

## ***Course-Specific Corpora in the Classroom: A News Media English Class in Taiwan***

**Shiauping Tian**

*National Taiwan University of Science and Technology, Taiwan*

**Limin Liu**

*Chung Yuan Christian University, Taiwan*

This study explores the feasibility of constructing a learning tool with the benefits of data-driven learning from a course-specific corpus and integrating it into a news media English class. Given the popularity of news media English classes in some Asian countries and the call in recent years for instructors to construct smaller and more specialized corpora to meet specific instructional objectives, an on-line system for assisting teaching and reading of English news was developed and used in a university-level news media English class in Taiwan. A lesson incorporating the system was designed and implemented. The lesson aimed at teaching the discourse features of news stories and improving English in general. After the lesson, students' responses to using the system were obtained by a questionnaire survey. Results of the questionnaire survey indicated an overall positive response to incorporating the learning tool into a news media English class along with suggestions for system design. The paper concludes with implications for future use of the system and possible ways to improve the design of systems for similar purposes.

For more than a decade now, corpora and concordance tools have been given increasing prominence in the foreground of language teaching and

learning. Although corpora concordancing had been used in the design of language teaching materials long before the widespread use of computers, it was not until recent years that enthusiasm for concordancing in language classrooms emerged. As computers become more powerful in processing and less limited in data storage, bringing corpora concordancing directly into the learning process of language learners has become feasible. Researchers and language teachers have proposed and implemented classroom concordancing in the teaching of lexical collocations, affixes, polysemous words, grammatical patterns, and discourse features (Ball, 1998; Dyck, 1999; Kettemann, 1995; Tribble, 1997). Given the potential for classroom concordancing in various areas of language teaching, it is worthwhile to explore the possibility of introducing concordancing into a popular course in Asian countries—News Media English.

## **DATA-DRIVEN LEARNING**

The nature of learning through concordancing is well summed up in the term “Data-driven learning” (DDL) proposed by Tim Johns (1991). DDL is basically learning by working through vast amounts of authentic language data to discover linguistic patterns. The computer provides multitudes of organized language data on demand, and students analyze the data to figure out patterns of language use. This approach enhances the effectiveness of language learning by providing a few advantages. First, students have quick and easy access to plentiful examples of authentic language use. With concordancing software, students can easily obtain a list of contextualized examples of the investigated feature.

Second, students take on the active roles of researchers, sorting through massive language data to discover rules and patterns embedded in the data. Declarative knowledge about the target language becomes the result of hypothesis testing and inference drawing instead of rule-application exercises or spoon-feeding. Therefore, inductive reasoning is encouraged in the process.

Third, learning becomes more learner-centered with the learner “tak[ing] part in his or her own profiles of meanings and uses” (Johns, 1991, p. 30). In other words, learner autonomy is cultivated by encouraging students to be responsible for their own learning.

Finally, the focus of learning is shifted from the product of analysis toward the process of analysis. As Aston (1998) noted, in addition to developing linguistic competence, working with corpora also fosters development of procedural knowledge—knowing how to interpret language data and generalize from them to solve problems. This last aspect is particularly significant to language teaching and learning. Since the ultimate goal of teaching is to cultivate independent learners with strategies to deal with and learn from new materials in new contexts, any learning tool with the capacity of fostering such skills would be highly valuable.

With corpora concordancing, the computer provides organized instances of real language use and facilitates learning by making linguistic patterns more easily detectable for learners to work out solutions to real language problems. As expressed by Tribble (1990, p. 11), “what the concordancer does is to make the invisible visible.” As a result, concordancers are considered by many language teachers to be “the type of software that most closely approaches fulfilling the potential of computer in language learning” (Stevens, 1995, p. 2), and a number of studies have been conducted to provide empirical evidence to the effectiveness of corpora concordancing in language classrooms (Hadley, 2000; Ilse, 1991; Lee & Liou, 2003; Someya, 2000; St. John, 2001).

Though most researchers see the instructional potentials in corpora concordancing, it is not presented as a perfect solution without limitations or challenges. Among frequently mentioned challenges are the linguistic complexity of authentic materials and the overwhelming amount of data generated by the concordancer. As much as we like to embrace the use of authentic materials in language learning contexts, it is highly possible that they present language far beyond EFL learners’ proficiency and create difficulty in processing. The challenge of concordancing with authentic materials is that the linguistic

difficulty and data load could lead to frustration deep enough to block the learners from further attempt in corpus work, which would then completely undermine the merits of using authentic materials (Tribble, 1997).

Another concern is the cognitive approach entailed in corpus work. As mentioned earlier, corpora concordancing fosters mostly an inductive approach to learning that is usually ignored in traditional instructional practices in Asia. By introducing concordancing into the classroom, instructors wish to cultivate a cognitive approach that would allow more learner participation. Nevertheless, different claims and mixed results from previous studies (Gross, 1991; Richards & Lockhart, 1994; Shaffer, 1989) might have suggested that the inductive approach of discovery-learning may not be the best choice for all language learners.

Other challenges fall under the technical aspect of concordancing, such as the availability of corpora suitable in size and content for various instructional objectives and the speed of concordancing. Issues like copyrights and fees often obstruct individual attempts to construct corpora for specific instructional purposes. In addition, the speed of concordancing is often found to plague on-line concordancers, since learners in this high-tech era have grown accustomed to instant retrieval of information from the Internet (Fletcher, 2004; Rundell, 1996; Stevens, 1995).

Despite the challenges and concerns, concordancing in instructional contexts is generally acknowledged as a valuable approach that promises more possibilities with a wider variety of resources as technology advances over time.

## **THE USE OF SMALLER AND MORE SPECIALIZED CORPORA**

As the computer gets more powerful and sophisticated, allowing the creation and processing of larger corpora, it seems only natural to assume that concordancing with larger corpora creates better results for language learning. Large general corpora with hundreds of millions of words, however, can be

unwieldy, and the size of concordance output may be unmanageable for EFL learners. As stated by Tribble (1997), “large corpora can present significant methodological problems for less experienced users—the risk of drowning in data is high!” Therefore, the potential of using smaller and more specialized corpora with texts in specific domains or genres has been explored (Aston, 1997, 1998; Rundell, 1996; Tribble, 1997). Smaller and more specialized corpora are easier to construct and manage, more relevant to the specific needs and interests of students, more effective in calling attention to features that repeatedly appear in a particular text type or genre, and can be used to train skills in using corpora.

Tribble (1997) advocates creation of micro corpora with easily obtainable texts that are tailored to the purposes of specific language courses. According to him, smaller corpora with texts in a single domain are conducive to identification of high-frequency words in the domain, increasing awareness of collocation and colligation, and investigation of grammatical features and discourse organization of a particular genre. And micro corpora with multiple themes have the additional advantage of facilitating contrasting of lexical use in different domains.

## **READING TO LEARN AND LEARNING TO READ**

As mentioned earlier, news media English and other similar courses are offered in many colleges and universities in Asia. News materials from various sources are frequently incorporated into the class because they provide easily accessible authentic materials about life events that can be more motivating than regular textbooks. In addition, news materials are usually considered an effective medium for learners to develop language skills (especially reading skills), critical thinking, cultural awareness, and life skills (Aiex, 1988, 2000; Harumi, 1993; Huang, 2003). As Internet use became more widespread in recent decades, instructors have been exploring the possibility of utilizing news resources on-line.

Considering the call in recent years for smaller and more specialized corpora created by instructors for specific teaching objectives, classroom concordancing with a corpus of authentic on-line English news for news media English classes appears to be an idea worth pursuing. But before any attempt in creating an English news corpus is made, it is imperative to establish pedagogical objectives for such courses. A quick survey of materials published for similar courses would lead to the finding that most materials in this field seem to be developed in the framework of learning English by reading the news. Very few of them reveal sufficient effort in guiding students on how to approach English news. It goes without saying that newspapers can be a valuable resource for EFL learners to improve their English; however, assuming that reading the news would be as instinctive as reading a storybook or other materials written in English can be somewhat risky. For most EFL learners, developing enough comfort with reading English news to include it in their daily activities does require some understanding of the characteristics of English news writing and training in strategies of reading the news. Therefore, while learning English through reading the news may well be a major objective of a course in news media English, familiarizing students with the structural and discourse features of English news reports and enhancing skills in reading them seem no less valid as an aim of the course.

Having made the case for including distinctive features of English news and related reading skills in the course, it is necessary to build a tool with functions to facilitate improvement of not only linguistic competence, but also skills in approaching English news.

The present study seeks to explore the potential of incorporating a data-driven learning tool with a course-specific corpus in news media English classes. To achieve this aim, a lesson was devised and implemented using a system developed by the authors. The lesson intends to develop students' English proficiency, awareness of distinctive features of English news reports, and skills of reading the news. The students' responses to this new tool were investigated after the lesson. The study examines specifically the students'

responses regarding skills that benefited from using such a system, problems encountered in the process of using the system, and suggestions for future development of similar systems.

The study was undertaken with the hope that it would bring into focus the benefits of concordancing with smaller and more specialized corpora for students with specific interests and needs. More importantly, it is hoped that the study would encourage development of effective tools to make reading English news a less intimidating and more rewarding experience for EFL learners.

## **METHOD**

### **Participants**

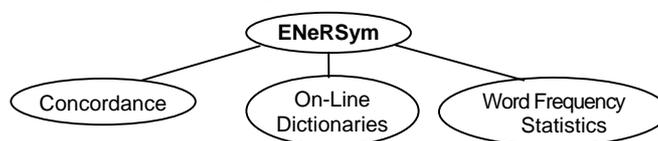
70 students enrolled in a news media English class in a university in Northern Taiwan participated in the study. The students were all non-English majors from various departments in the university, since the course was an elective one open to all students in their second to fourth years. For data analysis of the study, 65 valid data sets, excluding 5 incomplete data sets, were used.

### **The On-Line System**

In an attempt to meet the instructional objectives of the news media English course, an on-line learning system for reading English news was developed. The system consisted of three major components—an on-line concordancer with a course-specific corpus of on-line English news, word frequency statistics, and free on-line dictionaries. As claimed by Fletcher (2004), documents on the World Wide Web constitute a rich source of digital texts that are “unparalleled for quantity, diversity, and topicality” (p. 2). To locate materials that fit our course content, we turned our attention to a great

resource that had been widely available but often ignored—news reports on various news media websites. The on-line materials are especially valuable for the course because materials on these websites are updated regularly and therefore closely related to students’ knowledge and concerns. Figure 1 and Figure 2 present the structure and screenshot of the system.

**FIGURE 1**  
System Structure



**FIGURE 2**  
Screenshot of the System

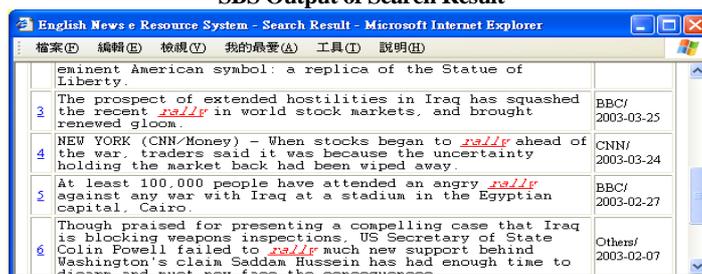


At the time of the study, the corpus consisted of 1,724 news headlines and news leads from major on-line news media, including news stories in two categories— world news and business news. The system provides two types of search output—the traditional KWIC (keyword in context) and SBS (sentence by sentence), as shown in Figure 3 and Figure 4.

FIGURE 3  
KWIC Output of Search Result



FIGURE 4  
SBS Output of Search Result



The word frequency statistics function allows searches for word frequency in a particular news category over a given period of time (by month, quarter, or year). The system administrator can decide what words to be left out in the calculation to exclude function words or basic vocabulary items from the list of frequently appearing words. For this particular lesson, we commanded the system to leave out the 1,056 junior high school basic vocabulary items as identified by the Ministry of Education in Taiwan. Word frequency statistics data also allow users to quickly get a glimpse of major news developments in a particular category over a set period of time (see Figure 5).

The system was developed with the following expectations in mind:

- (1) The system would help identify high frequency words in English news

by categories.

- (2) The concordance output would assist learning of grammar and collocations.
- (3) The concordance output would encourage inductive teaching of structural and discourse features of English news stories.
- (4) The corpus data would support teaching and development of skills in reading English news.
- (5) The corpus data would facilitate contrasting of lexical meaning/use in news of different categories.

The system was still in its initial stage of development at the time of the study; therefore, the news stories in the database were limited in number and category.

**FIGURE 5**  
**Word Statistics Search Result**

No.	Word	Total	Apr.	Mar.	Feb.	Jan.
1	Iraq	332	13	106	111	102
2	war	156	7	76	44	29
3	weapons	109	2	23	40	44
4	UN	99	7	34	32	26
5	pneumonia	82	7	75	----	----
6	United	77	6	29	27	15
7	Baghdad	76	5	29	21	21
8	Iraqi	75	3	25	29	18
9	Virus	70	22	42	6	----
10	Saddam	69	3	32	18	16
11	Blix	69	----	21	30	18
12	President	65	2	26	14	23
13	Bush	63	5	20	19	19
14	inspectors	61	----	8	22	31
15	chief	51	----	18	21	12
16	Hussein	50	2	22	14	12
17	nations	50	2	16	22	10
18	Council	46	----	18	23	5
19	deadly	45	14	27	1	3
20	mystery	43	7	34	2	----
21	officials	43	4	23	11	5
22	resolution	43	----	13	26	4
23	security	42	1	16	20	5

### Design of the Lesson

The lesson began with an introductory session on the major functions and resources provided by the on-line system. Oral explanation, printed handouts, and demonstration were also given on the different modes of search and output presentation provided by the system. Several examples were used in the demonstration session to familiarize the students with the system. Students' attention was also directed to some basic searching techniques to

facilitate operation of the system. Then the students were given ten tasks on hard-copy worksheets and asked to utilize the resources available on the system to complete the tasks. The tasks included 22 questions requiring the students to investigate features in language use and news content, such as analyzing the features of news headlines, finding out the meaning of abbreviations/acronyms commonly used in the news, identifying the Ws (Who, What, When, Where, etc.) in news leads, finding details while tracking the development of particular news events, analyzing word usage according to search results, and investigating the relationship between news events. The students were encouraged to take turns and work in groups, due to the limited number of computers available in the room and the large class size. The students were instructed to figure out answers to these questions using information available from the system, support their answers with examples from the search output, and to record their search procedures for validation and discussion. The students were also required to access the system at home, complete homework assignments, and present to the class the answers they worked out and demonstrate how they used the system to figure out the answers.

### **The Questionnaire**

To elicit students' responses to the on-line system, a questionnaire with five-point Likert scales and open-ended questions was administered after the lesson to investigate the students' perception of the benefits of using the system as a teaching/learning aid in news media English classes (see Appendix). The items on the questionnaire examined the students' views on five issues: effectiveness of using the system to assist reading of English news and learning English, future use of the system, preference in search output presentation, problems and difficulties with using the system, and system design in terms of corpus content and functions. The items regarding effectiveness of the system focused on two broad areas: learning of English in general and learning skills in reading English news. The items were

written in this way because these were two important aspects in a course that aimed to guide students on how to read English news and to improve students' English by reading the news. The study lasted for four weeks, including introduction of the system and administration of the questionnaire, with the class meeting two hours each week. Since this study was exploratory in nature, we consider the lesson sufficient for trying out the system and the instructional practice.

## **FINDINGS AND DISCUSSIONS**

### **Perceived Benefits of the System**

Overall, strongly positive responses were reported on the use of the on-line system as a learning aid. The students considered the system helpful for learning English and enhancing skills in reading English news stories. A mean score of over 4 points was reported on a five-point scale for both items (see Table 1).

**TABLE 1**  
**Reported Benefits of the System on Learning**

<b>Perceived benefit</b>	<b>Mean Score</b>
Learning of English in general	4.06
Learning of skills in reading English news	4.05

A closer look at the items investigating the beneficial effects in individual aspects reveals that the system was reportedly more effective in fostering skills related to better understanding of English news than those related to improving English in general. Table 2 lists individual aspects benefited from using the system and the perceived effectiveness of the system in each aspect.

**TABLE 2**  
**Reported Benefits of the System**

Rank	Benefit	Mean Score
1	Learn keywords and phrases in English news	4.34
2	Identify important news in a given period of time	4.28
3	Analyze syntactic features of news headlines	4.22
4	Understand commonly used abbreviations/acronyms	4.17
5	Track development of news events	4.12
6	Locate the Ws in news leads	4.09
7	Understand relationships between news events	4.03
8	Analyze lead sentences	3.95
9	Figure out meaning of vocabulary from context	3.92
10	Learn word usage	3.68
11	Improve grammar	3.55

Although all of the aspects reported above are related to successful reading of English news, the first eight of them contribute more directly to smoother reading and better comprehension of English news. The last three aspects are more related to the improvement of English proficiency in general.

Among the first eight aspects, a few of them are closely linked to the objectives of a news media English course and address major difficulties of reading English news. For example, the difficulties in reading English news are believed to lie in keywords/phrases, abbreviations/acronyms in the news, and the unusual syntactic features of news headlines. Some keywords/phrases that repeatedly occur in the news may not be encountered often in everyday conversations. Moreover, new keywords/phrases and abbreviations may be developed every day in news stories as new events come up. Therefore, it is necessary to familiarize students with those keywords/phrases and abbreviations to facilitate their understanding of the news. Since the system provides word frequency statistics and authentic contexts, the students considered it most effective in helping them identify and learn those keywords/phrases in the news (Mean=4.34).

Scanning the news is usually considered an important skill in reading newspapers, since people don't usually read newspapers the way they read a

textbook or a novel. Instead, they simply go over the headlines, looking for stories they are interested in and quickly get the most important information in the news stories. Therefore, the syntactic features of news headlines are often one of the first few lessons in reading English news. News headlines are usually written in a very condensed form, leaving out certain syntactic elements or using punctuation marks to substitute certain words (Fredrickson & Wedel, 1984). Students need to be aware of the special syntactic features of the headlines to successfully scan the news.

Another skill involved in scanning the news is the ability to quickly locate essential information of the news event (the Ws), which is generally found in the first paragraph, the lead, of a news story. From the results of the questionnaire, the system, with large numbers of authentic news headlines and leads in its database, was reported to be effective in assisting analysis of headline features (Mean=4.22) and practicing locating Ws in news leads (Mean=4.09).

Some other aspects contribute to more efficient reading of English news by facilitating understanding of news content, including identifying important news events in a given period of time (Mean=4.28), tracking development of particular news events (Mean=4.12), and understanding relationships between news events (Mean=4.03). For a course that attempts to encourage regular contact with English news to sustain the beneficial effects on linguistic competence, such assistance in comprehension of news content is highly desirable.

### **Future Use of the System**

To assess the system's potential role in the students' future learning process, the students were asked whether they planned to continue using the system or recommend the system to others to improve their English and skills in reading English news. The students reported positive responses to both items (see Table 3).

**TABLE 3**  
**Use of the System in Future Learning**

<b>Item</b>	<b>Mean Score</b>
Plan to incorporate the system in future learning	4.02
Plan to recommend the system to others	4.12

### **Preference in Output Presentation**

The system provides two modes of output presentation—the traditional KWIC (keyword in context) and SBS (Sentence by sentence). KWIC output, in which the keyword is presented in the center of the line with incomplete contexts on each side, is currently used more often in concordancers than SBS, in which the entire sentence of the keyword is presented. These two types of concordance output have different strengths and limitations. To investigate the students' perception of these two modes of presentation, the students were asked which type of output presentation was more conducive to learning and why. From the results of the study, it seems that, despite the current popularity of KWIC, more students (69.2%) consider presenting the entire sentence in which the keyword appears as more conducive to their learning. They reported that incomplete contexts reduced the possibility of inferring the meaning or analyzing the usage of the keyword successfully. They often needed to switch to SBS concordance to confirm their hypothesis about the keyword. The reported preference for SBS may imply the need for concordancers to provide users with the alternative of reading complete contexts, despite the current popularity of the highly-acclaimed KWIC search output.

**TABLE 4**  
**Reported Preferences Regarding Output Presentation Modes**

<b>Preferred mode of presentation</b>	<b>Number</b>	<b>Percentage</b>
KWIC (Keyword in context)	8	12.3%
SBS (Sentence by sentence)	45	69.2%
Both helpful	12	18.5%

### **Problems Encountered When Using the System**

Difficulties and frustrations encountered by the students during the lesson were also investigated with open-ended questions. The students reported difficulties in three domains—searching techniques, linguistic competence, and technical features.

Some students reported that they often needed to key in a few different keywords before they found out which keyword would return the best search results for a given task. This may have indicated that more tips on searching techniques should have been provided to the students either during the introductory session or along with the tasks.

Other students reported difficulties related to their proficiency of English. Some students expressed frustrations with analyzing the search results. Others mentioned that too many unfamiliar words made reading the search results a daunting task. These linguistic difficulties made the search results seem overwhelming to them. Difficulties like these are expected because authentic materials, despite the benefits often cited with their use, are estimated to be more challenging for EFL learners. And that is also why the system provided on-line dictionaries to make the task seem less intimidating.

The third domain of difficulties includes frustrations with the technical features of the system. The issue of speed that afflicts most on-line concordancers was also reported by the students, but in our case the issue arises from hyperlinks provided by the system. As introduced previously, the system allows the students to access the archive of the source media to read the news stories in full length. However, connection to these news sites was reported to be slow and unstable. Other students stated that lack of advanced search functions prevented them from further narrowing down the search results, and that too many windows produced by the system on the desktop may confuse them after making a few different searches.

**TABLE 5**  
**Reported Difficulties While Using the System**

Difficulty	Frequency of report
Selection of keyword	10
Speed of hyperlink connection	8
Analysis/interpretation of search results	6
Too many windows	3
Limited search functions	2

### **Suggestions for Future Development of the System**

In addition to reporting problems, the students also made several suggestions regarding the development of similar systems, including improving graphic layout of the system, expanding the corpus, providing additional functions, and advertising the system to benefit a larger population with similar needs.

As mentioned previously, the system was far from being fully developed at the time of the study, with only two categories of news from a few major on-line news media. The students suggested that more news categories from a larger number of sources would enhance the effectiveness of the system. Some students expressed the need for more domestic news that may be of more direct interest to them. Indeed, contrasting of word meaning/use across domains would be greatly facilitated if more news categories were included in the corpus.

The students also recommended a variety of additional functions as summarized in the following.

**TABLE 6**  
**Recommended Functions**

1	On-demand Chinese translation
2	Advanced search functions
3	Help menu or on-line technical assistance
4	Grammatical analysis of search output
5	Audio/video output of the news content
6	Record search history
7	Bulletin boards or chat rooms
8	On-line Q&A for English learning

Some of these recommended functions may not be directly pertinent to the purpose of the course, but the others merit attention as a reference for future development of systems for similar purposes. The recommended functions of providing grammatical analysis and translation (1 and 4) can be considered if seen fit to individual instructional objectives. Video/audio output of the news might well expand the scope of the system beyond reading, but negotiation with the source media shall precede implementation, given the concerns about copyrights. Finally, maintaining a Q&A section is also worth serious thoughts, since it would greatly enhance the interactive nature of the system and facilitate self-study at home but will require constant human effort in handling the questions.

Another finding worth mentioning is that the students seemed to find the system more helpful than the textbook assigned for the course (mean=3.92), which introduces vocabulary in the news with examples from authentic news reports. They found the system more motivating because the news events were up-to-date and more closely related to their lives. Some of them also reported that working out answers to questions seemed to be more exciting than just memorizing vocabulary definition or rules about word usage. This clearly confirms the claim that concordancing empowers the learners by helping them take on the more active role of researchers.

## CONCLUSION

This study explores the potential of developing an on-line system with the benefits of data-driven learning and integrating it into news media English classes. A lesson incorporating a system with news concordancing, word frequency statistics, and on-line dictionaries was conducted in a news media English class, and the students' perceptions of the benefits of such a system as a learning aid were investigated. Responses collected from the students show that the system was perceived to be effective in assisting reading of English news stories and learning of news media English, especially in the

learning of keywords/phrases in English news and headline features. The students also expressed strong willingness to incorporate the system in their future learning process. Many of them found the system valuable and expressed a keen interest in promoting the system because assistance in reading English news was not easy to find. These results seem to indicate a strong potential for development and use of similar systems.

On the other hand, the students' responses also include messages that future developers of similar systems may wish to take into account. First of all, a user-friendly interface should be a priority to reduce frustrations caused by technical complexities and encourage frequent use of the system. An efficient help menu or on-line tutorial can be created to provide support for self-access outside the classroom.

Second, advanced search functions need to be provided to make search results more accurate and less overwhelming. The corpora can be properly tagged to allow part of speech specification of a searched word and thus make the pattern of collocation more apparent. Functions that allow more complex searches, such as search by date of news or morphological forms of a certain word can be added to reduce irrelevant hits and facilitate interpretation of search results.

Third, additional resources to assist linguistic analysis of search results may be provided as an optional function depending on the purpose of the lesson or task. The system used in the study provides access to on-line dictionaries to compensate for the difficulty of authentic language data. Other resources to assist parsing of the linguistic data may be considered to serve different pedagogical purposes.

The present study represents an initial undertaking to explore the use of on-line resources to attain the benefits of concordancing with course-specific corpora for news media English classes. To further assess the effectiveness of such learning tools, more studies need to be conducted with a more fully developed system on larger samples over longer periods of time.

## THE AUTHORS

Shiauping Tian is an assistant professor at National Taiwan University of Science and Technology. Her research interests include L2 reading processes, reading strategy training, and CALL.

Limin Liu is an assistant professor at Chung Yuan Christian University. His research interests include database systems, natural language processing, and CAI.

## REFERENCES

- Aiex, N. K. (1988). Using newspapers as effective teaching tools (ED 300847). Retrieved March 12, 2004 from [http://www.ericfacility.net/databases/ERIC\\_Digests/ed300847.html](http://www.ericfacility.net/databases/ERIC_Digests/ed300847.html).
- Aiex, N. K. (2000). Newspapers as a teaching resource for adult learners (ED 440428). Retrieved March 12, 2004 from [http://www.ericfacility.net/databases/ERIC\\_Digests/ed440428.html](http://www.ericfacility.net/databases/ERIC_Digests/ed440428.html).
- Aston, G. (1997). Small and large corpora in language learning. In B. Lewandowska-Tomasczyk & J. Melia (Eds.), *Proceedings of the first international conference on practical applications in language corpora* (pp. 51-62). Retrieved March 24, 2003 from [http://www.uni.lodz.pl/iangl/g\\_a.htm](http://www.uni.lodz.pl/iangl/g_a.htm).
- Aston, G. (1998). What corpora for ESP? Retrieved March 24, 2003 from <http://www.sslmit.unibo.it/guy/pavesi.htm>.
- Ball, C. N. (1996). Tutorial notes: Concordances and corpora. Retrieved March 24, 2003 from <http://www.georgetown.edu/cball/corpora/tutorial.html>.
- Dyck, G. N. (1999). Concordancing for English teachers. Paper presented at TESL Manitoba, Canada. Retrieved March 21, 2003 from <http://home.cc.umanitoba.ca/~gdyck/conc.html>.
- Fletcher, W. H. (2004). Facilitating the compilation and dissemination of ad-hoc web corpora. Retrieved March 12, 2004 from [http://miniappolis.com/KWiCFinder/Facilitating\\_Compilation\\_and\\_Dissemination\\_of\\_Ad-Hoc\\_Web\\_Corpora.pdf](http://miniappolis.com/KWiCFinder/Facilitating_Compilation_and_Dissemination_of_Ad-Hoc_Web_Corpora.pdf).
- Fredrickson, T. L., & Wedel, P. F. (1984). *English by newspaper*. Boston, MA: Thomson.
- Gross, D. (1991). *A practical handbook of learning teaching*. London: Cassell.

- Hadley, G. (2000). Sensing the winds of change: An introduction to data-driven learning. *RELC Journal*, 33(2), 99-124.
- Harumi, I. (1993). A simple but effective use of newspaper articles in advanced EFL reading classes. *Guidelines*, 15(1), 37-49.
- Huang, Y. K. (2003). Easing into E-news: An EAP course for advanced readers. In *Selected papers from the twelfth international symposium on English teaching* (pp. 414-423). Taipei: Crane.
- Ilse, W. R. (1991). Concordancing in vocational training. *ELR Journal*, 4, 103-113.
- Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *ELR Journal*, 4, 27-45.
- Kettemann, B. (1995). On the use of concordancing in ELT. *TELL & CALL*, 4, 4-15.
- Lee, C. Y., & Liou, H. C. (2003). A study of using web concordancing for English vocabulary learning in a Taiwanese high school context. *English Teaching and Learning*, 27(3), 35-56.
- Richards, J. C., & Lockhart, C. (1994). *Reflective teaching in second language classrooms*. Cambridge University Press.
- Rundell, M. (1996). The corpus of the future, and the future of the corpus. Talk at Special Conference on New Trends in Reference Science, Exeter. Retrieved March 24, 2003 from <http://www.ruf.rice.edu/~barlow/futcrp.html>.
- Shaffer, C. (1989). A comparison of inductive and deductive approaches to teaching foreign languages. *Modern Language Journal*, 73, 395-403.
- Someya, Y. (2000). Online business letter corpus KWIC concordancer and an experiment in data-driven learning/writing. Paper presented at the 3rd Association for Business Communication International Conference, Doshisha University, Kyoto, Japan. Retrieved July 10, 2003 from [http://www.kamakuranet.ne.jp/~someya/DDW\\_Report.html](http://www.kamakuranet.ne.jp/~someya/DDW_Report.html).
- St. John, E. (2001). A case for using a parallel corpus and concordancer for beginners of a foreign language. *Language Learning & Technology*, 5(3), 185-203.
- Stevens, V. (1995). Concordancing with language learners: Why? when? what? *CAELL Journal*, 6(2), 2-10.
- Tribble, C. (1990). Concordancing and an EAP writing programme. *CAELL Journal*, 1(2), 10-15.
- Tribble, C. (1997). Improvising corpora for ELT: Quick and dirty ways of developing corpora for language teaching. In B. Lewandowska-Tomasczyk & J. Melia (Eds.), *Proceedings of the first international conference on practical applications in language corpora* (pp. 106-117). Retrieved March 24, 2003 from <http://web.bham.ac.uk/johnstf/palc.htm>.

**APPENDIX**  
**Questionnaire for Student Feedback (English translation)**

*Please read each of the following statements and check the box that corresponds to your perception.*

	5	4	3	2	1
1. I consider the system helpful for learning English in general.	Strongly Agree <input type="checkbox"/>	Agree <input type="checkbox"/>	Mildly Agree <input type="checkbox"/>	Disagree <input type="checkbox"/>	Strongly Disagree <input type="checkbox"/>
2. I consider the system helpful for learning vocabulary.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. I consider the system helpful for figuring out the meaning of vocabulary from context.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. I consider the system helpful for learning word usage.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. I consider the system helpful for learning grammar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. I consider the system helpful for identifying keywords that appear frequently in the news.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. I consider the system helpful for learning keywords/phrases in the news.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. I consider the on-line dictionaries provided by the system helpful.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. I consider the system helpful for learning skills in reading English news.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. I consider the system helpful for learning to locate the most important information in news stories (the Ws).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. I consider the system helpful for learning to find the subject and main verb of a news story.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. I consider the system helpful for learning the language of news headlines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- |   |                          |                          |                          |                          |                          |
|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 13. I consider the system helpful for understanding commonly used abbreviations/acronyms.     | <input type="checkbox"/> |
| 14. I consider the system helpful for tracking the development of news events.                | <input type="checkbox"/> |
| 15. I consider the system helpful for understanding the relationship between news events.     | <input type="checkbox"/> |
| 16. I consider the system helpful for finding out important news in a given period of time.   | <input type="checkbox"/> |
| 17. Overall, I find the system layout clear and easy to read.                                 | <input type="checkbox"/> |
| 18. Overall, I find the system easy to operate.   | <input type="checkbox"/> |
| 19. I think the system provides enough functions.   | <input type="checkbox"/> |
| 20. I think the functions provided by the system are helpful for learning news media English. | <input type="checkbox"/> |
| 21. I plan to use the system in my future effort in learning English.                         | <input type="checkbox"/> |
| 22. I plan to recommend the system to other students to help them learn English.              | <input type="checkbox"/> |
| 23. I consider the system more helpful than the textbook we used in class.                    | <input type="checkbox"/> |

*Please answer the following open-ended questions.*

24. Is there anything that you think the system helped you in learning but is not mentioned in the questions above? What is it?
25. What difficulties/frustrations did you encounter while using the system?
26. Is there any additional function that you think should be included in the system?
27. Is there any function provided by the system that you think is useless and should be removed?

Course-Specific Corpora in the Classroom: A News Media English Class in Taiwan

28. The system provides two types of search output (*KWIC* and *SBS*), which one do you think is more helpful?
29. Do you have any other suggestions/ideas for the development of the system? Please share them with us.
30. Is there any other free on-line dictionary that you would like to recommend to us?  
Name of the dictionary\_\_\_\_\_ Website\_\_\_\_\_