AUTOMATIZED MEMORY TECHNIQUES FOR VOCABULARY ACQUISITION IN A SECOND LANGUAGE

Gözde Özbal and Carlo Strapparava
FBK-Irst, Via Sommarive 18, 38050, Povo-Trento, Italy

Keywords: Second language learning, Vocabulary acquisition, Keyword method, Latent semantic analysis, Levenshtein distance, Text animation.

Abstract: Vocabulary acquisition is a very important step for learning a new language. On the other hand, many learners find this step difficult and time consuming. Several vocabulary teaching methods try to facilitate this step with various verbal and visual tips. However, the preparation of these tips generally necessitates a huge amount of time, money and human labor. In this paper, we propose to exploit Natural Language Processing driven creativity to develop an automatic system for task of vocabulary teaching. This system can automatically generate memorization tips for the words which users would like to memorize. The preliminary results are promising and motivating for further investigation, since they show that our approach can be quite effective on the related task.

1 INTRODUCTION

Linguistic creativity is a characteristic property of a natural language and it can be defined in many ways. However, we can focus on the definition of (CALC-09, 2009): "Creative language usage at different levels from the lexicon to syntax, to discourse and text".

Although only very little primary research was conducted up to the beginning of 21st century (Zawada, 2005), linguistic creativity is recently a popular and challenging research topic for several Natural Language Processing (NLP) tasks including sentiment analysis, text summarization, information retrieval, machine translation, and question answering. Creativity-aware systems are expected to enhance the contribution of Computational Linguistics to several practical areas such as education, engineering, and entertainment (CALC-09, 2009). Throughout this paper, we will focus on developing a system which automatically applies linguistic creativity to the education area, or more specifically, vocabulary acquisition in a second language.

Many books, online courses and software programs provide a long vocabulary list to teach the vocabulary of a language. Learners usually spend a lot of time and effort to memorize these lists by heart, and they find this process tedious and boring since learners find it difficult to fix all the words in memory one by one. For these reasons, these kinds of vocabulary lists are not very effective in many cases.

Various methods aim to facilitate the memorization process to teach the vocabulary of a foreign language. A very commonly used method is representing the meaning of the new word with related images and/or animations to help the learner to build a connection between the visual and verbal memory. Another popular method called the *keyword or linkword method* links the translation of the target word to one or more keywords in the native language which are phonologically or lexically similar to the target word.(Sagarra and Alba, 2006) To illustrate, for teaching the Italian word *tenda* which means *curtain* in English, the learners are asked to imagine *"rubbing a TENDER part of their leg with a CURTAIN"*.

These methods have been proven to be successful in many cases and helpful for learners. Accordingly, a considerable number of language books and software systems use them. However, since all visual and verbal tips are designed manually, a huge amount of time, labor and creativity is required for their preparation.

In this paper, we will present a fully automatized system which automatically produces memorization tips for each target word the user wants to memorize. These tips consist of keywords, sentences, colors, animations and images.

Özbal G. and Strapparava C..
AUTOMATIZED MEMORY TECHNIQUES FOR VOCABULARY ACQUISITION IN A SECOND LANGUAGE.
DOI: 10.5220/0003330300790087
In *Proceedings of the 3rd International Conference on Computer Supported Education* (CSEDU-2011), pages 79-87
ISBN: 978-989-8425-49-2
Copyright © 2011 SCITEPRESS (Science and Technology Publications, Lda.)

The rest of the paper is structured as follows. In Section 2, we will provide an overview of the current approaches using the keyword method to teach the vocabulary of foreign languages. We will also summarize state of the art in the research areas related to our study. In Section 3, we will make a detailed description of our system. Finally, in Section 4, we will outline the possible future work and draw our conclusions.

2 STATE OF THE ART

In this section, we will focus on the non-automatic methods used for vocabulary teaching, as well as state of the art in research areas related to our study.

