Are we asking the right questions? Understanding which tasks lead to the robust learning of the English article system

Ruth WYLIE
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213, USA
rwylie@cs.cmu.edu

Abstract. Motivated by both classroom needs and learning science questions, we developed two computer-based systems to help students learn the English article system (*a*, *an*, *the*, null). The first system, a menu-based task, mimics cloze activities found in many ESL textbooks. The second, a controlled-editing task, gives students practice with both detecting errors and producing the correct response. Results from a think-aloud study show significant performance differences between the two tasks. Students find the controlled-editing task more challenging but appear more motivated and engaged when using the system. Current studies plan to address whether this added difficulty produces greater learning gains on robust learning measures such as transfer to writing production and long-term retention.

Keywords. English as a Second Language, CALL

Introduction

Article usage has been cited as one of the most frequent causes of error and one of the most difficult grammar points to teach English language learners (ELL) [1]. It is not uncommon for students to complete worksheets successfully on the topic but continue to make errors during authentic production. This paper presents work that attempts to bridge the link between practice and production using computer-based tutoring systems. Motivated by a strong classroom need, we built two systems to help students learn to make distinctions regarding English article use (*a*, *an*, *the*, and null). The first is a menu-based tutor that mimics cloze, or fill-in-the-blank, exercises found in many textbooks. The second helps students learn both error detection and error correction skills via a controlled-editing tutor. Unlike in the production-only task where students know exactly where changes need to be made, in the controlled-editing task, students must first identify where the error exists and then take steps to correct it. However,

students are limited in the type of changes they can make. Both tutors provide immediate feedback and detailed hints to help students complete the task. Preliminary findings show that students find the controlled-editing task more challenging but appear more motivated and engaged when using the system than when using the menu-based tutor. Future studies plan to examine whether the added difficulty of the controlled-editing tutor results in larger and more robust learning gains as measured by transfer to writing production and long-term retention.

1. Motivations

1.1 Classroom Motivations

One of the main motivations for developing and studying tutors specifically for English article use is that students have great difficulty learning these skills. English Second Language (ESL) instructors report that teaching the English article system is one of their biggest challenges [1]. This was supported in an analysis of student essays collected from the Pittsburgh Science of Learning Center's English LearnLab during Summer 2005. Error analysis of the essays revealed that article errors represent approximately 15% of all the errors present. However, many English Language teachers are reluctant to spend much time on teaching articles since article errors do not often lead to communication failure (e.g. People generally understand the phrase *"Yesterday, I went to store"* even though *"Yesterday, I went to the store"* is correct). However, this is not to say that teachers do not recognize the harmful effects that grammar mistakes have on the credibility of written work [6], simply that the ratio of time spent to communication gains does not justify spending large amounts of time on the topic in the typical ESL classroom. Thus, a tutoring system that complements in-class instruction, but which students access in their own time, is likely to be well-received by both over-committed teachers and students looking for additional practice opportunities.

1.2 Learning Science Motivations

Since instructional time is at a premium, it is important to understand how students learn and retain information in order to design efficient and effective ways to teach them. This work compares the effects of two different tasks on student performance. In one condition, students are asked only to select the correct response (menu-based task). In the second condition, students are asked to correct errors by first detecting the error and then selecting the correct response (controlled-editing task). The trade-off between explicitly teaching both error-detection and production skills versus fully concentrating on production skills remains an empirical question.

One thought is that if students are highly skilled in production then they will get error detection for free. For example, if students, while reading a text automatically produce the correct article for each noun phrase, they will be able to recognize errors as those where the text differs from their generated response. However, as many teachers have suggested, perhaps the goal of completely error-free production is unrealistic and students need to have explicit practice both detecting errors and producing the correct response. This is supported by studies that show students who were responsible for

producing answers *and* detecting their mistakes made greater learning gains than students who were responsible for generation alone [5].

2. Tutoring Interfaces

2.1 Menu-Based Tutor: Production-Only Task

The menu-based tutor, built using the Cognitive Tutoring Authoring Tools (CTAT) [4], is similar to the cloze, or fill-in-the-blank, activities found in many ESL textbooks. Using this interface, students select an article from each drop-down menu in order to complete the paragraph. Students do not need to identify where errors exist; they simply choose a response for each box (Figure 1).

