A Divergent-Style Learning Support Tool for English Learners Using a Thesaurus Diagram

Chie Shimodaira[1], Hiroshi Shimodaira[2], and Susumu Kunifuji[1]

[1] School of Knowledge Science
Japan Advanced Institute of Science and Technology
Asahidai, Nomi, Ishikawa, 923–1292 Japan
csim@jaist.ac.jp
[2] Centre for Speech Technology Research
School of Informatics, The University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, United Kingdom
h.shimodaira@ed.ac.uk

Abstract. This paper proposes an English learning support tool which provides users with divergent information to find the right words and expressions. In contrast to a number of software tools for English translation and composition, the proposed tool is designed to give users not only the right answer to the user's query but also a lot of words and examples which are relevant to the query. Based on the lexical information provided by the lexical database, WordNet, the proposed tool provides users with a thesaurus diagram, in which synonym sets and relation links are presented in multiple windows to help users to choose adequate words and understand similarities and differences between words. Subjective experiments are carried out to evaluate the system.

1 Introduction

Language learning with the use of computers has been gaining greater attention due to the increasing demand for English learning as a foreign language (EFL). Some of such examples are, computer assisted language learning (CALL), automatic language translation, and translation supporting tools, many of which are made with free use of latest natural language processing techniques. Moreover, various types of teaching materials, computer software for English learners have been developed, and they are widely available at shops.

In spite of the availability of a large number of software and hardware mentioned above, it is interesting to see that English dictionaries and thesauruses including the electronic ones are still widely used and maybe the most popular. There seems to be some reasons for this: (i) dictionaries are handy, meaning they are portable and they contain richer information than software tools; (ii) software tools have limited uses, i.e. they are designed to specific purposes, whereas dictionaries can be used in various ways. In addition, it is believed that dictionaries help learners to improve their English skills.

* This is a draft paper for KES2006. The final version will be available at the LNCS website.

Dr. Mochizuki reported his empirical findings through his experience of teaching English [1]:

- Rich vocabulary can not be acquired without consulting dictionaries.
- Teaching English words by means of the learner's native language does not help a lot.
- Words should be learnt with the contexts.

He also remarked that it would be not until the learner has learnt the pronunciation, spelling, meaning, concept, associated words, grammar, collocation, frequencies and so on that he/she really understands the word.

When we compose documents in English with the aid of dictionaries, we try to choose the right words from numbers of possible candidates and examples in order to reflect the meaning which we want to convey actually. This process of seeking and selecting the right words is not straightforward but rather of trial and error, because it involves a number of backtracking. One of the great benefits with the process is that we can learn not only the target word or phrase in English but also other expressions intentionally or unintentionally in the process.

In the present study, we at first hypothesise that learners improve their language skills through the two types of activities, *convergent-style* and *divergent-style* learning. In the former activity, learners try to narrow down the possible candidates to find the right words and expressions, i.e. the answers, by the aid of various types of resources available. On the other hand, in the latter activity, learners explore various resources not only to have a better idea of the target word they are interested in but also to acquire other words, knowledge and information that are related to the word.

Dictionaries allow learners both types of the activities, whereas most of the learning support tools allow the convergent-style activity only. These tools are designed to show the specific information for users query ignoring other additional information as noises, which might have been useful for learners to improve their skills.

Based on the hypothesis, we predict that learning support tools which enable divergent-style learning and provide users with not only the target word but also other various information relevant to the word would be useful for English learners.

As the first step to develop such a support tool, the present study at first carries out behaviour analysis of English learners to identify functions which are required for the tool.

This paper is organised with four sections. The next section describes the behaviour analysis and system design, and the section three reports experimental evaluations. The final section is devoted to the conclusions.

2 Behaviour analysis and system design

Behaviour patterns of English learners were collected at an English language school in Scotland. A class of daily English composition was chosen for the

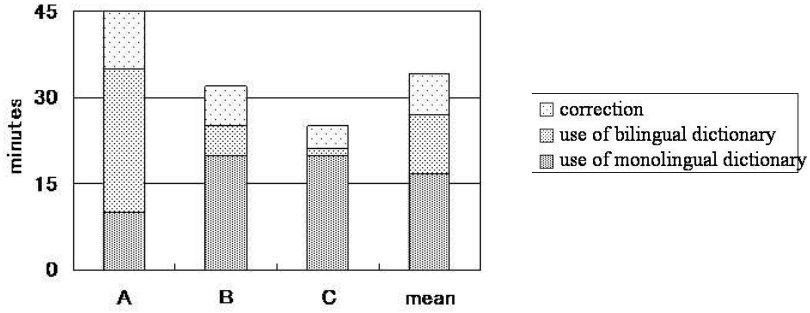

Fig. 1. Behaviour patterns in English compositions

analysis. In the class, there were twelve students, who were studying EFL at the school. The class was divided into three groups according to their English skills: A (beginner), B (elementary), and C (intermediate). Each group consists of four students.

