British Journal of Educational Technology *Vol 34 No 5 2003* 601–613

Learning process, strategies and web-based concordancers: a case study

Yu-Chih Sun

Dr Yu-Chih Sun is currently an assistant professor at the Language Teaching and Research Center at National Chiao Tung University, Taiwan. Her research interests include computer-assisted language learning and TESOL methodology. Address for correspondence: Language Teaching and Research Center, National Chiao Tung University, No 1001, University Road, Hsin-Chu, Taiwan, 300, ROC. Email: sunyc@mail.nctu.edu.tw.

Abstract

The recent widespread use of web-based concordancers seems to provide a promising mode for language teaching and learning, especially in the English as a foreign language (EFL) setting, because through concordancers students can easily gain exposure to a huge number of authentic and sorted language examples. This paper describes a case study of the learning process and strategies used by three Taiwanese college students in the concordancer setting. A web-based concordancer was used to assist the participants while undertaking a proofreading activity. Think-aloud protocol was used to collect their data. The results showed that the following four factors have influenced learners' learning process and strategies in use: (1) prior knowledge, (2) cognitive skills, (3) teacher intervention and (4) concordancer skills.

Introduction

The use of authentic linguistic examples is claimed to better help those learning English as a second or foreign language than invented or artificial examples (Johns, 1994). In the past teachers usually presented students with examples they had made up themselves. The weaknesses of such a method are that these examples lack authenticity and their number is limited. Students often have difficulty exploring underlying patterns from these poor examples. Underlying patterns refer to language rules that are not obvious to learners or that may be difficult for learners of English as a foreign language to discover. The unique sorting functions in a concordancer can help learners to uncover those rules. One effective and efficient solution that researchers have proposed is that students can benefit from a concordancer, concordancing or concordance gen-

Published by Blackwell Publishing Ltd., 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

erators (Johns, 1990; Levy, 1990; Sun, 2001). A concordancer allows users to interact with a selected corpus. When a keyword or phrase is typed, the system will search for and then display occurrences of the word or phrase in its immediate context and allow learners to discover patterns and adjust their misconceptions by observing extensive naturally occurring examples in real texts (Hill, 2000). These patterns refer to language rules that are grammatical or lexical. Since a concordancer can extract numerous examples of a particular language usage, it helps learners discover rules by retrieving instances of similar language usage.

Concordancer use has been energised by the advent and popularity of computing power and the development of computer concordancer programs (Steven, 1995). It has been regarded as one of the most promising recent ideas in computer-assisted language learning (Cobb, 1997; Levy, 1990, 1997; Owen, 1996; Steven, 1991, 1993; Tribble and Jones, 1997). In the late 1980s several concordancer programs for personal computers became available. Some of the well-known programs are MicroConcord, Micro-OCP and Mini-concordancer. The development of the Internet has led to another leap forward in terms of easier access to wider applications in language learning and teaching. Some of the well-known web-based programs are ConApp from the Virtual Language Center at Hong Kong Poly University, CoBuild Sampler by Collins Cobuild and the Gutenburg Web Concordancer.

The basic operation of most concordancers includes the following five types of search method: (1) only exact matches, (2) all instances containing the search string as a prefix, (3) all instances containing the search string as a suffix, (4) all instances containing the search string and (5) all instances containing both of the search words. Sort Method menus are usually displayed next to the keyword window. In order to reduce the loading of reading, some concordancers (eg, Sun's) also allow users to sort by order of increasing sentence difficulty. Figure 1 illustrates the search page of Sun's concordancer.

Johns and King (1990) have advocated the use of concordancers as a tool for data-driven learning (DDL). In this process learners play the role of a researcher who explores the use of language through investigating an authentic corpus. He identifies two approaches in concordancer-based DDL. One is the inductive approach, whereby learners identify patterns observed from the evidence of the concordancer output. The other is the deductive approach, whereby learners attempt to find examples to verify grammar rules.

To date there have been several studies exploring approaches to integrating concordancing tools into language learning and teaching (Chen, 2000; Cobb, 1997; Conrad, 1999; Flowerdew, 1993; Kettemann, 1994; Turnbull and Burston, 1998). However, very little research, if any, aims to examine the process that learners follow when using concordancing tools as a learning aid. Therefore, the goal of this study is to identify the learning process and strategies used by EFL learners in the concordancer setting and to examine the factors that impact on learner behaviour.

