ISSN : 2350-0743

www.ijramr.com

International Journal of Recent Advances in Multidisciplinary Research
Vol. 06, Issue 03, pp.4684-4689, March, 2019

RESEARCH ARTICLE

NEAR SYNONYMS AND THE FEATURE OF WORD SKETCH DIFFERENCE OF THE SKETCH ENGINE: A CASE STUDY OF THE ENGLISH ADJECTIVES BRAVE: COURAGEOUS

***Susi Yuliawati and Dian Indira**

Faculty of Cultural Sciences, Universitas Padjadjaran, Indonesia

ARTICLE INFO

Article History:

Received 27th December, 2018
Received in revised form
04th January, 2019
Accepted 10th February, 2019
Published online 28th March, 2019

Keywords:

Corpus, near synonyms, word sketch Difference, frequency, and collocation.

ABSTRACT

The study examines a linguistic phenomenon of near synonyms investigated by one of the features in a corpus tool of Sketch Engine namely Word Sketch Difference. The feature functions to compare and contrast two words through collocation analysis by presenting their collocates categorized on the basis of grammatical relation. Using the data derived from the English Web Corpus 2013 (enTenTen13), the present research aims to explore near synonyms by observing the usage of the English adjectives Brave: Courageous by employing the analyses of collocation. The study employs a mixed-method design in which quantitative and qualitative analyses are combined. In the quantitative analysis, the frequency of word usage and the significance of collocation are identified. The collocational behaviour and semantic categorization of collocates are then described qualitatively to interpret the degree of semantic similarities occurred between the two adjectives. The result of analysis strongly suggests that word frequency, collocational behaviour, and semantic categorization of collocates can be used as an indicator to examine near synonyms. Furthermore, the feature of Word Sketch Difference in the corpus tool of Sketch Engine can assist not only those who study language, but also teachers of English as a foreign language in explaining the near synonyms based on the big quantity of real language use.

INTRODUCTION

Students of English as a Foreign Language (EFL) often find difficulty in determining the right word to represent what they mean. The confusion specifically occurs when they find several choices of words with similar meaning. In the study of semantics, such relation is widely known as synonym. Teachers of EFL are certainly required to be able to teach vocabulary knowledge appropriately and effectively to their students. The skill of comprehending word meaning is generally trained in two ways: guessing meaning based on the context of their use in sentences and using dictionary effectively. In addition to EFL teachers, lexicographers or those who compile dictionaries are also required to have the capability of describing word meanings based on their use in natural settings. Thus, the norm of language use investigated from a large collection of the actual language usage in a society is very important. Synonym is obviously a familiar term for language learners. In brief, it is defined as the sameness of meaning. For a linguist, however, the definition is considered unsatisfying because in fact synonymous words do not entirely have identical meaning. Cruse (2000) argued that synonymous words are words which semantic similarities are more salient than differences. Based on this definition, the exploration of synonyms open up as it raises several questions such as what semantic differences do not destroy the intuition of sameness? Why are synonymous relations so frequent? How to detect semantic differences in synonymous relation? Furthermore, Cruse (2000) distinguished three different synonyms, which are absolute synonymy, proportional synonymy, and near-synonymy. According to Edmonds and Hirst (2002), absolute synonym is rarely found. If two lexical items recognized as absolute synonym, they are able to be substituted one for the other in any contexts in which their common sense is denoted with no change to truth-value, communicative effect, or meaning. In other words, meaning in the absolute synonym is completely identical. Cruse (2000) suggested some word pairs as the candidates of absolute synonyms, such as SOFA: SETTEE, which refer to a piece of furniture for seating typically with two arms and a back, and PULLOVER: SWEATER, which refer to a knitted garment covering the top half of the body. Differentiating the contexts of use of the word pairs is hard to find. Nevertheless, some philosophers such as Quine (1951) and Goodman (1952) argue that true synonym is impossible because it is really hard to defined. Even if absolute synonyms were conceivable, pragmatic and empirical arguments show that it would be very rare. Even though the word pairs of SOFA: SETTEE and PULLOVER: SWEATER, for example, are very close in meaning, they are found to have different meaning in certain contexts. The main difference of Pullover: Sweater lies in the way they put on or take off. People usually put on it over the head, while a sweater is not always put on over the head, but it may open and fasten down the front because it can have front buttons. Some also distinguish sofa from settee based on the size; settee defined as a small sofa with two arms and a back. Propositional synonyms defined in terms of entailment refer to the relation between two lexical items interchangeable in any expressions with truth-conditional properties without effect on those properties. Cruse (2000) suggested that differences in meaning of propositional synonyms involve one or two aspects

