Computational Linguistics and Chinese Language Processing
Vol. 13, No. 3, September 2008, pp. 279-306

A Language Information Retrieval Approach to Writing Assistance

Jyi-Shane Liu*, Pei-Chun Hung*, and Ching-Ying Lee+#

Abstract

We observe that current language resource tools only provide limited help for ESL/EFL writers with insufficient language knowledge. In particular, there is no convenient way for ESL/EFL writers to look for answers to the frequent questions of correct and appropriate language use. We have developed a language information retrieval method to exploit corporal resources and provide effective referential utility for ESL/EFL writing. This method involves the sequential operation of three modules, an expression element module, a retrieval module, and a ranking module. The primary design purpose is to allow flexible and easy transformation from questions to queries and to find relevant examples so that uncertainty of language use can be quickly resolved. We implemented the method and developed a prototype system called SAW (Sentence Assistance for Writing). Simulated language use problems were tested on SAW to evaluate the system's referential utility. Experimental results indicate that the proposed language information retrieval method is effective in providing help to ESL/EFL writers.

Keywords: Language Information Retrieval, Language Resources, ESL/EFL Writing

1. Introduction

Writing is a significant form of expression for conveying information and experiences. Effective writing demands appropriate articulation of meaning and prudent selection of words. For many ESL/EFL (English as Second Language/English as Foreign Language) learners, writing in English is an especially difficult task in which authors constantly face uncertainty as far as how to convert their thoughts into the second language. Writing production of ESL/EFL learners is also error-prone due to insufficient knowledge of the second language

* Department of Computer Science, National Chengchi University, Taipei, Taiwan
E-mail: jsliu@cs.nccu.edu.tw
+ Department of English, National Taiwan Normal University, Taipei, Taiwan
# Department of Applied Foreign Languages, Kang Ning Junior College, Taipei, Taiwan

[Received January 13, 2008; Revised June 20, 2008; Accepted August 8, 2008]

and the interference of the native language [Kobayashi and Rinnert 1992]. At the same time, writing in English (as a second language) is a very time-consuming process. ESL/EFL writers usually take considerable time in searching for the right ways of language expression with help from various language resources and tools. However, the substantial amount of time spent by the authors often results in less than satisfactory improvement in writing quality. This has frustrated many ESL/EFL writers and has become a major obstacle in effective writing production. In view of English writing as an information processing task by ESL/EFL writers, the difficulty of the task seems to be related to a number of factors, such as language information insufficiency, costly language information acquisition, and limited gain in language information use.

Recent developments in corpus linguistics have offered a new aspect for the study of language teaching and learning. A corpus is a large collection of sampled texts from existing written pieces or spoken records presented in electronic form [Marcus *et al*. 1993]. Subsequently, a corpus can be compiled to reflect the natural and actual occurrence of language use; thus, it provides abundant language resources for linguistics study and language acquisition. For instance, corpus linguistics has exploited corporal data for statistical analysis and comparative interpretation of language use [McEnery and Wilson 1996]. Efforts have also been directed to utilize corporal data for language learning references [Conrad 1999; Tsui 2005]. Indeed, corporal data allow many forms of data processing tasks to derive language information for various purposes. In particular, the idea of using a corpus as a book of reference is especially appealing to second language writers who beg for in-context examples of actual usage, so that uncertainty can be resolved and text fragments can be verified and re-used. To this end, an effective tool must focus on second language writers' special needs and provide referential corporal cases that fill in their unknown gaps of language use. The benefit of such a solution to ESL/EFL writing includes increased writing efficiency, improved writing quality, and better writing experience.

One of the primary tools for corpus exploration is concordancing. A concordance is a list of occurrences of a particular word with their immediate context drawn from a collection of texts. The targeted word is referred to as keyword. A concordance is usually displayed as a series of text lines in which the keyword is centered in its context. Concordancing has been frequently used as an analytic tool in linguistics for various issues of language study, such as analyzing word usage, computing word frequencies, finding and analyzing collocational units [Sinclair 1991]. The use of concordances in English Language Teaching (ELT) has also been advocated so that students learn certain language phenomena in an inductive way [Weber 2001; Sun 2003; de O'Sullivan and Chambers 2006]. However, Yoon [2005] pointed out that a concordance is not particularly helpful for ESL/EFL writers. Unlike language study and learning in which adequate exploration is encouraged, second language writing is a

time-constrained problem-solving task. Language information critically needed and directly useful in the ESL/EFL writing process is not readily available through concordancing.

From the point of view of information processing techniques, concordancing is a basic level application of information retrieval. As discussed above, concordancing employs a simple query form, straightforward hit or miss decision, and no ranking on retrieved results. In essence, a concordance is intended to be used as an overall observation tool on language phenomena of specific target words. The approach is less than satisfactory, and sometimes is even incapable of supplying useful information for the guidance and referential help needed by ESL/EFL writers. After all, writing is a production process that is based on language knowledge. In order to successfully convey the intended message, authors must contemplate the correct use of vocabulary and arrange appropriate word combinations and sequences, so as to construct concrete and coherent text content. Being language deficient in the text content construction process, ESL/EFL writers constantly face uncertainty and need to look for answers. Sometimes, they may even get off on the wrong foot due to misconceptions and spend much time in vain. Many ESL/EFL writers have been struggling with these problems and have not received sufficient help from the currently available language resources and tools. We stress that a new corpus exploration tool must be developed to better assist ESL/EFL writers. Such a writing assistance tool would need to offer more flexible types of queries and retrievals and would need to present the results that best suit users' needs in the writing process.

We propose a language information retrieval method to address the problem of corpus utilization. The approach is intended to help language deficient ESL/EFL writers find useful language use examples from a corpus and assist their decision making on correct language use. Our language information retrieval method includes three modules. The first module is a flexible expression model to allow a variety of combinations of semantic elements that reflect users' partial language knowledge. The semantic elements may include words, collocations, phrases, and formulaic expressions in complete or partial forms. Users can form a query by combining these semantic elements that are partly known and partly unknown. The second module is a selective retrieval mechanism with search options of exact match and partial match. Both options are supported, so as to retrieve adequate and useful examples according to the user’s confidence level on the query. The third module is an evaluation and ranking mechanism. After a submitted query and a selected search option, the retrieved results are evaluated for their consultation values and ranked accordingly. The final output is a list of exemplar sentences with decreasing consultation values so that users can receive help quickly. We implemented the approach and developed a corpus utilization tool for the purpose of ESL/EFL writing assistance. The tool is named SAW (Sentence Assistance for Writing). We used both objective measures and human subjects to gauge SAW's performance and observed

that SAW is capable of providing adequate references and satisfactory guidance for users' language information need in the ESL/EFL writing process. The results attest to the efficacy of the proposed language information retrieval method in exploiting corpus to assist ESL/EFL writing.

2. ESL/EFL Writing and Language Resource Tools

Compared to other language skills such as listening, speaking, and reading, writing is usually considered to be more difficult to develop and requires more in-depth language cognition to perform. High-quality written texts are founded on comprehensive language knowledge and its skillful utilization in production. Besides being language deficient, ESL/EFL writers are also subject to interference from their native languages. Both factors lead to text production that may contain incorrect words, grammar, and structures. For instance, to express the notion of music composition in a verb-noun pair, Chinese students tend to use "make", "create", or "produce" as the verb with "music" as the noun. Previous research indicated that ESL/EFL writers need three types of pre-requisite knowledge - words, collocations, and grammatical structures [Shei and Pain 2000; Altenberg and Granger 2001]. Among them, collocations are particularly unfamiliar to ESL/EFL writers. Collocations are a small group of words that co-occur with high frequency and become fixed word combinations. For example, the word "problem" as a noun usually goes with "cause", "create", and "solve" as verbs in an English verb-noun pair. In contrast, the combination of "make" as a verb and "problem" as a noun is rare, yet it is a straightforward translation from Chinese.

Collocations are commonly accepted agreements and habits in language use and are not transitive from language to language. Ilson *et al*. [1997] classified collocations into two types - grammatical collocations and lexical collocations. Grammatical collocations refer to compositions of a dominant word and a preposition, an article, or a conjunction, such as "decide on" and "determined by". Lexical collocations are frequently used combinations of noun, verb, adjective, and adverb, such as "strong tea", "absolutely not", and "notoriously difficult". Due to the interference of native languages, ESL/EFL writers are particularly prone to collocation errors, which may find resolution from better utilization of language resources.