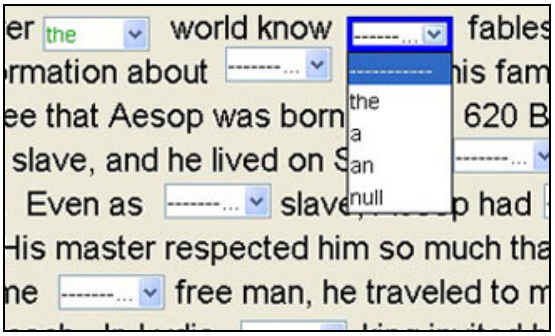

Fig. 1. Menu-based interface for Production Only Task

The hints available while doing this task guide students in the article selection process. When students first for a hint, they are presented the basic dichotomy between the definite (*the*) and indefinite (*a*, *an*, or *null*) articles. The second hint provides further guidance with respect to the specific noun phrase with which the student is working, while the third, and final, hint gives the student the answer to the problem as well as an explanation of why that answer is correct. For example, given the noun phrase, *People all over ___ world*, the final hint reads, "Please insert *the* into the blank. The definite article is used here because there is only one *world*, so *world* is a unique noun. Examples of other unique nouns are *the moon*, *the internet,* and *the sky*."

2.2 Controlled-Editing Tutor: Detection + Production Task

The controlled-editing tutor is also implemented via CTAT. The controlled-editing tutor allows students to insert, remove, or change articles anywhere in the text. However, only articles can be edited thus preventing students from completing rewriting sentences in order to avoid certain grammar constructions. In this interface, students must both detect the error and produce the correct response (Figure 2).

Students can also ask for hints while using this system. Since students are now asked to identify as well as correct the errors, the first two levels of hints aid in the detection process by narrowing the search space in which students look for an error (e.g. Look for an error in the second sentence.). The subsequent three levels of hints are identical to those hints given in the production-only task.

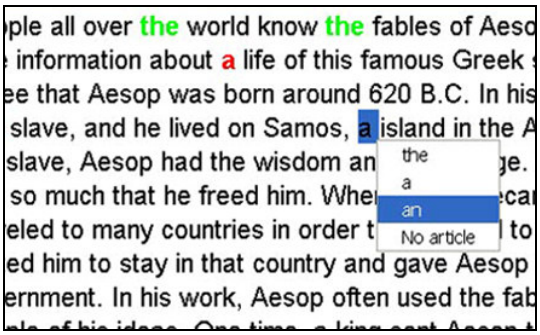

Fig. 2. Controlled-Editing interface for Detection + Production Task

3. Think-Aloud Study

A study was conducted in order to examine performance differences between the two task types as well as gather verbal protocols in order to understand which rules and heuristics students employ when solving these problems. If students performed equally well (or poorly) using the different interfaces, one would not expect to see differences in learning. In answering these questions, we performed a think-aloud study in which we invited ELL participants into the lab and asked them to complete tasks using the two systems. Students were told beforehand that the only errors present within the paragraphs were article errors.

3.1 Task Content

The problem paragraphs came from intermediate and advanced-level ESL textbooks (e.g. [2], [3]). The first problem was shorter in length and used simpler vocabulary than the second problem did. Students were randomly assigned to one of two groups: those in Group 1 completed the intermediate level problem using the editing (production and detection) interface and the advanced level problem with the menu (production-only) interface. Students in Group 2 did the opposite (See Table 1).

Table 1. Type of interface used for each problem

	Intermediate Level Problem	Advanced Level Problem
Group 1	Production + Detection Task	Production Only Task
Group 2	Production Only Task	Production + Detection Task

3.2 Participants

Participants were recruited from Carnegie Mellon University's InterCultural Communication Center's Academic Culture and Communication (ACC) program. The program is a six week summer program designed for newly admitted non-native English speaking students. Traditionally, students entering the program have high TOEFL scores (greater than 580) and at least an intermediate level of spoken fluency. ACC students are a highly educated and highly motivated group.

There were six participants total (three female, three male), and all were native Chinese speakers. The average age was 28 years old and all students had been learning English since middle school (average number of years = 14.4 years). Using self-report scales, participants gave an average proficiency score of 3.4 for reading, 3.0 for writing, and 2.5 for speaking (where 1 represents absolute beginner and 5 fluent).

The participants of this study represent a population that is more advanced than the overall target population of these systems. This is due to the think-aloud methodology employed for the pilot study. We needed students whose English ability was high enough that they were able to verbalize what they were doing. However, because English articles are often one of the last grammar points for students to master, we were able to use texts that were challenging even for these advanced participants.

