# Student Use of Concordancers for Grammar Error Correction

***Jaci Mull, Research Institute of English Language Education***
***Susan Conrad, Portland State University***

Emmy and Maria, two intermediate-level ESL students, are given a peer essay and asked to correct five errors highlighted in it. When they discuss the error, *They can save money for better future*, they can't agree on what to do. Emmy believes *better* should be changed to *best*. Maria believes they need to insert *the* before *better*. At this point, they reach for a concordancer and search a corpus. They do a search for *best future* and *better future* with no results. Then Emmy suggests they search for *better* alone, and they find many samples. Emmy notes that several samples, including *a better chance* and *a better understanding*, use the indefinite article. She suggests, "Put an article in front of better. So, 'a better'." The pair corrects the error and moves on to the next.

This scenario shows the potential power of students using concordancers to correct written errors. In the example above, the students would have falsely corrected the error had they been editing the paper with only their prior knowledge. Instead, the pair conducted their own small research study of English and noticed a gap between their hypothesized corrections and the target language. They then used the authentic target language samples they had found to establish an accurate and appropriate language pattern which they applied to the error in order to correct it. No teacher was needed to guide them through this process; the students were in full control of their language learning.

In fact, Emmy and Maria were real students in a small case study conducted to investigate what upper-intermediate level ESL students with little prior training could do with a concordancer when correcting grammar errors. In this article, we first provide some background information about student use of concordancers. We then describe the methods and results of the study, and discuss suggestions for future research and how to teach concordancing to students.

## Background

Concordancers are software programs that, when connected to a corpus, allow users to search for all occurrences of a word or sequences of words. Early descriptions of the potential for student concordancing were given by Tim Johns, who called it Data Driven Learning (DDL). Johns (1991) explains DDL as an investigation allowing students to form their own questions about the language, ask for authentic language samples via a concordancer, and use those language samples to help them find an answer to their question. DDL relies on the students to form their own generalizations about the language once they are presented with samples of the target language. Johns strongly connects DDL to induction, in which the language "rule" students are trying to learn is unknown to them and they must discover it for themselves by making generalizations from target language samples. The only difference Johns notes between DDL and traditional induction is that, for DDL, the teacher need not know exactly what patterns await, and may be equally surprised as the students by the results. This sense of

mutual exploration gives the learners a sense of autonomy and power in their language learning. Showing similar enthusiasm for students as researchers, Tribble and Jones (1990) suggested a variety of ways that teachers could prepare concordance lines for students to analyze or allow students to conduct their own searches for lexical or structural patterns.

While some educators may still be wary about focusing on form in more communicative or task-based classrooms, it is important to note that DDL has nothing to do with progressive grammar teaching. Johns (1991) himself notes that educators should not be trying to teach students a structured set of grammar. Instead, he sees DDL as an unparalleled opportunity to explore difficult areas of language that progressive grammar teaching traditionally avoided. He champions allowing students to find answers to their questions free of teacher intuition, which may not be based on close analysis of data and can be fallible. Johns suggests that through DDL, students are able to come to a far more subtle understanding of the target language than through other methods.

Several studies following Johns (1991) and Tribble and Jones (1990) focused special attention on students' ability to learn lexical patterns with concordancers (Cobb, 1997; Todd, 2001). Lexical patterns are fairly easy to spot in a Key Word in Context (KWIC) concordance search. The student simply types in one word and sees all instances of that word lined up with context on either side. It is easy to see what words associate with the search term. For grammar structure patterns, however, the task is more challenging. Students must be aware of parts of speech and not simply exact words. Even so, several studies have found that students can make generalizations about structure as well as lexis from concordance results (Gaskell, 2002; Gaskell and Cobb, 2004; Kennedy and Miceli, 2001 & 2010). Of course, when students are engaged in error correction with a concordancer, the task is slightly different from what Johns (1991) was describing. With error correction, the teacher has already identified an error and therefore has more control over what the student will discover.