Corresponding author:* **Susi Yuliawati
Faculty of Cultural Sciences, Universitas Padjadjaran, Indonesia.

of non-propositional meaning. The most important aspects are 1) differences in expressive meaning; 2) differences in stylistic level such as formal and non-formal situations; and 3) differences in the fields of discourse. The propositional synonymis reflected in the word pair SHIN: FIBULA, which both refer to the front of the leg below the knee. People usually use the word SHIN as a general term without expressive or certain stylistic loading, whereas medical specialists acting in that role often use FIBULA. Near-synonyms, also known as plesionyms, are generally said to have a very close meaning, but the meaning is not identical. If two lexical items are near-synonyms, they cannot fully substitute one for the other. They can be discriminated based on the aspects of denotation, connotation, implicature, or register (DiMarco, Hirst, and Stede, 1993). Edmonds and Hirst (2002) argued that semanticists such as Ullmann (1962), Cruse (1986), and Lyons (1995) tend to define near-synonyms on the basis of propositional meaning. Cruse (1986), for instance, distinguished cognitive synonyms from near-synonyms. Cognitive synonyms are words that preserve their truth-conditional when intersubstituted in a sentence. However, there may be changes in their expressive meaning, style, or register. In the meantime, near-synonyms are typically words that when intersubstituted in a sentence, they change the truth-conditional. In spite of that, they still produce semantically similar sentences. Unlike semanticists, lexicographers are frequently associated synonyms with near-synonyms. They commonly define synonyms on the basis of the sameness of meaning; the differences only lie in how broad the definition ought to be. Roget (in Chapman, 1992), for example, suggested an extensive definition because he followed the vague principle: "the grouping of words based on idea". Word senses are finally classified according to proximity of meaning in the hierarchical structure of *Roget's Thesaurus*. In the *Webster's New Dictionary of Synonym,* synonyms are only defined as two words or more that have the same or very close essential meaning. In line with that, *Webster's Collegiate Thesaurus* suggests the limitation of synonyms on the basis of the likeness of meaning. The synonyms here are defined as objective denotation uncoloured by peripheral aspects such as connotation, implication, or idiomatic usage. Obviously, the main point of these definitions is that near-synonyms must be specified by two important elements, which are the sameness of essential meaning and the difference of peripheral meaning (Cruse, 1986).

There have been several approaches used to study near-synonyms. They are not only from the perspectives of linguistics, particularly from semantic study, but also from computational linguistics. For example, Sun, Huang, and Liu (2011) suggested a method to identify near-synonyms and similar-looking words (NSSL) developed from the application system that can learn words with semantic relations of NSSL. The system was tested by three different experimental models to the subject of Chinese students of EFL. On the other hand, Edmond and Hirst (2002) developed a new computational model to demonstrate the meaning of near-synonyms in a detailed way by displaying its differences. This computational model was designed in such a way with the purpose to be applicable for the machines of translation as well as text crawling. In relation to this previous research, the present research is basically aimed to investigate near-synonymous words from the perspective of corpus linguistics by making use of a product created by experts in the field of computational linguistics. As mentioned earlier, the description of meaning including near-synonyms is better investigated according to the norms of language use prevailing in a society. Therefore, using corpus linguistics to study near-synonyms is regarded to be the most appropriate due to some of the advantages possessed by the approach. Corpus linguistics is an empirical approach with a set of method and procedure to analyse the big quantity of linguistic data. According to McEnery and Hardie (2012), corpus linguistics is a field of study that focuses on a set of method or procedure to investigate language. The word corpus itself is not a new term because the method to search meaning by involving many various contexts in a large collection of texts had been conducted since the 13th century.

