Using Corpora in Autonomous Study of English Vocabulary: A Case Study of STORM

ZUO Qingzhao

Department of Translation, The Chinese University of Hong Kong

zuoqingzhao0511@163.com

Abstract

Vocabulary is central to language and of great importance to the Chinese English Foreign Language (EFL) learners. At present, it is often the case that many Chinese EFL learners are very weak in vocabulary acquisition. Different reasons account for this situation faced by Chinese EFL learners such as the examination-oriented teaching, cultural differences, large classes and insufficient time available.

The purpose of this paper is to show the importance of autonomous and data-driven learning in vocabulary learning, especially with the use of corpora. This research takes STORM as a case study. Data are gathered from COCA (Corpus of Contemporary American English). The paper has conducted lexicon-grammatical analysis such as collocation, colligation and semantic prosody of STORM and a new entry of storm is designed. Results indicate the relevance of the use of data-driven learning as a way of enriching the knowledge of word and improving students' autonomous learning. This study may raise pedagogical implication in English vocabulary learning and lexicography.

Keywords

Chinese EFL learner, data-driven learning, autonomous learning, vocabulary acquisition

Introduction

Vocabulary is a basic element of a language. British linguist David Wilkins stated that "Without grammar, very little can be conveyed. Without vocabulary, nothing can be conveyed" (1972:111). This significant comment stressed the indispensable role played by vocabulary in a language. Therefore, vocabulary acquisition is one of the fundamental tasks facing *English as Foreign Language* (EFL) learners and is arousing more and more attention around the word.

Scholars in China have attached importance to the research of vocabulary acquisition and teaching in EFL classroom since 1980s. The research has been strongly influenced by the development of other subjects. It has been suggested that the application of corpus linguistics (e.g., Yang, 2002) and cognitive linguistics in the practice of English vocabulary teaching (e.g.,Gu, 2002). Although much attention has been paid to the corpus-based approach to the vocabulary teaching from teachers' point of view, few studies has explored the use of corpus from the perspective of Chinese EFL learner. Motivated by the above research gap, this paper attempts to investigate how the corpus can help Chinese EFL learner in autonomous learning of English vocabulary.

1 Theoretical framework

There are many strategies for vocabulary acquisition. This paper adopts corpuses-based approach to study vocabulary. The following is the theoretical background of data-driven learning autonomous learning and lexicon-grammatical analysis.

1.1 Data-driven learning

Data--driven learning (DDL) was put forward by Tim Johns in 1990s and defined by Johns and King (1991:iii) as " the use in the classroom of computer-generated concordances to get students to explore regularities of patterning in the target language, and the development of activities and exercises based on concordance output". Johns believes that "research is too serious to be left to the researchers": that the language-learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data" (1991:2).

Charles (2002) points out that this method of teaching allows students to investigate a corpus of native-speakers speech or writing with a concordancing program and exposes them to authentic examples of language usage rather than contrived examples found in grammar textbooks.

Many have advocated that DDL "favor learning by discovery" (Tribble and Jones, 1990: 12). It implies that the study of grammar, collocation or vocabulary is based on the research and investigation rather than learning by rote or spoon feeding. Scholar like Kettemann (1995:4)

argues that a concordancing program is "an extremely powerful hypothesis testing device" that allows the student to conduct inductive explorations of grammatical constructions "on vast amounts of data" (as cited in Charles, 2002).

1.2   Autonomous learning

However excellent a teacher may be, students will never learn a language well without extra-class study and self-instruction. As Nunan suggests that, not everything can be taught in class (1998a:3). Effective EFL learning relies heavily on learners' autonomous learning. Holec (1981:3) defined autonomy as "the ability to take charge of one's own learning". He developed this definition further in 1985 by talking about autonomy as a conceptual tool. According to this definition, learners of autonomous study shoulder the responsibility for all decisions in regard to one's learning such as objective determination, methods and techniques selection, and self-evaluation.

Based on this theory, teachers at home and abroad have implemented teaching practices by adopting modern technology. For example, Mariangela (2005:145) presents the importance of computer use in extra-class study and self-instruction in foreign language learning.

1.3   Lexicon-grammatical analysis

Although the basic element of language is lexis, word shall not be studied in isolation. As J.R. Firth stated, "You shall know a word by the company it keeps" (1957: 179). Part of the meaning of a word is that it collocates with other words (Halliday,1966). Thus collocation, "the occurrence of two or more words within a short pace of each other in a text" (Sinclair, 1991a: 170), needs to be studied. Furthermore, "many uses of words and phrases show a tendency to occur in a certain semantic environment" (Sinclair 1991:112). This particular collocational fact were analysed in detail and given a name as "semantic prosody" by Louw in 1993. Louw argues that when writers violate the habitual use of semantic prosody, readers can detect a difference between what the writer says and what he or she really believes. Semantic prosody can be used to create an ironic atmosphere and thus add vividness to the article. Stubbs (1996) classifies semantic prosody into negative prosody, positive prosody and neutral prosody.