Figure 1: Illustration of Sun's concordancer web page

Method

Participants

Three college students (Students C, H and R) participated in the current case study. All were enrolled in the Department of Applied Foreign Languages at a university located in central Taiwan. Each student represents a particular program within the department. Student C is in the two-year regular daytime program whereas Student H is in the three-year night school program. Both programs award students with a Bachelor's degree upon graduation. These undergraduate programs are subject to two different nation-wide entrance examinations. The daytime program is more competitive than the night school program, thus students' language proficiency also varies between the two. Student R is from the TESOL (Teaching English as a Second Language) graduate program. Both Students C and H were seniors at the time whereas Student R was in her first year of graduate studies. Since this is a pilot study on learning process and strategies in the concordancer context, only three students were chosen to participate.

Procedure

The study began with an individualised tutorial. Each student spent about 30 minutes undergoing the following process:

1. Received a brief overview of the purpose and function of the concordancer.
2. Observed the teacher's demonstration on how to use the concordancer tool.
3. Gained hands-on experience on how to use the concordancer.
4. Demonstrated capability in concordance research.

Students were then given an eight-question proofreading list with different types of grammatical errors. They were asked to check the sentences one by one to see if they were correct or not, and were encouraged to use the concordancer to look for evidence to support their answers.

There are two main types of verbal report to examine process-oriented description. One is a retrospective report, which has the advantage of not interfering with the process itself, as well as the potential disadvantage of being distorted or inaccurate, depending on the best recollection of the subject (Hare, 1981). The other type of description is an introspective or 'think-aloud' protocol. This protocol is a kind of verbal report in which participants comment on their thoughts and behaviour while undertaking a process-oriented task. Think-aloud reports provide a more direct view of the problem-solving process. They allow researchers or teachers to learn about what is actually going on in a learner's mind and provide insight into his/her cognitive problem-solving activity. Therefore, in the current study, data were gathered using a think-aloud procedure; the students were asked to verbalise their thoughts as they attempted the proofreading exercises.

In order to familiarise the students with the 'think aloud' technique, the researcher conducted an initial pilot test by giving them a short orientation to thinking aloud and by asking them a few questions. They were encouraged to think aloud on their own as they digested the questions, until they were able to operate in this way with relative ease.

For detailed data-analysis, each student's think-aloud oral report was recorded and transcribed immediately after the session. The duration of the think-aloud process ranged from 50 to 70 minutes, depending on the student. Tentative criteria for the measurement of both mental activity and learner strategies were developed in the early stage of data coding and were later refined. The criteria included a timeline for each student's reported mental activities, keyword searches performed with the concordancer, concordancing skills used, reflection on learner's prior knowledge, learner's cognitive stages, rule formulation, circumstances and timing of teacher intervention, mode of response, and sections of the erroneous sentences that drew the learner's attention. These timelines facilitated the analysis of the learner's strategies and process.

Results and discussion

Research question 1: What are the learning process and strategies used by EFL learners in the web-based concordancing setting?

Of the eight proofreading questions, Student R got half of them correct at the very beginning, while Students C and H got only two correct. The prior knowledge modes