2.1 Non-automatic Vocabulary Teaching Methods

The study introduced by (Sagarra and Alba, 2006) compares the effectiveness of three learning methods including the semantic mapping (i.e. creating a diagram with words in the first language which are semantically related to the new word in the second language), rote memorization (i.e. memorizing the translation of the new word in the second language by rehearsal) and keyword method on beginner learners of a second language. Their results show that using the keyword method with phonological keywords and direct links between the keyword and translation leads to better learning of second language vocabulary for beginners.

(Sommer and Gruneberg, 2002) introduces a software using the keyword method for teaching French to 13-year-old students as a complementary learning aid. The main finding of this study is that students find this method easier and faster than conventional methods.

Currently, the keyword method is still very commonly in use by many vocabulary teaching systems:

As an example, "The Italian for FROG is RANA. Imagine you RAN A mile after seeing a scary FROG." can be visualized on http://www.linkwordlanguages.com/software_demos/italian/it-example01.htm. Here, the two keywords *ran* and *a* are used to represent the target word *rana* whose translation is *frog*. To help the learner build a semantic relationship between the keywords and the translation, a sentence is also provided.

On http://www.italianlanguagesecrets.com/memory-techniques-italian-vocabulary.html, it is asserted that if a funny association is created between the translation and the keywords, the target word will be remembered forever. As an example, for the Italian target word *pollo* of which English translation is *chicken*, a funny association is built with *polo* such as "Imagine a group of *chicken*s playing *polo* together, listen to them cackling (fare coccode', we say in Italian), imagine them moving forth and back in a *polo* stadium.".

On http://www.learning-at-home.co.uk, the pronunciation of the Italian word *uccello*, which means *bird* in English, is divided into two smaller pieces: *you* and *cello*. Then, the association between the keywords is established with the sentence "Imagine the conductor in an animal orchestra saying to a *BIRD*: *YOU CELLO*, me conductor."

Lastly, (Duyar, 2001) builds homogeneous associations from English to English and associates words from Turkish and English by using the keyword method. For instance, the sentence "*A sausage* can *lessen* my hunger" is created for the target word *assuage* where the visual link words are *a* and *sausage*.

In this paper, our main focus will be the task of *automatizing* the generation of all these kinds of associations.

2.2 Latent Semantic Analysis

While generating keywords, we should take into account the semantic similarity of words as well as the lexical and pronunciation similarity. In addition, we need to consider meanings of keywords to be able to generate sentences containing them. Furthermore, we should calculate affective meanings of words to build colorful animations. Lastly, for the retrieval of an image for a target word, we have to build semantic associations with this word and candidate images. For all these reasons, we have the necessity to analyze semantics of texts at different levels of granularity.

Measures of word and text similarity have been used for a long time in NLP applications (Budanitsky and Hirst, 2006). Many knowledge-based and corpus-based methodologies and metrics have been developed. For our research, we exploit latent semantic analysis (LSA) (Deerwester et al., 1990) which is a corpus-based measure of semantic similarity. The main idea behind LSA is that two concepts co-occur in the same texts with high frequency when they are associated in common sense knowledge.

LSA uses a sparse document-term matrix whose rows correspond to documents and whose columns correspond to terms. The value of each entry is typically a term frequency-inverse document frequency (tf-idf), which is proportional to the number of times the term appears in the matrix and offset by the frequency of the term in the corpus, so that rare terms

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

are promoted to reflect their relative importance.

LSA applies the singular value decomposition (SVD) technique on the original matrix to obtain a low-rank approximation. At the end of this process, each document and term are represented by a vector with a much lower dimension than the total number of words, so that documents and terms with similar meaning are close in the low-dimensional space. Thus, LSA can be viewed as a way to overcome some of the drawbacks of the standard vector space model (sparseness and high dimensionality).

The similarity in the resulting vector space is measured with the standard cosine similarity. It can be also noted that LSA yields a vector space model that allows for a homogeneous representation (and hence comparison) of words, word sets, and texts. For example, in order to represent word sets and texts by means of an LSA vector in the present work, we have used a variation of the pseudo-document methodology described in (Berry, 1992).

