

# Developing a Vocabulary Learning System with Near-Synonyms and Similar-Form Words

Ming-Chi Liu<sup>a</sup>, Koun-Tem Sun<sup>b</sup>, Kun-Te Wang<sup>c</sup>, Yuen-Min Huang<sup>d</sup>

<sup>acd</sup>*Department of Engineering Science, National Cheng Kung University, Taiwan*

<sup>b</sup>*Institute of Computer Science and Information Education, National University of Tainan, Taiwan*

huang@mail.ncku.edu.tw

**Abstract:** Various vocabulary learning systems have been developed to assist students of English as a Foreign Language (EFL) in acquiring the knowledge of vocabulary. By using the electronic dictionary, people only acquire vocabularies on a superficial. Most corpus-based system focuses on lexical explanation, such as giving synonyms and antonyms. Only a few studies have attempted to use lexical structures. However, these systems do not take into consideration students of EFL who emphasized English words recognition, such as Chinese students. Therefore, the near-synonyms and similar-form (NSSF) words are difficult for them to learn. In this study, we propose a system to assist students of EFL in learning the NSSF words. Results of a pilot study showed that students regarded this NSSF word learning system to be of great assistance in learning English vocabulary. This encourages us to promote the NSSF words learning system to students of EFL in other countries.

**Keywords:** Vocabulary learning, Discovery learning, WordNet, Near-synonyms, Similar-form words

## Introduction

Two NSSF words can be defined as two similar-form words which have a similar meaning. However, they are not synonyms, and thus cannot be substituted one for the other. For example, the two words “transform” and “transfer” are NSSF words. First, they have the same prefixes “transf.” Second, they share the same meaning of “change”. But a more precise look reveals that they are only near-synonyms and have different concepts. The word “transform” is defined as “change somebody or something completely”. And the word “transfer” is defined as “move from one place to another.” The Chinese learner will likely confuse their meaning due to learning groups of NSSF words together. The focus, then, of this research is to develop a semantically-related NSSF word learning system, and then to help Chinese students practice English vocabulary.

### 1. Method for finding near-synonyms and similar-form words

The recommended method is shown in Figure 1.a. (1). Part of speech tagging (POS tagging): A POS tagger will mark up each word to a particular part of speech. Brill’s tagger[1] is used in this procedure. (2). Stop word check: this procedure removes words that have no particular meaning in the text. (3). Stem: this procedure uses Porter Stemming algorithm[3] to reduce the derived word to its stem. (4). Word sense disambiguation

(WSD): this procedure finds out the most appropriate sense of a word that is used in a given sentence. Lesk's algorithm [2] is often used in WSD. (5). Finding near-synonym: there two step in this procedure. One is breadth-first search (BFS) that to find neighboring words around the querying word. The another one is Wu and Palmer's semantic similarity [4] that to measure the similarity between the querying word and the neighboring words. (6). Finding similar-form words: the two most similar-form words are found by computing the Levenshtein distance. (7). Match examples and definitions of NSW: this procedure retrieves definitions and examples of these two words from WordNet, and displays them as matching items for the learner (Figure 1.b).

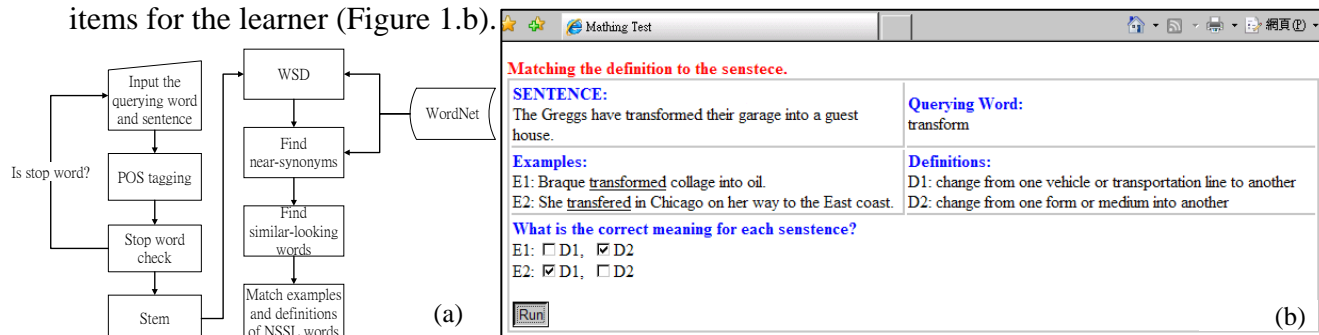


Figure 1: (a). The process of NSSF words learning system. (b). Screens of the NSSF words learning system

## 2. Conclusions

The benefits of the developed system fall into three points:

- This system proposes implementing a WordNet-based method to finding NSSF words.
- The NSSF words learning system not only helps to distinguish some confusing concepts, but also considers the difference in learning style between Chinese learners and English speakers.
- This system uses the matching of items to improve the vocabulary acquisition.

The pilot study showed that students generally regarded the NSSF word learning system as a great assistance in learning English vocabulary. It encourages us to promote the NSSF words learning system to other learners of English as a Foreign Language.

## Acknowledgements

This work was supported in part by the National Science Council (NSC), Taiwan, ROC, under Grant NSC 95-2221-E-006-307-MY3, NSC 95-2221-E-006-306-MY3, and NSC 96-2524-S-032-001.

## References

- [1] Eric, B. (1992). A simple rule-based part of speech tagger. The proceedings of the third conference on applied natural language processing (pp. 152-155), Trento, Italy.
- [2] Michael, L. (1986). Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. The proceedings of the 5th annual international conference on systems documentation (pp. 24-26). Toronto, Ontario, Canada.
- [3] Porter, M. F. (1980). An algorithm for suffix stripping. Program-Automated Library and Information Systems, 14(3), 130-137.
- [4] Zhibiao, W., & Martha, P. (1994). Verbs semantics and lexical selection. The proceedings of the 32nd annual meeting on association for computational linguistics (pp. 133-138). Las Cruces, New Mexico.