A recent trend in language resources development has been to exploit a corpus for language use information in practical contexts [Biber *et al*. 1998]. Given the need to investigate the language phenomena of a specific word, corpus tools are used to provide both a quantitative assessment and actual examples of the different usage situations. Take Collins-Cobuild English Dictionary, published by Haper-Collins, as an example. The tool leverages an in-house corpus to provide collocational information on the target word. The collocational information includes a list of co-occurring words that are immediately before or after the lookup word. The co-occurrence list is also statistically assessed and ranked by

T-scores.

The collocational information of words has also been of interest to the lexicography community. In general, the purpose is to exploit corporal resources to investigate lexical behavior with the use of statistical measures of word co-occurrence. One of the representative studies is the work of Word Sketches, developed by Kilgarriff and associates at the University of Brighton [Kilgarriff and Rundell 2002]. Word Sketches is lexical profiling software designed for lexicographers to uncover the key features of a word's behavior. An inventory of grammatical relations is adopted to provide target types for developing collocation lists. Given lexicographic interest in a particular word, the output of Word Sketches is a list of words categorized by grammatical relations to the input word and associated by a statistical measure of its collocational significance.

Another type of corpus tool emphasizes offering reference examples of the lookup item and allows more variety of lookup items. For instance, VIEW, developed by Mark Davies at Brigham Young University, accepts lookup items in the type of a word, a partial word, a part-of-speech, or a phrase [Davies 2005]. The tool works on the British National Corpus as its language resource and produces a keyword in context as the primary output format. The output is a list of contexts, a specified window of words around the target, in which the lookup item appear. No filtering or ranking is attempted. Users are expected to look for useful information among the overloaded list of appearance.

Our position on language resources are in line with the current trend of corpus tools development. However, we argue that, for the purpose of ESL/EFL writing assistance, a specialized corpus tool should be developed to attend to the immediate need of language use decision in the writing task. Due to the difference in language cognition and in the levels of language knowledge, ESL/EFL writers often use a variety of query items in hope of obtaining potential references for the same intended expression element. However, current corpus tools are restricted in the types of allowed query; therefore, they are unable to provide help when users cannot come up with the appropriate queries. For example, some ESL/EFL writers are able to use the exact phrase "by and large" as a query to obtain its usage references. Other ESL/EFL writers who are not familiar with the phrase may try to use "large" or "by large" as query items according to their vague cognition. These incorrect or ambiguous queries would not lead to direct and useful references with the current corpus tools. This is exactly the dilemma of conflict between ESL/EFL writing need and current corpus tools. The lower the language knowledge of the ESL/EFL writer, the more language use help he or she needs. Yet, the usefulness of current corpus tools seems to hinge on the language level of the users. This usage obstacle has excluded many low to middle level ESL/EFL writers from getting help from corpus resources. Much of the problem can be attributed to a lack of proper design in the particular method of assisting less language skillful users.

3. Language Information Retrieval and Recommendation

We propose a method for language information retrieval to tackle the problem of ineffective corpus utilization for ESL/EFL writing. The language information retrieval method allows incorrect or ambiguous queries and tries to find the best reference examples so that users can explore and confirm the correct expression elements to convey their intended messages. Retrieved reference examples are evaluated in relevance to the user's query that indicates the language information needed. The final set of references are ranked and recommended in a sequential relevance order to provide users an efficient way to decide appropriate use of language. The method is designed to anticipate users with various levels of language knowledge and provide flexible query forms to cope with partial language cognition of users. The data source for retrieval is a corpus (or a set of corpora) and the retrieved results are ranked and displayed with sentences as a unit. This will facilitate users' ability to observe, learn, and confirm usage of certain expression elements in a complete exemplar sentence.

The language information retrieval method is a process that involves receiving a user query as an information need, retrieving sentences from data sources, and recommending a set of relevant sentences in ranked order. The method is implemented with three modules: expression element, retrieval, and ranking. The expression element module enables users to convey their language information need with a set of flexibly combined expression elements. The design is to allow a variety of query forms that come from different levels and aspects of language cognition. The retrieval module converts the expression element combination into its corresponding query condition and selects a set of sentences from the data source that match the query condition. The ranking module evaluates the set of selected sentences in relevance to the user's information need and recommends referential sentences to users in a ranked order.

3.1 Expression Element Module

The expression element module is designed to offer query flexibility for users to represent their various language cognition. Our implementation currently includes the following expression elements:

1. **Exact words:** Exact words are used when users are able to provide correct and complete spelling of a word or multiple words. For instance, users can specify a query with the word "asleep" and expect to obtain referential sentences that contain "asleep". Alternatively, multiple exact words, such as "fall asleep" can also be specified as a query.
2. **Prefix and Suffix:** Prefix (or suffix) of a word allows users to specify partial spelling of a word to represent their incomplete word memory. ESL/EFL writers often encounter the problem of inefficient reference lookup when they forget the exact spelling of the target word. Suppose a user needs usage references of the word "determine", yet without the

complete memory or cognition of the target word. Prefix (or suffix) can be used in this situation to initiate the query. We use the symbol "%" to represent uncertain or unspecified part of a word. For instance, users specify "deter%" to include all possible words that begin with "deter" and "%mine" to include all possible words that end with "mine".

3. **Wildcard:** A wildcard allows users to include an uncertain or unspecified single word as part of the query. We use the symbol "#" to represent a wildcard word which can match with any single word. Suppose a user needs to look up references for the 2-gram "responsible for" but is not sure about the word after "responsible". He/she can use the query "responsible #" to represent his/her referential need.
4. **Part-of-Speech (POS):** When the corpus includes POS tags, they can be used as part of the query condition. POS tags in English come with many types. The current study focuses on six primary types that include preposition (P), adjective (J), noun (N), adverb (D), verb (V), and other (O). POS tags provide an additional constraint in the query when users possess word class knowledge about the target word. Suppose a user is in need of usage examples of "native on", but is not sure about the preposition after "native". He/she can formulate the query as "native P".
5. **Subsequence:** Subsequence represents a sequence of zero to multiple unspecified words. It is designed to allow convenient inquiry for some phrasal structures and word combination that include various contexts in the middle, such as "rather ... than" and "exercise ... right". We use the symbol "*" to represent a subsequence. For example, usage examples of "rather ... than" can be looked up by the query "rather * than".

The set of expression elements constitute a space of language usage constraints in which users can select appropriate expression elements to represent a particular language information need based on existing language knowledge. The most common language reference need of ESL/EFL writers corresponds to a query type that includes both a specific part and an unspecified or constrained part. In a scenario where a user wants to express the approximate meaning of increasing the capacity of knowledge but is not sure whether to use "extend" or "expand" as a verb before "knowledge" as noun, he/she can specify the query "ex% knowledge" to obtain relevant usage examples for comparison and decision. For users who are not even familiar with both of the two words and just roughly know there should be a proper verb before the noun "knowledge", they can specify the query "V knowledge" to initiate a constrained exploration. The set of expression elements can also be combined flexibly into a sequence to form a more constrained query, such as "a pro% P" and "would rather V than V". When users are more knowledgeable on the language use for their intended meaning, a more constrained query can be used to provide a better focused retrieval.

3.2 Retrieval Module

Retrieval module performs the task of matching query with data items (sentences) in the corpus and selecting proper candidates that may provide useful language information. We adopt search rules of exact match and partial match. Exact match requires that all the occurrence conditions specified in the query are satisfied by the candidate sentences in the corpus. Partial match allows the selection of a candidate sentence in which some of the constraints in the query are not met. We anticipate that ESL/EFL writers will always have the problem of insufficient or even incorrect language knowledge. If a language tool provides only exact match, the effectiveness of the tool may largely depend on users' language knowledge levels. A less capable language user will find it difficult to initiate a successful search and often fail to obtain useful results in many tries. This will become a contradiction of the purpose that the tool should be designed to assist users who really need help. Partial match provides more flexible selection and unlocks the potential of deriving useful information even when users' queries are ill-formed.

Consider a scenario in which a low level ESL/EFL writer has only a partial idea of the two words "only" and "also" in the complete phrasal structure of "not only ... but also". The user can specify a query in the form of "only also" and partial match will retrieve sentences that contain the phrasal structure of "not only ... but also". This will offer the user an opportunity to recognize the complete phrase and learn its proper usage. In the situation where the user's language use in the query is incorrect, partial match is also useful in retrieving potentially relevant sentences and providing a chance for users to recognize and correct their errors. For instance, a user specifies "an university" as a query. The query has a wrong article "an" in front of the noun "university". A search will most likely be concluded with empty result if it is conducted by exact match. In contrast, partial match will be able to retrieve sentences that contain both words of "an" and "university" with some unconstrained words in between. The result of many partially matched examples and no (or little) exactly matched examples may allow users to recognize their errors and deduce the correct usage. In essence, we consider partial match as a necessary component for providing indirect referential help.