Apart from collocation and semantic prosody, colligation also shows how a word behaves grammatically in the text. Hoey defines colligation as 'the grammatical company a word keeps and the positions it prefers'; in other words, a word's colligations describe what it typically does grammatically' (2000:234). Colligation is a similar idea to collocation, but has a different emphasis. While collocation emphasized the words that belong to these grammatical classes, colligation focuses the relationship among different grammatical classes

2   Methodology of the study

The corpus used for this research is Corpus of Contemporary American English (COCA). It is the largest corpus of American English currently available. It was constructed by Mark Davies, a professor of corpus linguistics in Brigham Young University. COCA consists of more than 425 million words from 1990-2011. The corpus is updated once or twice a year. Texts from different genres are included in this corpus such as spoken, fiction, popular magazines, newspapers, and academic journals. It is suitable and convenient for conducting lexicon-grammatical analysis and looking at current ongoing changes in the language.

Aaccording to a survey, many students only know STORM as a natural phenomenon and are unfamiliar with other usages of this word. This causes difficulties in their reading and translation. Therefore, storm is chosen as a case of study. STORM is selected as KWIC (Keyword in the context) for generating concordance lines and collocation patterns. A collocation span of three to both sides of the node word STORM was set. The concordance lines can be sorted by left or right. From the contexts of concordance lines, other possible prosodic meanings were observed. In addition, the colligation can be determined from the collocation of STORM.

3   Analysis of the result

3.1   Collocation and colligation

As defined before, collocation is two or three word clusters which occur with a more than chance regularity throughout spoken and written English. Semantic prosody refers to the positive and negative environment of a word. Colligation is an abstracted description of the collocational class.

According to the concordance lines, STORM tends to have the following collocation and colligation patterns when it is used as a noun.

    1. det adj N:storm
    a tropical/perfect/big/electrical/political/violent/ storm
    2. det N: storm prep
    a storm of controversy/negative publicity uncertainty/ hostility/ fury/bullets/violence
    a storm of applause, emotion, tears, laughter
    3. V det N: storm prep
    ignite/cause/ lead to/touch off/draw/set off/arouse/ unleash/provoke a storm of

4. V up det N: storm

cook/stir/ talk/smoke/ talk/ cry/ dance/ sweat/ argue/ chat/ boil / beat up a storm

5. Det N N: storm

a snow/ rain storm, brain storm

6. prep N: storm

by storm

7. det N: storm V

a storm hits/damages

Based on the analysis of COCA, the most frequently verbs that follow storm can be see in the following table.

Table 1: verbs collocates with storm

Is	was	had
has	hit	come
would	will	came
chase	could	pass
blew	broke	broke
surge	comes	began
made	struck	struck
moved	brought	continue

Similarly, the collocation and colligation of verb *storm* can be listed as follows.

1. Verb+preposition (used as intransitive verb) storm out of/off to/up to/over to/ in with/back into/back to/ away from/ off toward/in with/ ashore in/back from/down to.

2. Verb+noun (used as a transitive verb) strom the building/ the field/ the gates/ the cockpit/ the stage/ the barricade/ the city/ the home/ the palace/ the theater/ the parliament/ the hospital

3.2 Semantic prosody

Just as what Louw (1993) suggests, in the past large number of semantic prosodies remained inaccessible and were not studied satisfactorily. With the development of corpus linguistics and lexicography, the analysis of semantic prosody becomes practicable.

STORM when used as a noun has the following possible semantic prosodies derived form its left and right collocates.

Table 2: Semantic prosodies of *storm* used as a noun from its left collocates

No.	Semantic prosody	Left collocates	Examples
1	degree	*severe, massive, dangerous, great, bad, deadly, devastating*	A *huge* storm hits.
		severe, huge, massive	
2	location	*tropical, coastal, American*	a *tropical* storm
3	time	*approaching, coming, current, sudden*	The *approaching* storm
4	color	*blue, white, red, black, dark*	the great *white* storm clouds
5	business	*financial, economic*	a *financial* storm
6	military	*military, Deserted*	*Deserted* Strom Operation
7	science	*solar, electric*	the *solar* storm

Table 3: Semantic prosodies of *a storm of* from its right collocates

Semantic prosody	Right collocates	Examples
disapproval	*controversy outrage protest criticism anger hostility reprimand*	a storm of *uncertainty*
emotion	*anguish fury laughter applause tears*	a storm of *applause*
nature	*bullets snow*	a storm of *snow and ice*

According to Table 2, the adjectives used before STORM show the degree of storm. More often than not, the degree is severe. Table 3 shows that "a storm of" typically collocates with c*ontroversy/media attention/outrage/ hostility/ fury/bullets/violence*. This phrase refers to that a lot of people express their strong feelings out of upset and their emotions are agitated. The words that collocate with "a storm of" refer something bad. On this occasion, it is a negative semantic prosody. This is probably because that When storm hits/ strikes an area, it usually brings trouble to people with fall of rain, wind, snow or sometimes thunder or lightening. It may also cause damage. Generally speaking, when people come up with storm, it has a negative meaning.

