Design and Development of the Sentence-based Collocation Recommender with Error Detection for Academic Writing

Chih-Chien Kao[1], Jia-Wei Chang[1,2], Tzone-I Wang[1], Yueh-Min Huang[1], Po-Sheng Chiu[3]

[1] Department of Engineering Science, National Cheng Kung University, Taiwan
[2] Department of Computer Science and Information Engineering, National Taichung University of Science and Technology, Taiwan
[3] Department of E-learning Design and Management, National Chiayi University, Taiwan
tnjimgau@gmail.com, jiaweichang.gary@gmail.com, wti535@mail.ncku.edu.tw, huang@mail.ncku.edu.tw, chiups@mail.ncyu.edu.tw

Abstract

Appropriate collocations in writing of research papers in English can make the context smoother and expression of ideas more precise. In consequence, it is easier for the reader to understand and the purpose of sharing the outcome of research is accomplished. However, for non-native English speakers, the choice and use of collocations is very difficult. For this reason, this study is intended to refer to a large amount of high-quality academic literature to establish a collocation corpus and adopt natural language processing techniques and statistical methods to develop a collocation recommendation system. The system will allow users to enter sentences, automatically detect the locations and types of collocations and recommend synonymous collocations in accordance with the semantics and frequency of use. The fitness of collocations in the system for beginning sentences achieves 73.1%. Writers of academic papers can use the system to select appropriate collocations, reduce erroneous use of collocations and improve the quality of their papers.

Keywords: Collocation recommender, Collocation corpus, Academic writing assistance

1 Introduction

Universities and graduate institutes in Taiwan have realized the importance of assuring the academic writing ability of their students and researchers can meet international standards and have therefore established related indicators. For non-native English speakers, however, producing a succinct and comprehensible academic paper is by no means an easy task. Choice of vocabulary has an immense effect on the fluency of writing in English [1]. It is a big challenge for non-native English speakers to be articulate when writing in English [2]. Above all, use of collocations is very important. Besides connecting the words and contents in a piece of writing, collocations can also help enhance the clarity of contents and the ideas the writer wishes to express when they are used properly [3].

Presently, most collocation recommendation systems recommend collocations in accordance with the keywords entered, including COCA [4] (http://corpus.byu.edu/coca/), Just-The-World (JTW) [5] (http://www.just-the-word.com/), Tango [6] (http://candle.cs.nthu.edu.tw/collocation/) and OZDIC [7] (http://www.ozdic.com). Nonetheless, many non-native English speakers are likely to misuse collocations or unable to identify appropriate collocations. Therefore, in this study, a sentenced-based collocation recommendation system equipped with the function to detect errors will be developed to help writers of academic papers use correct and suitable collocations to improve the quality of their papers. Users only need to enter an English sentence and the system will automatically mark out all the collocations and offer a set of synonyms for each collocation for users to understand which collocations can be applied as well as choose correct collocations from the synonyms recommended.

The collocations commonly adopted in different fields may differ. In this study, the focus is set on the field of engineering and collocations used in high-quality journal papers collected in the Science Citation Index are adopted to be the language material for academic papers. The collocations applied in such papers are extracted and classified and the frequency of use of each collocation is calculated to build a collocation corpus based on which the collocation recommender will be developed.

*Corresponding Author: Po-Sheng Chiu; E-mail: chiups@mail.ncyu.edu.tw
DOI: 10.3966/160792642019012001021

2 Background

2.1 Definition of Collocation

Lewis defines collocation as "certain words co-occur in natural text with greater than random frequency" [8] whereas the definition in the Oxford Collocations Dictionary suggests, "collocation is the way words combine in a language to produce natural-sounding speech and writing" [9]. Based on the abovementioned definitions, it can be generalized that a collocation is a combination of words that tend to appear together more frequently than other word combinations.

According to definitions, a collocation can normally be divided into a base word and collocate(s). In a collocation dictionary, many sets of collocates can be found to match a base word [10]. For this reason, most collocation recommendation systems will recommend several search results in accordance with the base word entered. In recent years, there have been many discussions about use of collocations in teaching and learning of writing in English. More and more researchers have come to be aware of the importance of collocations in academic writing [11] because academic researchers can end up writing equivocal sentences if they lack the ability to use collocations correctly and this will have a deep impact on dissemination of knowledge. Hence, appropriate use of collocations is very important in academic writing when a researcher wishes to give precise descriptions.