The students were asked to write a formal letter of about 300 words in length within two and a half hours. During the task, each student can ask a teacher for advice twice. Students are allowed to use any paper dictionaries to complete their task only after they are advised by the teacher twice.

Fig. 1 shows how long each group spent with respect to the three behaviour patterns: (i) correcting their compositions, (ii) consulting bilingual dictionaries (i.e. English - native-language), and (iii) consulting monolingual (English) dictionaries. It can be seen from the figure that monolingual dictionaries are more frequently used than bilingual dictionaries in the higher classes. In addition, further analysis revealed that the students in group C (intermediate) used dictionaries mainly to examine the phrasal verbs, words, and prepositions which they used in the current texts. They also used a thesaurus to find more suitable words.

After the class, we carried out a subjective evaluation on several types of dictionaries in terms of (i) frequency in the use of the dictionary, (ii) usefulness of the dictionary for learning English, and (iii) the degree of satisfaction. Subjects were asked to evaluate each type of dictionary and give a score between 0 (lowest) and 5 (highest). Fig. 2 shows the summary for each group.

It can be found from the figure that the frequencies in the use of a thesaurus and collections of examples increase as the English skills improve. Further survey which was based on questionnaires to the students revealed that the purposes of using dictionaries change depending on their English skills. Beginners mainly use bilingual dictionaries to find the right words. On the other hand, students of upper classes use dictionaries for the purpose of confirming the words which they already know or finding more suitable words or expressions. This suggests that a divergent style of learning is taking place at this stage.

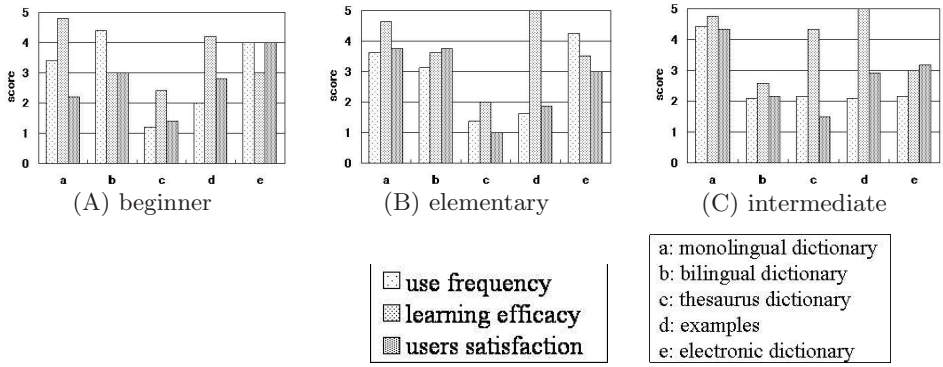

Fig. 2. Subjective evaluation in terms of dictionary types

It should be noted that the users regard a thesaurus potentially useful for language learning, whereas they do not use it that much and they are not much satisfied with it. This is because that a thesaurus does not provide users with detailed information on each word, and users are forced to consult an English dictionary to get further information.

Some of the crucial drawbacks of existing dictionaries can be summarised as follows:

- Users are forced to repeat searching for a word, keeping the search result in mind and backtracking until a suitable expression is found.
- It is not easy for users to compare similar words and find differences between them because it's difficult to see more than one entry at the same time.
- There are not many examples given.

If we could develop a support system which resolves the above problems without losing the merits of paper dictionaries and thesauruses, such a system is expected to be more useful and effective for language learning than the paper ones.

As the first step, the present study focuses on enhancing a thesaurus by solving some of the problems above. To be specific, the following functions are to be implemented.

1. Provides a thesaurus diagram, in which more than one synonym set can be displayed at a same time in multiple windows and collocations are shown as the links between collocated words if existing.
2. Presents examples which are obtained from an internet search engine.
3. Provides a clipboard window to take notes of the search results.

3 System Implementation

In order to implement the first function described in the previous section, we've employed the lexical database, WordNet [2], in which English nouns, verbs, ad-

jectives and adverbs are organised into synonym sets, each representing one underlying lexical concept. Different relations link the synonym sets.

In the present system, words which belong to a same synonym set are graphically shown in a circular window (*synonym-set window*, here after) by means of a magnetic spring model [3]. Users are allowed to have a look at more than one synonym set at a same time. In such a case, each synonym-set window displays a different synonym set from each other.

