in China English (i.e. *supply*, *present*, *advise*, *inform*, *explain*, *provide*, *put*, *remind*). It is argued that such new uses of verbs cannot be attributed to the influences from Inner Circle varieties, where these verbs generally are not used in a ditransitive manner; nor can such new uses be attributed to Outer Circle varieties, such as Indian English, because very few Chinese speakers of English have had exposure to these varieties or speech communities. With respect to verb-complementation, a strong preference for the Type I pattern of the verb *give* among China English speakers was found. This preference was discussed in terms of the perception of verb transitivity—that is, *give* can be perceived to be highly transitive by China English speakers. Additionally, the analysis presented here has revealed a certain level of bi-/multilingual creativity, realised through the replacement of constituents of

give with phrases written in Chinese characters, and through drawing on the similarities between the verb *give* and Chinese word *gěi* phonetically, semantically, and grammatically. Finally, the analysis of the collocates in the direct object slot of *give* shows a strong preference among Chinese English speakers for the Type I complementation pattern as well as the simple present verb tense. These findings suggest that there are certain concrete associations or co-selections between lexis and grammar in a range of aspects in China English. These locally emerged linguistic patterns, through close comparison with those found in Inner Circle and Outer Circle varieties (i.e. British English and Indian English), are attested to be distinctive and characteristic of the Chinese variety of English. While some China English scholars (Gu and Xiang 1997; Kirkpatrick and Xu 2002; Li 1993) have maintained that an essential component of China English is the presence of Chinese characteristics, which can be seen at the level of lexis, sentence structure, and discourse, this study shows that the presence of Chinese characteristics is also evident at the lexis-grammar interface. It is the view here that these lexicogrammatical features can and should be regarded as a concrete instantiation of structural nativisation in China English.

In the field of world Englishes, this study has contributed new knowledge to the interdependent relationship between lexis and grammar in studying structural nativisation in local varieties of English. While this study has focused on a single verb *give*, it has uncovered intriguing aspects of association or co-selection between specific lexical items and grammatical constructions in China English. In addition, the interdependence between lexis and grammar can be explored not only through the three major areas examined in this study, but also from an array of other perspectives. For example, Mukherjee (2009) showed that new prepositional verbs (e.g. *discuss about something*, *visit to somebody*) have also manifested as a new form of structural nativisation at the level of lexicogrammar in Indian English. It should be noted that research on lexicogrammar should not prioritize lexis at the expense of grammar or marginalize lexis in favor of grammar (Römer & Schulze 2009). The link between lexis and grammar can be profitably explored and should be given more attention to in research on local varieties of English.

Additionally, this study has made methodological contributions to the study of world Englishes. First, while previous studies on lexicogrammatical features of English varieties have used data from the ICE series of corpora or corpora compiled from online newspaper articles, this study shows that online discussion forums can also be a valuable resource for data collection and analysis. As more people are connected to the Internet and conducting an increasingly large proportion of their lives in digital domains, it is important that world Englishes scholars turn to new media for data collection and analysis, such as social networking (e.g. Facebook) and microblogging (e.g. Twitter) sites (see Al-Sa'di & Hamdan, 2005; You 2008, 2011; Seargeant & Tagg 2011). More importantly, data collected from online discussion forums and other new media outlets are qualitatively different from those collected from online newspaper articles. While the latter are most often written by a small group of journalists and have typically undergone editorial reviews and revisions, often done by Inner Circle speakers of English as a way of 'quality control', the former are indicative of the actual language used by the speakers of the local varieties of English. Second, this study demonstrates that the notion of web-as-corpus (cf. Kilgarriff & Grefenstette 2003; Hoffmann 2007) can be more fruitfully explored for studying world Englishes by writing computer programs. Contrary to general belief, writing programs to retrieve webpages (and for linguistic studies in general), as Biber et al. (1998: 256) pointed out, 'does not require a special aptitude in computer science or mathematics.' Moreover,

many open source packages and libraries of code have been developed for common tasks such as web scraping. For example, in compiling the China English corpus, this study has made use of a HTML scraping package called *BeautifulSoup*, which is capable of taking an URL, retrieving the content of the webpage and returning specific section of the webpage according to particular HTML/CSS tags (e.g. *id*, *class*) the researcher assigns. Once the programs have been developed and tested, one could easily retrieve a large amount of data from new media outlets and compile corpora with size well in excess of millions of words, thus forming a large-scale corpus from which more generalizable claims can be made about various aspects of an English variety.