(Strapparava and Valitutti, 2004) proposes a variation of Latent Semantic Analysis (LSA) to calculate the similarities between terms and documents, in which a term-frequency/inverse document frequency (tf-idf) weighting schema is used. For the representation of a document, the normalized LSA vectors of all the terms inside are summed up.

(Strapparava and Mihalcea, 2008) compares several algorithms for the SEMEVAL 2007 task on affective learning. According to the proposed approach, it is possible to represent an emotion at least in three different ways. The first one is the vector of the word that denotes the emotion itself (shortly named as *LSA single word*). The other is achieved by representing the synset of the emotion (shortly named as *LSA emotion synset*). The last one also adds the words in all the synsets which are labeled with the emotion in question, in addition to the previous set (shortly named as *LSA all emotion words*). Among these three ways, *LSA all emotion words* model provides the highest recall and F-measure, in terms of coarse grained evaluations.

(Strapparava et al., 2007b) automatically detects the affective meaning of texts in order to animate the words inside. WordNet-Affect is used for obtaining direct affective words and synsets. An LSA mechanism is used as a selection function, which provides the semantic affinity between a concept and an emotion. After obtaining the affective load of a sentence, the words which are most similar to emotional concepts are selected. The resulting affective meaning is conveyed through an animation.

(Strapparava et al., 2007a) extends this study by exploring its effectiveness in advertisement production. The proposed system converts familiar textual expressions to affective variations based on lexical semantic techniques, and animates them according to their affective contents.

(Valitutti et al., 2008) focuses on generating affective advertising headlines. The proposed work is mainly based on creative variations of familiar expressions with LSA.

2.3 Text based Image Retrieval

(Borman et al., 2005) underlines the fact that associations between meanings of words and their visual representations can have a positive impact on language learning especially for children and people with language disorders. Accordingly, it focuses on adding visual representations to machine readable dictionary entries in order to build illustrated semantic networks which encode word/image associations. To achieve this, images collected from and validated by online users are integrated with the use of automated image extraction techniques through image meta-search. Synset/image associations are based on user uploads, user free association, system guesses, or initial automated seeding.

(Hayashi et al., 2009) proposes a cross-language image search system which exploits Google image search to collect images for translation candidates. The main goal of this study is to improve the selection of correct translations for a query term. Even though the method proposed did not conduct any image sense disambuigiation, an improvement was obtained on the task of query term selection.

(Mihalcea and Leong, 2008) underlines the benefits of the usage of pictorial representations to convey information for people who study a foreign language especially for children. Accordingly, a system is proposed to automatically produce pictorial representations for simple sentences. This is achieved by determining meanings of words with a sense tagger and identifying pictorial representations for each noun and verb with the system proposed in (Borman et al., 2005).

(Fujii and Ishikawa, 2005) states that existing search engines such as Google cannot handle polysemy for image retrieval. It proposes a method to associate images on the Web to encyclopedic term descriptions for specific word senses. This method uses text in HTML files as a pseudo-caption of the image based on a term-weighting method.

(Fujita and Nagata, 2010) proposes a method to provide appropriate images for each word sense. Candidate images are collected from the web and queries for minor senses are expanded using synonyms and hypernyms.

3 SYSTEM DESCRIPTION

The vocabulary teaching system we propose is called MEANS, which stands for "Moving Effective Assonances for Novice Students". The interface of MEANS can be implemented in different ways. In the simplest case, users can make a query through the user interface to request memory tips for a word. After the submission of this query, the system provides several memorization tips including keywords, sentences, colorful animations and images, together with the translation and pronunciation of the target word. As an alternative, it can also be implemented as an add-on for a web explorer, so that users can right click on any word on a web page for which they want to retrieve memory hints. In the current implementation, the system supports teaching Italian to English speaking students. However, it can easily be extended to support other language pairs.

The system mainly consists of four modules: i) Selecting keywords ii) Generating sentences iii) Building and selecting colorful animations iv) Retrieving and selecting images. In this section, we will give information about the current design and implementation of these modules.