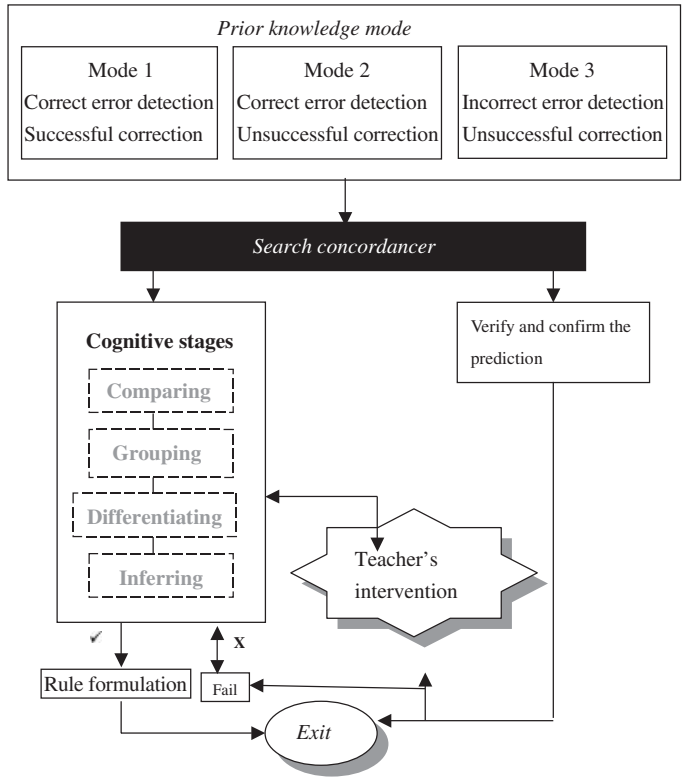

Figure 2: Flow chart of Concordance Learning process

that the students worked through to decode and investigate the sentences could be classified into three typical varieties. While learners undertook analysis of concordancer outputs, they usually worked through four cognitive stages, namely comparing, grouping, differentiating and inferring. A flow chart of the three modes and four stages of cognitive activity is detailed in Figure 2. Examples of the three typical modes are presented below. The categories describe mental activity during several stages of proofreading with concordancer help.

Mode 1

For mode 1 students detected and corrected the error successfully, and then searched the concordancer to verify their prediction and obtain confirmation. This mode often occurred when students had complete prior knowledge and used deductive thinking skills to search for supporting examples from the concordancer. For example, after reading the second question:

*Let's keep this information just between you and *I.*

Student R was able to point out that *I* was wrong and should be changed to an objective, *me*. To make sure, she verified her prediction by searching for the keyword *between* and scanning the examples that were followed by two personal pronouns. In this way she found two similar examples as listed below.

940 Sudden quarrels and rage flashed between my husband and me, as the underta

1790 ouflage. The big difference between them and us, between the continenta

With these supporting examples, she confirmed that the sentence should be revised to

Let's keep this information just between you and me.

Mode 2

The second mode applied when students could identify the error correctly but were not very sure about how to correct it. From concordancer outputs they could find some examples to verify their predictions and provide confirmation. For example, student C used mode 2 to work on the following question:

*The software runs *good on this computer.*

After reading the question Student C was suspicious about *good*, so she conducted a concordancer search on *run* and received the following outputs:

555 ight to horses who have previously run well around the National course a

556 es a winner over these fences, has run well both times this campaign after

1266 alled Octopush News with a print run well into three figures. Wha

1267 world championships, he will have run well. He likes tough going. Just

547 onship events. We're hopeful he'll run well". If we had a fiver for

She looked at the outputs and found some supporting examples that were helpful. She then confirmed that the sentence should be modified to

The software runs well on this computer.

Mode 3

The last mode applied when the students could neither detect nor correct the error. This usually occurred when there was a large gap between the target knowledge and the students' prior knowledge. If prior knowledge of a certain concept was insufficient, the student might be very confused by the teacher's interventions and make false predictions repeatedly. Take Student H, for example. Her prior knowledge about the grammatical concepts in the fifth question was incomplete.

*The only *criteria for the job is English proficiency.*

Using a concordancer to solve the problem inductively became a hindrance for her. Therefore, she tended to rely constantly on the teacher's intervention. After reading the sentence Student H thought that it was wrong. She believed *an* was needed in front of *English proficiency* because she thought "*English proficiency* is a countable noun" thus it had to be preceded by an indefinite article.

*The only *criteria for the job is *an English proficiency.*

To confirm her prediction, she searched for *is a* and got the following outputs:

15	me through. Yet a royal marriage is an act of state; it involves the Gover
17	ty that it is hard to believe he is an actor reading a script. Howe
67	limatation in the Bois de Boulogne is an amusement park with swings, round
81	aton, managing director, said: "It is an anxiety, wondering whether we wil
83	s spokesman, said. "What we want is an apology and the Pope to withdraw wh

With this indirect evidence she confirmed that a countable noun should follow an indefinite article but assumed incorrectly that *proficiency*, like *act*, is countable. Since Student H's prior knowledge about countable/uncountable nouns was insufficient, it was very difficult for her to figure out the right answer by inductively decoding data from the concordancer. The teacher then intervened by providing deductive instruction, telling her that *proficiency* is an uncountable noun and explaining the rules governing its use. Using this deductive explanation Student H left this dead end and started to look for other possible errors. She read the sentence again and became suspicious about spelling mistakes. To verify the spelling of *criteria* she used the concordancer again to search for the word. From the concordancer outputs listed below she confirmed that *criteria* was spelled correctly but did not notice the singular/plural issue.

1 ntages in stressing these economic criteria and leaving governments to decide

2 animals share the same standards, criteria, attitudes and fears that we do,

3 the evidence in the light of their criteria, Drs Pope and Hudson made an asto

4 of clarity, seeks to interpret the criteria flexibly, so that all but the mos

The teacher intervened again by asking her to focus on the singular/plural problem. Student H then figured out from the following examples that *criteria* is a plural form, due to the occurrence of the verb *were*.

27 piness was the cause, but strict criteria were drawn up which made it the mos

28 knew what a particular lender's criteria were, everyone would get through."

However, she was still confused about the sentence correction and predicted wrongly again that *only* and *is* should be modified to accord with the plural *criteria*.

*The ~~only~~ *criteria for the job *are English proficiency.*

The teacher then intervened again to remind her that *English proficiency* is uncountable, and Student H finally figured out that the subject should be singular and noticed that there might be a singular form of *criteria*. However, she could not think of the word. With the teacher's suggestion of searching for *criterion*, she received the following outputs and confirmed that the plural form of *criteria* should be changed to a singular form, *criterion*.

1 , demand was not merely a relevant criterion but that there was additionally a

2 tritional status". "By this criterion," he says,"UK children have never

3 ent in 1994, when the Maastricht criterion is 60 per cent. In the Commission's

Research question 2: What are the factors that influence student learning outcome?
The observations detailed above revealed that four main factors tend to influence learners' concordancer behaviour, including (1) learners' prior knowledge, (2) learners' cognitive skills, (3) teacher intervention and (4) concordancer skills. Each factor

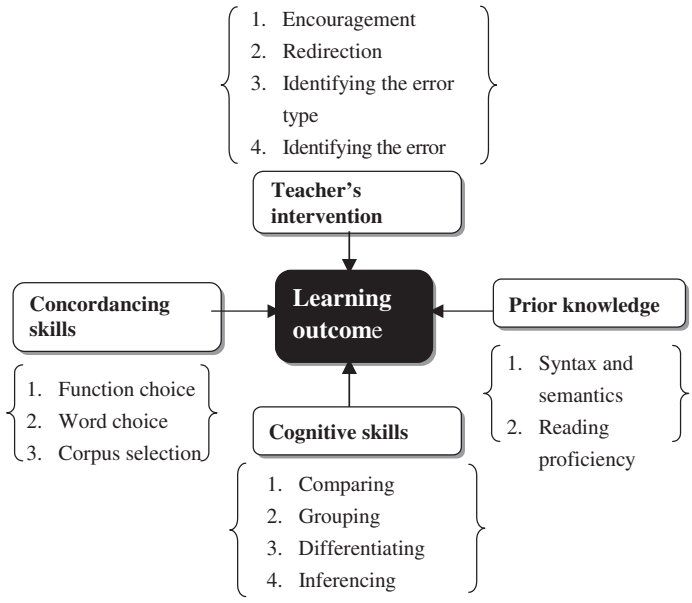

Figure 3: Conceptual framework of concordancer learning

will be discussed in the following paragraphs. Illustration of the relationship between learning outcomes and the four factors is presented in Figure 3.

Prior knowledge

In this case study students' prior knowledge could be classified into three levels, including complete, incomplete and lack of prior knowledge. With different degrees of prior knowledge students undergo different learning stages. For students who have complete prior knowledge the concordancer plays the role of a supporting tool, which can help them to verify their hypotheses and affirm their intuition and prior knowledge. However, for students whose prior knowledge is incomplete the concordancer could encourage them to explore and build their own structural knowledge. From the process of predicting, looking for the evidence, analysing concordancer outputs, verifying, and making inferences, students could not only clarify certain language concepts, but also be trained to think and learn inductively. For students who have limited or no relevant prior knowledge the concordancer can also provide them with a tool to explore, sort out certain linguistic rules and practise inductive thinking skills. The learning curve for this level is arduously steep, in that students tend to get confused easily about the concordancer outputs; thus, they need either a stronger degree of teacher involvement, or to learn in a more structured environment.

