

Title	母国語の語義干渉の低減による適切な語義選択支援に関する研究
Author(s)	Kritsuthikul, Nattapol
Citation	
Issue Date	2017-09
Type	Thesis or Dissertation
Text version	ETD
URL	<a href="http://hdl.handle.net/10119/14830">http://hdl.handle.net/10119/14830</a>
Rights	
Description	Supervisor:長谷川 忍, 情報科学研究科, 博士

氏名	NATTAPOL KRITSUTHIKUL		
学位の種類	博士(情報科学)		
学位記番号	博情第 374 号		
学位授与年月日	平成 29年 9 月 22 日		
論文題目	Improving proper word meaning selection by reducing interference of word sense in native language (母国語の語義干渉の低減による適切な語義選択支援に関する研究)		
論文審査委員	主査	長谷川 忍	北陸先端科学技術大学院大学 准教授
		東条 敏	同 教授
		Nguyen Minh Le	同 准教授
		Thepchai Supnithi	National Electronics and Computer Technology Center Principal Researcher
		Cholwich Nattee	Sirindhorn International Institute of Technology, Thammasat University 准教授

## 論文の内容の要旨

In Thailand, English is learned in a school for foreign language. It has been taught since elementary school and so on. Despite of learning for several years, English proficiency of Thais is still in unacceptable level. From all communication skills, writing skill is the most important skill since it is a key indicator for official communication. As writing tasks become more difficult in higher education level, students must apply a wide range of skills to write legibly and logically. This combination of requirements makes writing the most complex and difficult use of language. From observation, English writing skill of Thai students is one of the most troublesome in learning. Their writings contain a grammatical error and misleading vocabulary. However, grammatical errors in written text are explicitly detectable and easier to comprehend comparing to vocabulary issue. Hence, this thesis studies on reasons for the vocabulary error of Thai students in English learning and proposes a method/tool to improve the problem.

To find a cause of vocabulary errors, a corpus of English essay written by Thai students is used as source for analysis. From accessible resources, the corpus from the New EAGLE v3 was selected as a base for the error analysis. It contains 320 essay documents, and incorrect parts in the documents were manually tagged with error type. From my own analysis, the issues are grouped into three types as global comprehension, syntactic and semantic issue. Although the tag set is heavily focused on syntactic level, I found that the documents with less comprehensible are those with the tag of “Wrong words” and “L1 thought patterns (word-by-word translation)”. The incorrect part in these documents is about selecting an inappropriate word to represent an intended concept and leading to misunderstanding or incomprehensible. From deeper analysis, I found that the incorrect words of these types from most documents were not totally off when translated back to their native

language, Thai. The reverse-translation to Thai of the text is understandable though it is hardly understandable in English counterpart, as the chosen translated word does not semantically fit in the context. This shows that English learners did become not aware or realize about homonym (different meaning with the same surface form) in their native language (L1) while trying to translate to target language (L2). Therefore, it conclusively indicates that homonymy of L1 word is one of the cause for misusing English words.

To help students understand about L1 homonym, a confusing word from L1 is a major key. Since this issue is language-dependent, Thai language is specifically investigated in this work as it is a students' native language. It is noted that the confusing words of L1 homonym in this work is not the same as common confusing word in English learning English because the cause of confusion is not from L2 (the target language, English) but multi-sense from L1 part in students' thought in writing process. The words in the list are collected from a frequently used Thai word containing many senses and appeared in the above-mentioned corpus causing the error.

The tool is implemented to give hints on L1 homonym once students use a word from the list of confusing words of L1 homonym. The hints in this work include possible translations of the L1 words and collocation statistic of the L2 words. The possible translations of the L1 words are generated from an existing Thai-English bilingual dictionary, LEXiTRON. The confusable word in English show all possible Thai translation words to help them realize that the word is ambiguous in their native language. In addition, those Thai words are attached with their English translations as a guidance for other possible words in a case that the students do not know about the English translation of a given new sense. This hint is generated in graphical style as word translation relation. Another hint is given in a Pie Chart to inform collocation statistic of the L2 word. This Pie Chart will hint students about how many the chosen word compositions are used in general. It is the fact that if students compose a sentence containing a word with wrong sense, the frequency of those consecutive words will be noticeably low. The Pie Chart is generated in focusing on the confusable word and their surroundings. The statistics are from the hit rate of search engines including Google and Bing. The last hint is Text Hint which generates from suggestions between the word relation of hypernym and hyponym in WordNet. The Text Hint represents to the students as a text in order to provide the most appropriate for the students not to be too general and specific. With these three hints, it is expectable that students will realize and learn the correct word sense from L1 to L2 and select proper word in their English writing.