2.2 Techniques of Natural Language Processing

Natural language processing techniques need to be applied in analysis of language materials and establishment of a corpus. In this study, the Stanford Parser [12] developed by the Stanford Natural Language Processing Group is adopted to obtain the syntactic information of sentences, as shown in Figure 1 in which "dependencies" represents the relations between words. There are 48 types of such relations [12]. This study is conducted based on the assumption that such relations can be utilized as the basis of collocation classification. In other words, identification and extraction of collocations are performed in accordance with the various relations formed as a result of dependencies. Meanwhile, the Stanford Parser also indicates the part of speech of each word and this is useful in stemming and semantic analysis [13].

The purpose of stemming is to reduce the three parts of speech and plural form of a word to the root. This can make calculation of collocations more accurate. For example, "went" and "goes" come from the word "go." The three words have the same meaning but appear in varied forms. In calculation of word appearance frequency, however, they are three different words. In consequence, errors can occur in calculation of collocations. Therefore, "went" and

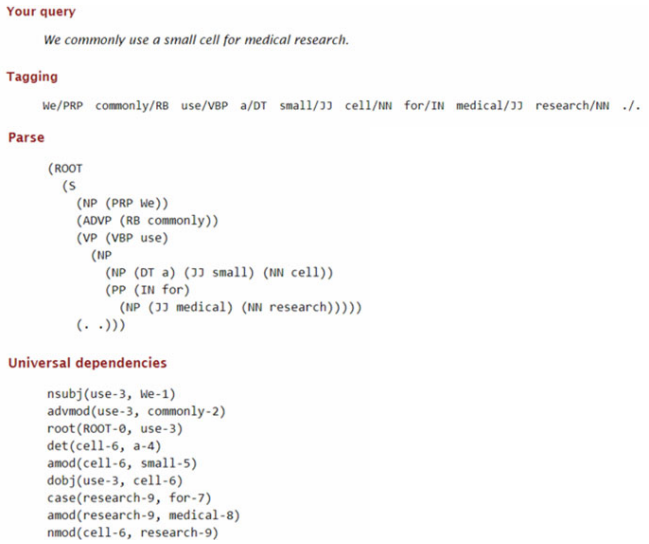

Figure 1. Syntactic Information from Stanford Parser

"goes" must be reduced to the root "go," so that all three of them will be considered the same word and calculation of collocations will become more accurate.

WordNet [14] is a large electronic dictionary developed by a research team from Princeton University. The vocabulary collected in the dictionary is categorized according to parts of speech, namely noun, verb, adjective and adverb.

2.3 Collocation Retrieval Methods

Use of n-grams is a rather classical approach to find collocations. In the Xtract collocation search system developed by Smadja [15], for example, means and variance are adopted. In the English language materials, the text files in a defined range are scanned and three filter conditions are applied to extract bi-gram vocabulary combinations from combinations in non-continuous locations. Initially, the five words before and after (-5 and +5 around the target word) are respectively combined with the target word to form 10 bi-gram collocations as well as calculate the frequency of their co-occurrence. Then the bi-gram combinations with frequency higher than the threshold value are extracted while the distance between the two words of each word combination is also recorded and word combinations with low frequency of co-occurrence are filtered out.

Next, the distribution of the locations of collocates in all the word combinations before and after the target word is calculate to establish the variance of the collocates and retain only collocates with larger variance values. This means these collocations are grammatically sound. If the collocates are too far apart and the frequency of appearance is average, it suggests such words may appear together sometimes, but it can be coincidental and such combinations are not common grammatical structures. Hence, such word combinations can also be filtered out.After filtering, the remaining bi-gram word combinations are extended to the n-gram range to test whether these word

combinations are collocations with high co-occurrence. In the end, based on the part of speech of the two words in the collocations, several collocations can be formed, including verb-object, subject-verb, adjective-noun, etc. In other words, the collocation to be adopted is the word combination with the highest co-occurrence.