Fig. 3 shows a screenshot of the system, where a multi-word query "invade violate infringe law" was given in the query window, which is just under the function buttons located at the the top part of the screen. Since the query consists of four words in this case, there are four synonym-set windows displayed, each of which corresponds to each word in the query. Fig. 4 displays an enlarged view of the synonym-set window for word "law". In the synonym-set window, each non-terminal (green) node denotes a "sense" of the word, i.e. a meaning of the word. When the word has more than one meaning, more than one sense node appears in the window. By clicking a sense node with the mouse left-button, a small rectangular window appears to show a brief description of the word and examples in that context. The contents of the window can be stored in a "clipboard", whose window is displayed when the "CLIP BOARD" button in the top menu is pressed.

An user interface for manipulating the windows has been implemented as well: users can move, rotate or delete a diagram and change its size for a better view.

When more than one synonym-set window is displayed, users can check whether some collocations exist between the synonym sets by pushing the "Join Senses" button in the top menu. Collocations are shown by red straight lines in Fig. 3.

The second function, i.e. showing examples retrieved from an internet search engines, was implemented using the Google search engine. The whole-set or subset of the query can be employed as the key words for the internet search. To this end, "List" button is used to select words in the query, and the words selected are shown in the key-word window, which is located just under the query window (Fig. 3). "Google" button is used to invoke the search engine for the key words, and the search result is shown in the Google search window, which is just under the key-word window. At the same time, the hit count of the search is shown in the key-word window.

4 Evaluation and Discussions

4.1 Evaluation

A preliminary subjective evaluation by ten EFL students was conducted to assess the prototype system. Fig. 5 shows the proportions of target words and non-target words which were consulted by the users through the experiment, where a target word denotes the one which the user intended to search, whereas non-target words are the others. It can be seen from the figure that the users of the

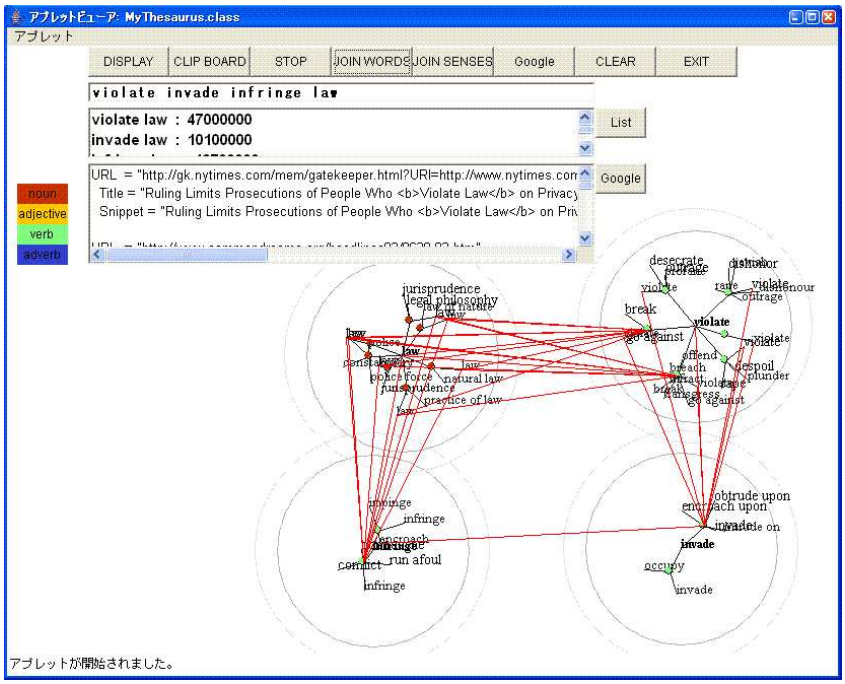

Fig. 3. A screenshot of the proposed system

proposed system checked the meanings and examples of not only the target words but also the non-target words. This suggests that divergent style of activity took place.

Another subjective evaluation using the same EFL students was carried out to compare user's preferences for (i) paper dictionary, (ii) translator (Google), (iii) Visual Thesaurus 3 [4], (iv) Expert System (IdeaFisher) [5], and (v) Thesaurus Diagram (proposed system). As is shown in Fig. 6, when compared in terms of usefulness, the proposed system showed higher score than the other systems and comparable score to paper dictionary.

4.2 Related works

There are two systems which are relevant to the proposed system. One is 'Visual Thesaurus" [4] and the other is "Visual Browser" [6].

The Visual Thesaurus has implemented similar functions with those of 1 and 2 described in section three, however, one of the crucial differences is that Visual Thesaurus accepts only a single-word query at a time while our system accepts a query of multiple words. The multiple word query helps users to check whether the combination of words in a sentence is adequate or not. It also helps users to find more suitable combination of words by showing collocations by lines between words each of which belongs to the synonym set of each word of a user's query.