To evaluate the tool, an experiment to see students' improvement in vocabulary selection and a questionnaire about using the tool were conducted. 210 Thai students in Grade 8 in the same school participated in the experiment. All the students were asked to answer the vocabulary knowledge test which consists ten questions with given multiple choices for the confusable words before doing the main test. The Thai word for the missing part was given to represent intended concept meaning. The Thai word, though, was a confusable L1 homonym, and the choices were the possible translations of the confusable Thai word. When the students conducted the test once, the proposed tool was given to them to change the answer freely. A comparison between the pre-using and with-tool was collected to see the improvement of the answer results. In the testing, the samples

were grouped to 14 testing groups. The groups were to find the effect on different hint types from the tool. Moreover, the questionnaire was given after the with-tool experiment for the students to inform feeling about the tool.

The experiment results in overall showed that around 30% of the total answers were adjusted while they used the tool, and about 78% of the changed answers were to change from incorrect to correct answer. Regarding the hint types, the word relation hint performed best to convince them to correct the wrong answer the most, and the Pie Chart hint came in second. A one-way ANOVA was conducted to compare the effect of the hint types. There was a significant effect of amount of hint types at the  $p < 0.05$  level for the seven conditions [*Welch's F*(6, 89.723) = 5.199,  $p < 0.001$ ]. Post hoc comparisons using the Games-Howell test indicated that the mean score for the word relation hint condition was significantly different. The results can also be implied that the students with moderate proficiency or lower were more affected by the hints than those with high proficiency. In addition, no case was found in the experiments that the students chose to change from a correct answer to incorrect answer with the given hints. From the results, it showed that the tool can help the students to realize a different in word sense selection based on ambiguity of L1 homonym effectively. Moreover, it can be used to give vocabulary information for the students to learn new meanings of the confusable word.

From the questionnaire, the students mentioned that the hints were effective to give them a clue in their word selection. The Relation of word graph received the best favorite while the word relation came in secondly. 77% of the students found the tool was useful, and their main reason was that it could give extra information of ambiguity in words across languages. Moreover, they stated that the hints of the tool were easy to understand.

In conclusion, this study reveals that homonym from L1 can greatly cause a misusing L2 word in English writing. To solve the issue, the tool is implemented to clarify the ambiguity from L1 with the hints including all word relations of confusable L1 word and Pie Chart of collocation statistic of L2. The word relation hint helps Thai students to be aware of possible meanings of L1 and realize their intended meaning before selecting an English word. The collocation hint assists them to crosschecking a possibility of word compositions and to observe context more carefully after choosing the English words.

Keywords: English as a Foreign Language (EFL), L1 interference, Meaning Selection, Vocabulary Acquisition, Computer-assisted Vocabulary Instruction (CAVI)

## 論文審査の結果の要旨

本論文は、ESL (English as Second Language)における学習支援の一環として、タイの中学生を対象に、適切な意味の英語を選択する際にしばしば起こる母語の干渉を低減するための学習

支援システムを提案するとともに、開発したプラットフォームの有効性について評価・考察したものである。

ESL の学習における課題は従来から様々なものが挙げられているが、本論文では、特に文法構造や語彙が大きく異なるタイ語における誤りを分析し、スペルミスに続く大きな誤りとして母語の干渉があることを実証的に示した。さらに、一つのタイ語に対して複数の英単語に翻訳可能な語を集めた **Confusing Word List** と検索エンジンのヒットレートを利用して、それらの語に対する学習者の誤りを検出するとともに、(1)正解と間違いの語の間の概念間の上下関係のテキストによる提示、(2)**Tri-gram** を利用したヒットレート結果の円グラフによる表示、(3)双方向辞書によるタイ語-英語-タイ語の関係チャート表現、をヒントとして提示できる学習支援プラットフォームを開発した。

また、提案手法の評価として、タイの中学 2 年生 15 名を対象として **Confusing Word List** に含まれるタイ語を含む英語のライティング課題を行わせ、誤りの検出率を調査したところ  $F$  値 = 0.91 という結果となり、一定の妥当性があることが示された。さらに、同じくタイの中学 2 年生 210 名を品詞情報の有無およびヒントの組み合わせ 7 パターンの 14 グループに分け、**Confusing Word List** の語を含む文章の翻訳時に適切な英単語を選択させる課題を実施したところ、品詞情報の有無が正解率の向上に大きな影響を与えないこと、(3)関係チャート表現を含むヒントグループの正解率向上がテキストのみのグループと比較して有意に向上することを示した。

以上、本論文は、ESL 学習者の単語選択において母語の干渉を低減させることを目的とした学習支援プラットフォームを開発することで、学習者単独では気が付くことが困難な課題を解決するための足掛かりを得たものであるといえる。本論文でまとめた **Confusing Word List** はタイの中学 3 年生までの英語教科書で現れ、4 語以上の翻訳可能性がある 84 セットに留まっているが、新たなリストを作成するだけでより同様のヒントを自動的に生成できる点で、ESL 語学学習支援研究の実質化の観点から学術的に貢献するところが大きい。よって博士 (情報科学) の学位論文として十分価値あるものと認めた。