Some studies [16] have used Mutual Information (MI-score) methods to extract collocations. The significance of the MI-score is to assess the correlation between every two random variables. It can be applied to analyze the frequency of co-occurrence of two words to evaluate the likelihood of such a word combination being a collocation. As shown in Formula (1), the *x* and *y* respectively represent the frequency of co-occurrence of the two words. *p(x)* is the probability of *x* appearing by itself and *p(y)* the probability of *y* appearing by itself, and *p(x, y)*stands for the probability of*x* and *y* appearance together. By using this formula, the correlation of the *x* and *y*words can be established.

$$I(X;Y) = \sum_{x,y} p(x,y) \ln \frac{p(x,y)}{p(x)p(y)} \quad \textbf{(1)}$$

If two certain words seldom appear in training language materials, but they often appear together, it will cause the MI-score to be high and lead to the misjudgment of their being a collocation. However, the t-score can be applied to test the intensity of correlation of the two words in a collocation to deal with situations in which two words seldom appear together in training language materials, but the correlation is high. The t-score formula is as follows:

$$t(u,v) = 1 - \frac{f_u \cdot f_v}{f_{uv}} \quad \textbf{(2)}$$

In Formula (2), f represents the frequency of appearance of a word in a combination. f_u and f_v are respectively the frequencies of appearance of words u and *v* in training language materials. f_{uv} is the frequency of *u* and *v* appearing together in training language materials [17]. In order to increase the accuracy in identification of collocations, the correlation ratings of word combinations with the MI-score and the t-score are consolidated [18] and a frequency threshold value is defined to filter out word combinations with correlation values that are too low.

3 System Design

This study proposes a set of methods to set up an academic collocation database and develop a collocation recommendation system that allows users to search for collocations by entering sentences. To set up the database, a large amount of academic paper language materials is analyzed and natural language processing techniques are applied to establish the relations between syntactic information and word combinations to acquire collocations from the language materials. The interface installed on the collocation recommendation system allows users to enter keywords or sentences to look for collocations.

3.1 Sources of Language Materials

The training language materials are selected from 3,164 papers accepted in the five recent years by a number of top journals collected in the Science Citation Index.

3.2 Establishment of the Collocation Database

First, Stanford Parser is adopted to perform syntactic analysis on every sentence in the papers to establish the dependencies in each sentence and the part of speech of every word in the sentence. The dependencies indicate the relations between the words in a word combination and the locations of the words in the sentence and there are 48 word combination relations. According to the definition from [19], there are eight types of common collocations: noun+preposition, preposition+noun, adjective+preposition, verb+noun, adjective+noun, noun+verb, adverb+adjective and verb+adverb. Therefore, in this study, the compositional relations of dependencies are adopted to correspond to these eight common types of collocations, as shown in Table 1.

Table 1. Corresponding Relations between Dependencies and Collocation Types

Dependencies	Description	Collocation
advmod (adverbial modifier)	An adverb modifying the word following it	Adv. + Verb
amod (adjectival modifier)	An adjective modifying the noun following it	Adj. + Noun
dobj (direct object)	A verb followed by noun	Verb + Noun
nsubj (nominal subject)	A noun or subject followed by a verb, adjective or another noun	Noun + Verb Noun + Adj. Noun + Noun
pobj (object of a preposition)	A preposition followed by a noun	Prep. + Noun
prep (prepositional modifier)	A preposition in collocation with another word or words	

According to Table 1, dependencies are used to extract collocations from sentences. Then, stemming is conducted on the words in each combination and the part of speech of each word is also indicated. Next, the information on the collocations and the frequency of their appearance are updated in the collocation database. This procedure is repeated to analyze the sentences in all the literature to complete establishment of the collocation database.

3.3 Use of Sentences to Seek Collocation Recommendations

The proposedsystem is a web-based tool and it accepts the use of sentences to run searches. It can automatically detect the collocations in the sentence keyed in and mark the locations. When the user clicks on a marked collocation, the system will display other collocations recommended as shown in Figure 2.

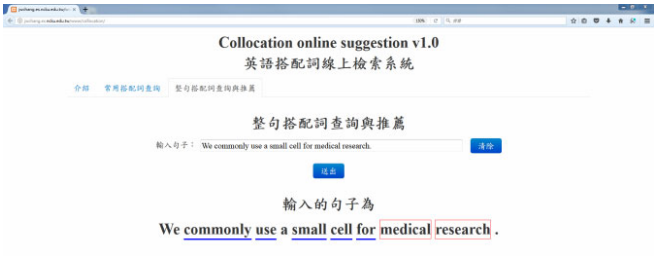

Figure 2. Web-based user interface of the proposed system

After an example sentence (e.g., "We commonly use a small cell for medical research.")is entered, the system will send it to Stanford Parser to analyze the dependencies and parts of speech. The information on the dependencies is enough for the system to identify the compositional relations in the collocations and the locations of the collocations and mark the collocations in the sentence for the user to click on any collocation to find other collocations recommended. The system can recommend collocations in three modes:

Mode 1: The word clicked on to be the base word (e.g., the word, "commonly", in theexample sentence) and the part-of-speech of the collocate are adopted to search for common collocates (commonly + V/Adv, for example, as shown in theleft part of Figure 3). However, the semantics are not taken into account in this mode. Only words that match the part of speech of the base word are recommended and the frequency of their use in collocation with the word commonly is applied as the basis of recommendation sequencing.

