

A Corpus-based Research on English for the Coverage of Earthquakes

Li Wei

Xijing University, Xi'an, Shaanxi Province, China, 710123

Keywords: Corpus; English for the coverage of earthquakes; news English

Abstract: Based on McCarthy's semantic field theory and Sinclair's corpus-based approach, lexical characteristics of English for the coverage of earthquakes are described from the respects of key words graphs, lexical chunks and the semantic preferences, hoping to provide guidance for teaching and researching news English.

1. Introduction

It is difficult for English majors to learn news English and take exam of it. English for the coverage of earthquakes is a significant branch of news English and it is the research object of this thesis. However, it is restricted to the net hard news of VOA. Adopting semantic field theory proposed by McCarthy and corpus-based approached put forth by Sinclair, it is designed to probe into lexical characteristics of English for the coverage of earthquakes from key words graphs, lexical chunks and the semantic preference of lexical chunks. It is expected that the research will provide reference for news English teaching and research.

2. Theoretical Basis

When words are used to show an individual's world perception, there is a hierarchical relationship among them. At the highest level, there is a semantic space composed of semantic fields. A semantic field consists of kinds of lexical fields derived from the vocabularies and phrases in some certain fields, and those vocabularies and phrases share in the concept and knowledge of the semantic field. Words in semantic field present a clustering relation (McCarthy, 1990). From the perspective of semantic field, key words, lexical chunks and collocations are designed to be analyzed so as to show how they embody pragmatic function of the coverage of earthquakes.

Key words are the ones with obviously higher occurrence frequency than a certain referential corpus, and the degree of the higher occurrence frequency is defined as keyness. Being composed of key words, the lexical field of the coverage of earthquakes is not solely a direct reflection of the semantic field of the coverage of earthquakes, but also a convenient method to study on the interactions among lexical groups. A collocation refers to the co-occurrence of two or more words within a short span in the text. It is clarified that the span studied in the thesis is between 4- (left) and 4 (right). Lexical chunks are the larger units of language use than collocations (Wei Naixing, 2002). It is not "open-choice principle" but "idiom principle" that is adopted by native speakers to choose suitable collocations or lexical chunks to express their intentions. As what pointed by Pawley & Syder in 1983, idiomatic lexical chunks were belonged to an important index which showed the language popularity. In the thesis, the operability of lexical chunks is defined as consecutive collocations with high frequency (≥ 3) in the corpus; semantic preference means semantic features and types of lexical items and common words.

3. Research Methods

After selecting the hard news by typing into key word "earthquake" on VOA website dating from 8, 2012 to 7, 2013, the author collated texts and set up a mini corpus about English for the coverage of earthquakes. PowerGREP (main index tool), AntConc (aided index tool), and Frown (referential corpus) are tools adopted in the thesis. With the guide of semantic field theory, key words were

extracted first, and then the semantic analysis of them was undergone. Next, the key words graph about English for the coverage of earthquakes was established. Finally, 2 or more lexical chunks were selected to analyze the characteristics and semantic preference of lexical chunks.

4. Lexical Characteristics of English for the Coverage of Earthquakes

4.1 Key words digraphs

As shown in Table 1, among 123 key words selected from the referential corpus Frown, earthquake and quake are the most key words. Meanwhile, the top 10 key words are relevant to earthquake elements (magnitude, kilometer), news sources and news terms (voa, geological, say).

Table 1 Key Words of English for the Coverage of Earthquakes (excerpts)

Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	44	511.007	earthquake	6	70	153.886	be
2	36	451.453	quake	7	12	150.484	geological
3	23	233.944	magnitude	8	17	143.172	strike
4	17	189.623	kilometers	9	10	125.404	voa
5	15	188.105	tsunami	10	29	105.981	say

From the above table, it can be seen that “earthquake” and “quake” are cores of those key words. The whole key words graph is composed of such 8 aspects as topic, time, place, organizations, people, influence, measures, and news sources. Those 8 factors demonstrate the overall structure of the coverage of earthquakes, for instance, when and where earthquakes break out, what key information delivered by Earthquake Monitoring Bureau, what earthquake influence reported by media, and people and work involved in disaster rescue.

Specifically, the research shows these characteristics of key words. For example, “earthquake / quake” is core; “magnitude”, “epicenter”, and “depth” are the reflections of earthquake elements. Key words about news sources include “VOA”, “State Media”, “news agency”, “Xinhua”, etc.; about “coverage” are “say” and “report”. Organizations, such as “The U.S. Geological Survey” and “The Pacific Tsunami Warning Center” are closely related to earthquakes. The factor “people” is made up of main characters concerning disaster relief, for instance, “Chinese Premier Wen / Li”, “officials”, “experts”, “spokesman”, and “injured people”. Key words about effect can be classified into 3 categories, which respectively are the ones about casualties: “casualties”, “death”, “the death toll”, “injured”, “killed”, and “wounded”; the verbs about earthquake occurrence and property damage are “strike”, “damage”, “shake”, and “destroy”, “rattle”; nouns about destructiveness and secondary disaster include “buildings”, “a large number of structures”, “impassable roads”, “tsunami”, “disaster”, “landslides” and so on. Keys words about rescue are either synonyms of “rescue” or its hyponyms, the former includes “recovery” and “relief efforts” and the latter includes “rescue / emergency workers”, “tents”, “quilts”, “coats”, “shelters”, “camps” and so forth.