It is also important to point out that native speaker intuition is not particularly useful for the study of the interdependent nature of particular lexical items and their grammatical environments, which are generally difficult to perceive, let alone predict (Biber et al. 1998). By contrast, the corpus-based approach is well suited for documenting and describing this intricate and interdependent relationship between lexis and grammar. However, the corpus size appears to be a factor for certain areas of study of lexicogrammar in world Englishes. For instance, while the China English corpus in this study consists of over 7 million words, results for new ditransitive verbs are still ‘a low-frequency phenomenon’ (Mukherjee 2009: 125). Future research on this issue would benefit from using larger corpora.

This corpus-based study of lexicogrammar in China English has provided some insights regarding how the lexis-grammar interface can be successfully explored by studying the structural nativisation of Expanding Circle varieties. However, the study is limited in scope as only written data has been analyzed. Analyses of spoken data in China English could prove equally insightful (Xu 2010). Additionally, this study has only examined data from an online discussion forum. Future research could use more balanced corpora (e.g. Mukherjee 2009) and explore the effects of register (Biber 2012; Balasubramanian 2012) on lexicogrammar in the structural nativisation of China English. Finally, future studies can also examine the verb-complementation and collocation patterns (cf. Koch & Bernaisch 2013) of the eight new ditransitive verbs unveiled in China English.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Fengyang Ma, Brooke Ricker, Kingsley Bolton, and the anonymous reviewers for their constructive feedback on earlier versions of this paper.

NOTES

1. The 21st Century online discussion forum can be accessed at <http://bbs.i21st.cn/>.
2. The software developed in this article will be made freely available for research purposes upon publication of this study. Interested researchers can contact Haiyang Ai for the software program at aihaiyang@gmail.com.
3. Information in parenthesis indicates the file from which an example is drawn. Each file corresponds to a forum thread from which the data is collected.

REFERENCES

- Al-Sa'di, Rami A. & Jihad Hamdan. 2005. ‘Synchronous online chat’ English: Computer-mediated communication. *World Englishes* 24(4). 409–424.
- Balasubramanian, Chandrika. 2009. *Register variation in Indian English*. Amsterdam: John Benjamins.
- Bautista, Maria Lourdes S. 2004. The verb in Philippine English: A preliminary analysis of modal would. *World Englishes* 23(1). 113–128.