3.3 Ranking Module

The purpose of the ranking module is to improve the effectiveness of the reference consultation so that users may obtain necessary language information in a short list. The ranking module evaluates the selected candidates in relevance to the user query and sorts the reference list in a decreasing relevance order. We adopt the Multiple Sequence Alignment (MSA) technique [Needleman and Wunsch 1970], previously developed in bioinformatics, to perform the relevance evaluation between the user query and sentence candidates. The MSA technique has been used to evaluate the similarity between biological (such as DNA or protein)

sequences and produce sequence alignments for three or more sequences since they are difficult and time-consuming to align by hand. Given a set of sequences S_1, S_2, ..., S_n, n>=3, with different length, an MSA on the set of sequences is a set of aligned sequences of the same length, A_1, A_2, ..., A_n, where A_1 corresponds to S_1, A_2 corresponds to S_2, ..., and A_n corresponds to S_n. The aligned sequences allow gaps represented by the symbol "–" between elements.

Suppose we have three sequences S_1 = CCAATA, S_2 = CCAT, S_3 = CAATA, where the element set is {A, C, T}. Two of the possible MSAs are A_1 = CCAATA– – – – – – – – –, A_2 = – – – – – – CCAT – – – – –, A_3 = – – – – – – – – – – CAATA, and A_1 = CCAATT, A_2 = CCA–T–, A_3 = –CAATA. The fitness or goodness of all possible alignments can be computed as a summed score of all match scores between elements at the same location. We usually use a substitution matrix to indicate the match scores between all possible pairs of elements. There have been a number of algorithms, notably the Needleman-Wunsch algorithm [Needleman and Wunsch 1970], to search for the optimal MSA. However, the primary task of the ranking module is to derive a relevance order of the reference list with respect to the user query. We turn to the center star algorithm which takes a sequence as the center to align with all other sequences. Given a set of *n* sequences of length *k*, the complexity of the center star algorithm is $O(k^2n^2)$ and the goal is to derive a satisficing solution [Francis *et al*. 2003].

We proceed to formulate the symbolic sequences of sentences and queries. Elements of symbolic sequences include six POS tags (P, J, N, D, V, O), three other expression elements (exact word, prefix/suffix, and wildcard), and two alignment elements (don't care and gap). Symbolic representation for each element are "P" for preposition, "J" for adjective, "N" for noun, "D" for adverb, "V" for verb, "O" for other (types of POS tag), "=" for exact word, "%" for prefix/suffix, "#" for wildcard, "X" for don't care, and "–" for gap. Wildcard is a constraint specified by users in a query to accept any words at a designated location. Both don't care and gap are assigned to words and locations in the retrieved sentences for relevance evaluation.

Consider a scenario where a user specifies a query "a pro% P" and is converted as S_0. The symbolic sequence of S_0, noted as SS_0, is "= % P". Assume that three sentences, S_1, S_2, and S_3, in the corpus are retrieved by partial match and their corresponding symbolic sequences are noted as SS_1, SS_2, and SS_3. Words printed in bold and italic are those matched with the expression elements in the query: "a" (exact word), "pro%" (prefix), and "P" (preposition).

SS_0 = % P

S_1. This posed ***a*** particular ***problem for*** an agent.

SS_1. X X = X % P X X

S_2. Listening to all these personal accounts has had ***a profound*** effect ***on*** us.

SS_2. X X X X X X X = % X P X

S_3. Increasingly acid rain is ***a problem in*** Europe too.

SS_3. X X X X = % P X X

Match scores between pairs of elements can be set to reflect certain relevance policy in order to favor some types of language use over the others. In the current study, we used a preliminary setting of match scores between element pairs without deliberative policy design. The match score between two identical exact words is the highest and is set to 100. The next highest match score is between the pair of two identical prefix words (or suffix words) and is set to 50. A pair of two identical POS tags has a match score of 25. The match score of a pair of wildcards is set to 5. Both pairs of don't cares and gaps have a match score of 0 since they have little effect on the relevance. Match scores of all other pairs are set to -1 to reflect some extent of deviation. The current setting of match scores is for the purpose of demonstration and is not intended to obtain better performance. Table 1 lists all the symbolic elements and their match scores in a substitution matrix.

Table 1. The Substitution Matrix of Symbolic Elements in MSA

	P	J	N	D	V	O	=	%	#	X	–
P	25	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
J	-1	25	-1	-1	-1	-1	-1	-1	-1	-1	-1
N	-1	-1	25	-1	-1	-1	-1	-1	-1	-1	-1
D	-1	-1	-1	25	-1	-1	-1	-1	-1	-1	-1
V	-1	-1	-1	-1	25	-1	-1	-1	-1	-1	-1
O	-1	-1	-1	-1	-1	25	-1	-1	-1	-1	-1
=	-1	-1	-1	-1	-1	-1	100	-1	-1	-1	-1
%	-1	-1	-1	-1	-1	-1	-1	50	-1	-1	-1
#	-1	-1	-1	-1	-1	-1	-1	-1	5	-1	-1
X	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	-1
–	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0

Taking the symbolic sequence SS_0 as center, the center star algorithm produces a set of aligned sequences A_0, A_1, A_2, and A_3, with gaps inserted so that they are of equal length. The relevance of A_1, A_2, and A_3 with respect to A_0 can be computed as sum-of-pair scores and are denoted as C_1, C_2, and C_3. Let S(x, y) represents the match score between element x and element y. We have the following results.

A_0	–	–	–	–	–	–	–	–	=	%	P	–	–	–
A_1	–	–	–	–	–	–	X	X	=	X	%	P	X	X
A_2	X	X	X	X	X	X	X	X	=	%	X	P	X	–
A_3	–	–	–	–	X	X	X	X	=	%	P	X	X	–

$C_1 = S(–,–) + … + S(=,=) + S(\%,X) + S(P, \%) + S(–,P) + ... + S(–,X) = 93$

$C_2 = S(–,X) + … + S(=,=) + S(\%,\%) + S(P, X) + S(–,P) + ... + S(–,–) = 139$

$C_3 = S(–,–) + … + S(=,=) + S(\%,\%) + S(P, P) + S(–,X) + ... + S(–,–) = 169$

According to the relevance scores, the ranking module will recommend the reference list in the order of S_3, S_2, and S_1.

S_3. Increasingly acid rain is ***a problem in*** Europe too.

S_2. Listening to all these personal accounts has had ***a profound*** effect ***on*** us.

S_1. This posed ***a*** particular ***problem for*** an agent.

Given a user query that retrieves n sentences from the corpus, the computational time required for the ranking module is $O(k^2n^2)$, where k is the length of the longest sentence in the set of retrieved sentences. Suppose $n = 100$, $k = 20$, and one algorithmic step corresponds to 100 microprocessor instructions, the ranking module will consume up to 400 M microprocessor instructions. Given the computational speed of 200 Mips (million instructions per second) on an average 1 GHz personal computer, the execution time of the ranking module will be about 2 seconds.

3.4 System Implementation and Exemplar Process

We implemented the language information retrieval method in a prototype system called Sentence Assistance for Writing (SAW). The expression element module in SAW provides five types of expression elements. The retrieval module offers the options of exact search and partial search to select relevant sentences from corpus. In the ranking module, we use the MSA technique to align the retrieved sentences with the user query and produce quantitative assessment of their relevance. The final reference list of sentences is recommended to users in a decreasing relevance order in hope of assisting the user's language use decisions in an efficient way. The system architecture of SAW is shown in Figure 1.

Assume that a user is looking for usage references of the phrase "not only...but also". Without exact memory of the phrase, he/she specifies the query in the form of four sequential words "not only but also". Suppose that there are three relevant sentences, S_1, S_2, and S_3, in the corpus. S_1: We must also make sure that future generations not only read, but also have a real enthusiasm for visiting bookshops and libraries. S_2: This was not only humiliating but also very awkward for Baldwin. S_3: This is not only easier, but also more fun.