Cognitive skills

The process of analysing concordancer examples consists of four stages—comparing, grouping, differentiating and making inferences. First, students briefly observe the con-

cordancer outputs, identify the general differences in structures and try to identify the data relevant to the target question. Second, after students have some ideas about the distinctions in the concordancer outputs, they try to classify them into groups. For example, keywords followed by the same preposition may be grouped into the same category. Third, when the subjects finish grouping, they identify the critical semantic or structural differences between each category. Sometimes the stage of differentiating occurs while students are doing grouping. That is, the two stages occur simultaneously. Finally, after students figure out the main differences of each group, they might try to formulate the hypothesis and then apply the inferences they have made to predict the answer. Although the process could be divided into four stages, some effective learners might combine certain stages into one. That is, they do not follow the four distinct stages one by one but combine some of them while analysing concordancer outputs.

Teacher intervention

Students with limited prior knowledge tend to get stuck often while performing concordancer tasks and need teacher intervention to break the ice. Teacher intervention can be classified into four levels in this case study. The intervention given here was based on the principles of scaffolding. That is, intervention by the teacher proceeds from general to specific intervention, depending on the student's current cognitive levels. The most general intervention occurred when students tried to give up or believed that there was nothing wrong with the proofreading sentence. If this happened the teacher would encourage them to conduct a more detailed investigation. Take question 4, for example.

We had wanted to wait until September to make a decision, however, we decided to act now.*

Student H could not determine whether a semicolon or a comma should be placed before *however*. The teacher then encouraged her to examine closely the punctuation aspect of the concordancer outputs. Thus, this level of intervention is more general and serves as a stimulus for learners.

The second general intervention takes place when students are misled by a false hypothesis. The teacher would try to orient them in other possible directions by telling them that the hypothesis might be wrong. Take question 6 as an example:

*The reason is *because we were late.*

Student C searched for keyword *the reason is because* but could not find out any useful cues to revise the sentence. The teacher then intervened by asking if it would be better to use *the reason* as a search keyword, to expand the search outputs. So, in this type of case, the teacher not only reorients the learner with a new direction but also suggests a keyword for a concordancer search. The above-mentioned types of intervention still leave students with plenty of space to predict, explore, compare and think inductively. However, for students with more limited prior knowledge, more specific and deductive guidance might be called for.

The other two levels of intervention are more specific. One is to specify the error type, which would help students narrow down the options. Take question 4 as an example:

We had wanted to wait until September to make a decision, however, we decided to act now.*

When Student R could not detect any error and thus searched the concordancer at random, the teacher told her that the errors in the sentence could involve vocabulary, grammar or punctuation to help her narrow down the possible error type.

However, sometimes students have very limited or no prior knowledge of certain concept and cannot make inferences even with closely related concordancer outputs. In such a case the teacher would intervene to a more specific extent by specifying the error type or even providing the answer.

Concordancer skills
The students' skill in using concordancer tools also influenced their learning efficiency. For example, if they were unclear about how to select an appropriate 'search string', they would retrieve a huge number of irrelevant concordancer outputs which would take them a long time to analyse, or even lead them in the wrong direction. On the other hand, if they knew how to select effective 'search strings', 'alphabetic lists' or 'sort types', they would get more relevant and organised outputs that facilitated searching, analysing and making inferences.

Conclusions

Concordancers provide a new way of teaching and learning language. Unlike traditional rule-based instruction where students are usually provided with scarce chance to test their hypotheses on a particular linguistic usage, language research is learner-initiated in the concordancer-based learning environment. Therefore, both learning and concordancer output correspond to students' needs/or wants (Flowerdew, 1993). The experience helps learners develop 'language awareness' and empowers them to be autonomous and responsible for their own learning.