- Bautista, Maria Lourdes S. 2008. Investigating the grammatical features of Philippine English. In Kingsley Bolton & Maria Lourdes S. Bautista (eds.), *Philippine English: Linguistic and literary perspectives*, 201–218. Hong Kong: Hong Kong University Press.
- Berns, Margie. 2005. Expanding on the Expanding Circle: Where do WE go from here? *World Englishes* 24(1). 85–93.
- Bianco, Joseph Lo, Jane Orton, & Yihong Gao. (eds.). 2009. *China and English: Globalization and the dilemmas of identity*. Bristol: Multilingual Matters.
- Biber, Douglas. 2012. Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory* 8(1). 9–37.
- Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.
- Bolton, Kingsley. 2002. Chinese Englishes: From Canton jargon to global English. *World Englishes* 21(2). 181–199.
- Bolton, Kingsley. 2003. *Chinese Englishes: A sociolinguistic history*. Cambridge: Cambridge University Press.
- Chen, Meilin, & Hu, Xiaoqiong. 2006. Towards the acceptability of China English at home and abroad. *English Today* 22(4). 44–52.
- Collins, Peter. 2007. Can/could and may/might in British, American and Australian English: A corpus-based account. *World Englishes* 26(4). 474–491.
- DeKlerk, Vivian. 2005. The use of *actually* in spoken Xhosa English: A corpus study. *World Englishes* 24(3). 275–288.
- Du, Ruiqing & Yajun Jiang. 2001. Jinershi nian 'Zhongguo Yingyu' yanjiu shuping [An overview of studies of 'China English' in the last two decades]. *Waiyu Jiaoxue yu Yanjiu [Foreign languages teaching and research]* 33(1). 37–41.
- Ersso, Olavarria & Philip Shaw. 2003. Verb complementation in Indian Standard English. *English World-Wide* 24(2). 137–161.
- Firth, John R. 1957. *Papers in linguistics: 1934–1951*. Oxford: Oxford University Press.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Granger, Sylviane & Stephanie Tyson. 1996. Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes* 15(1). 17–27.
- Gu, Guanjie, & Mingyou Xiang. 1997. Wei zhongguo yingyu yibian [An argument for China English]. *Waiyu yu Waiyu Jiaoxue [Foreign languages and foreign languages teaching]* 5. 11–13.
- Halliday, Michael A.K. 1966. Lexis as a linguistic level. In Charles E. Bazell, John C. Catford, Michael A. K. Halliday & Robert H. Robins (eds.), *In Memory of J. R. Firth*, 148–163. London: Longman.
- He, Deyuan & David C. S. Li. 2009. Language attitudes and linguistic features in the 'China English' debate. *World Englishes* 28(1). 70–89.
- He, Deyuan & Lindsay Miller. 2011. English teacher preference: The case of China's non-English-major students. *World Englishes* 30(3). 428–443.
- Ho, Mian Lian, & Irene F.H. Wong. 2001. The use of *ever* in Singaporean English. *World Englishes*, 20(1). 79–87.
- Hoffmann, Sebastian. 2007. From web-page to mega-corpus: The CNN transcripts. In Marianne Hundt, Nadja Nesselhauf & Carolin Biewer (eds.), *Corpus Linguistics and the Web*, 69–85. Amsterdam: Rodopi.
- Hu, Xiaoqiong. 2004. Why China English should stand alongside British, American, and other 'World Englishes'. *English Today* 20(2). 26–33.
- Hundt, Marianne & Ulrike Gut (eds.). 2012. *Mapping unity and diversity world-wide: Corpus-based studies of new Englishes*. Amsterdam: John Benjamins.
- Jackson, Howard. 1990. *Grammar and meaning: A semantic approach to English grammar*. London: Longman.
- Jiang, Yajun. 2003. English as a Chinese language. *English Today* 19(2). 3–8.
- Jung, Kyutae & Su Jung Min. 1999. Some lexico-grammatical features of Korean English newspapers. *World Englishes* 18(1). 23–27.
- Kennedy, Graeme. 1998. *An introduction to corpus linguistics*. London: Longman.
- Kilgariff, Adam & Gregory Grefenstette. 2003. Introduction to the special issue on the Web as corpus. *Computational Linguistics* 29(3). 333–347.
- Kirkpatrick, Andy & Zhichang Xu. 2002. Chinese pragmatic norms and 'China English'. *World Englishes* 21(2). 269–279.
- Koch, Christopher & Tobias Bernaisch. 2013. Verb complementation in South Asian English(es): The range and frequency of 'new' ditransitives. In Gisle Andersen & Kristin Bech (eds.), *English corpus linguistics: Variation in time, space and genre*, 69–89. Amsterdam: Rodopi.
- Li, Wenzhong. 1993. China English and Chinese English. *Waiyu Jiaoxue yu Yanjiu [Foreign languages teaching and research]* 25(4). 18–24.
- McEnery, Tony, Richard Xiao & Yukio Tono. 2006. *Corpus-based language studies: An advanced resource book*. London: Routledge.
- Matthews, Peter H. 1981. *Syntax*. Cambridge: Cambridge University Press.

- Mukherjee, Joybrato. 2009. The lexicogrammar of present-day Indian English: Corpus-based perspectives on structural nativization. In Ute Römer & Rainer Schulze (eds.), *Exploring the Lexis-Grammar Interface*, 117–135. Amsterdam: John Benjamins.
- Mukherjee, Joybrato & Sebastian Hoffmann. 2006. Describing verb-complementational profiles of new Englishes. *English World-Wide* 27(2). 147–173.
- Mukherjee, Joybrato & Marco Schilk. 2008. Verb-complementational profiles across varieties of English: Comparing verb classes in Indian English and British English. In Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta & Minna Korhonen (eds.), *The dynamics of linguistic variation: Corpus evidence on English past and present*, 163–181. Amsterdam: John Benjamins.
- Newman, John. 1996. *Give: A cognitive linguistic study*. Berlin: Mouton de Gruyter.
- Pandharipande, Rajeshwari. 1987. On nativization of English. *World Englishes* 6(2). 149–158.
- Römer, Ute. 2005. *Progressives, patterns, pedagogy: A corpus-driven approach to English progressive forms, functions, contexts and didactics*. Amsterdam: John Benjamins.
- Römer, Ute, & Schulze, Rainer. 2009. *Exploring the lexis-grammar interface*. Amsterdam: John Benjamins.
- Rong, Pei. 1991. Zhongguo yingyu shi keguan cunzai [China English is real]. *Jiefangjun Waiyu Xueyuan Xuebao [Journal of the PLA Institute of Foreign Languages]* 48(1). 1–8.
- Sand, Andrea. 2004. Shared morpho-syntactic features in contact varieties of English: Article use. *World Englishes* 23(2). 281–298.
- Schilk, Marco. 2011. *Structural nativization in Indian English lexicogrammar*. Amsterdam: John Benjamins.
- Schilk, Marco, Tobias Bernaisch & Joybrato Mukherjee. 2012. Mapping unity and diversity in South Asian English lexicogrammar: Verb-complementational preferences across varieties. In Marianne Hundt & Ulrike Gut (eds.), *Mapping unity and diversity world-wide*, 137–165. Amsterdam: John Benjamins.
- Schneider, Edgar W. 2004. How to trace structural nativization: Particle verbs in world Englishes. *World Englishes* 23(2). 227–249.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Scott, Mike. 2004. *WordSmith Tools (Version 4.0) [Computer software]*. Oxford: Oxford University Press.
- Sergeant, Philip, and Caroline Tagg. 2011. English on the Internet and a 'post-varieties' approach to language. *World Englishes* 30(4). 496–514.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stubbs, Michael. 1995. Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language* 2(1). 23–55.
- Wei, Yun & Jia Fei. 2003. Using English in China. *English Today* 19(1). 42–47.
- Wong, Jock. 2004. Reduplication of nominal modifiers in Singapore English: A semantic and cultural interpretation. *World Englishes* 23(3). 339–354.
- Xu, Zhichang. 2010. *Chinese English: Features and implications*. Hong Kong: Open University of Hong Kong Press.
- Yang, Jian. 2005. Lexical innovations in China English. *World Englishes* 24(4). 425–436.
- You, Xiaoye. 2008. Rhetorical strategies, electronic media, and China English. *World Englishes* 27(2). 233–249.
- You, Xiaoye. 2010. *Writing in the devil's tongue: A history of English composition in China*. Carbondale, IL: Southern Illinois University Press.
- You, Xiaoye. 2011. Chinese white-collar workers and multilingual creativity in the diaspora. *World Englishes* 30(3). 409–427.
- Zhang, Hang. 2002. Bilingual creativity in Chinese English: Ha Jin's *In the Pond*. *World Englishes* 21(2). 305–315.

(Received 6 July 2014)