When a user makes a word query through the user interface, the system first selects the most appropriate keywords from a corpus consisting of the most frequent words in the language already known. Accordingly, another query including these keywords and the translation of the target word is made against the web. Then, a selection mechanism is carried out to select the best among the candidates. The keywords and translation are animated with appropriate colors mainly according to the emotion they convey. Lastly, the most suitable image representing the target word is retrieved from the web.

3.1 Selecting Keywords

The main role of this module is to automatically create appropriate keywords for each target word the learner is willing to memorize. In order to achieve this, both the lexical and pronunciation similarity of candidate words with the target word are taken into account. After the list of candidate keywords are obtained, a selection mechanism is applied on this list to determine the most appropriate keyword(s).

For the task of building the list of keywords for a target word, we preferred to restrict candidate keywords to be as simple and frequent as possible instead of using a whole dictionary. In fact, providing much more difficult keywords than the target word would not make much sense for the sake of facilitating the memorization process. Accordingly, we have used a collection of 6000 most frequently used English words (Insightin, 2010). This collection basically uses search engine index databases and it was created by calculating the ranks of word frequencies by running the list of words in the WordNet dictionary database against words frequently used in search engine queries from 2002 to 2003.

3.1.1 Lexical Similarity

Lexical similarity can be defined as a measure of the degree to which the word sets of two given languages are lexically similar (Lewis, 2009). Among the different possible distance metrics which can be applied for calculating the lexical similarity between the target word and the candidate words, we have chosen the *Levenshtein distance*. Note that this notion is different from the semantic similarity related to the LSA methodology, which is a metric to measure the likeness of the meaning/semantic content of terms.

The Levenshtein distance is a metric for measuring the amount of difference between two sequences, and the Levenshtein distance between two strings is defined as the minimum number of edits required for the transformation of one string into the other. The allowable edit operations for this transformation are insertion, deletion, or substitution of a single character (Levenshtein, 1966). For example, the Levenshtein distance between "kitten" and "sitting" is 3, since the following three edits change one into the other, and there is no way to do it with fewer than three edits: kitten $\rightarrow$ sitten (substitution of 'k' with 's'), sitten $\rightarrow$ sittin (substitution of 'e' with 'i'), sittin $\rightarrow$ sitting (insertion of 'g' to the end).

The process of finding lexically similar keyword candidates for a target word can be summarized as the following:

First, a list of substring pairs are produced by splitting the target word at all possible positions. Thereby, a total number of configurations equal to the length of the target word are produced. Each configuration contains either one or two substrings. (e.g. for the target word *libro*, possible substrings would be: *libro*, *l* and *ibro*, *li* and *bro*, *lib* and *ro*, and lastly *libr* and *o*.)

Afterwards, the set of words in the corpus which minimize the Levenshtein distance is calculated for each substring. Among this set, the system selects the word which maximizes the total number of common consecutive letters at the beginning if the calculation is for the first part, at the end if the calculation is for the second part, and both at the beginning and end if

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

Table 1: Example keywords obtained using lexical similarity.

Substrings		Keywords	
First	Second	First	Second
-	libro	-	liar
l	ibro	la	into
li	bro	lip	brown
lib	ro	lip	rot
libr	o	liar	of

the calculation is for the whole target word.

This last criterion regarding the number of common consecutive letters is based on the *locomotive factor* defined by (Duyar, 2001) which states that the beginning and ending sounds of a word behave like locomotives, and the sounds in the middle are like wagons following the locomotives. Therefore, the storage of a foreign word in the memory is mostly related to triggering the locomotive-like parts of that word.

At the end of all these processes, either a keyword or a keyword pair is obtained for each configuration. All these keywords have the common property that they are lexically the most similar words either to the substrings of the target word or to the target word as a whole. Following the previous example, the keywords obtained for the target word *libro* can be found in Table 1.

3.1.2 Pronunciation Similarity

In addition to lexical similarity, the system also takes into account *pronunciation similarity* for the selection of keywords. In order to achieve this, two phonetic recources consisting of one for English and one for Italian are used.