However, the process was still done manually (McCarthy and O'Keeffe, 2010). Now, corpus linguistics is better known as the study of linguistic data (both in written or spoken forms) by utilizing computer to store and process the data and thus it allows researchers to collect and analyse data in relatively big scale. As a result, the researchers do not rely on their intuition when investigating and making claims about the linguistic data. One of the previous studies examining near-synonyms by employing corpus linguistic approach was done by Xiao and McEnery (2006). They used contrastive study by comparing near-synonyms in English and Chinese. From the observation of collocational behaviour and semantic prosody, they examined three case studies of near-synonyms on the noun groupCONSEQUENCE, CAUSE, and PRICE/COST. Based on the analysis, they revealed that even though Chinese and English are clearly two different languages which are not related, the collocational behaviour and semantic prosody of the near-synonyms are nearly similar. The other research is from Uba (2015) who discussed the internal semantic structure of the near-synonym adjectives: IMPORTANT, ESSENTIAL, VITAL, NECESSARY, and CRUCIAL from a corpus-based behavioural profile approach. The research used the data which was taken from the British National Corpus (BNC) and nine traditional references (five dictionaries and three thesauruses) to describe the distributional patterns especially types of nouns that each adjective modified. The analysis showed that the traditional references were not given a satisfying description of the internal semantic structure of the near-synonym adjectives. As a result, the application of the corpus-based behavioural profile approach was proved to be significance to gain an in-depth account of the internal semantic structure of near-synonyms. The other important study related to the present research is from Yang (2016). In his research, Yang (2016) compared the usage of verbs LEARN and ACQUIRE by using of concordance analysis, word sketch, and word sketch difference from a corpus tool namely Sketch Engine. Using the data taken from the British National Corpus (BNC), the research showed that the collocational analysis resulted from the features of word sketch and word sketch difference could be used to discriminate the synonymous verbs, which are LEARN and ACQUIRE. The research also revealed that native speakers did not only memorize words in isolation, but also chunks of words regarded as the building blocks of a language. The native speakers acquired the chunks of words as ready-use units that contributed to the fluency and the naturalness of their utterances. Thus, it is highly recommended that teachers of EFL who aim to help their students achieving English fluency and accuracy to teach the language skills by applying collocational patterns. Based on the previous research on near-synonyms, the present research focuses on the investigation of near-synonyms by utilizing the feature of Word Sketch Difference of the corpus tool Sketch Engine. Unlike the

previous research conducted by Yang (2016), this research examines near-synonym adjectives instead of verbs, BRAVE and COURAGEOUS which according to Cruse (2000) they are near-synonymous words. We use *enTenTen13* English corpus that is also accessed through the Sketch Engine as the data. The present study is essentially aimed at exploring the feature of Word Sketch Difference to investigate near-synonyms based on a large collection of natural texts and proving the opinion of Yang (2016) that the feature is a very fruitful tool to especially demonstrate differences between near-synonym words. The difference between this study and the study conducted by Yang (2006) particularly lies in the class of word that become the focus of analysis; Yang (2006) examined two near-synonym verbs (LEARN and AQCUIRE) and this examines two near-synonym adjectives (BRAVE and COURAGEOUS). The research done by Uba (2015) actually studied on a group of near-synonym adjectives as well, but he applied different technique and procedure of analyses. Xiao and McEnery (2006) also studied near-synonym words, but their investigation focused on a group of near-synonym nouns and was done comparatively by comparing their usages in English and Chinese.