In order to guide students through an inductive process and consolidate their learning, Gaskell (2002) created a set of five steps for students to complete while correcting grammar errors (see Figure 1). These five steps instruct students to first identify the error they wish to correct, and then to search with the concordancer for samples of English that show more accurate uses of the grammatical form. Students are then asked to summarize the pattern they see emerging from the concordance samples, followed by stating a grammar rule based on the samples they see and finally correcting the error. The hope with guiding steps like those in Gaskell (2002) is that students will follow the steps in order and by doing so induce the structure patterns needed to complete the editing task. At the end of the project, students also have a record of the grammar errors they have corrected. However, up to this point there is little evidence about whether or not students actually follow this process when they use a concordancer and

1) Example of the error
2) How the word/structure was used in the concordancer
3) What was learnt from the concordancer
4) How to fix the error
5) Correction of the error

*Figure 1: Gaakell (2002) steps for correcting grammar errors*

whether or not students can actually generalize structure patterns solely from looking at concordance lines.

### *Research Question*

As teachers, we know that students do not always complete tasks using the processes that we anticipate. While we might look at their end product and assume they have, for instance, induced patterns in the target language, we cannot be certain without direct observation. They may, after all, simply be guessing or relying on previous knowledge. We were therefore interested to investigate what students are actually doing with the tools we gave them. With this in mind, we asked the following questions:

1. How do students with little training employ a concordancer to investigate and correct errors in written work?
2. To what degree does a worksheet help guide the students through the inductive process for grammar error correction with a concordancer?

## Methods

Four ESL students participated in this case study. All were enrolled in upper-intermediate English classes in the intensive English language program at Portland State University, had achieved a score of 520 or higher on the TOEFL, and hoped to enroll in the university eventually. All the participants were comfortable with computers, but none had used corpora or concordancers for English language learning previously.

The participants completed the project in pairs to encourage verbalizing their thought process as they worked with the concordancer. All pairs had studied together in the same class and were familiar with each other before the study. The first pair consisted of Edison, a 27 year-old male student from South Korea and Y, a 23 year-old female student from Thailand. The second pair consisted of Maria, a 23-year-old female student from South Korea, and Emmy, a 36-year-old female student from Thailand (all names are pseudonyms). The diverse first language backgrounds encouraged students to verbalize their ideas in English.

The study was broken up into four sessions (see Figure 2). In the first session, the pairs of students were introduced to a concordancer and allowed 30 minutes to experiment with it using a guiding worksheet.

| Session | Activity | Materials | Time |
|---|---|---|---|
| 1 | Training with concordancer and time to experiment | Essay with 10 errors | 30 min. |
| | Journal writing | Journal | 10 min. |
| 2 | Interview about session 1 | NA | 10 min. |
| | Work session: Error correction time with concordancer and required worksheet | Essay with 5 errors<br>Worksheet (required) | 20 min. |
| | Journal writing | Journal | 10 min. |
| 3 | Interview about session 2 | NA | 10 min. |
| | Work session: Error correction time with concordancer and optional worksheet | Essay with 5 errors<br>Worksheet (optional) | 20 min. |
| | Journal writing | Journal | 10 min. |
| 4 | Final Interview | NA | 30 min. |

*Figure 2: Sessions in the study*

The training session was kept short because average classrooms do not have an abundance of time with which to train students on new technologies and we wanted to see what students could do given modest training. Students were taught searches only with words and word sequences, not with grammatically "tagged" sequences (which make grammar searches easier but require more advanced training). In the second session, students were briefly interviewed about the training session and then were given 20 minutes to edit grammar errors in a prepared essay using a concordancer. They were also given a guiding worksheet to aid with induction. The third session was set up exactly as the second session, including the brief interview and work time, but the worksheet was optional. Finally, in the last session, students met with the researcher for a lengthier interview.