As the English phonetic resource, the CMU Pronouncing Dictionary (Lenzo, 2010) has been chosen. This dictionary is a machine-readable pronunciation dictionary for North American English which contains over 125,000 words and their transcriptions. It has mappings from words to their pronunciations in the given phoneme set. The current phoneme set contains 39 phonemes, for which the vowels may carry lexical (primary and secondary) stress. From this dictionary, we retrieved the pronunciation of the most frequent English words as found in the aforementioned collection.

As for the Italian phonetic resource, we obtained the phonetic transcription of Italian words through the use of a morphological engine developed by our research unit. This engine is able to decompose a given word into sequences of morphemes: more than 100k morphemes assure a good coverage of Italian texts.

Table 2: Example keywords obtained using pronunciation similarity.

Substrings		Keywords	
First	Second	First	Second
-	libro	-	lip
l	ibro	low	hero
li	bro	lip	broke
lib	ro	lip	row
libr	o	lip	oh

Each morpheme is associated to its meta-transcription, which is an intermediate representation that can evolve in different ways, depending on the adjacent morphemes. Words which cannot be decomposed are processed by a rule-based grapheme-to-phoneme engine.

In order to calculate the distance between pronunciations of words, Levenshtein distance is used in a similar way to the calculation of the lexical similarity, except with some relaxations. For the lexical similarity, Levenshtein distance expects any two letters being compared to be exactly the same for a score of 0, whereas for pronunciation similarity a set of letter pairs including b-p, d-t, v-f, g-k, s-z are also considered as a match. In addition, since the two dictionaries use different standards for representing phonemes, we had to apply several mappings between them. It must be also noted that, the information regarding syllables and stresses are ignored in the current version of the system.

At the end of the calculations, either a keyword or a keyword pair is obtained for each configuration. To illustrate, the keywords obtained for the target word *libro* can be found in Table 2.

3.1.3 Selection Mechanism

After all candidate keywords are calculated by using either lexical or pronunciation similarity, a preparation step takes place. During this step, a morphological analysis is conducted to construct a list of all possible part of speech (POS) information for each keyword. Afterwards, for each possible POS, a list of possible domains are found. We exploit WORDNET-DOMAINS as a resource to have a feasible list of semantic domains (Magnini and Cavaglià, 2000). In particular, we use a subset of the domain labels. This subset was selected empirically to allow a sensible level of abstraction without losing much relevant information, overcoming data sparseness for less frequent domains (Magnini et al., 2002). In the LSA space we check the similarity among keywords and domain names. After this step, a latent semantic analysis is performed to find the LSA similarity of the

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

keyword with the translation of the target word. In the present work, we have used an LSA space built on the full British National Corpus limiting the number of dimensions to 400 for the dimensionality reduction. BNC is a very large (over 100 million words) corpus of both spoken and written modern English, (BNC-Consortium, 2000). Other more specific corpora could also be considered to obtain a more domain oriented similarity.

After the LSA similarity between the keyword and the translation is calculated, the following selection mechanism is carried out among the candidates to choose the most appropriate one(s) for each target word:

1. According to whether the candidate has been created by considering lexical or pronunciation similarity, the same type of similarity between the target word and the concatenation of the keywords are calculated. Let us call this distance the *overall distance*.
2. In order to select keywords with better quality among the ones having the same *overall distance*, a higher priority is given to the ones whose distances to the related target substrings are more homogeneous. To this end, each keyword distance is normalized dividing it by the length of the corresponding substring. Then, the *standard deviation* of the sequence of normalized distances is measured.
3. To differentiate the keywords which are semantically very similar to the translation of the target word, a threshold distance value (0.75), which has been determined empirically, is used. As a future work, we plan to use a more systematic approach for setting this parameter by employing a labeled set of data. If one of the keywords has an LSA distance larger than this threshold, a flag called *big lsa* is set to true.
4. The number of common domains between the keyword(s) and the translation of the target word is found. Let us call this number the *number of common domains*.
5. In accordance with the previous calculations, the order of priorities used while analyzing each configuration is the following: *less overall distance*$\gg$*pronunciation similarity*$\gg$*less standard deviation*$\gg$*big lsa*$\gg$*number of common dommains*

At the end of this process, the most appropriate keyword(s) for representing the target word is (are) obtained.