However, it is not fair to consider all "a storm of +noun" phrases as negative semantic prosodies. From the concordance lines of COCA, there are some collocations like *a storm of laughter/*

applause. These phrases may show the neutral or positive prosodic features.

In addition, when STORM used as a verb, the semantic prosody can also be analyzed. For example, *Storm out/off/into* implies that somebody goes somewhere quickly because he or she is angry or upset. This is a negative semantic prosody.

3.3 Design a new entry of STORM

storm[1] noun

1. [C] very bad weather with heavy rain, snow, strong winds and THUDER and LIGHTENING *tropical/violent/ severe storms* A powerful storm **hit/ struck** the city. The storm **began** as we came home.
2. (to form compounds) a thunderstorm/ snowstorm / sandstorm/ brainstorm
3. [C, usually singular] a situation many people suddenly express very strong feelings about something **a storm of sth** *arouse, cause/provoke a storm of protest/question*
4. [C, usually singular] a sudden loud noise that is caused by emotion or excitement **a storm of sth** *a storm of noisy weeping/ laughter/ A storm of applause burst from the crowd.*
5. IDM **a storm in a teacup** unnecessary worry or anger about something that is not important *These debates are something of a storm in a teacup.*
6. **cook/dance/talk/smoke+ up a storm** *informal* to do something with all your strength and energy
7. **take sth/sb by storm** a) to be successful in a particular place *The play took London by storm.*b) to attack a place successfully by using a large number of soldiers.

storm[2] v 1. [T] to attack a place suddenly and take control of it by using a lot of force *An angry crowd stormed the police.*

2. [I] to go somewhere fast because you are angry and upset **storm+ out of, off to, over to** *She stormed out of the home. The guests stormed off to another casino.*

3. [I,T] literary to shout something angrily

4 Conclusion

The paper presents an example of data-driven learning in autonomous study, which emphasizes that learners can become researchers and discover new possibilities to learn vocabulary and thus improve English proficiency. The present study has analyzed the lexicon-grammatical pattern of STORM such as collocation and semantic prosody by using COCA. The findings indicate that information provided by dictionaries may not be sufficient. The use of corpus in autonomous learning can supplement the drawbacks of dictionaries. This study has pedagogical implications and may give some insights to the dictionary compilation.

References

Charles, F. (2002). *English Corpus Linguistics*. New York: Cambridge University Press.

Davies, Mark. (2008). The Corpus of Contemporary American English (COCA): Available online at http://www.americancorpus.org.

Firth, J.R. (1951/57). *Papers in Linguistics*. London: Oxford University Press.

Firth, J. R. (1957) A Synopsis of Linguistic Theory, 1930-55. In Palmer, F.R. (ed) (1968). *Selected Papers of J.R. Firth 1952-59*. London and Harlow: Longmans.

Halliday, M.A.K. (1966). Lexis as a Linguistic Level. In Bazell, C.E., Catford, J.C.Halliday, M.A.K. and Robins, R.H. (Eds) *In Memory of J.R. Firth*. London: Longmans.

Holec,H. (1981). *Autonomy and Foreign Language Learning*. Oxford: Pergamon Press.

Hoey, M. (2000). A World Beyond Collocation: New Perspectives on Vocabulary Teaching. In Lewis, M. (ed) *Teaching Collocations*. Hove: Language Teaching Publications.

Johns, T. and King, P (eds) (1991) *Classroom Concordancing*, ELR Journal (New Series) 4.

Louw, B. (1993). Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies. In Baker, M., Francis, G. & Tognini-Bonelli, E. (Eds) *Text and Technology. In Honor of John Sinclair*. Philadelphia/Amsterdam: John Benjamins.

Mariangela, B.N. (2005). Self-access study and cooperative foreign language learning through computers. *Linguagem&Ensino*, 2, 145-169.

Nunan, D. (1988a). *Syllabus Design*. Oxford: Oxford University Press.

Partington, A. (1998). *Patterns and meanings: using corpora for English language research and teaching*. Amsterdam: Benjamins.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J. (1991). *Reading concordances: an introduction*. Harlow: Person/Longman.

Stubbs.M. (1996). *Text and Corpus Analysis*. Oxford: Blackwell.

Tribble, C. and Jones, G. (1990). *Concordances in the Classroom*. Longman: London.

Wilkins, David A. (1972). *Linguistics and Language Teaching*. London: Edward Arnold.