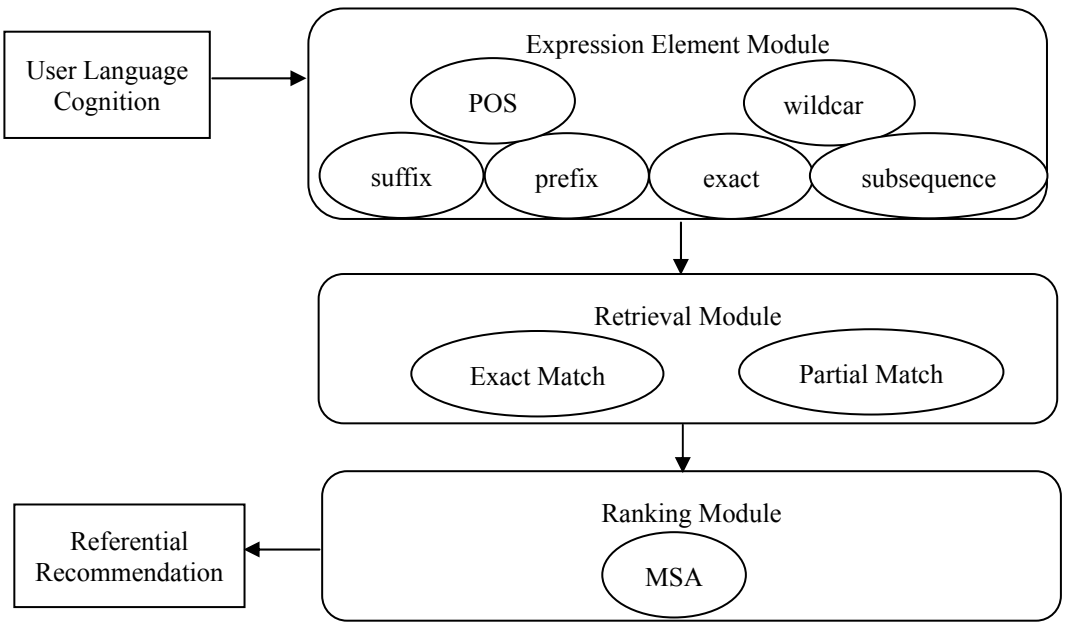

Figure 1. System Architecture of SAW

At first, the expression element module converts the user query to a symbolic sequence of four exact words. Next, the retrieval module will carry out the search option selected by the user. In the case of the exact search, the retrieval will result in an empty set since no sentence in the corpus contains the exact subsequence of "not only but also". In the case of the partial search, gaps between exact words in the query are allowed. As a result, the three relevant sentences, S_1, S_2, and S_3, will be matched and retrieved. The ranking module, then, proceeds to evaluate their relevance and produces an ordered list according to a relevance policy embedded in the substitution matrix. If the relevance policy simply favors shorter sentences, the ranked reference list will be S_3, S_2, and S_1. The above scenario is illustrated in Figure 2.

We use some query cases to demonstrate SAW's actual results. Four queries, "deter% by", "native P", "responsible #", and "either * or", are used as test examples. They cover the expression elements of prefix, exact word, POS, wildcard, and subsequence. Both search options for retrieval are tested to compare their effects. In particular, partial search is used in the query of "either * or" to retrieve relevant sentences, while exact search is used in the other

three queries. The actual system images are shown in Figures 3, 4, 5, and 6.

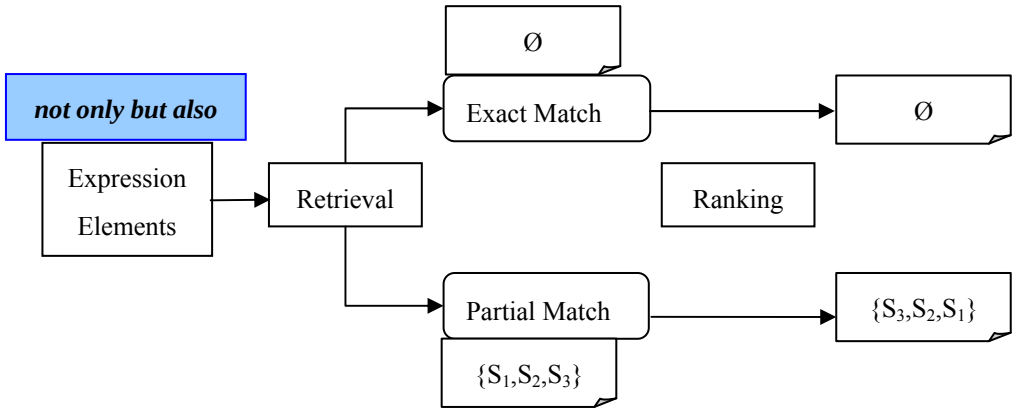

Figure 2. An Exemplar Language Information Retrieval Process

SAW allows users to select a particular corpus and specify a search option by clicking on certain buttons. The search result is a recommended reference list of complete sentences in the corpus. Ten sentences, as a convenient set size, are displayed in each result page. Users can ask for more references by clicking the "next page" button. In addition, SAW also provides statistics of the search result. For example, in the search result of "deter% by", there are 676 cases of "determined by", 27 cases of "deterred by", 12 cases of "determination by", 2 cases of "deterrence by", 2 cases of "determine by", and 1 case of "determining by". These statistics provide hints for better decisions as well as clues for further exploration, if needed.

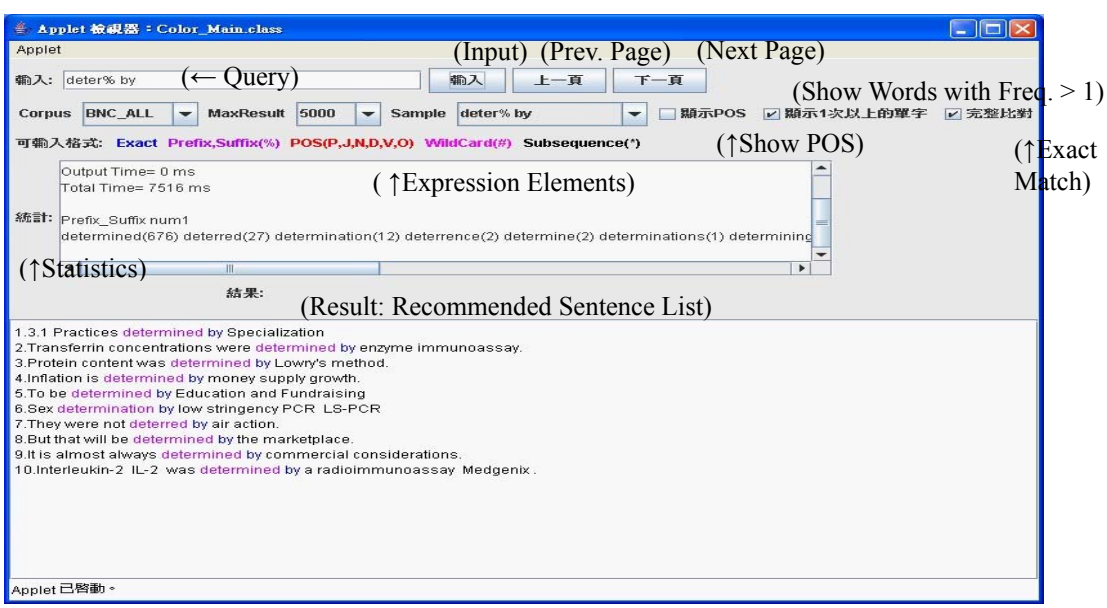

Figure 3. SAW's Output with Query "deter% by"