3.2 Generating Sentences

This module is responsible for providing the learner with a sentence containing both the set of keywords and the translation of the target word. The main goal here is helping the user to build a semantic relationship between the keywords and the translation.

Following that intention, two types of queries are executed against Google by using Google API. The first aims to retrieve the sentences containing the keywords and the translation whose positions in the sentence are all next to each other by trying every possible permutation. If no results are retrieved after the first attempt, a more relaxed query without the position restriction is executed.

The system uses another selection mechanism to provide the most appropriate sentence among all candidates. As a preparation process, all sentences are cleaned and POS-tagged, and then the LSA similarity of the translation word with the candidates are calculated. The highest priority is given to the sentences retrieved by the first query type. In case of an equivalence, the one in which the keywords and the translation are less distant from each other in total is preferred. If another equivalence occurs, the shorter sentence is selected.

Below, we list examples of system decisions for a sample of target words: *porta, cane, tirare, occupato*.

Target Word: Porta
Translation: Door
Keywords: port - a
Sentence: To establish a connection (to get into your house) intruders need to find A PORT (DOOR) that is open.

Target Word: Cane
Translation: Dog
Keywords: can - a
Sentence: CAN A DOG smile?

Target Word: Occupato
Translation: Busy
Keywords: occupy - to
Sentence: She has returned to work and I encourage her to keep herself BUSY TO OCCUPY her time.

Our current approach has a low recall since it is not always possible to retrieve a sentence matching all the constraints. We discuss the possible solutions to overcome this limitation in Section 4.

Figure 1: An example text animation representing *anger*.

3.3 Building and Selecting Colorful Animations

For the sake of attracting attention and helping learners to build a connection between words and their affective meanings, we use text animation. This module is generally inspired by the study represented in (Strapparava et al., 2007b). However, the design of the animations and the package used to build these animations differ from this study. In addition, while (Strapparava et al., 2007b) can be considered as a proof of concept which underlines the fact that words can be vitalized automatically, to our knowledge, we are the first to explore this idea from a more applicative point of view.

During the design, we have created five different animations in total: four for representing affective texts and one for neutral (non-affective) texts. The emotional categories which we have focused on are anger, sadness, joy and fear. The design of the animations has been inspired by animated emoticons representing these emotions on the web.

Accordingly, *anger* is built on *jittering* and *scaling up* both in the x and y coordinates, whereas *fear* only consists of a continuous *jittering* in the x coordinate. *Joy* is a combination of a *hop* in the y coordinate, a *scale up* both in the x and the y coordinates and an *oscillation* accompanied by a *rotation*. *Sadness* is based on a slow negative *change* in the y coordinate in addition to a *scaling up*. Lastly for animating a *neutral text*, a continuous *scaling up and down* is used. Although the whole effect of an animation cannot be conveyed by a static image, Figure 1 shows an example of a text animated with *anger*.

In order to represent emotions of texts with appropriate colors, we rely on the results of the psycholinguistic experiments reported in (Kaya, 2004). This study investigates and discusses the associations between colors and emotions by conducting experiments where college students are asked to indicate their emotional responses to principal, intermediate and achromatic colors, and the reasons for their choices. We have analyzed the frequencies of emotional reactions given to each color and picked the most frequent one for each color. Accordingly, we have chosen the color *red* for representing *anger*, *black* for *fear*, *yellow* for *joy*, *gray* for *sadness*, and *dark gray* for the *neutral* text.

After a sentence is selected for a specific target word, the lexical affective semantic similarity metric is used to find out which emotional categories the translation of the target word and the keywords belong to. Affective similarity is measured using the 'LSA Emotion synset' method proposed in (Strapparava and Mihalcea, 2008). In this method, the LSA vector representing an emotion is simply the sum of the vectors which correspond to the words in the same WordNet synset.