MATERIALS AND METHODS

For the present research, we use English Web Corpus 2013, also well known as enTenTen13 corpus, as the data. The corpus size is around 19 billion words that was constructed from English texts crawled from the internet. The samples of data were compiled from five sub-corpuses that are the collection of texts taken websites with Australian domain (.au), Canadian domain (.ca), UK domain (.uk), US domain (.us), and Wikipedia. This indicates that the data at least involves the language varieties of Australian English, Canadian English, and British English. The corpus itself is available and accessible from Sketch Engine, a corpus tool that has been widely used to explore how language works by many different people such as linguists, lexicographers, translators, teachers and students (Kilgariff, et al. 2004 and 2014). The tool provides not only English corpora, but also other language corpora. There are 400 ready-to-use corpora in more than 90 languages, each having a size of up to 20 billion words to provide a truly representative sample of language use. In addition to the corpora, the Sketch Engine also supplies the users with several features for data processing such as Word List to generate a word frequency list of a corpus, Concordance to generate a list of examples of the search word or phrase in a corpus usually in a format of KWIC (key word in contexts), Word Sketch to create a one-page summary of the search word's grammatical and collocational behaviour, and Word Sketch Difference to compare and to contrast to words by analysing their collocations and displaying collocates divided into categories based on grammatical relations. By exploring several features of the Sketch Engine, particularly the feature of Word Sketch Difference, we examine a word pair of BRAVE: COURAGEOUS. To study the near-synonym adjectives, we use a mixed-method research design in which quantitative and qualitative approaches are combined to provide an in-depth understanding of the research topic than either approach alone (Creswell, 2014). The present authors use three procedures of corpus analysis to investigate the similarities and differences of meaning in the word pair of BRAVE: COURAGEOUS. The first is frequency analysis that is used to identify the occurrence of each word in the corpus by using the feature of Word List. The second is collocation analysis aimed to identify significant collocates and to classify the semantic categories of the search words by utilizing the feature of Concordance. Third, the analysis of collocational behaviour by observing the collocates of the search words on the basis of the grammatical relation with help of the feature of Word Sketch Difference. To measure the strength of collocations, the study uses the significant test of Log Dice within the span of 5:5 which the calculation is generated by the features of Concordance and Word Sketch Difference. We use USAS (Ucrel Semantic Analysis System) to categorize the significant collocates semantically. The result of corpus analysis is subsequently interpreted qualitatively by applying the analysis of semantic preference to determine the main differences and similarities of meaning in near-synonym words.

RESULT AND DISCUSSION

On the basis of the word frequency list, the occurrences of BRAVE and COURAGEOUS in the corpus of enTenTen13 can be identified. The adjective BRAVE is more frequently used than COURAGEOUS. The occurrence of BRAVE is 73%, while COURAGEOUS is only 27%. As presented in Table 1, it means that the frequency of BRAVE is nearly 3 times of COURAGEOUS. The frequency analysis of word usage also reveals that people are likely to prefer the word BRAVE rather than COURAGEOUS. Using the Word Sketch Difference, we can compare and contrast the usages of BRAVE and COURAGEOUS from the analysis of collocation because the feature displays the summary of collocates, words that always tend to co-occur with BRAVE or/and COURAGEOUS. In this case, the feature does not only present the collocates, but also classifies them into categories based on the grammatical relations. Thus, we can demonstrate the similarities and differences of the adjectives BRAVE and COURAGEOUS from the collocational behaviour. Despite that the two adjectives BRAVE and COURAGEOUS share a number of syntactical patterns, the collocates in each pattern apparently differs. Firstly, in the 'and/or' pattern, the collocation tokens for BRAVE are 68,653 and 32,021 for COURAGEOUS. It indicates that there are more words in 'and/or' pattern with BRAVE. In other words, many words co-occur more frequently with BRAVE than with COURAGEOUS. Among those collocates (see Table 2), some words tend to co-occur only with BRAVE, which are *courageous* and *foolish,* and the other is only with COURAGEOUS, which is *visionary*. In the meantime, it is found that are more words that co-occur with both BRAVE and COURAGEOUS, such as *young, noble, loyal, adventurous, strong, bold, wise, honest, resourceful, honorable, fearless, heroic, valiant, daring, selfless, faithful, determined, outspoken, tenacious, brave, principled, compassionate*. With regard to this, it actually has given an indication of semantic differences between BRAVE and COURAGEOUS, i.e. unlike COURAGEOUS, BRAVE does not only encompass positive meaning, but also negative one. Based on this pattern, we further investigate the co-occurrence of BRAVE and COURAGEOUS in the construction of phrase, since it is found that COURAGEOUS is the most frequent collocates of BRAVE, which the co-occurrence is 1,248. The most prominent finding is that the frequency of 'and' pattern is higher than 'or' pattern. As shown in Table 3, both *brave and courageous* and *courageous and brave* occur more frequently than *brave or courageous* and *courageous or brave.* The frequency of *brave and courageous* is even nearly 27 times of

Table 1. The word frequency list of the adjectives *BRAVE: COURAGEOUS*in the English corpus enTenTen13

	Brave	Courageous	Brave + Courageous
Frequency	226.027	82.280	308.307
Percentage of occurrence	73%	27%	

Table 2. The Collocates of BRAVE and COURAGEOUS in 'and/or' pattern

Brave	Courageous	Brave & Courageous
courageous, foolish	visionary	young, noble, loyal, adventurous, strong, bold, wise, honest, resourceful, honorable, fearless, heroic, valiant, daring, selfless, faithful, determined, outspoken, tenacious, brave, principled, compassionate