A number of conclusions can be drawn from this case study. Results show that concordancing tools alone do not necessarily lead to productive induction. A number of factors could also contribute to the concordancing outcome, including prior knowledge of a specific linguistic form or usage, concordancing search skills, inductive thinking strategies and teacher intervention.

Learners with limited prior knowledge are more likely to be confused about the selection of search words, and the loading of data analysis tends to be heavier due to the lack of clear direction. To accommodate this limitation timely teacher intervention or the design of a tutorial program with appropriate guidance could be helpful. Furthermore poor reading proficiency can also limit the learning outcome due to the need for massive reading of concordancer data. In order to cater to students' different reading abilities the development of an extensive corpus with varying levels of reading

difficulty can better serve the needs of learners with different levels of reading proficiency.

Unfamiliarity with inductive learning strategies also tends to make exploration of linguistic patterns a struggle. Traditionally language is usually taught deductively in many EFL settings. Instruction usually starts from exposing students to the rules of syntactic and semantic forms, and then provides them with a number of applications. Learners tend to have scarce chance, if any, of exercising inductive thinking strategies. Therefore extensive guidance in using inductive learning strategies is highly recommended in future system design.

Inability to perform concordancer searches effectively also appears to be another problem. With ineffective search skills learners tend to receive many indirect and irrelevant concordancing outputs that are likely to make analysis more complex and frustrating. Thus modelling concordancing search functions and practising in a more extended context become necessary. In addition, guidance from the system or instructor becomes important in some cases.

In summary, the results of the current study provide us with a more in-depth understanding of EFL students' learning processes and strategies in the web-based concordancer environment, and identify some crucial factors that have impact on the effectiveness of learning. In terms of further research it is important to examine the relationship between reading proficiency and performance in concordancing activities in order to shed light on the effectiveness of concordancers in language learning. In addition, learners' attitudes toward concordancers are also worth investigating in order to provide some affective insights for program developers and classroom teachers.

References

Chen H (2000) Developing a web concordancer for English as foreign language learners *in* Proceedings of ICCE/ICCAI 2000 Conference, Taiwan, 340–341.

Cobb T (1997) Is there any measurable learning from hands-on concordancing? *System* **25**, 3, 301–315.

Conrad S M (1999) The importance of corpus-based research for language teachers *System* **27**, 1, 1–18.

Flowerdew J (1993) Concordancing as a tool in course design *System* **21**, 2, 231–244.

Hare V C (1981). Readers' problem identification and problem solving strategies for high and low knowledge articles *Journal of Reading Behavior* **13**, 4, 359–365.

Hill J (2000) Revising priorities: from grammatical failure to collocational success *in* M Lewis (ed) *Teaching collocation: further development in the lexical approach* Language Teaching Publications, Hove, UK.

Johns T (1994) From printout to handout: grammar and vocabulary teaching in the context of data-driven learning *in* T Odlin (ed) *Perspectives on pedagogical grammar*. Cambridge University Press, Cambridge.

Johns T and King P (1990) Classroom concordancing *English Language Research Journal*, University of Birmingham, Birmingham.

Kettemann B (1994) Concordancing in stylistics teaching Festschrift zum 30-jährigen Bestehen des Instituts für Anglistik der Universität Salzburg, Salzburg, in print.

Levy M (1990) Concordances and their integration into a word-processing environment for language learners *System* **18**, 2, 177–188.
—— (1997) *Computer-assisted language learning: context and conceptualization* Oxford University Press, New York.
Owen C (1996) Do concordances require to be consulted? *ELT Journal* **50**, 3, 219–224.
Steven V (1991) Classroom concordancing: vocabulary materials derived from relevant authentic text *English for Specific Purpose* **10**, 1, 35–46.
—— (1993) Concordances as enhancements to language competence *TESOL Matters* **2**, 6, 11.
—— (1995) Concordancing with language learners: why? when? what?" *CAELL Journal* **6**, 2, 2–10.
Sun, Y C (2001) E-learning in the EFL classrooms. Paper presented at the First English Teaching Conference. Taichung, Taiwan.
Tribble C and Jones G (1997) *Concordances in the classroom: A resource book for teachers* Athelstan Publications, Houston, TX.
Turnbull J and Burston J (1998) Towards independent concordance work for students: lessons from a case study *ON-CALL* **12**, 2, 10–21.