Among the possible emotional categories, the most dominant one (i.e. the one with the highest score) is selected. Then, the animation and color designed for the dominant emotion is used to animate the translation and keywords. Thereby, we try to trigger both the verbal and visual memory of the learner, thus increasing the chance that the target word and its translation will be learned more easily and in a shorter time.

3.4 Retrieving and Selecting Images

Other than providing a set of keywords and a sentence with animated text, the system also displays an image representing the target word. This image is retrieved from the web automatically by using Google API to query the image service of Google with the translation of the target word.

With the execution of each query, 24 results are obtained. The most suitable one among them is chosen with the help of a selection mechanism based on text-based image retrieval. First the content information of the image is POS tagged. Then, in order to select the most similar image semantically, the LSA similarity of each image with the translation is calculated. Accordingly, the image with the highest score is presented to the learner.

4 CONCLUSIONS AND FUTURE WORK

In this paper, we have described a vocabulary teaching system which automatically produces memorization tips using state of the art NLP and IR techniques. Since all visual and verbal aids used to teach a vocabulary are designed manually with current approaches, building the necessary memorization tips for each new word requires many resources such as time, labor and creativity. Our main goal is to develop a real-world application available as a web-service to make the preparation of such resources much easier.

In its current design and implementation, given a target word, our system is able to produce keywords,

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

retrieve related sentences and images from the web and generate colorful animations.

The preliminary results show that our approach can be quite effective on the related task. This study can be considered as a proof of concept for the idea and we are encouraged to further explore several issues for future work.

As technical improvements, we would like to exploit resources using the same standard for the phoneme representation, so that we do not have to apply a mapping mechanism for the calculation of pronunciation similarity. As an other improvement, we want to take into consideration phonemes instead of letters, and the information regarding syllables and stresses, and investigate whether the performance can be improved in this way. In addition, we can investigate the effect of using interval values for relaxed matches during the calculation of pronunciation similarity. We are aware that sentences retrieved from the web with our current technique are not reliable all the time. For instance, they might contain inappropriate content for students. However, we are assured that reliability is a crucial issue in an e-learning environment especially for children. As a possible solution, we would like to explore the impact of conducting a domain control with LSA.

Next, for handling the cases in which no sentences containing the keywords and the translation are retrieved, we plan to explore the effect of applying lexical substitution on sentences containing one or two of the query words. However, this is a challenging problem since we have to make sure that the new sentence conforms to the grammar rules.

Regarding images, we would like to improve our method to discriminate images with different senses for the same query word. Since LSA uses a low-dimensional representation for terms, terms with similar meanings are close in the low-dimensional space, and the representation of meaning is with better quality in comparison to a traditional vector space method. Accordingly, we can handle polysemy by using the synset information in the query to disambiguate the text information of images in the low-dimensional space. Second, we plan to conduct experiments on the effect of using different texts related to the image such as the title, content or other pieces of text occurring in the page containing the image.

To evaluate the performance of the overall system, we are going to convert our prototype to an online service and collect user feedback for further improvements. More specifically, we will ask users whether the memory tips have been useful for the memorization so that we can find out which modules have a bigger impact on the learning process. Additionally, users will be able to rate the tips. For instance, they will be able to state whether the selected keywords are appropriate, or the sentence is meaningful and/or humorous, or the displayed image conveys the meaning of the target word. In addition to the online feedback, we are also considering to conduct more specific experiments in which we will host subjects in a closed environment. We will provide a subset of these subjects with memorization tips for a set of words in a language which they were not exposed to before, while traditional methods will be used to teach the same vocabulary to the rest. At the end of this process, we will make a vocabulary test to investigate the impact of our method and make a comparison with traditional methods.

REFERENCES

Berry, M. (1992). Large-scale sparse singular value computations. *International Journal of Supercomputer Applications*, 6(1):13–49.