4.2 Lexical chunks and semantic preferences

In the following part, the commonly used lexical chunks and their auxiliary meanings in English for the coverage of earthquakes are to be discussed.

With the aid of AntConc, lexical chunks with 2 or more words are extracted. In total, there are 29 lexical chunks with complete structures and actual meanings are selected to analyze semantic preferences of common chunks.

There are a certain number of lexical chunks in English for the coverage of earthquakes and they are made from 9 words at most, a case in point is “There were no immediate reports of casualties or damage.” Among those lexical chunks, some of them are composed of from 4 to 6 words and others from 2 to 3 words; the number of the former is 9 and the latter is 19. From the perspective of structure, those lexical chunks are characterized by sentences, clauses, verb phrases, noun phrases, preposition phrases, and adverb phrases (shown in Table 2). In contrast to Nai Weixing’s study on the acquisition of lexical chunks for English learners in China, it is found in that the lexical chunks

used by native speakers are characterized by more frequent use, longer length, and more various structures.

In respect of pragmatics, lexical chunks in English for the coverage of earthquakes center on two aspects. The first focus is about news sources, such as “The U.S. Geological Survey” which appears 11 times. The second focus is about factors concerning earthquakes, including earthquake magnitude, occurrence of earthquakes, casualties, epicenter, and focal depth. It’s required that news about the earthquake should be proved by the authority. Meanwhile, it is the idiomatic expressions that can be helpful for readers being equipped with schema about earthquake discourse to know the earthquake. To conclude, in specific context, the lexical chunks related to the coverage of earthquakes have such pragmatic function as delivering message accurately and promptly.

Table 2 Lexical Chunks (2 to 9 words) in the Coverage of Earthquakes (excerpts)

	lexical chunks	Freq	collocation patterns of lexical chunks
sentence	There were no immediate reports of casualties or damage.	4	
clause	China’s official Xinhua news agency says that	3	+ casualties / damage
noun phrase	The U.S. Geological Survey	11	+ reports / earthquake situations
preposition phrase	at a depth of	4	+ # kilometers
noun phrase	magnitude earthquake	12	a # - +
verb phrase	has struck	4	+ place
adverb phrase	at least	3	+ #

In the following part, the collocation pattern of “The U.S. Geological Survey” will be discussed as a case in point. Among collocations of “The U.S. Geological Survey” (shown in Table 3), at column 1R, the top word is “say”, and then other words meaning “coverage”, including “reports”, “registered”, “put”, and “estimated”. Among those words, 6 ones are belonged to general present tense and 5 to general past tense. Column 2R is dominated by “the” and column 3R by earthquake terms. At column 4R, collocations are 5 “earthquake”. From the above analysis, it can be easy to find the common sentence patterns of “The U.S. Geological Survey” are the followings.

- 1) The U.S. Geological Survey says the # magnitude earthquake ...
- 2) The U.S. Geological Survey said the quake ...

Table 3 Statistical Table about Collocations of “The U.S. Geological Survey” 1R – 4R

Rank	Freq	Keyness	Keyword	Rank	Freq	Keyness	Keyword
1	44	511.007	earthquake	6	70	153.886	be
2	36	451.453	quake	7	12	150.484	geological
3	23	233.944	magnitude	8	17	143.172	strike
4	17	189.623	kilometers	9	10	125.404	voa
5	15	188.105	tsunami	10	29	105.981	say

5. Conclusion

Key points of teaching English are the most common word forms, the key use patterns of them (collocations and semantic prosodies), and the typical combinations of them (lexical chunks) (Sinclair & Renouf, 1988). Key words, collocations, lexical chunks, and semantic preferences mentioned in the thesis are key words of English for the coverage of earthquakes, and they can be used not only as teaching materials but also as study method sample of data-motivated study.

According to the study on inter-language used by English learners in China by Wei Naixing (Wei Naixing, 2004), the learners either overuse typical lexical chunks of common words or underuse

them. Additionally, whether the learners can successfully use the words is first dependent on whether they can set up a words net focusing on the theme, association and collocation and then on whether the relation can be accurately reflected in using native language. It's the author who holds the idea that in teaching foreign language vocabulary, the method focusing on the theme and in-depth vocabulary knowledge is beneficial to promote students' abilities of applying English and communicating with others.

References

- [1] McCarthy, M. Vocabulary [M]. Oxford: Oxford University Press, 1990.
- [2] Wei Naixing. Collocation's Definition and the Research System of It [M]. Shanghai: Shanghai Jiao Tong University Press, 2002.
- [3] Pawley, A., H. Syder. Two Puzzles for Linguistic Theory: Native-like Selection and Native-like Fluency [A] // J. Richard, R. Schmidt (eds.). Language and Communication [C]. New York: Longman, 1983.
- [4] Wei Naixing. The Initial Research Report of COLSEC [J]. Modern Foreign Languages, 2004 (2).
- [5] Sinclair, J., Renouf A. A Lexical Syllabus for Language Learning [A] // M. J. McCarthy, R. A. Carter (eds.). Vocabulary in Language Teaching [C]. London: Longman, 1998: 141-160.