Figure 3. Recommendations from the proposed system

Mode 2: To recommend collocates that can be applied to replace the original collocate, the meanings of words have to be considered. Under such circumstances, the system will recommend synonyms of the original collocate (e.g., the combination, "commonly+use", in the example sentence) to be applied with the word commonly to form synonymous collocations(see the middle part of Figure 3). The frequency of such words appearing with commonly will be indicated to help users make their choices.

Mode 3: To provide more synonymous collocations that can be adopted to replace the original combination(e.g., "commonly+use"), the synonyms of "commonly" and "use" are both selected to form more recommended synonymous collocations while the frequency of appearance of each combination in the language materials is also indicated for the user's reference in choice of substitutive collocations(see the right part of Figure 3).

Correct use of collocates is very important for learners of English as a foreign language, but they often find it difficult [20] because combinations of English words are very diverse and complicated. Hence, it is indeed a tough challenge for them to use collocations correctly in sentences they write. For this reason, in this study the frequency of collocated word combinations is calculated to assess the reliability of the collocations in the collocation database of this system. In other words, by setting a co-occurrence frequency threshold, it is possible to judge whether the collocated combinations appear in sentences are rare usages or misuses. As shown in Figure 3, the system marks out rare and misused collocations with red frames to suggest the frequency of co-occurrence of such word combinations is low and the correctness may be doubtful. Therefore, collocations recommended offered in this system are arranged in descending order, from high to low, according to the frequency of co-occurrence of the word combinations found.

4 Experiments

To define the system's error-detection threshold value and test the results of collocation recommendation, this study needs data to run tests on sentences entered for collocation searches and also identify wrong collocations. In the following, the sources of data adopted for testing, the performance evaluation indicators defined and the experiment conducted to verify the results of recommendation will be explained.

4.1 Tests of the Threshold Value in Detection of Erroneous Collocations

Artificially created sentences with wrong collocations are used in this study as the data to be tested. The frequency of appearance of wrong collocations in the collocation database is calculated to determine the threshold value in assessment of wrong collocations in order to accurately detect rare or misused collocations. The data used for testing in the

experiment are selected from 15 papers accepted by top journals, including the Knowledge-based Systems, Neurocomputing and Expert Systems with Applications. 1,000 sentences are chosen for five experienced English teachers to produce sentences containing misused collocations and a total of 1,874 wrong collocations are generated. The frequency of appearance of these wrong collocations in the collocation database established in this study and the percentages they account for are as shown in Figure 4 and Figure 5.