***They had to determine for themselves what kind of error they faced.***

***Materials***

In their ESL class, the students were working on writing persuasive essays, so we designed the study to ask students to correct errors in similar essays. A corpus was specially created for them using 80 persuasive essays written by native speakers of English. The essays were found in the free use sections of two essay sharing websites: www.123helpme.com, and www.allfreeessays.com. Students used a copy of MonoConc 2.2 (Barlow, 2002) as their concordancer to search through the essays.

To develop the error correction task, three persuasive essays were created to match the style taught in upper-intermediate writing classes. One essay was for the training session, and the other two were for the work sessions. In order to use errors appropriate for these students, incorrect sentences were selected from diagnostic essays written by students of the same class level from a previous term and incorporated into the created essays. In order to be selected for incorporation into the new essays, the errors had to satisfy two criteria. First, they had to fit one of five grammatical categories students at this level typically struggle with: articles, prepositions, tense, clause construction, or agreement. Finally, they had to be searchable in the corpus. Specifically, by searching for words in the highlighted sentences, the students had to be able to find at least 10 samples that could help them correct the error. The training essay included 10 of these errors to ensure students had plenty to practice with the concordancer. Five errors were then incorporated into each of the work session essays, and the sentences were highlighted. One representative error from each grammar category was included in each work session essay. Students were aware that the highlighted sentences had grammar errors, but aside from that, they had to determine for themselves what kind of error they faced.

Rather than simply handing students the Gaskell (2002) steps as seen in Figure 1, a guiding worksheet was written to walk students through the five steps (Figure 3). It posed the steps in the form of directions. For example, instead of "example of the error" students were instructed to *1) Choose an error to correct* and *2) Write the error below*.

***Data Collection Instruments and Analysis***

During the first three sessions, video recorders were pointed at the students' computer screens to capture the searches they made with the concordancer and to record their voices as they negotiated how to cor-

rect the error and use the concordancer. At the end of each error correction session, students were asked to write a journal entry about their experience using a concordancer. Throughout the sessions the students were also interviewed about their opinions and to clarify points made in their journals.

For the analysis, we qualitatively described the process used by each pair for each error. We used evidence in the journals and interviews to interpret the processes more fully, and identified common opinions and attitudes about using the concordancer.

## Results and Discussion

Each student pair had 10 errors to correct (5 in each work session). Both pairs thought that some of the errors had obvious corrections that they could make based entirely on their previous knowledge. Feeling confident, they did not see a need to use the concordancer to confirm their knowledge of how the structures worked. The fact that time was short – only 20 minutes of work time for correcting 5 errors in each work session – might also have made them feel they needed to focus on the answers they felt least sure about. In fact, all of the corrections made without the concordancer were accurate (3 for Emmy and Maria, and 6 for Edison and Y).

### *How students employed the concordancer*

In the end, a total of 11 errors were addressed using the concordancer (7 by Em-

***Worksheet***

Please answer the questions below, in order, while you work on the computer to find corrections for the error you are investigating.

1. Choose an error from the essay.

2. Write the sentence containing the original error below:
   _____
   _____

3. Use MonoConc (the computer program) to find examples of better grammar to correct the error in the essay.

4. Look for a grammar pattern in the examples you find. Write down three examples you see that have a grammar pattern similar to the error you are correcting:
   A. _____
   B._____
   C._____

5. Based on the grammar pattern you have found, think of a rule to explain why these examples are correct (and the original error is incorrect).

Rule: _____
_____

6. Use the rule you just wrote to correct the original error that was in the essay. Write the corrected sentence:

Corrected Sentence: _____

*Figure 3: Student worksheet*

my and Maria and 4 by Edison and Y). With regards to the first question, "How do students with little training employ a concordancer to investigate and correct errors in written work?" we found only one instance of a full induction process as set out by the worksheet, which asked students to look at examples before generating a rule about the structure. Instead, students used shorter processes that focused on hypotheses they already held, with the pairs having mixed success when they found a lack of evidence for their hypotheses. While analyzing the students' processes, we found the following three points especially important.