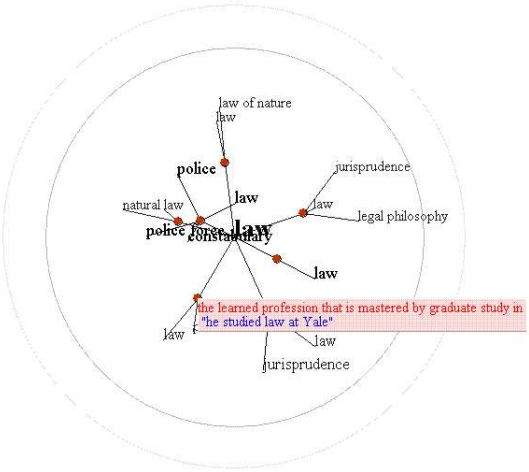

Fig. 4. A screenshot of the synonym-set window for word "law"

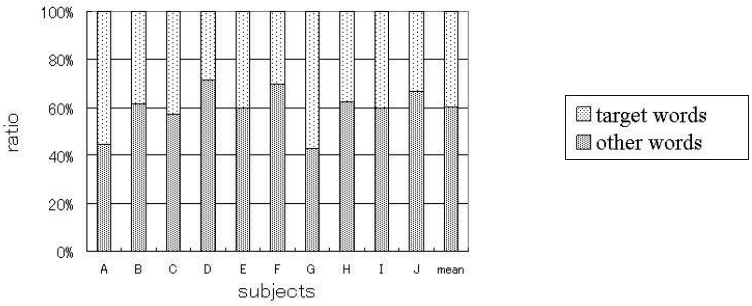

Fig. 5. Proportions of target and non-target words

The Visual Browser is a Java application for visualising the data in RDF (Resource Description Framework) schema [7]. The main difference between this tool and our tool resides in the fact that the former is designed for general applications, whereas the latter is designed for the specific application. As the result, although the Visual Browser can visualise more than one synonym-set of the WordNet, it just displays the synonym-sets which are linked with each other. On the other hand, in our system, it is the user who specifies what synonym-sets to display, and those sets which are not linked with each other can be displayed as well. The user can check whether the two words has some relationships or not. This user initiative interface provides the user with more flexible use of the tool than the Visual Browser.

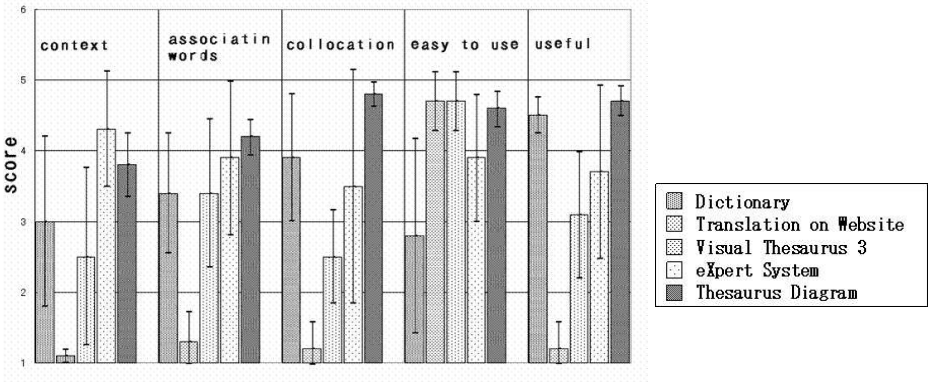

Fig. 6. Subjective evaluation on different systems

5 Conclusions

The present study aimed to develop a software tool which supports a divergent style of English learning. Subjective evaluation of the prototype system revealed that users used the system not only to find or check the right words but also to find other words which are relevant to the right words. It was also found that the clipboard function was supported by the users since it reduces the user's load of searching the right words.

However, several problems with the system have been pointed out: it is not easy for the user to find the right word if the user does not know the word at all; it is difficult to identity each synonyms in the window when too many synonyms are displayed at the same time; users sometimes get lost when they have followed relevant words too many times. Future works include attempts to resolve these problems and further evaluation of the system.

References

1. Mochizuki, M., Aizawa, K., Tono, Y.: Manual for English Vocabulary Teaching. Taisyukan (1999) (in Japanese).
2. Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.J.: Introduction to WordNet: An on-line lexical database. International Journal of Lexicography **3**(4) (1990) 235–312
3. Sugiyama, K., Misue, K.: Graph drawing by the magnetic spring model. Journal of Visual Languages and Computing **6**(3) (1995) 217–231
4. Visual Thesaurus: http://www.visualthesaurus.com/.
5. IdeaFisher: http://www.ideacenter.com/.
6. Visual Browser: http://nlp.fi.muni.cz/projekty/vizualni_lexikon/.
7. RDF schema: http://www.w3.org/TR/rdf-schema/.