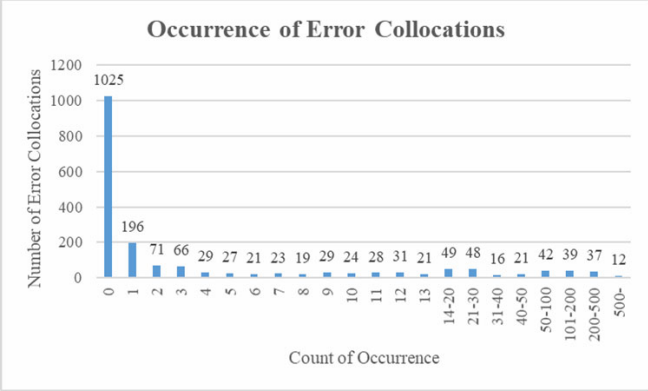

Figure 4. Occurrence of Error Collocations

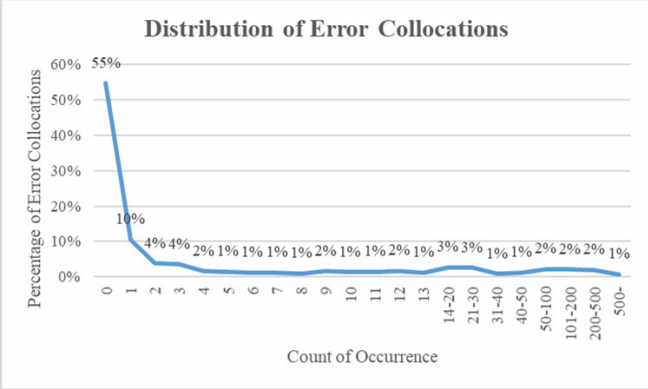

Figure 5. Distribution of Error Collocations

Figure 4 shows most of the wrong collocations never appear in the collocation database developed in this study. In Figure 5, the percentages of the wrong collocations are all lower than 4% after they appear three times and the total wrong combinations appearing between 0 and three times account for 73% of the total wrong combinations. Hence, three times of appearance is deemed an appropriate threshold in detection of wrong collocations in this study.

4.2 Tests of Sentence-based Collocation Recommendation

In this study, an experiment is designed to understand users' satisfaction with the substitutive collocations recommended and the system's accuracy after they enter sentences. In the experiment, five experts analyze and assess the correctness of the collocations recommended after they enter sentences in the system.

To start, each of the five experts selects three academic papers from the journals specified and picks out two hundred sentences from the three papers. One thousand sentences are collected in total as test data which include 2,463 collocations. Next, each expert enters the two hundred sentences one after another to search for recommended collocations and also record the results and their quality. The five experts need to evaluate the results achieved through the three different recommendation modes.

1. Is the first recommended collocation obtained through Mode 1 the best collocation recommendation?

In Mode 1 of this system, the search for collocations is conducted according to the word clicked on by the user as the base word and the part of speech of its collocate, such as the commonly + V/Adv shown in Figure 3. The five experts must evaluate whether the first collocation, compared to the other collocations recommended, is the best one. In the experiment, the five experts evaluate the search results for the 2,463 collocations in the one thousand sentences entered. It turns out that the first collocation in 1,801 searches is the best recommendation.

Table 2 shows the rate of optimal fitness achieves 73.1% (1,801/2,463). Since users often choose collocations in accordance with the recommendations from such systems, the correctness of the first collocation recommended is extremely important. Therefore, the higher the correctness of the fist recommendation, the more reassured academic writers will be to use the first collocation recommended by the system to improve the quality of their writing.

Table2. Results of First Recommendation

	Number of Accepted First Candidate	Number of Test Collocations	Percentage (%)
Expert 1	342	536	63.8%
Expert 2	376	515	73.0%
Expert 3	277	454	61.0%
Expert 4	466	493	94.5%
Expert 5	340	465	73.1%
	1,801	**2,463**	**73.1%**

2. How many synonymous collocations found through Mode 2 can be applied to replace the original collocation in the sentence keyed in?

In Mode 2 of this system, collocation recommendation is executed according to the word clicked on by the user as the base word and the synonyms of the collocate. The five experts have to evaluate how many combinations can be formed with the synonymous collocates found and applied directly in the sentence entered. There are one thousand sentences serving as the test data and they contain 2,463 collocations in total. In the 2,463 searches, Mode

2 offers 7,782 synonymous collocated combinations. The outcome of the experiment shows the five experts accept 3,062 of the 7,782 synonymous collocations recommended by the system as correct and applicable to replace the original collocations in the sentences keyed in without changing the semantics. In other words, about 60.6% (3,062/7,782) of the collocations generated by the system with the synonyms of the original collocates in the sentences entered are correct.

However, 60.6% means there is still room for improvement. The outcome is not so great mainly because some of the words have multiple definitions and the numbers of synonyms are large. In consequence, the system may come up with collocations that appear semantically inconsistent in the context of the sentence. For instance, the word "have" has several different meanings and the synonyms found total around 30. Therefore, using the collocations generated with the synonyms of "have" can cause semantic inconsistency. In other words, if the contextual meaning of the word in the sentence is defined first, collocated combinations semantically consistent with the sentence entered will be accurate and the correctness of the collocations recommended in Mode 2 will be improved.

3. How many synonymous collocations found through Mode 3 can be applied to replace the original collocation in the sentence keyed in?