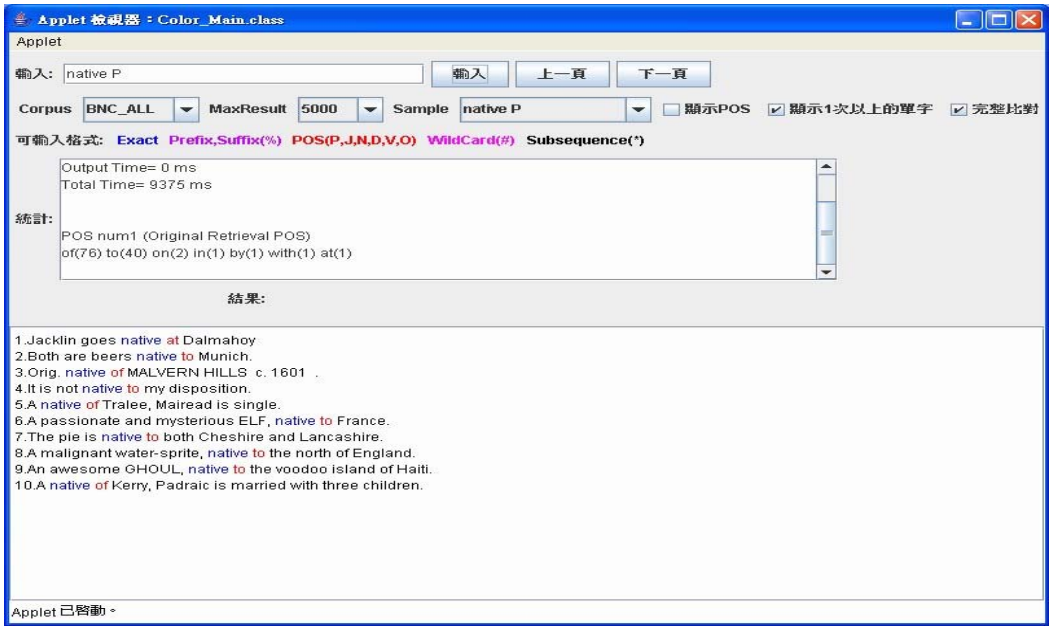

Figure 4. SAW's Output with Query "native P"

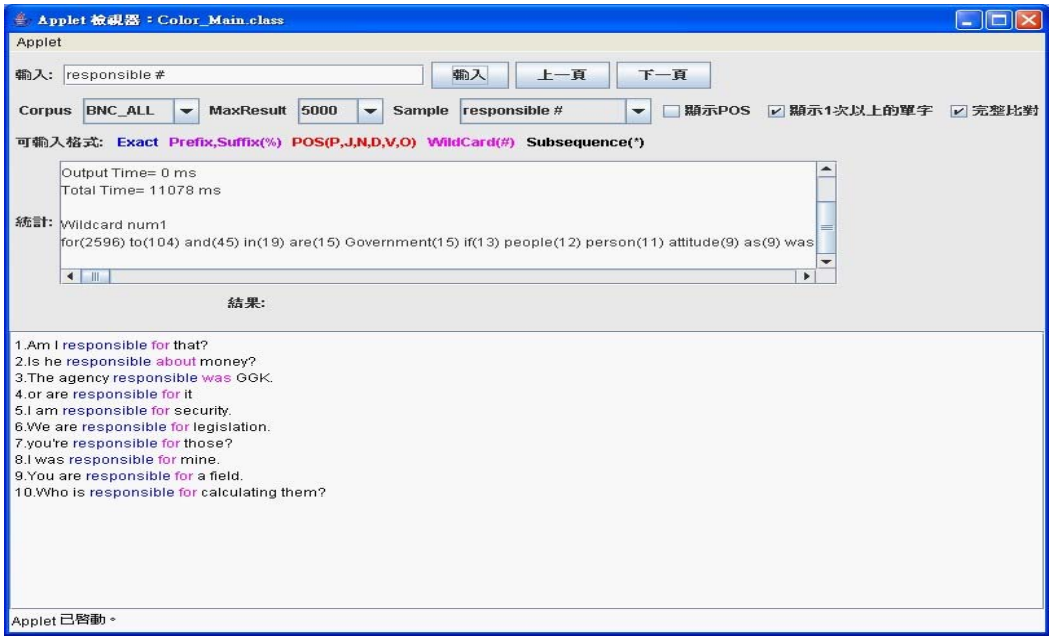

Figure 5. SAW's Output with Query "responsible #"

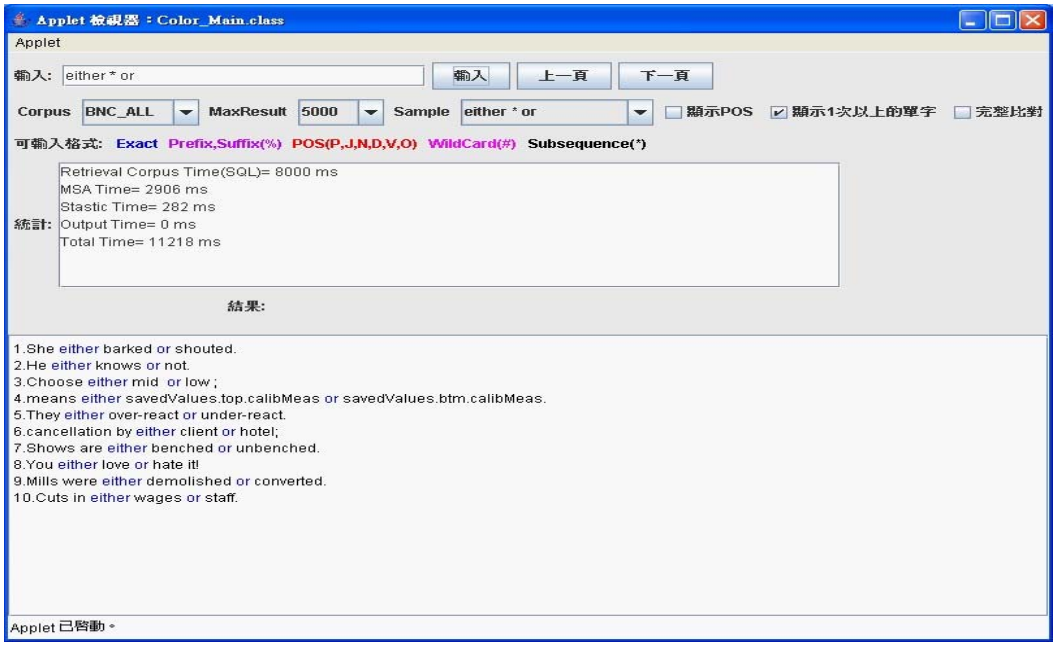

Figure 6. SAW's Output with Query "either * or"

Currently, when SAW is run on an average personal computer (2 GHz microprocessor), the system response time is approximately 10 seconds. The statistics box of SAW in Figure 6 shows recorded computational time of each major activity. While the ranking module (MSA time) takes less than 3 seconds, most of the computational time, *i.e.*, 8 seconds, is consumed by the retrieval module. This is due to the enormous size of the BNC corpus. SAW's response time can be further sped up by running on a more powerful machine and using a smaller corpus.

4. Experimental Evaluation

We used the British National Corpus (BNC) in the experiment. BNC is a balanced synchronic English text collection of samples of written and spoken language from a wide range of sources dating from 1974 to 1994. The corpus contains approximately 3.5 million sentences. POS used in the corpus are classified into sixty-two types. We adopted only six primary types (noun, verb, adjective, adverb, preposition, and other) and mapped BNC's all sixty-two types into one of the six primary types. The mapping rules are constructed manually such that a group of detailed POS types in BNC are mapped to a primary POS type in SAW. For example, a total of 25 verb-related POS types (VBB, VBD, VBG, VBI, VBN,…) in BNC are all mapped to the verb type in SAW. Actual POS mapping and conversion is automatically performed by programs.

4.1 Evaluation Method

The purpose of the experiments is to evaluate the performance of SAW in assisting ESL/EFL writers with language use references. We conducted two sets of experiments for both objective and subjective evaluations. The first set of experiments was a simulated English test for objective evaluation. Question items from accredited English capability tests were selected to simulate language use problems encountered by ESL/EFL writers. The second set of experiments was a subjective evaluation with human subjects based on test scores and questionnaires.

The sources of the English capability tests were College Entrance Exams in Taiwan, a total of 16 test sets from 1994 to 2005, and TOFEL, a total of 11 test sets from 2000 to 2002. Among the test sets, a subset of question items concerning the use of 45 phrases and 12 syntactic structures were selected as samples of language use problems. The experiment is designed to provide two angles of analysis on SAW's referential utility. The first angle is on the collocation use. Using 18 selected collocations as the sample group, we used four types of expression elements (exact word, prefix, POS, and wildcard) in SAW to simulate a variety of users' partial language knowledge. Recommended results for each query type were analyzed and compared based on their referential utilities. The second angle considers the condition in which users' queries contain partial errors. We selected 12 syntactic structures and simulated users' queries with incorrect language knowledge. Recommended results from SAW were observed and analyzed to see if information leading to correct language use can be derived.

The second set of experiment attempts to assess users' subjective responses to SAW's referential utility. We set up an English test by selecting a set of question items from the sourced test sets. The English test includes 16 single-choice questions and 3 question items of Chinese-to-English translation. Test scores from different experimental setups are compared and analyzed to evaluate SAW's performance. A set of questionnaires is also designed to gauge users' subjective perception of SAW's referential utility.

We consider a set of performance measures to conduct a quantitative evaluation. Some of the performance measures are adopted from the research area of recommendation systems [Sarwar *et al.* 2000] and are intended to evaluate SAW's helpfulness in obtaining language use information. Others are designed to solicit users' subjective response. The set of performance measures are as follows.

1. **Test score**. We use a set of English tests to simulate language use problems encountered by ESL/EFL writers. Test score is an objective measure of how well a subject can derive correct answers to question items either by his/her own language knowledge or by the referential information provided by SAW. Therefore, we can use subject's test score as a performance measure and compare test scores under different conditions to evaluate SAW's referential utility.

2. **Usefulness level.** Usefulness level is the user's subjective perception on whether the recommended results of SAW actually help solve the language use problems. We use a four-level measure that includes high, middle, low, and none. Usefulness level is a subjective indicator of SAW's performance and is solicited by the questionnaire.
3. **Satisfaction level.** Satisfaction level is the user's subjective perception of the overall language information relevance of SAW's recommended results with respect to users' query. We use a five-level measure that includes highly satisfied, satisfied, neutral, unsatisfied, and highly unsatisfied. Satisfaction level is also a subjective indicator of SAW's performance and is solicited by the questionnaire as well.
4. **Matchness.** Matchness is a relevance indicator of SAW's recommended results in response to the input query. We compute matchness as the ratio of the number of recommended sentences that are relevant to the user's expected language use to the number of sentences recommended by SAW. The definition is adopted from the performance measure of precision in information retrieval. Matchness is designed to reflect the proximity of recommended results to the user's language information need as indicated by the query. Matchness is an objective performance measure evaluated by the researchers and has a numeric value between 0 and 1.
5. **Reference cost.** Reference cost refers to the user's reading cost of the recommended results. In the current study, we use the number of words in a sentence as the reading cost. It is assumed that short sentences convey better referential effects on the targeted language use in simple context.

4.2 Simulated Language Use Problems on Collocations

We used question items in English tests to simulate language use problems of ESL/EFL writers in regards to collocations. Assuming a set of testing queries from specific to vague, recommended results were observed and analyzed to evaluate the referential utility to users with different levels of language knowledge. The English test conducted in the experiment consisted of question items concerning 10 two-word collocations and 8 multi-word phrases. Four types of expression elements, exact word, prefix, POS, and wildcard, were used to form the group of testing queries for each question item. Except queries by exact words, all other types of queries also contain combinational variations. Recommended results were evaluated by the performance measure of "matchness" and were averaged over the same type of query variations.

The 10 two-word collocations are "specialize in", "responsible for", "familiarity with", "relevance to", "deal with", "essential to", "native to", "composed of", "determined by", "guard against". The first type of query was composed of exact words of the target collocation,

such as "specialize in", and contains no variation. The second type of query was formed by the combination of an exact word and a prefix. One of the two words in the collocation was given as an exact word, and the first letter of the other word was given as prefix. For the collocation "specialize in", two variations, "s% in" and "specialize i%", were tested. The third type of query assigned a POS tag to one of the two words in the collocation. For the same collocation "specialize in", two variations, "V in" and "specialize P", were formed. The fourth type of query used a wildcard in one of the two words in the collocation, such as two variations "# in" and "specialize #" for the collocation "specialize in". Although some of the query variations are unlikely to be used by actual users, they are systematically formed as vague queries for the purpose of obtaining an approximate lower bound matchness performance. In the experiment, SAW was instructed to provide the top ten referential sentences in the ranked list for users. The matchness of the recommended results was computed over the selected ten sentences. The average matchness of exact words, prefix, POS, and wildcard over the ten collocations are 1, 0.45, 0.305, and 0.14, respectively. The results indicate that SAW is able to provide at least one out of ten matched referential sentences even with vague queries.

In addition to two-word collocations, we also tested language use reference needs of phrases composed of three and four words, which include “a proportion of”, “in danger of”, at one time”, “as a result of”, “play a virtual role”, take the form of”, “in the presence of”, and “make it impossible to”. A set of testing queries was constructed in the same way as the queries for two-word collocations. The only difference is that the number of variations increases as the length of the target phrases increases. Again, some of the query variations may not be actually used by users and are adopted for system evaluation to estimate lower bound performance. Experimental results confirm that the average matchness of the referential information provided by SAW increases as the number of words in the query increases. This is due to the stronger structureness of the multiple-word phrases. Even when one of the words in the query is vague or absent, users still can obtain useful language use information in the recommended results. The average matchness of exact word, prefix, POS, and wildcard over the ten collocations are 1, 0.78, 0.51, and 0.42, respectively. Again, the results indicate that SAW provides at least four out of ten matched referential sentences in multiple-word phrase queries.

4.3 Phrasal Structure and Incorrect Query

Another subset of experiments is concerned with the referential utility of the expression element of subsequence. We selected a set of 12 phrasal structures with indefinite numbers of words between fixed words, including "enable...to", "derive...from", "expose...to", "not only...but also", "would rather...than", "distinguish...from", "expand...into", "provide...with", "the same...as", "either...or", and "so...that". The expression element of subsequence can be

used to form the proper query, such as "enable * to", and "not only * but also", to retrieve referential sentences containing the use of the target phrasal structure.

We simulated users' incorrect language knowledge with partial errors in two testing queries for each target phrasal structure. For example, testing queries of "enable * and" and "enable * but" were deliberatively tried to expect the retrieval of the target phrasal structure "enable...to". In preparing testing queries with partial errors, we replaced the correct prepositions in the target phrasal structures with "and" and "but". The only exception was for the phrasal structure of "not only ... but also" in which "but" was replaced by "and" and "or". The results of the partially erroneous queries were evaluated by the average matchness for each target phrasal structure. Again, these false queries may not be practical but can serve as an estimation purpose.

The experimental results show that the utility of recommended results varies with the coupling strength of keywords in the target phrasal structure. For example, the coupling strength between "enable" and "to" are the strongest among the 12 phrasal structures in the experiment. Even when the preposition was replaced with an incorrect one in the testing query, recommended referential examples retrieved with partial match still contained sufficient and specific usage examples of the target phrasal structure "enable...to". In contrast, the coupling strength between keywords in the phrasal structures "either...or" and "so...that" was too weak for SAW to provide referential examples under incorrect queries. However, a set of recommended results with very low matchness from the user's incorrect query seems to carry an effect of indirect references. When users see no usage examples of their incorrect queries, they may induce the possibility that their queries contain errors and further recognize the correct phrasal structure from the common usage pattern in the recommended results.

The experiment demonstrated that SAW's mechanism is capable of allowing false queries or vague queries and providing users with sufficient referential utility to obtain useful language use information. Such a capability is distinguishable from most corpus tools and is intended to break up the conflicting dilemma of requiring sufficient language knowledge to successful usages of language resource tools. In other words, SAW offers flexibility in retrieving referential examples, which is especially important to low to middle level ESL/EFL writers.

4.4 Questionnaire and Test Score

The purpose of conducting an English test and a questionnaire is to measure SAW's referential utility based on users' subjective experiences. The English test consisted of 19 question items, which included 16 single selection items and 3 Chinese-English translation items. We used SAW to prepare referential examples from the BNC corpus for each question item. Depending on the experimental setup of the subject's group, these referential examples for question items

may be provided to some of the subjects during their test taking. The subjects also evaluated the usefulness level and the satisfaction level of the provided referential examples when answering the question items. After completion, test sheets were graded and test scores of different groups were compared to indicate the level of help SAW provided for the subjects in simulated writing conditions.

We recruited a total of 98 students in a local junior college as the experimental subjects. An accredited English vocabulary test was conducted to measure the vocabulary level of the subjects, which generally corresponds to their overall language proficiency. According to Nation's research [Nation 1993], the percentage of 1000 most common vocabulary words in general English articles is about 74% and the coverage percentage of 2000 most common vocabulary words grows to 81%. We conducted a "1000 English Word Test" [Cobb 2007] on our experimental subjects. Based on the vocabulary test, the average recognition percentage of the experimental subjects on the 1000 most common vocabulary words is 80.44%, and the average recognition percentage reduces to 46.10% on the 2000 most common vocabulary words. In other words, the average subjects are on a word recognition level of less than the 2000 most common vocabulary words, which may be classified as low to middle level ESL/EFL writers.

We divided the subjects into three groups by random selection from the population of the subject pool. It is assumed that variation of English proficiency in each group is normally distributed and the average language proficiency level is approximately the same among the three groups. The experimental setup is described as follows. Group 1 took the test in the condition of a general English capability test without any help. Group 2 took the same test with referential examples of high matchness on each question items. Group 3 took the same test with referential examples of low matchness on each question items. A total of 98 questionnaires and test sheets were collected, in which 12 responses were nullified due to no response or obvious faults, such as invalid selection and singular selection. Among the validated 86 responses, 44 responses came from Group 1, 20 responses came from Group 2, and 22 responses were from Group 3.

Table 2. Comparison of Selection Test Scores under Different Help Condition

	Test Score		Performance Improvement
	Mean	Standard Deviation	
Group 1 (no help)	4.42	1.95	(baseline)
Group 2 (high matchness help)	6.69	2.23	51.4%
Group 3 (low matchness help)	5.97	1.85	35.1%

Test responses were evaluated by comparing test scores under different conditions of referential help. The test contains a set of 16 question items of single selection. Each question item is credited with 0.625 points and the full score is 10 points. The results, summarized in Table 2, show that test scores (both Group 2 and Group 3) with referential help by SAW are higher than those without help (Group 1). The average test score of Group 1 is 4.42 points with a standard deviation of 1.95. Group 2 took the test with referential examples of high matchness (0.63) and got an average score of 6.69 points with a standard deviation of 2.23. Group 3, with referential examples of low matchness (0.35), obtained an average score of 5.97 points and a standard deviation of 1.85. The performance improvement rates of 51.4% and 35.1% suggest that SAW's referential utility leads to better language use decisions in ESL/EFL writing.

We calculated the significance level of the statistical hypothesis test to further verify that the increase of test scores is not an event by chance. Assuming that subjects' test scores are normally distributed, we formulated the statistical significance test as follows. The null hypothesis (H_0) states that SAW's referential utility has no effect on improving subjects' test scores. The alternative hypothesis (H_1) states that subjects' test scores increase due to SAW's referential utility. The control group is the subjects in Group 1 who took the test without any referential help and resulted in an average test score of 4.42 points. The treatment groups are the subjects in both Group 2 and Group 3 who were offered referential examples by SAW during their tests. We first conduct the statistical hypothesis test between the control group of Group 1 and the treatment group of Group 2 (20 subjects with referential examples of high matchness). With the significance level set at $\alpha = 0.05$, the threshold to refute the null hypothesis is computed by $4.42 + (1.65*1.95/(20)^{1/2}) = 5.14$, where 1.65 is the Z-score at $\alpha = 0.05$ and 1.95 is the standard deviation of the control group. The average test score of Group 2 is 6.69 points, which is higher than the threshold of 5.14 points. Similarly, for the treatment group of Group 3 (22 subjects with referential examples of low matchness), the threshold to refute the null hypothesis is computed by $4.42 + (1.65*1.95/(22)^{1/2}) = 5.10$, where 1.65 is the Z-score at $\alpha = 0.05$ and 1.95 is the standard deviation of the control group. The average test score of Group 3 is 5.97 points, which is higher than the threshold of 5.14 points. Therefore, with the comparison of both treatment groups to the control group, we can refute the null hypothesis and support the alternative hypothesis, at the statistical significance level of 0.05, that SAW's referential utility helps increase the subject's test scores.

Besides question items of single selection, the test also contains three question items of Chinese-to-English translation. This part of the test is a closer simulation to ESL/EFL writing problems. Among the 86 test responses from Groups 1, 2, and 3, 44 test responses are from subjects of Group 1 who performed the translation without any referential help, while 42 test responses are from subjects of Groups 2 and 3 who were provided with the same set of

referential examples. The translation answers are graded based on three components: syntactic structure 50%, vocabulary 25%, and grammar 25%. The full score of the translation test is 10 points and the test sheets were graded by an experienced English teacher. The results are shown in Table 3. The average score of Group 1 is 4.20 points with a standard deviation of 2.56. The average score of Groups 2 and 3 is 5.83 points with a standard deviation of 2.46. Subjects who performed the translation with referential examples obtained an average score of 1.63 points higher than subjects without referential help. This score increase represents a 38.8% rate of performance improvement. Again, in the statistical hypothesis test, the threshold to refute the null hypothesis is computed by $4.20 + (1.65*2.56/(44)^{1/2}) = 4.97$, where 1.65 is the Z-score at $\alpha = 0.05$ and 2.56 is the standard deviation of the control group (Group 1). The average test score of Groups 2 and 3 is 5.83 points, which is higher than the threshold of 4.97 points. Therefore, we can refute the null hypothesis and support the alternative hypothesis, at the statistical significance level of 0.05, that SAW's referential utility helps improve the subject's performance on translation.

Table 3. Comparison of Translation Test Scores under Different Help Conditions

	Test Score		Performance Improvement
	Mean	Standard Deviation	
Group 1 (no help)	4.20	2.56	(baseline)
Group 2 & 3 (referential help)	5.83	2.46	38.8%

Usefulness is a performance measure designed to solicit a user's evaluation of how useful the referential examples are for making language use decisions. While responding to the test, our experimental subjects also rated the referential examples provided by SAW on a scale of 1 to 4, with 4 being the highest level of usefulness. As a result, Group 2 rated the referential examples for single selection questions with an average usefulness of 2.78, while the number given by Group 3 is 2.70. Referential examples for translation questions were rated by the two groups with an average usefulness of 2.80. Overall, SAW's referential utility was given usefulness in the range of 2.7 to 2.8, which may be judged as slightly above-average usefulness. We also solicited subjects' responses to sentence length (with 1 being too short and 3 being too long) and satisfaction level (on a scale of 1 to 5) of the referential examples. The responded numbers indicate that the sentence length is appropriate (2.09) but the subjects are less than satisfied (2.76) with the referential examples. Sampled interviews revealed that, while referential sentences did help users make better language use decisions, some of the sentences from the BNC corpus may contain words that are too difficult for low to middle level users and cause frustration. This problem can be solved by selecting an appropriate corpus for users of different levels of language capability.

4.5 Usage Scenarios based on Actual ESL/EFL Writing Samples

As part of the field study in the research, we collected a set of ESL/EFL students' writing samples from a local junior college. Due to their insufficient English proficiency, many errors can be found in their composition assignments. We selected a few representative sentences written by the students for running a scenario of using SAW in realistic ESL/EFL writing by low to middle level writers.

The first exemplar sentence is "I would rather going shopping than staying home." It is clear that the student who wrote the sentence did not have a complete knowledge of the phrasal structure "would rather...than". The correct sentence should be "I would rather go shopping than stay home." We used SAW with two types of queries "would rather V than V" and "would rather * than *" to simulate users' language use problems. Users who have better knowledge of the phrasal structure but are not absolutely sure may use the query "would rather V than V" to verify their choices. SAW recommended five usage examples in the form of complete sentences. The matchness of the recommended results is 100% since all five sentences contain the target phrasal structure. Alternatively, the less constrained query "would rather * than *" may be used to explore the options in the unspecified parts. Among the top ten referential sentences recommended by SAW, seven of them present correct forms of verbs. The matchness of the recommended results is 70%. In both cases, an average user should be able to derive correct language use decisions from the referential examples provided by SAW.

The second exemplar sentence is "I need a nest of glasses." The student who wrote the sentence was mistaken on the article for the noun "glasses". A straightforward query to explore a set of possible articles before the noun "glasses" would be "a # of glasses". SAW's recommended results include the use of "tray", "set", "couple", and "pair", while the student's first guess "nest" does not appear. By observing the referential examples, the student is likely to make a correct selection on the appropriate article for his/her expressional purpose.

The third exemplar sentence is "My mon always gets sick so my father tells me you have to care of her." The novice writer made errors in the spelling of "mom" and the phrase "take care of". Two separate queries were used to simulate the student's possible guesses to looking for answers. The first query was "mo% * father". Among the top ten recommended referential examples, seven sentences include the correct spelling of "mother". The second query was "to care of". Again, eight of the top ten recommended sentences show the use of "take care of". Based on the three scenarios above, it is reasonable to suggest that, with the help of SAW, many of the errors made by apprentice ESL/EFL writers can be avoided and self-corrected.

5. Related Research

Recent research in facilitating English learning with computational techniques has received much interest. We discuss some of them that are relevant to our research ideas. One of the

notable studies in facilitating ESL/EFL learners is the REAP project at CMU. The project developed computer-assisted techniques to improve reading comprehension and vocabulary learning with adaptive text passages selection [Heilman *et al*. 2006]. A lexical acquisition model for an individual reader is used to provide student-specific reading practice and remediation. The project conducted a series of experiments to compare students' reading ability and vocabulary knowledge before and after the assistance of REAP. The experimental evaluation includes the use of questionnaires as well as comprehension and vocabulary test to examine students' progress. Quantitative results are analyzed to determine the statistical significance. In the same vein of assisting users to improve their language performance, our research focuses on developing a mechanism to suggest referential examples and help resolve the uncertainty of language use in the writing context. We also apply the same principles of instructional and statistical validation in evaluating our research results.

Another related research is the Gsearch tool developed at the University of Edinburgh [Corley *et al*. 2001]. The tool is designed to facilitate the selection of sentences from text corpora by syntactic criteria. The query notation allows flexible grammatical and lexical constraints to be specified. A search can be formulated by four components - corpus, grammar, goal, and option, in a command-like form. The output is a set of sentences either in SGML format or in visualized syntax trees. The target users are linguists who wish to investigate lexical and syntactic phenomena in unparsed corpora. The tool can also be used in an instructional context for advanced students in linguistics study. While the operational purpose of retrieving a subset of sentences from a corpus is the same, our research emphasizes a query notation that is simpler for less language knowledgeable users to express their questions of language use. Referential sentences selected by SAW are ranked and recommended to users so that they can quickly resolve their uncertainty of language use in a short list. Our research goal is to remove the obstacles for low to middle level ESL/EFL students to tap into the corporal resources.

Another type of corpus resource application is the work of [Chung *et al*. 2005], in which a method is proposed to extract bilingual collocations from a parallel corpus. Both statistical and linguistic information are integrated and trained to identify collocation types and instances in each monolingual corpus. The method, then, adopts a statistical word alignment technique and dictionaries to extract collocation translation equivalents from the parallel corpus. A collocation reference tool, called TANGO, is built to support searching for collocations and translations of a given word. The tool is suggested to be applicable to machine translation, cross language information retrieval, and computer assisted language learning. From the point of view of ESL/EFL users, TANGO is useful for seeking collocation information and instances with a given word in either the native language or the second language. In contrast, SAW is designed to allow more flexible types of query and provide effective usage references

for second language writing. SAW also pays special consideration to referential help for ill-formed queries due to the user's insufficient language knowledge. Both tools can complement each other to enable better access to language information in corpus resources.

6. Conclusion

The central notion of our research is that the use of language resources for writing assistance must consider the language barriers of ESL/EFL writers. We have proposed a method to address the issue and provide evidence to support the value of an easily-accessed referential utility. A set of experiments was conducted to simulate language use problems and evaluate SAW's referential utility. Based on the experimental results, we conclude that the proposed language information retrieval method is effective in providing help to ESL/EFL writers.

We believe that the proposed method can be further extended to cater to different levels of ESL/EFL writers. One way of providing the right level of referential examples is to tap into an appropriate corpus that provides materials of a suitable language level. For example, SAW will best serve a child/adolescent ESL/EFL user with a corpus of beginners' reading materials. Another extension is to use a secondary filtering step after the relevant examples have been retrieved from the corpus. Sentences that are composed of vocabulary outside of a specified range of familiarity, such as the 2000 or 5000 most common words, can be filtered out. We will make the extensions in our future work and conduct more experiments with different settings. For example, we plan to conduct an online English test with online referential retrieval and record subjects' actual response time in getting answers. We also plan to investigate the effects of incorporating additional language resources, such as dictionaries and POS tags.

The method and the implemented system, SAW, can also be used as a pedagogical tool for writing practice in self-learning and in instructed assignments. We hope to have shown that, by the proposed method, a significant step can be taken to reduce the obstacles of language communications and to increase the capability and productivity of ESL/EFL writers.

References

Altenberg, B., and S. Granger, "The Grammatical and Lexical Pattern of MAKE in Native and Non-native Student Writing," *Applied Linguistics*, 22(2), 2001, pp. 173-195.

Biber, D., S. Conrad, and R. Reppen, *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge University Press, 1998.

Chung, T. C., J. Y. Jian, Y. C. Chang, and J. S. Chang, "Collocational Translation Memory Extraction Based on Statistical and Linguistic Information," *International Journal of Computational Linguistics and Chinese Language Processing*, 10(3), 2005, pp. 329-346.

Cobb, T., Compleat Lexical Tutor - 1,000 English Word Test, http://www.lextutor.ca/, 2007.

Conrad, S., "The Importance of Corpus-based Research for Language Teachers," *System*, 27, 1999, pp. 1-18.

Corley, S., M. Corley, F. Keller, M. Crocker, and S. Trewin, "Finding Syntactic Structure in Unparsed Corpora: The Gsearch Corpus Query System," *Computers and the Humanities*, 35(2), 2001, pp. 81-94.

Davies, M., "The Advantage of Using Relational Databases for Large Corpora," *International Journal of Corpus Linguistics*, 10(3), 2005, pp. 307-334.

de O'Sullivan, I., and A. Chambers, "Learners' Writing Skills in French: Corpus Consultation and Learner Evaluation," *Journal of Second Language Writing*, 15(1), 2006, pp. 49-68.

Francis, Y. L., N. L. Ho, T. W. Lam, W. H. Wong, and M. Y. Chan, "Efficient Constrained Multiple Sequence Alignment with Performance Guarantee," *IEEE Computational Systems Bioinformatics*, 2, 2003, pp. 337-346.

Heilman, M., K. Collins-Thompson, J. Callan, and M. Eskenazi, "Classroom Success of an Intelligent Tutoring System for Lexical Practice and Reading Comprehension," In *Proceedings of the 9th International Conference on Spoken Language Processing*, 2006, Pittsburgh, U.S.A., paper 1325-Tue1WeS.4.

Ilson, R., B. Morton, and B. Evelyn, *The BBI Dictionary of English Word Combinations*, John Benjamin Publishing Company, Amsterdam, 1997.

Kilgarriff, A., and M. Rundell, "Lexical Profiling Software and its Lexicographic Applications - A Case Study," In *Proceedings of the 10th European Association for Lexicography (EURALEX) International Congress*, 2002, Copenhagen, Denmark, pp. 807-818.

Kobayashi, H., and C. Rinnert, "Effects of First Language on Second Language Writing: Translation versus Direct Composition," *Language Learning*, 42(2), 1992, pp. 183-209.

Marcus, M. P., B. Santorini, and M. A. Marcinkiewicz, "Building a Large Annotated Corpus of English: the Penn Treebank," *Computational Linguistics*, 19, 1993, pp. 313-330.

McEnery, T., and A. Wilson, *Corpus Linguistics*, Edinburgh University Press, Edinburgh, 1996.

Nation, I. S. P., "Measuring Readiness for Simplified Material: A Test of the First 1,000 Words of English," In M. L. Tickoo (Ed.), *Simplification: Theory and Applications*, *RELC Anthology Series*, 31, 1993, pp. 193-203.

Needleman, S. B., and C. D. Wunsch, "A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins," *Journal of Molecular Biology*, 48, 1970, pp. 443-453.

Sarwar, B., G. Karypis, J. Konstan, and J. Riedl, "Analysis of Recommendation Algorithms for E-Commerce," In *Proceedings of the 2nd ACM Conference on Electronic Commerce*, 2000, Minneapolis, Minnesota, U.S.A., pp. 158-167.

Shei, C. C., and H. Pain, "An ESL Writer's Collocational Aid," *Computer Assisted Language Learning*, 13(2), 2000, pp. 167-182.

Sinclair, J., *Corpus, Concordance, Collocation*, Oxford University Press, 1991.

Sun, Y. C., "Learning Process, Strategies and Web-based Concordancers: a Case Study, " *British Journal of Educational Technology*, 34(5), 2003, pp. 601-613.

Tsui, A. B. M., "ESL Teachers' Questions and Corpus Evidence," *International Journal of Corpus Linguistics*, 10(3), 2005, pp. 335-356.

Weber, Jean-Jacques, "A Concordance and Genre-informed Approach to ESP Essay Writing," *English Language Teaching Journal*, 55(1), 2001, pp. 14-20.

Yoon, H., "An Investigation of Students' Experiences with Corpus Technology in Second Language Academic Writing," PhD thesis, Ohio State University, U.S.A., 2005.