3.3 Results

The data reveal a strong performance difference between the two tasks. Again, the difference in the two interfaces was that in the production-only version, students only had to choose which article to insert for each given box. In the production and detection condition, students had to identify *and* correct the errors. If students were able to locate errors correctly but had difficulty correcting them, we would expect for the results to be the same regardless of the interface. However, if error identification is the true obstacle, we would expect students in the detection and correction condition to perform worse than students using the production-only tutor.

For both problems, the production-only task resulted in higher accuracy, as measured by the percent of necessary changes that were correctly made. While the difference between the two interfaces was not significantly different for the easier, intermediate-level problem ($t=1.45$, $p = 0.13$), the performance results were significantly different for the advanced problem ($t=2.13$, $p = 0.05$), suggesting that when students are presented with level-appropriate texts, detecting errors is a formidable challenge (Figure 3).

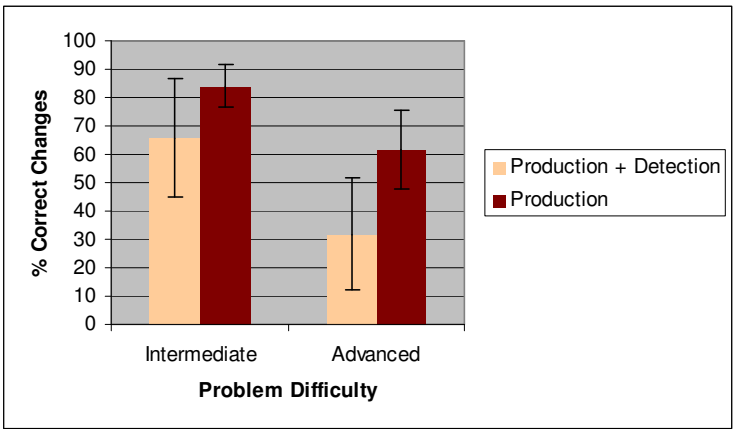

Fig. 3: Results comparing student performance on the two tasks

4. Conclusions and Future Work

Learning the English article system is a challenging but necessary process for students writing academic and professional texts. Thus, it is important to understand which tasks best support knowledge acquisition, retention, and transfer. This work marks the first steps of this process: understanding the educational and learning science motivations, developing the tutoring systems, and conducting initial student studies. While the results show significant performance differences between the production and detection and the production-only tasks, the question, and focus of current work, is whether this added difficulty will produce greater learning gains. If so, these findings would resonate well with work done by Schmidt and Bjork [7] which shows that increased difficulty during acquisition can result in greater post-training retention. Current plans also include moving the tutors out of the lab and into the classroom in order to gather authentic data from a more diverse student sample. Finally, learning measures will be expanded to include more robust items such as student performance on authentic writing production as well as long-term retention tests in order to answer learning science questions that align with the true educational objectives of the unit.

Acknowledgements. This work was supported in part by the Pittsburgh Science of Learning Center which is funded by the National Science Foundation award number SBE-0354420 and by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B040063 to Carnegie Mellon University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Special thanks to Teruko Mitamura and Ken Koedinger, Jim Rankin, for his work in developing the tutoring systems, and the three anonymous reviewers for their suggestions and feedback.

References

[1] Celce-Murcia, Marianne, and Larsen-Freeman, Diane. (1983). *The Grammar Book: An ESL/EFL Teacher's Course*. Rowley, Massachusetts: Newbury House Publishers.

[2] Fuchs, M., Bonner, M., & Westheimer M. (2006). *Focus on Grammar 3: An Integrated Skills Approach, Third Edition.* White Plains, NY: Pearson Education, Inc.

[3] Huckin, T. & Olsen L. (1983). *English for Science and Technology: A Handbook for Nonnative Speakrs*. McGraw-Hill, Inc.

[4] Koedinger, K. R., Aleven, V., Heffernan. T., McLaren, B. & Hockenberry, M. (2004). Opening the Door to Non-Programmers: Authoring Intelligent Tutor Behavior by Demonstration. In the Proceedings of 7th Annual Intelligent Tutoring Systems Conference. Maceio, Brazil.

[5] Mathan, S. & Koedinger, K. R. (2003). Recasting the feedback debate: Benefits of tutoring error detection and correction skills. In U. Hoppe, F. Verdejo, & J. Kay (Eds.), Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies, Proceedings of AI-ED 2003 (pp. 13-18). Amsterdam, IOS Press.

[6] Master, P. (1997). The English Article System: Acquisition, Function, and Pedagogy. *System*. 25,(2) 215-232.

[7] Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. Psychological Science, 3(4), 207–217.