Table 3. BRAVE and COURAGEOUS as a phrase

...and/or/,/Ø...	frequency	...and/or/,/Ø...	frequency
brave and courageous	746	courageous and brave	136
brave, courageous	233	courageous, brave	56
brave or courageous	27	courageous, and brave	10
brave courageous	25	courageous brave	8
brave, and courageous	7	courageous or brave	8
brave,or courageous	2	courageous, or brave	1

Table 4. The modifiers of BRAVE and COURAGEOUS

Brave	Courageous	Brave & Courageous
recklessly, mighty, foolishly, stupidly, insanely, outrageously, very, incredibly	politically, morally, oddly, stunningly, supremely, intellectually	exceptionally, extraordinarily, amazingly, remarkbly, unbelievably, astonishingly, uncommonly, outstandingly, admirably, immensely

Table 5. Nouns modified by BRAVE and COURAGEOUS

Brave	Courageous	Brave/courageous
face, world, demon, soul, soldier, knight, man, warrior, firefighter, adventurer	witness, leadership	hero, deed, pioneer, patriot, fight, journalist, heroine, stance, whistleblower, stand, battle

Table 6. 'The BRAVE /COURAGEOUS in' pattern

Brave in	Courageous in	Brave/Courageous in
combat, war, pursuit, sense, land, way, attempt	effort, action, decision, approach, faith, witness, willingness	situation, battle, face, choice, fight

Table 7. 'The BRAVE /COURAGEOUS for' pattern

Brave for	Courageous for
good, sake, someone	lord

Table 8. The 20 top significant collocates of BRAVE

Rank	Collocates	Frequency	LogDice	Rank	Collocates	Frequency	LogDice
1	souls	4,977	8.1	11	men	11,630	6.2
2	brave	2,394	7.1	12	Huxley's	672	6.1
3	courageous	1,704	7.1	13	cold	4,292	6.1
4	bold	2,165	6.7	14	world	6,034	6.0
5	soldiers	2,914	6.7	15	crowds	1,007	6.0
6	warriors	1,410	6.7	16	Kaa	620	6.0
7	warrior	1,224	6.5	17	noble	891	5.8
8	Aldous	1,825	6.4	18	illusions	628	5.8
9	soldier	1,341	6.3	19	fearless	626	5.8
10	Brave	856	6.3	20	heroes	906	5.8

Table 9. The 20 top significant collocates of COURAGEOUS

Rank	Collocates	Frequency	LogDice	Rank	Collocates	Frequency	LogDice
1	brave	1,582	7.3	11	battle	292	5.7
2	bold	1,213	6.5	12	visionary	1,621	5.5
3	principled	240	6.2	13	resourceful	206	5.4
4	fearless	342	6.1	14	Conversations	153	5.3
5	Captains	267	6.0	15	valiant	137	5.3
6	Courageous	205	6.0	16	tenacious	129	5.3
7	compassionate	169	5.9	17	honorable	118	5.2
8	selfless	364	5.9	18	intelligent	170	5.1
9	heroic	213	5.8	19	noble	525	5.1
10	daring	289	5.7	20	Sailing	278	5.0

brave or courageous, while the frequency of *courageous and brave* is 17 time of *courageous or brave.* This evidence essentially demonstrates that most of the English language users has an intuition to distinguish BRAVE from COURAGEOUS. In other words, they notice that BRAVE and COURAGEOUS cannot fully substitute one for the other. Secondly, in the 'modifier' pattern, the collocation tokens of BRAVE are 27,080 and 6,896 of COURAGEOUS, suggesting that there are more words used as modifier for BRAVE than COURAGEOUS. The compelling evidence is that some words are found to collocate only with one of the adjectives. As presented in Table 4, words that collocate only with BRAVE are *recklessly, mighty, foolishly, stupidly, insanely,* and *outrageously* and words that collocate only with COURAGEOUS are *oddly, stunningly, supremely,* and *intellectually.* In the meantime, the words *very* an *incredibly* are more likely to co-occur with BRAVE; they also co-occur with COURAGEOUS, but the frequency is very low.