1. A full induction process occurred for one pair but not the other.

As described in the opening of this article, Emmy and Maria began one of their concordance searches to correct the error *they can save money for better future* with two incorrect theories on how to change the error. Once they realized there was no support for either of their theories, they had to reconsider. In an interview Emmy admitted she first doubted the evidence from the concordancer, but, knowing the corpus was composed of all native speaker essays, she decided to trust the results and look for another solution. After searching for *better* alone, she was able to spot the regular use of an indefinite article and create a generalization from that.

***Once they realized there was no support for either of their theories, they had to reconsider.***

We were surprised that there was only this one case of induction. It occurred when Emmy and Maria acknowledged the lack of evidence for their previous hypotheses and had no other hypothesis to check, but found a useful search to gather new evidence. Edison and Y found the same error challenging, but they were not able to create a new generalization using evidence from their concordancing. Just like Maria, Y believed the correction should be *they can save money for <u>the</u> better future*. When she found no results to support her answer, she abandoned her hypothesis, but she was unable to conduct a new search and generate a new hypothesis. Her partner Edison suggested a completely different but inappropriate correction (discussed in point 3 below). Y did not accept this correction, but she offered no alternative. This pair appeared unable to figure out how to search for results that would generate a new hypothesis rather than just confirm what they hypothesized. They may also have felt constrained by the short amount of time for the task. In any case, this pair's inductive process never fully developed, and they never successfully corrected the error.

2. Students used the concordancer to support or challenge their hypotheses from previous knowledge.

The most common way that these students used the concordancer was to confirm patterns that they expected but were not entirely confident about. For every error correction attempted, students started with their own hypothesis (or competing hypotheses) about the correction and then consulted the concordancer. In most cases (7 out of 11), students found support for the first or second hypothesis they checked and could then confidently correct the errors.

Two of the other four cases where students used the concordancer are covered by points 1 and 3 (concerning full induction and generalizations that change meaning). The final two cases add more support to the interpretation that students can find it difficult to search with a concordancer when they

do not already have a clear hypotheses to test. For example, Emmy and Maria could not figure out a useful search when they were working on a structure that had an omitted relative pronoun: *if you take two different people have the same commitment*. They did not know what the error was and did four searches that focused on the verb *take*, which did not help them. Eventually Emmy remembered the need for a relative pronoun, based solely on her previous knowledge, not on anything she saw in the searches. Both students then felt confident about the correction and did not feel the need to confirm it.

3. Students could make generalizations that were grammatically accurate but changed the original meaning.

***The students were keen to use the concordancer in this way.***

For the phrase *for better future*, Edison decided the best way to correct the error was to use the chunk *in the future*, which he had heard many times. He did a search and found plenty of evidence in concordance lines to support use of this phrase. He therefore corrected the sentence to read *they can save money in the future*. While this sentence is grammatically accurate, the meaning had changed significantly from the original text. This example illustrates one of the limitations of concordancing: if a reader does not already understand meaning differences of similar wordings, it is difficult to see the difference without more extensive analysis of the context in the corpus. For students at this level, teacher input would likely be a more efficient means of helping the student understand the meaning change than corpus analysis would.

***The guiding worksheet***

Our second research question asked, "To what degree does a worksheet help guide the students through grammar error correction with a concordancer?" In general, the answer appears to be that the worksheet was not helpful. Even when the worksheet was required, the students usually completed the steps in a different order or back-filled the worksheet after they were satisfied they knew how to correct the error. During the one full process of induction, Emmy and Maria had the guiding worksheet with them, but - even though they eventually went through all the steps in the order listed - they didn't refer to the worksheet until after they had analyzed the concordance lines and decided on a correction. During the third session, when students were free to choose whether they used the worksheet or not, both pairs ignored it completely.