In Mode 3 of the system, collocation recommendation is done according to synonyms of the word clicked on by the user as the base word and the collocate in the sentence entered to offer more diverse collocations. The five experts need to evaluate how many combinations can be formed with the synonyms found and applied directly in the sentence. There are one thousand sentences serving as the test data and they contain 2,463 collocations in total. 18,625 synonymous collocations were found through Mode 3. The outcome of the experiment shows the five experts accept 9,638 of the 18,625 synonymous collocations recommended by the system as correct and applicable to replace the original collocations in the sentences keyed in without changing the original semantics. In other words, about 51.7% (9,638/18,625) of the collocations generated by the system based on the synonyms of the base words and the original collocates in the sentences entered are correct.

The ratio of synonymous collocations that are generated through Mode 3 and can be applied to replace the original collocations is only 51.7%. The reason, as concluded in this study, is the same as in Mode 2. Since the synonyms of two words are used to produce more synonymous collocations and each of the two words can have several definitions, it is likely that more collocations contextually inconsistent in the sentences can be formed. Therefore, to improve the correctness of the collocations generated through Mode 2 and Mode 3, it is necessary to define the semantic meaning of the word in the context first in order to find collocations semantically consistent with the rest of the sentence entered.

5 Conclusion

There are many collocation search systems available, but most of these systems have not been specifically designed for the academic sector. Plus, most of them require keywords to be entered. So far, there is no system that accepts users to search for collocations by entering sentences.

This is why a system able to allow users to enter sentences, mark the collocations and offer collocation recommendations is developed in this study. In the meantime, training language materials are also collected from academic papers in the engineering field while Stanford Parser is adopted for syntactic analysis, indication of part of speech and WordNet is also applied to find roots and synonyms to establish a collocation database to develop an online collocation search system.

The collocation search system developed in this study is also equipped with the function to detect wrong collocations in sentences entered. If a sentence includes any collocation with excessively low frequency of use, the system will mark it out to remind the user to do something about it. In addition, this system also recommends collocations that can be used to replace the original collocation in the sentence entered. Writers can choose from the collocations recommended according to the context of the sentence to polish their wiring.

However, there is still room for improvement in some areas that further studies can specifically work on. The following are some suggestions for researchers performing similar studies in the future:

Expansion of the corpus. Compared to other collocation search systems, the collocation corpus established in this study has fewer training language materials. This can make the reliability of the recommendations offered questionable. Meanwhile, the contents of the corpus mainly come from academic papers in the engineering field which is not the only field in the academic sector; therefore, academic papers from more fields can be collected to expand the scale of the corpus, so that such systems can offer more correct and diverse collocations.

The correctness of synonymous collocations recommended. It is discovered in the aforementioned experiment that synonymous collocations can be semantically inconsistent because the words in a word combination may have multiple definitions and searches for synonyms are launched in accordance with all the meanings of such a word. If the meaning of a word in the sentence entered is clearly defined, the system will be able to come up with correct synonymous collocations.

Enhancement of the error detection mechanism. In the system developed in this study, error detection is conducted based on the frequency of co-occurrence of words. If the frequency of appearance of a collocation is lower than the threshold value established, it will be considered a rare or wrong collocation. Nonetheless, whether there is something grammatically wrong or illogical is not taken into consideration. If the functions of syntactic detection and semantic logic assessment can also be incorporated to help users modify their sentences, such a system will be able to help academic writers express themselves more precisely and fluently.

Acknowledgements

This work was supported by the Ministry of Science and Technology, Taiwan, R.O.C., under contract no. MOST105-2511-S-415-017-MY3.Special Thanks are expressed to the participating experts.

References

[1] R. L. Leed, A. D. Nakhimovsky, Lexical Functions and Language Learning, *The Slavic and East European Journal*, Vol. 23, No. 1, pp. 104-113, Spring, 1979.

[2] J. Bahns, Lexical Collocations: A Contrastive View, *ELT Journal*, Vol. 47, No. 1, pp. 56-63, January, 1993.

[3] P. Howarth, Phraseology and Second Language Proficiency, *Applied Linguistics*, Vol. 19, No. 1, pp. 24-44, April, 1998.

[4] M. Davies, *The Corpus of Contemporary American English, Brigham Young University*, http://corpus.byu.edu/coca/.

[5] Just-The-World, http://www.just-the-word.com/.

[6] J.-Y. Jian, Y.-C. Chang, J. S. Chang, TANGO: Bilingual Collocational Concordancer, *Proc. of the ACL 2004 on Interactive Poster and Demonstration Sessions*, Barcelona, Spain, 2004, Article No. 19.