In the same fashion, the words *politically* and *morally* are more likely to co-occur with COURAGEOUS. Meanwhile, there are many words collocate with both BRAVE and COURA GEOUS, such as *exceptionally, extraordinarily, amazingly, remarkbly, unbelievably, astonishingly, uncommonly, outstandingly, admirably, immensely*. Something worth noting here is thatin the 'modifer' pattern, BRAVE tends to collocate with words having negative meaning as they relate to foolishness and insanity while COURAGEOUS tends to collocate with words having positive meaning as they relate to moral and intelectuallity. Thirdly, in the 'nouns modified by BRAVE/COURAGEOUS' pattern, the collocation tokens for BRAVE is 119,776 and 46,759 for COURAGEOUS, indicating that there are more words which are modified by BRAVE than by COURAGEOUS. As presented in Table 5, words that collocate only with BRAVE are *face, world,* and *demon.* Some words that are more likely to collocate with BRAVE are *soul*, *soldier, knight, man, warrior, firefighter,* and *adventure,* while those that are more likely to collocate with COURAGEOUS are *witness* and *leadership.* Meanwhile, some words that collocate with both BRAVE and COURAGEOUS are *hero, deed, pioneer, patriot, fight, journalist, heroine, stance, whistleblower, stand,* and *battle.* The most compelling difference to note here is that BRAVE is more likely used to modify physical things, whereas is more likely used to modify something involving intellectual and moral. Fourthly, in the 'BRAVE/COURAGEOUS in' pattern, the collocation tokens of BRAVE are 1,437 and 902 of COURAGEOUS.

As presented in Table 6, some words collocate only with BRAVE in this pattern such as *combat, war, pursuit, sense, land,* and *way* and the other words collocate only with COURAGEOUS such as *effort, action, decision, approach, faith, witness,* and *willingness*. In the meantime, some words that collocate with both BRAVE and COURAGEOUS are *situation, battle, face, choice,* and *fight.* Unlike in the 'BRAVE/COURAGEOUS in' pattern, there are only a few words collocate either with BRAVE or COURAGEOUS in the 'BRAVE/COURAGEOUS for' pattern. As seen in Table 7, there were some words collocate only with BRAVE such as *good, sake,* and *someone* and only with COURAGEOUS, which is *lord.* However, there is no word collocate both with BRAVE and COURAGEOUS in this pattern. In addition to this, there are actually other patterns of the occurrences of BRAVE/COURAGEOUS with preposition that are the patterns of 'BRAVE/COURAGEOUS of', 'BRAVE/ COURAGEOUS on', and 'BRAVE/COURAGEOUS at'. However, there are only a few collocation tokens for COURAGEOUS and thus the adjective does not have significant collocate in the pattern 'BRAVE/COURAGEOUS of' and 'BRAVE/COURAGEOUS at'. The last analysis of the near-synonym BRAVE: COURAGEOUS is conducted by determining significant collocates for each of them by using a statistical significant test with the aim to investigate their semantic categories. The present research uses Log Dice function within a span of –/+5 with the minimum frequency of each collocate being set 5 and the minimum frequency in given range 3. The significant test is done automatically with the help of Concordance feature of the Sketch Engine. The following tables present the 20 top significant collocates of BRAVE and COURAGEOUS. As shown in Table 8, the significant collocates of BRAVE can be grouped into four categories based on the USAS:

- Fear/bravery/shock: *courageous, brave, fearless, bold*
- Warfare, defence and the army: *soldiers, warriors, warrior, soldier*
- People (male): *men, heroes*
- Proper names: *Huxley, Brave, Aldous, Kaa*

Based on the semantic categories, we can subsequently determine the semantic preference of BRAVE. Stubbs (2002) defines that semantic preference as a relation, not between individual words, but between a word and sets of collocates that share the same semantic feature. Thus, we can identify that there are four sets of semantic preference of BRAVE. They are fear/bravery/shock; warfare, defence and the army; people especially males; and proper names consisting of male names. It apparently suggests what senses make up the meaning of BRAVE: it is not only constituted by the sense relating to emotional states/actions of bravery, but also by the sense relating to physical strength. As seen in Table 9, the significant collocates of COURAGEOUS can be grouped into four categories:

- Fear/bravery/shock: *brave, fearless, courageous, heroic, daring, valiant, bold.*
- Personality traits: *courageous, compassionate, selfless, tenacious.*
- Ability: ability, intelligent: *visionary, resourceful, intelligent*.
- General ethics: *principled, noble.*