The usefulness of a worksheet, of course, depends on the design of the worksheet. The worksheet in this study did provide a structure for the training session and forced the students to record what they had learned. In trying to encourage a full inductive process, however, the worksheet did not allow enough of a role for students' use of their previous knowledge. It did not allow them a quick way to state and confirm their hypotheses when they were correct. At the same time, it did not guide students through what appeared to be the most challenging aspect of concordancing: determining new searches to try to identify patterns when they didn't already have a clear hypothesis to test. If a guiding worksheet is to be used in the future, it seems more appropriate to follow the students' natural inclinations. Rather than start with the concordance lines, the revised worksheet could, like many other corpus-based materials, ask students to use previous knowledge to state a hypothesis, and then prompt students to search the concordancer for support (or a lack of support)

for the hypothesis. Students could revise their hypothesis as many times as they wanted until they had found appropriate support. At that point they could state a rule and give a final correction for their error. If the students ran out of their own hypotheses to try, the worksheet could guide them through steps for new searches that might reveal patterns they had not thought about previously.

## Student enthusiasm

This study was not set up to investigate student use of concordancers during writing production, but the students in the study were keen to use the concordancer in this way, even after only 90 minutes with the concordancer spread over three days. In interviews, all four students compared the concordancer to other reference tools they already used at home and saw the concordancer filling a gap in their resources. They trusted the concordancer and corpus because they saw native speaker examples, and they felt it could give them ideas for language patterns they might not find in a standard dictionary search.

Bernardini (1996) argues student concordancing should focus more on discovery learning in which students search a target language corpus openly in a more exploratory fashion. Giving students concordancers as exploratory tools to inform their writing is not unprecedented. Kennedy and Miceli (2010) have produced two studies investigating just that and have found that students were able to create personal strategies for using the concordancer. Their students used the concordancer to hunt openly for new words and expressions that they could use in their writing, to generalize lexical and structure patterns when they had a specific question in mind, and to find target language equivalents for what they wanted to say.

In our study, the students themselves suggested that they might have been better able to induce language patterns had they been investigating the corpus freely as opposed to correcting errors. Both pairs of students came to this conclusion independently and with no prompting from the researchers. Interestingly, Edison and Y, who never showed a tendency towards induction both agreed that the concordancer would be more useful when exploring it openly rather than correcting grammar errors. Both pairs also suggested during interviews that they thought the concordancer would be best used as a reference at home where students have more time to explore and feel more relaxed. It is interesting to note that by following the students' suggestions, teachers would be able to use their valuable class time to focus on their regular lesson plans, while still allowing students the chance for autonomous language research outside of class.

## Conclusion

### *Concordancing and error correction*

The results of this study were mixed with respect to using concordancing for error correction. In the majority of cases, students felt so confident correcting the given errors that they did not use the concordancer to find supporting evidence. The chosen errors thus appeared to be too easy for the students in the study. Concordancers are most useful when the students need more evidence about how a structure works – not when students just need more time to apply their declarative knowledge to correct an error – but this can be a difficult judgment to make when planning an activity for numerous students. At the same time, when one pair of students did have the opportunity to use new evidence to revise their inaccurate hypotheses, they did not use the concordancer effectively.

Despite the shortcomings found in this study, however, the students also demonstrated some productive uses of concordancing that facilitated their autonomy as language learners. They used the concordancer to find evidence that confirmed their hypotheses about how to correct errors. In some cases, they also recognized evidence that their hypotheses were wrong, and in one case, induced a new, accurate generalization that allowed them to correct an error.

***Student ability to use the concordancer***

Many studies have argued in favor of gradual, structured training activities in the classroom before allowing students to work more independently with a concordancer. For example, Kennedy and Miceli (2010) created an "apprenticeship" program to train their students in concordancing and dedicated roughly 30% of their writing class, to it. Chambers (2005) argued against training students as if they were future corpus linguists, but still had students in her study complete 9 hours of corpus linguistics training.

***The students demonstrated some productive uses of concordancing that facilitated their autonomy as language learners.***

Realizing that the average instructor is not in a position to squeeze hours of concordancing training into their already tight curricula, this study took a different tack and examined how students used a concordancer after only 30 minutes of introduction to basic word searches. While the students in this study expressed some frustration with the software initially, and would likely have benefitted from more training focused on how to design searches, they did not seem discouraged by their lack of training. They used the concordancer when they felt unsure about how to correct errors. In addition, the more the students used the concordancer, the more ideas they reported for applying it to their regular studies and doing investigations when they had more time at home. Thus, even with minimal training, concordancing might be a useful activity for some students. Even instructors who do not want to spend much time with concordancing in the classroom might consider introducing their students to concordancing at least as a reference tool. Students can then judge for themselves whether it is compelling enough to pursue at home or in extra study time in a computer lab.

***Introducing students to concordancing***

Introducing students to concordancing requires access to two things: a corpus and a concordancer. In this study the researchers chose to create a corpus tailored to the type of essays the class was working on, and if teachers are able to do the same, it appears to be quite beneficial. Simply being able to tell students "this is a collection of writing just like what you are trying to do" seemed to add a level of credibility and authority to the corpus in the students' eyes. There are now a number of free or relatively inexpensive concordancers that can be used with your own corpus (e.g. AntConc is available free at www.antlab.sci.waseda.ac.jp, MonoConc is available for order at www.athel.com). However, teachers need not make their own corpus or invest in a concordancer. The Corpus of Contemporary American English (corpus.byu.edu/coca) contains 450 million words of American English with a free searchable interface. Searches are possible in categories such as academic texts, newspapers, and popular magazines – thus allowing students to focus on a more specific type of writing than, say, searching for a common

word using Google. The corpus also includes grammatical tags if teachers want to introduce students to more advanced grammar-related searches.

Student use of concordancers has probably not yet lived up to the ideal Johns envisioned when he first wrote about DDL in the late 1980s and early 1990s. Much research is still needed into the best ways to use concordancers, both when applied to error correction and during language creation. Longitudinal research into the development of student skills with concordancers is also needed. It is still largely unknown whether students will slowly develop strategies on their own over time, or whether training is the only way to improve students' skills. Even while research continues, however, there is reason to believe that students can get some benefit from consulting a concordancer, whether they fully induce patterns from the results or simply confirm what they already hypothesize.

## References


Barlow, M. (2002). MonoConc Pro (Version 2.2) [Computer software]. Houston, TX: Athelstan.

Bernardini, S. 1(996). Corpora in the classroom: An overview and some reflections on future developments. In J.McH Sinclair (Ed.), *How to use corpora in language teaching (pp.* 15-36). Amsterdam: Benjamins.

Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning and Technology 9*(2), 111-125.

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

Gaskell, D. (2002). *Can and will elementary ESL learners use concordance information to correct their own written errors?* Unpublished master's thesis, Universite du Quebec a Montreal, Quebec.

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors. *System 32(3),* 301-319.

Johns, T. (1991). Should you be persuaded: two samples of data-driven learning materials. *English Language Research Journal 4,* 1-16.

Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology, 5(3),* 77-90.

Kennedy, C., & Miceli, T. (2010). Corpus-assisted creative writing: introducing intermediate Italian learners to a corpus as a reference resource. *Language Learning & Technology 14*(1), 28-44.

Todd, R.W. (2001). Induction from self-selected concordances and self-correction. *System, 29(1),* 91-102.

Tribble, C., & Jones, G. (1990). *Concordances in the classroom: A resource book for teachers.* Harlow, England: Longman.



***Jacqueline Mull*** *is the Vice Dean of the Research Institute of English Language Education in Kobe, Japan and is a full-time instructor at Kobe Shukugawa Gakuin University. She is a native of Oregon and graduated from Portland State University with an M.A. in TESOL.*

***Susan Conrad*** *is a Professor of Applied Linguistics at Portland State University, where she regularly teaches the class Corpus Linguistics in Language Teaching. She is co-author of several books that use corpus linguistics, including Real Grammar: A Corpus -Based Approach to English (Pearson, 2009).*