[7] OZDIC, http://www.ozdic.com/

[8] M. Lewis, Implementing the Lexical Approach: Putting Theory into Practice, *The Electronic Journal for English as a Second Language*, Vol. 3, No. 1, pp. 223-232, November, 1997.

[9] P.-S. Chiu, Y.-M. Huang, The Development of a Decision Support System for Mobile Learning: A Case Study in Taiwan, *Innovations in Education and Teaching International*, Vol. 53, No. 5, pp. 532-544, September, 2016.

[10] Y.-C. Chen, T.-X. Yen, J. S. Chang, Associating Collocations with WordNet Senses Using Hybrid Models, *Proc. of the Twenty-Fourth Conference on Computational Linguistics and Speech Processing*, Chungli, Taiwan, 2012, pp. 47-60.

[11] J.-C. Wu, Y.-C. Chang, T. Mitamura, J. S. Chang, Automatic Collocation Suggestion in Academic Writing, *Proc. of the ACL 2010 Conference Short Papers*, Uppsala, Sweden, 2010, pp. 115-119.

[12] D. Klein, C. D. Manning, Accurate Unlexicalized Parsing, *Proc. of the 41st Annual Meeting on Association for Computational Linguistics*, Sapporo, Japan, 2003, pp. 423-430.

[13] J. W. Chang, M. C. Lee, T. I. Wang, Integrating a Semantic-based Retrieval Agent into Case-based Reasoning Systems: A Case Study of an Online Bookstore, *Computers in Industry*, Vol. 78, pp. 29-42, May, 2016.

[14] G. A. Miller, WordNet: A Lexical Database for English, *Communications of the ACM*, Vol. 38, No. 11, pp. 39-41, November, 1995.

[15] F. Smadja, Retrieving Collocations from Text: Xtract, *Computational Linguistics*, Vol. 19, No. 1, pp. 143-177, March, 1993.

[16] K. W. Church, P. Hanks, Word Association Norms, Mutual Information, and Lexicography, *Computational Linguistics*, Vol. 16, No. 1, pp. 22-29, March, 1990.

[17] A. Thanopoulos, N. Fakotakis, G. Kokkinakis, Comparative Evaluation of Collocation Extraction Metrics, *Proc. of the Third International Conference on Language Resources and Evaluation*, Las Palmas, Canary Islands- Spain, 2002, pp. 620-625.

[18] K. Ackermann, Y.-H. Chen, Developing the Academic Collocation List (ACL)– A Corpus-driven and Expert-judged Approach, *Journal of English for Academic Purposes*, Vol. 12, No. 4, pp. 235-247, December, 2013.

[19] M. Benson, E. Benson, R. F. Ilson, *The BBI Combinatory Dictionary of English: A Guide to Word Combinations*, John Benjamins, 1986.

[20] H.-J. H. Chen, Developing and Evaluating a Web-Based collocation Retrieval Tool for EFL Students and Teachers, *Computer Assisted Language Learning*, Vol. 24, No. 1, pp. 59-76, January, 2011.

Biographies

Chih-Chien Kao received the M.S. degree in Department of Engineering Science from National Cheng Kung University, Taiwan. He is a Ph.D. student on Engineering Science at National Cheng Kung University, Taiwan. He is also the chairman of Information Center, Bureau of Education, Tainan City Governmen.

Jia-Wei Chang received his Ph.D. degree at the Department of Engineering Science in the National Cheng Kung University in 2017. In 2012, he graduated with a M.S. in engineering science from National Cheng Kung University. His research interests include natural language processing, data mining and e-learning technologies.

Tzone-I Wang is a professor of engineering science at National Cheng Kung University, Taiwan. He held B.S. of engineering science at National Cheng Kung University and Ph.D. degree at Imperial College, University of London, UK. His major research areas include artificial intelligence, web and network services, and e-learning technologies.

Yueh-Min Huang is a Chair Professor in the Department of Engineering Science, National Cheng-Kung University, Taiwan. His research interests include e-learning, wireless networks, artificial intelligence. He received his M.S. and Ph.D. degrees in Electrical Engineering from the University of Arizona in 1988 and 1991 respectively.

Po-Sheng Chiu is currently an assistant professor of the Department of E-learning Design and Management, National Chiayi University, Taiwan. He received his Ph.D. degree in Department of Engineering Science from National Cheng Kung University, Taiwan, in 2013. His research interests include Educational Technology, Mobile Learning.