Based on the semantic categories, it can be seen that there are four sets of semantic preference of COURAGEOUS. They are fear/bravery/shock; personality traits; ability relating to intelligent; and general ethics. The semantic preference likely indicates that the meaning of COURAGEOUS is made up not only by the sense of the emotional states or actions of bravery, but also the senses of personality traits, intellectual ability, and ethics. The analysis of semantic preference clearly distinguishes BRAVE from COURAGEOUS. Although BRAVE and COURAGEOUS share the same meaning, their differences are clearly revealed here. Both of the adjectives share the same basic meaning viz. relating to the actions/states of bravery, whereas the differences lie in the prototype centre i.e. BRAVE is prototypically physical and COURAGEOUS is prototypically involving intellectual and moral aspects. As a result, the present research apparently has proven Cruse's argument (2000) that the minor discrimination between

BRAVE and COURAGEOUS lies in the difference of prototype centre.

Conclusion

This research has introduced the corpus tool Sketch Engine, specifically the feature of Word Sketch Difference, and its advantages in investigating synonymous adjectives. The results show that Word Sketch Difference generates collocation analysis that does not present the significant collocates of the adjectives BRAVE and COURAGEOUS, but also classify the collocates based on the grammatical relations. As a result, we can compare and contrast the two adjective not only based on the semantic pattern, but also on the syntactical pattern.

REFERENCES

Chapman, R.L. (Ed.). 1992. Roget's International Thesauruss.

Creswell, J.W. 2014. Research design: Qualitative, quantitative, and mixed methods approaches. USA: Sage Publications, Inc.

Cruse, D. A. 1986. Lexical Semantics. UK: Cambridge University Press.

Cruse, D. A. 2000. Meaning in language: An introduction to semantics and pragmatics.UK: Oxford University Press.

DiMarco, C., Hirst, G. and Stede, M. 1993. The semantic and stylistic differentiation of synonyms and near-synonyms. Working notes of the AAAI Spring Symposium on Building Lexicons for Machine Translation. Standford University, March 1993 pp.114–121.

Edmonds, P. and Hirst, G. 2002. Near-synonymy and lexical choice. Computational Linguistics. Vol. 8, 2, pp. 105–144.

Goodman, N. 1952. On likeness of meaning in semantics. In L. Linsky (Ed.). Semantics and the philosophy of language. University of Illinois Press. pp. 67–71.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíč ek, M., Kovář , V., Michelfeit, J., Rychlý, P. and Suchomel, V. 2014. The Sketch Engine: Ten years on. Lexicography, 1 (1), 7-36.

Kilgarriff, A., Rychly, P., Smrz, P. and Tugwell, D. 2004. The Sketch Engine. Proc Euralex, Lorient, France.

Lyons, J. 1995. Linguistic semantics: An Introduction. UK: Cambridge University Press.

McCarthy, M. and O'Keeffe, A. 2010. Historical perspective: What are corpora and how have they evolved?. In O'Keeffe, A. and McCarthy, M. (Eds.). The Routledge handbook of corpus linguistics. USA and Canada: Routledge.

McEnery, T. and Hardie, A. 2012. Corpus Linguistics. Cambridge.

Quine, W.V.O. 1951. Two dogmas of empiricism. Philosophy Review (60) 1, pp. 20–43.

Stubbs, M. 2002. Words and phrases: Corpus studies of lexical semantics. Blackwell Publishing.

Sun, K., Huang, Y. and Liu, M. 2011. A WordNet-based near-synonyms and similar looking word learning system. Educational Technology and Society, 14(1), pp. 121–134.

Uba, S.Y. 2015. A corpus-based behavioural profile study of near-synonym: Important, Essential, Vital, Necessary, and Crusial. *International Journal of English Language and Linguistics Research*, Vol. 3, No. 5, pp. 9–17.

Ullman, S. 1962. Semantics: An introduction to the science of meaning. Oxford: Blackwell.

Xiao, R. and McEnery, T. 2006. Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. Applied Linguistics 27/1: hlm. 103–129. Oxford University Press.

Yang, B. 2016. A corpus-based comparative study of Learn and Acquire. English Language Teaching Vol 9, No.1, pp. 209–220.

****