BNC-Consortium (2000). British national corpus. http://www.hcu.ox.ac.uk/BNC/.

Borman, A., Mihalcea, R., and Tarau, P. (2005). Pic-net: Pictorial representations for illustrated semantic networks. In *Proceedings of the AAAI Spring Symposium on Knowledge Collection from Volunteer Contributors*.

Budanitsky, A. and Hirst, G. (2006). Evaluating WordNet-based measures of lexical semantic relatedness. *Computational Linguistics*, 32(1):13–47.

CALC-09 (2009). Workshop on computational approaches to linguistic creativity (calc-09). http:// aclweb.org/ aclwiki/index.php?title=CALC-09.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T., and Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391–407.

Duyar, M. S. (2001). *Accelerated Word Memory Power*. Kardelen Ofset Ltd. Sti.

Fujii, A. and Ishikawa, T. (2005). Toward the automatic compilation of multimedia encyclopedias: Associating images with term descriptions on the web. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence*, WI '05, pages 536–542, Washington, DC, USA. IEEE Computer Society.

Fujita, S. and Nagata, M. (2010). Enriching dictionaries with images from the internet - targeting wikipedia and a japanese semantic lexicon: Lexeed -. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pages 331–339, Beijing, China. Coling 2010 Organizing Committee.

Hayashi, Y., Bora, S. A., and Nagata, M. (2009). Utilizing images for assisting cross-language information retrieval on the web. In *Proceedings of the 2009*

IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 03, WI-IAT '09, pages 100–103, Washington, DC, USA. IEEE Computer Society.

Insightin (2010). 6,000 most frequently used english words. http://www.insightin.com/esl/.

Kaya, N. (2004). Relationship between color and emotion: a study of college students. *College Student Journal*, pages 396–405.

Lenzo, K. (2010). The cmu pronouncing dictionary. http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10:707—710.

Lewis, P. M. (2009). *Ethnologue: Languages of the World*. SIL International.

Magnini, B. and Cavaglià, G. (2000). Integrating subject field codes into WordNet. In *Proceedings of LREC-2000, Second International Conference on Language Resources and Evaluation*, pages 1413–1418, Athens, Greece.

Magnini, B., Strapparava, C., Pezzulo, G., and Gliozzo, A. (2002). The role of domain information in word sense disambiguation. *Natural Language Engineering*, 8(4):359–373.

Mihalcea, R. and Leong, C. W. (2008). Toward communicating simple sentences using pictorial representations. *Machine Translation*, 22:153–173.

Sagarra, N. and Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of spanish. *The Modern Language Journal*, 90(2):228–243.

Sommer, S. and Gruneberg, M. (2002). The use of linkword language computer courses in a classroom situation: a case study at rugby school. *'Language Learning Journal*, 26(1):48–53.

Strapparava, C. and Mihalcea, R. (2008). Learning to identify emotions in text. In *SAC '08: Proceedings of the 2008 ACM symposium on Applied computing*, pages 1556–1560, New York, NY, USA. ACM.

Strapparava, C. and Valitutti, A. (2004). WordNet-Affect: an affective extension of WordNet. In *Proceedings of LREC*, volume 4, pages 1083–1086.

Strapparava, C., Valitutti, A., and Stock, O. (2007a). Affective text variation and animation for dynamic advertisement. In *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction*, ACII '07, pages 242–253, Berlin, Heidelberg. Springer-Verlag.

Strapparava, C., Valitutti, A., and Stock, O. (2007b). Dances with words. In *IJCAI'07: Proceedings of the 20th international joint conference on Artifical intelligence*, pages 1719–1724, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Valitutti, A., Strapparava, C., and Stock, O. (2008). Textual affect sensing for computational advertising. In *Proceedings of AAAI Spring Symposium on Creative Intelligent Systems*, pages 117–122.

Zawada, B. E. (2005). *Linguistic Creativity and Mental Representation with Reference to Intercategorial Polysemy*. PhD thesis, University Of South Africa.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS