Running head: WORD PREDICTION PROGRAMS

Word Prediction Programs with Phonetic Spelling Support: Performance Comparisons and Impact on Journal Writing for Students with Writing Difficulties

Anna Evmenova, Marci Kinas Jerome, Michael Behrmann, & Heidi Graff

George Mason University

Abstract

This study examined the effects of current word prediction software programs available that support phonetic/inventive spelling on the legibility, fluency and length of journal writing by six students with severe writing and/or spelling difficulties in grades three through six during a month long summer writing program. A changing conditions single subject research design was used and replicated across six students. Using a daily writing prompt, students alternated between Co:Writer, WordQ, and WriteAssist word prediction programs. Results provide evidence for the effectiveness of various word prediction programs over word processing demonstrating improvement of the total number of words, word and sentence fluency, and composition rate in students' writing. Furthermore, qualitative interviews reveal that all students enjoyed the word prediction programs and found them beneficial.

Word Prediction Software for Students with Writing Difficulties

In recent years there has been an increasing interest in technology applications for students with high-incidence disabilities including those with writing difficulties. Several applications were discussed in the literature (Higgins & Raskind, 2004; Edyburn, 2005; Lewis, 1998, MacArthur, Ferretti, Okolo & Cavalier, 2001). Writing is a very complex skill and students may experience difficulties with any of its aspects: from mechanics to written content expression. However, it is evident that computer-related technologies may enable those students to bypass their deficits and support them though all stages of the writing process (Behrmann & Jerome, 2002; Lewis, 1998; Williams, 2002; Zhang, 2000). Research has been accumulated to determine the effectiveness of word processors for easier text alteration and manipulation (Lewis, Ashton, Haapa, Kieley, & Fielden, 1998; MacArthur & Graham, 1987; MacArthur & Schwartz, 1990; Zhang, 2000), as well as spell checkers and other aids for easier editing (Ashton, 1999; MacArthur, Graham, Haynes, & DeLaPaz, 1996; McNaughton, Hughes, & Ofiesh, 1997; Montgomery, Karlan, & Coutinho, 2001). Text-to-speech software programs allowing users to hear written products were found to be effective for accuracy monitoring (MacArthur, 1998, 1999; Raskind & Higgins, 1995). Outlining and brainstorming programs allowing visual representation of ideas were determined to support users in planning and organization of their writing (Anderson-Inman & Ditson, 1990; Blair, Ormsbee, & Brandes, 2002; Sturm & Rankin-Erickson, 2002), and speech recognition programs allowing for the transformation of spoken words into text were examined as an alternative way of writing for students with learning disabilities and/or writing difficulties (De La Paz, 1999; MacArthur & Cavalier, 2004; Quinlan, 2004; Raskind & Higgins, 1999; Higgins & Raskind, 2000).

However, the area of assistive technology for students with mild disabilities is still not fully developed (Edyburn, 2005). The information on the actual use of technology to support students specifically with high-incidence disabilities is limited (Blackhurst, 2005; Edyburn, 2001), including the area of utilizing word prediction software for students with learning disabilities. Historically, designed for users with physical disabilities, word prediction was determined to increase the typing rate and decrease spelling errors by that population (Tumlin & Heller, 2004). While the number of reduced keystrokes addressed the needs of students with physical disabilities, the features of word prediction software also found application in compensating for word recall, spelling, and handwriting difficulties of students with learning disabilities (Lewis, 1998). With a word prediction program, the user is offered a list of word choices as she/he begins to type the word. The suggestions may appear before or after the first letter of the word is entered. Word “prediction” features allow for the words to be generated based on the lexical and grammatical context, while word “completion” simply completes the word after some initial spelling attempts (MacArthur, 1999; Sitko, Laine, & Sitko, 2005). As other assistive technology solutions, word prediction may yield writing products of higher quality if the features of the program are coordinated with the user’s abilities and needs (Ashton, 2005; Sitko, Laine, & Sitko, 2005). However, despite being a promising application, the use of word prediction for students with writing difficulties is somewhat understudied (MacArthur, et. al, 2001).

Previous Research

The majority of word prediction studies were conducted almost a decade ago. Several of the most recent studies have revealed that writing readability/legibility and spelling of students with learning disabilities and writing difficulties improve with word prediction (Handley-More, 2003; MacArthur, 1998, 1999; Williams, 2002). In his original study, MacArthur (1998) investigated the effects of speech synthesis and word prediction software programs as compared to a word processor. For four out of five students the features offered by those programs resulted in improved legibility and spelling in dialogue journal entries. Later, MacArthur (1999) extended that study using more sophisticated word prediction programs. In this latter design students alternated between handwriting, a word processor, and a word prediction program. The results were quite modest yielding improvement in proportion of correctly spelled words for one out of three students and decreased composition rate for two students. Technology had no effect on proportion of legible words during journal writing. Furthermore, the follow-up study was conducted with the same students in the attempt to control for complex word prediction capabilities with the increased demands of the writing task (MacArthur, 1999). Students wrote from dictation, thus increasing vocabulary demands. The results demonstrated improvement across all variables for two out of three students with the decreased composition rate. Williams (2002) and Handley-More, et al. (2003) reported relative improvements in the number and a variety of words, as well as in the percentage of legible and correctly spelled words.

All aforementioned studies noted possible effectiveness of word prediction for students with writing difficulties. Still, it was noted that the impact of word prediction was quite limited. The fact that students had to know the exact beginning letters of the word without the possibilities of phonetic substitutions presented one of the major limitations. Thus, students with severe spelling problems did not benefit from word prediction programs because very often they did not know the correct initial letters (MacArthur, 1998). Word prediction technology has developed significantly since then. The major difference is that current software programs attempt to recognize phonetic spelling when words are spelled the way they sound as compared to a conventional spelling. This way software recognizes inventive spelling, so it is not necessary to enter the exact beginning letters to receive a legitimate prediction. Thus, current technology may be more beneficial for students with learning disabilities than the older versions.

Research Questions

The purpose of this study was to determine the effects of current word prediction software programs available that support phonetic/inventive spelling on the legibility, fluency and length of journal writing by students with severe writing and/or spelling difficulties. The initial comparison of three different word prediction programs including students' preferences was conducted. This study therefore, was intended to replicate and extend the work of previous researchers (MacArthur, 1998, 1999) by asking the following research questions:

1. Do the fluency and length of journal writing increase while students with writing difficulties use current word prediction software vs. word processing?
2. Do the length and rate of writing depend on a word prediction program and its features?
3. Do word and sentence fluencies depend on a word prediction program and its features?
4. What program do students find the most helpful and enjoyable?

Method

A changing conditions single subject research design was used and replicated across six students.

Participants

Students. Participants were students in grades 3 through 6 attending the CompuWrite summer camp. All students were referred as having writing difficulties by their parents and teachers. Some of them were also identified as having learning disabilities by their schools. Among all campers, 7 students were identified as potential candidates for participation based on the informal writing assessment as well as writing samples collected prior to the study. The possible benefits from word prediction features was determined based on the criteria including students' phonetic spelling, limited vocabulary, word recall, and keyboarding skills as well as the reading ability to differentiate between words on a prediction list.

Student 1 was a 12-year-old boy with specific learning disabilities in written expression and math. He was a rising 7th grader who received special education services in the inclusive general education classroom. According to his parents and a camp teacher, Student 1 had difficulties with planning and organizing his writing. It was observed that he fixated on the spelling and lost track of thought. He also tended to spell words phonetically.

Student 2 was a 9-year old, rising 4th grade boy with specific learning disabilities in written expression and reading comprehension. Student 2 received services in both special education resource classroom and inclusion settings. According to his mother, the writing process was “physically difficult for him to do and for others to read” the finished product. He had great ideas that came from his well-developed imagination but putting them down on paper was a struggle for him. Student 2 was also identified as having autism spectrum tendencies so he had very strong opinions about what technology he wanted and/or refused to use for his writing.

Student 3 was a 10-year old, rising 5th grade boy recommended for the summer camp because of his reluctance toward writing. He was also tested and found to have traits for autism spectrum disorder with a very high-functioning level. Student 3 was reported to be easily frustrated with the writing process when he could not think of the correct spelling. He found it hard to focus, organize, and convey his ideas into a cohesive document. He was considered a study participant for his phonetic spelling and much needed support with finding the right word to convey his ideas.

Student 4 was a 9-year old boy, a rising 4th grade student. He was identified as having attention deficit disorder (ADHD) and received up to 8 hours a week extra help in the general education classroom for difficulties with writing. His mother shared with the researchers that Student 4 “did not like to write and did so as little as possible.” His biggest problem was attending to the task; therefore, he was considered a candidate for word prediction use to provide him with additional word choice support.

Student 5 was an 11-year old boy, a rising 6th grade student. He received special education services in the inclusive settings for specific learning disabilities in written expression. Student 5 was reported to have severe spelling difficulty; therefore, he was considered a good candidate for the use of word prediction programs. His mother requested all writing assignments to be completed on a computer with an additional help for "spelling issues".

Student 6, a 9-year old boy, was identified as having specific learning disabilities in written expression and math. This rising 4th grade student was pulled out into a special education resource classroom to receive extra help in writing and math. After careful observations and consultations with his family, school teachers, and camp instructor, Student 6 was determined as a study candidate to help him overcome the hesitance to write. In addition, utilizing computer programs for writing met his challenges with fine motor skills and handwriting.

One more student was withdrawn from a study after the first week. It was determined that his needs in writing mechanics could sufficiently be met with the *Microsoft Word* spelling checker and a voice output software program.

Teachers. CompuWrite provides a unique internship opportunity for current and prospective teachers working on their master's and licensure in learning disabilities, emotional disturbance, and mental retardation at George Mason University. While gaining authentic experiences with the current technology options for students with mild disabilities, interns are responsible for working with children. Six interns participated in the study. The teacher/student ratio was 1:1. However, as explained later students completed activities for the word prediction study independently. Prior to the beginning of the CompuWrite camp all interns received training on the use and integration of major writing computer tools in their lesson plans. Thus, the utilization of three word prediction programs did not interfere with the CompuWrite camp activities.

Setting

The study took place during the CompuWrite summer camp at George Mason University (GMU). CompuWrite is a four-week long summer camp that uses technology and innovative computer software programs to enhance the writing process and improve written language skills for students experiencing difficulties with the writing process. Camp's activities were divided between two sessions that lasted 2.5 hours each. CompuWrite camp was located at the computer laboratory at GMU. The large room with 25 computers was organized into 5 stations containing 4-6 computers each. A majority of camp instructors had their own individual stations while two of them shared one station with 6 computers. Computer stations were spread throughout the room so that instruction at each station did not interfere with the rest of the campers. In addition, students using text-to-speech features in word prediction programs wore earphones at all times.

All instructional sessions were conducted in 1:1 instructional arrangements. Students were situated directly in front of the computer screen. Camp instructors sat beside the students and provided assistance as needed. During the observations, the researcher was located behind a student and a teacher, so that she could easily see the computer screen but not distract the students. The researcher began observations prior to the study rotating from station to station, so students had time to get used to being observed.

Materials

All conditions. In all conditions students were asked to write daily for 20 minutes in response to the journal entry prompt. The purpose of such journal writing is to provide students with more writing opportunities and daily practice (Reagan, 2005). It usually is free of any kind of evaluation (Williams, 2002). Personal narrative prompts were randomly assigned to students from a list of 30 pre-design prompts. They were interesting and unbiased based on gender, ethnicity and socio-economic status (e.g., "What is your favorite part of the day?", "What is something that makes you feel happy or sad?", etc.)

Baseline condition. In the baseline condition students used *Microsoft Word* for journal writing. Students were not able to use spell checkers and grammar checkers during writing.

All treatment conditions. In the treatment conditions students used three word prediction programs: *Co:Writer, WordQ*, and *WriteAssist*. A student types the word either in the separate program application or in Microsoft Word. As each letter is typed the list of predicted words appears in the small window located by the cursor. If the intended word appears in the list, a student can select the word by clicking on it or typing the number of that word. That selected word is automatically added to the sentence. If the intended word does not appear in the predicted list, a student continues to type. All three programs provide speech feedback so students have an option to hear predicted words before selecting one of them. These programs also have an option for the teacher to decide how many words will be predicted for the student (usually between 1 and 9). While the number of predicted words is usually based on individual student's needs, for this study the number of predicted words was limited to 5 in all the programs. All three programs have spell checkers built into them. However, for the sake of this study the spell checker option was disabled in all word prediction programs as well as in the word processing. While these three programs are somewhat similar in their features, they are slightly different in the level of sophistication and the size and diversity of the dictionary. Condition-specific materials are described next.

Co:Writer. Co:Writer SOLO Edition is the latest version of the *Co:Writer* word prediction program developed by Don Johnston Inc. It utilizes Linguistic Word Prediction intelligence. With that function, the word prediction list does not depend on the correct first letters. It is based on the phonetic, inventive spelling that is very typical for students with learning disabilities and writing difficulties. In addition, *Co:Writer* offers such functions as eWordBank and Topical Dictionary. Such features support student's writing on different topics and in different genres predicting the most appropriate words for the selected topic and/or genre.

WordQ. WordQ by Quillsoft is a word prediction tool used with a standard word processor. This program suggests words when students have trouble spelling or choosing the word just like other word prediction programs. However, unlike other two programs, *WordQ* does not correct grammar or punctuation so the quality of writing still depends on students.

WriteAssist. WriteAssist by Second Guess software is a dyslexia-oriented word predictor. Program features include context-depended prediction which ensures that a student will be offered suggestions even without typing anything. The program will make a prediction of the possible next word based on grammatical patterns and the context.

Experimental Design

Changing conditions single-subject design was chosen to investigate the effects of the various independent variables represented by different word prediction programs on students' writing (Alberto & Troutman, 2006). Prior to treatment, students' baseline level of writing was collected for a minimum of three data points across the first week of camp. Following the stabilization of baseline, the first treatment was introduced. Students were randomly assigned to different word prediction programs so that each student had an opportunity to try 3 programs by the end of the study. During each following week students wrote using various programs alternating the order across participants. The random assignment to each of the programs for particular weeks was used to control the influence of the increasing mastery and familiarity with word prediction skills (Table 1). Changing treatments sequentially allowed the examination of various programs before finding one that was the most beneficial for each particular student. The replication across students allowed for establishment of stronger functional relationships between various word prediction programs and students' writing as well as to control for confounding variables such as novelty of treatment (Clark, 1994; Weller, 1996) and acquisition of necessary skills (Alberto & Troutman, 2006).

<INSERT TABLE 1 HERE>

Dependent Measures

The dependent variables examined included: total number of words, word fluency, sentence fluency, and composing rate by written words per minute.

Total number of words. Total number of words was calculated in each of students' writing samples. The differences were calculated between the length of writing and the use of word processing vs. word prediction. In addition, the differences in length were compared among the three different programs.

Word fluency. Word fluency was calculated by dividing the number of legible and correctly spelled words by the total writing time in minutes. As defined by MacArthur (1998) legible words are those words that can be correctly decoded even when taken out of context. In order to identify the number of legible words, the independent scorer started with reading each word in isolation covering the rest of the words. The purpose for this was to identify if the words made sense separate from the context. In order to avoid guessing, such procedure began with the last word in the passage and went backwards towards the beginning. Then, reading the whole passage together, it was important to make sure that a student meant that word and not another one in the particular context. Homonyms were not considered legible words but considered as spelling errors.

Sentence fluency. Another measure, sentence fluency, was determined by counting word sequences that were grammatically correct and used a capital letter at the beginning of the sentence and the punctuation at the end. Then that number was divided by the total writing time. This measure, as with all the previous ones, was compared across the programs.

Composing rate. Composing rate of typed words per minute demonstrated how long it took each student to complete writing as well as the number of total words students used in their writing. It was calculated by dividing the total number of words in students' writing by the total composing time. Teachers were asked to record the beginning and ending writing time. In addition, the researcher observed all the students during writing.

Procedures

All conditions. Once student and parent permissions were obtained, students were engaged in journal writing at the beginning of each camp session. The study was conducted over a period of four weeks of writing camp. Teachers gave students personal narrative prompts for writing encouraging them to do their best. Students could take less or more time writing but the general time for this activity was approximately 20 minutes. Later during the day students were engaged in other writing activities including brainstorming, drafting, revising, editing and production. Thus, the purpose of journal writing activity was just to provide another opportunity to write without spending time on editing. However, if students wrote less than 3 sentences, teachers asked them to say more, provide more details on the topic. Students were also encouraged to figure out the spelling of the words or choose the word from the prediction list on their own, so that teachers were not helping with that verbally. The researcher was in the classroom during the journal writing activities to gather observational data on the writing speed as well as the students' use of word prediction functions.

In both conditions, after completing the journal writing activity, students saved the copy of the work on the computer. In addition, they printed two copies: one to include in their folder and the second one for the researcher.

Baseline condition. During the first week of the camp, students were writing their journal entries using a word processor. Depending on students' typing and computer skills, they received some instruction in typing and using the word processor if needed. Such instruction included one-on-one training from a teacher and/or using *Type2Learn* software program with interactive lessons to teach how to type and lessons to improve the speed. Comparing writing samples to the ones produced using word processing program created conditions for controlled baseline. The existing practice of handwriting was substituted with using technology for the baseline condition, thus less number of changes occurred when students were introduced to treatment conditions (Kennedy, 2005).

Treatment condition. Students and teachers received instruction on how to use each word prediction program. The researcher conducted a training session with students as well as a separate training session with teachers addressing the main features of each word prediction program. Participants learned how word prediction works. Training sessions included a short PowerPoint presentation addressing the basic information students and teachers had to know about word prediction features. Then, it was demonstrated that every time students did not know how to spell a word they could look at the predicted list and move the mouse over the words to hear them pronounced out loud. Students learned about the speech-feedback option where they could hear any word, phrase, and sentence as many times as they wanted. However, the use of this feature was optional. Students and teachers were encouraged to try and type a sentence using word prediction features.

The students were randomly assigned to the order, in which they would use programs weekly (see Table 1). Each week teachers modeled the particular program that was assigned to students for that week. Students learned how to start a program, enable the word prediction feature, utilize the speech feedback feature if chosen, and where to look for the predicted list. Due to the similarity of programs, teachers introduced students to each program on Monday of each week as opposed to a more formal training. Teachers modeled the journal-writing activity for students in order to address specific functions of the particular program. All teachers were extensively trained on the use of each program as part of their internship requirement. In addition, the researcher developed a short training manual including the handout for teachers to make sure that children were introduced to the program the same way. After that, students had some time to practice using each software program.

Interobserver Reliability and Fidelity of Treatment

Interobserver reliability was determined using printed writing samples as permanent products. All journal entries were scored on 4 dependent variables: total number of words, word fluency, sentence fluency, and the composing rate. In order to establish scoring reliability, random writing samples (33%) were distributed to the independent professional to ensure that she scored them the same way as the researcher. The permanent product agreement was calculated using the total agreement method (Kennedy, 2005). Each observer recorded and calculated total number of words, word and sentence fluency ratio as well as composition rate. The smaller total number of each variable recorded by the observers per writing sample was divided by the larger total, and multiplied by 100%. Interobserver agreement averaged 99 percent (ranging from 89-100%) for total number of words, word and sentence fluency as well as composition rate.

Fidelity of treatment data was collected during 33% of all sessions. The randomly observed activities were compared to the checklist to ensure that teachers and students were doing what they were supposed to do. Such observations occurred both during journal writing and word prediction program trainings for students. The number of correct behaviors performed by teachers and students were divided by the number of planned behaviors and multiplied by 100%. The operation of computer programs was checked prior to each session, and each session was observed to ensure that the programs ran properly. The sessions when programs experiences technical difficulties were excluded from the data analysis. Mean procedural reliability was 100% across programs and students.

Social Validity

Social validity of each intervention, each individual word prediction program, was examined through students' and teachers' interviews conducted throughout as well as at the end of the study. Social validity as defined by Kennedy (2005) is "the estimation of the importance, effectiveness, appropriateness, and/or satisfaction various people experience in relation to a particular intervention" (p.219). As with any assistive technology device or program it is very important to ensure a person's willingness to use it. Students' preferences of a program and its technological features very often play a more important role than their effectiveness. A very large percent of assistive technologies are abandoned because they do not meet person's needs and expectations (Scherer, 2005). It is critical to seek students' input when technology is selected (Parette, Wojcik, Peterson-Karlan, & Hourcade, 2005).

Data Analysis

All students' writing samples in both baseline and three treatment conditions were analyzed. Data on the number of total words, word fluency, sentence fluency, and composing rate were calculated. Visual analysis of graphed data points across all dependent variables in both baseline and treatment condition was conducted. Evaluation of changes was conducted in means estimating the average rate of change in students' writing across 4 variables (Alberto & Troutman, 2006). In addition, the effectiveness of different words prediction programs on students' writing was calculated using the percentage of non-overlapping data (PND) score revealing "the proportion of overlapping data displayed between treatment and baseline" (p. 27) (Scruggs, Mastropieri, & Casto, 1987). A larger PND indicates the higher effectiveness of an intervention.

Furthermore, randomization tests were conducted for the data collected across all measured dependent variables for each student. Randomization tests and analysis were conducted with the help of special software for single-subject designs (Todman & Dugard, 2001) and SPSS. Randomization tests present a solution for the data with great variability within each phase that confounds clear differences through visual analysis. These tests are especially recommended in cases when treatment effects may not be obvious (Scruggs, Mastropieri, & Regan, 2007). A single-subject design study utilizing randomization tests incorporate random assignment of treatments (Ferron & Onghena, 1996). Single-case Two Randomized Treatments test was chosen and conducted for the baseline and each of the word prediction programs separately for each student and dependent variable. The treatment sessions were randomly assigned for a single participant. This randomization test hypothesizes the difference in means between the baseline and each separate word prediction treatment. Thus, it tests the prediction that each program will help a single student to produce improved writing sample as compared to word processing. The prediction is based on 2,000 randomly selected mean differences between various sessions (Todman & Dugard, 2001).

Students' and teachers' interviews were conducted and analyzed to determine social validity of word prediction in general as well as each individual program. After the interviews were transcribed, the researchers conducted the preliminary analysis by dividing the text into segments and coding those segments using in vivo codes. Then the data was further inspected using a constant comparative method and was manually grouped together on similar dimension creating separate themes (Guba & Lincoln, 1989). Furthermore, relationships among codes within themes were explored through a concept map. It was anticipated that the data would yield multiple viewpoints so the researchers would have an opportunity to compare participants' perspectives on different word prediction programs.

Results

Word Processing Versus Word Prediction

As indicated in Figures 1, 2, 3, and 4, graphed data for each participant provided evidence for the effectiveness of various word prediction programs over word processing demonstrating improvement of the total number of words, word and sentence fluency, and composition rate in students' writing. The overall PND score across students and across all three programs demonstrates average 80 percent improvement in total number of words, 82 percent in word fluency, 63 percent in sentence fluency, and 84 percent in composition rate from word processing to word prediction. Tables 2 and 3 present individual PND scores and randomization test results for every student for each separate dependent variable when using three different word prediction programs. As indicated in Tables 2 and 3, different students performed better with different programs. Thus, it is critical to analyze the performance of each individual student across all dependent variables.

Total Number of Words with Different Word Prediction Programs

As can be seen from Figure 1, all students demonstrated relative increase in the total number of words from word processing baseline across three word prediction programs. Based on the visual analysis, Students 2, 3, 5, and 6 demonstrated greater improvements in the total number of words when using the *WordQ* word prediction program. While Student 1 showed progress with all three programs, he appeared to produce a slightly larger number of words particularly with the *WordQ* program. Student 4's graphed data indicated effectiveness of the *WriteAssist* software for that student.

<INSERT FIGURE 1 HERE>

The percent of non-overlapping data corroborates the results drawn from the visual analysis. The variety existed in the degree of improvement among the students and software programs corresponding to the range of PND changes. Thus, Students 2, 3, 5, and 6 scored 100% PND when using *WordQ* demonstrating 100% improvement in the total number of words as compared to the word processor. The same students' PND scores for other programs ranged from 20 to 80 percent yielding a smaller but still a moderate progress. The PND scores for Student 1 were 100 percent for each of the programs. Student 4's PND scores arrayed from 80% for *WordQ* and *Co:Writer* to 100% for *WriteAssist* word prediction program. Table 2 list specific PND scores for each student and each word prediction program across all four variables.

<INSERT TABLE 2 HERE>

The randomization tests further support the aforementioned results. According to randomization tests, the differences in the total number of words between the baseline and writing with *WordQ* program for Students 2, 3, 5, and 6 were statistically significant ($p<0.05$). Other programs did not produce statistically significant results despite the visual improvements for those students. Differences between the baseline and each word prediction program for Student 1 were statistically significant ($p=0.01$). Student 4 demonstrated the statistically significant improvement in the total number of words only when writing with *WriteAssist* software (p=0.02). Detailed randomization tests results can be found in Table 3. As mentioned before randomization tests are especially helpful when the visual analysis is hindered. Thus, despite the fact that Student 3 demonstrated a higher mean line as shown in Figure 1 for the total number of words while using the *WriteAssist* program, the randomization test yield no statistically significant difference between baseline and WriteAssist ($p=0.07$) with 80 percent of non-overlapping data. On the other hand, the *WordQ* program with lower mean line produced statistical significance ($p=0.02$) with a PND score equal 100 percent.

<INSERT TABLE 3 HERE>

Word Fluency with Different Word Prediction Programs

Visual inspection of the data points for word fluency across three different programs suggested the following results. Using *WordQ* software program resulted in larger improvement of word fluency for 4 students (Students 2, 3, 5, and 6) while Student 1 performed better with *Co:Writer* and Student 4 with *WriteAssist*.

<INSERT FIGURE 2 HERE>

All students increased in word fluency when using *WordQ* word prediction program by the average of 93% (ranging from 80% to 100%). The average PND score for *Co:Writer* program was 73% (20%-100%). In turn, students improved in word fluency by averaged 80% with *WriteAssist* software (60%-100%). Some differences were also statistically significant based on the randomization test results. Thus, *Word Q* program was associated with statistically significant differences for all students ($p<0.05$). The word fluency significantly increased for 2 students (Students 1 and 6) with the *Co:Writer* software ($p<0.05$). Students 1, 3, and 4 demonstrated statistically significant progress in word fluency between baseline and writing with *WriteAssist* ($p<0.05$).

Sentence Fluency with Different Word Prediction Programs

<INSERT FIGURE 3 HERE>

Sentence fluency with three different word prediction programs and word processing demonstrated great variability within baseline and treatment phases for all students. Visual inspection of the data was somewhat cumbersome due to such variability. The PND scores and randomization tests contributed to a complicated visual analysis. PND scores were averaged to 70 percent for *WordQ* and 60 percent for *WriteAssist* and *Co:Writer* demonstrating modest effects of word prediction programs on sentence fluency in journal writing for students participating in the study. Sentence fluency increase between baseline and *WordQ* program was statistically significant for Student 4 and 6 ($p<0.05$). *WriteAssist* produced statistically significant results only for Student 6 ($p<0.05$). *Co:Writer* was found statistically significant in sentence fluency for Student 1 ($p<0.05$).

Composition Rate with Different Word Prediction Programs

<INSERT FIGURE 4 HERE>

All three programs allowed students to increase their composition rate to a different degree. Graphic representation of data points in composition rate for Student 1 demonstrated increased performance with *Co:Writer*. One hundred percent PND score and statistically significant difference on randomization test ($p=0.02$) supported such conclusion. Despite the visual analysis, *WordQ* and *WriteAssist* also had 100 percent PND score and were statistically significant ($p<0.05$).

Visual analysis for Students 2, 5, and 6 indicated higher composition rate when writing with the *WordQ* word prediction program. PND scores (80%-100%) and statistically significant results of randomization tests ($p<0.05$) corroborated the effectiveness of *WordQ* software over other programs for these students. However, it is important to note that despite the fact that other programs appeared to result in lower composition rate, they were still moderately effective for these students. Thus, the difference between the baseline and *WriteAssist* for Student 5 was also statistically significant ($p=0.02$). In turn, *CoWriter* was statistically effective for Student 6 in composition rate ($p=0.01$).

Graphic representations for Students 3 and 4 suggested slightly higher composition rate with the *WriteAssist* program. However, while Student 3 improved significantly only with *WriteAssist* software ($p=0.02$), randomization test result for Student 4 yielded statistically significant difference for all three programs ($p<0.05$).

Social Validity

Overall, all students enjoyed the word prediction programs and found them beneficial.

They indicated that writing was much easier when they used word prediction. Student 3 noted that he did not have to write the whole word and the program would finish it for him. Another student mentioned that word prediction made him type words faster. One more example of the advantage of word prediction as reported by Student 1 was that it “helped find words and see if they were correct or not in order to use them.” In addition, Student 2 reported that word prediction made him think faster.

In regard to which program students found the most helpful and enjoyable, 4 out of 6 students preferred *WordQ* to other programs, while the other two other students liked *WriteAssist* and *Co:Writer* the most. Students who chose *WordQ* referred to it as the fastest and having “better words.” One student mentioned that “it was like telepathic” because “the words came up” as you were just thinking about them. Another student enjoyed that the program read the sentences “exactly as you read them.” Student 2 refused to use other programs because they did not have as many voices as *WordQ* did. As for those students who did not prefer *WordQ*, they noted that constant speech-feedback was annoying and that the window although smaller “moved around to places where they did not want it.” Both those problems were eliminated within program options as students expressed their opinions.

The main problem with the *WriteAssist* program as reported by students was the big window that “did not move and covered the words.” *WriteAssist* was found to have less word choices. In addition, the vocabulary was not appropriate for students as it offered “bad words” as reported by students and teachers. *Co:Writer* had more “technical glitches” within the software in comparison to other programs. It is fair to say that since the time of the study several glitches were addressed and solutions offered on the manufacturer’s website.

The teachers supported students’ opinions and preferred *WordQ* as their favorite program. It was found to be the easiest to use while offering a large choice of features. Fourbutton toolbar of W*ordQ* program was determined to be “very straight forward” and simple to handle. Teachers noticed the vocabulary issues with *WriteAssist* software program as it offered “profane words.” One teacher described it as the “most primitive of all”. In addition, one teacher with vision impairments noted that *WriteAssist* program “would be hard to use with a screen enlarging program” as the prediction window stays in the locked position on the screen as opposed to *WordQ* and the word window of *Co:Writer*. In turn, the *Co:Writer* program had “a cleaner language” and had more features for students. However, it was also more difficult to use for students in this study and presented more technological difficulties.

Discussion and Implications

The primary goal of this study was to explore the effects of various word prediction programs on students’ journal writing as compared to word processing. Consistent with previous research (Handley-More, 2003; MacArthur, 1998, 1999; Williams, 2002) the results of this study demonstrated word prediction effectiveness on various aspects of writing process for students with writing difficulties as compared to word processing. Students’ writing increased in the number of total words, word and sentence fluency as well as composition rate from the baseline differently across various word prediction programs. Due to the nature of the changing conditions single subject research design, these results should be interpreted with caution (Kennedy, 2005). The absence of the return to baseline prior to starting each new program alters the conclusions about functional relationships. However, this study presents unique information comparing different word prediction programs. External validity of this study is enhanced through replication across different participants (Horner, et al., 2005). Furthermore, random assignment of students to different order of programs’ implementation controls for vulnerability of findings to confounding effects of novelty (Clark, 1994; Weller, 1996).

Overall, regardless of the order in which it was introduced, 4 students performed better with *WordQ* word prediction program on the number of total words, word fluency and composition rate. One student demonstrated larger number of words, higher word fluency and composition rate with the *WriteAssist* software program. Student 1's performance improved equally with all three word prediction programs so the recommendation of one of them depended solely on his preferences. Sentence fluency did not increase dramatically with any of the programs. The explanation may be that the primary purpose of word prediction programs is to address spelling difficulties and it provides only limited support for grammatical structure. It was surprising that *WordQ* produced the largest number (3) of statistically significant improvements from the baseline in sentence fluency. As mentioned before WordQ word prediction program does not correct grammar or punctuation so the quality of grammatical aspects of writing still depends on students. Despite this finding, additional assessment is required before recommending this program to students who need additional help with punctuation and capitalization.

Social validity of the goals, procedures, and effects were examined through students' and teachers' interviews. Both students and teachers enjoyed using word prediction programs and found them helpful for the writing process. All students benefited from word prediction features that supported their writing difficulties. For example, a student with fine motor/handwriting difficulties mentioned the ability to type faster with word prediction. Furthermore, a student who experienced difficulties with putting his ideas on paper reported that word prediction made him "think faster."

A majority of students and teachers preferred the *WordQ* word prediction program to other two. Such decision was supported by the extended and appropriate vocabulary and ease of use. It was interesting that several students referred to WordQ "telepathic" abilities recognizing its contextual prediction. Although *Co:Writer* and *WriteAssist* also have similar features, *WordQ* predictions were determined to be more precise. One student, who liked *WriteAssist* the most, demonstrated a better performance with that program on all depended variables. The student who chose *Co:Writer* demonstrated equally significant improvement with all the programs so following his preferences, *Co:Writer* was recommended as a word prediction program for him.

Based on the results of this study, *Co:Writer* word prediction was reported as being more difficult to use due to extended number of features. That allows to suggest that this program would be a better choice for older students who could use and benefit from them. *WriteAssist* was the only program that offers users up to 30 choices in the prediction list. This program was also reported as having a more adult-oriented vocabulary. It is also important to note that a student with ADHD considered the prediction window moving with a cursor annoying and distracting. Thus, it is important to consider turning off such feature in *WordQ* and *Co:Writer* with “word window” prediction. In addition, for students with autism spectrum tendencies a program with larger selection of voices (*WordQ*) could be preferable.

Limitations

The results of this study should be interpreted without taking into consideration the following limitations. First of all, the short duration of the writing camp guided the number of journal writing sessions randomly assigned for each word prediction program. Furthermore, the length of the summer camp hindered the researchers’ ability to test maintenance and level of continuous improvement of students writing with a particular word prediction program. In addition, time also influenced the choice of the research design that excluded return to baseline, thus preventing establishment of strong functional relationships between different word prediction programs and improvements in students’ journal writing. Finally, the research setting at CompuWrite summer camp was different from a general education classroom where a majority of participants receive special education services. Generalization of the writing improvement with word prediction program was not assessed in realistic school environment.

Implications for Further Research

This pilot study provided several areas for future research. First, the study could be replicated allowing longer periods of time for each word prediction program. Additionally, the future researchers should employ different single subject research designs (e.g., multiple-baseline across participants) to establish stronger functional relationships between the writing improvement and each word prediction program.

Second, it would be interesting to examine the effectiveness of different word prediction programs on other more meaningful writing activities that require editing in more natural school settings. Such research may also focus on more extensive features of word prediction programs including prediction based on topic and genre.

Lastly, a majority of research studies examining the effectiveness of word prediction are single subject research studies (Siko, Laine, & Sitko, 2005). This may be explained by the specificity of word prediction programs that makes them appropriate for students with very specific abilities and needs. Such studies, including the present one, discourage interpretation and generalization of finding to a larger population. Group design experimental studies with a large number of participants would provide generalizable information on word prediction effectiveness.

References

Alberto, P. A., & Troutman, C. A. (2006). *Applied behavior analysis for teachers*, (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Anderson-Inman, L., & Ditson, L. (1999). Computer-based concept mapping: A tool for negotiating meaning. *Learning and Leading with Technology, 26(8)*, 6-13.

Ashton, T. M. (1999). Making technology work in the inclusive classroom: A spell CHECKing strategy for students with learning disabilities. *Teaching Exceptional Children, 32(2)*, 24-27.

Ashton, T. M. (2005). Students with learning disabilities using assistive technology in the inclusive classroom. In D.L. Edyburn, K. Higgins, & R. Boone (Eds.), *Handbook of special education technology research and practice* (pp. 229-238). Whitefish Bay, WI: Knowledge by Design, Inc.

Behrmann, M., & Jerome, M. K. (2002). *Assistive technology for students with mild disabilities: Update 2002*. Arlington, VA: ERIC Clearinghouse on Disabilities and Gifted Education.

Blackhurst, A. E. (2005). Historical perspectives about technology applications for people with disabilities. In D.L. Edyburn, K. Higgins, & R. Boone (Eds.), *Handbook of special education technology research and practice* (pp. 3-29). Whitefish Bay, WI: Knowledge by Design, Inc.

Blair, R. B., Ormsbee, C., Brandes, J. (2002). Using writing strategies and visual thinking software to enhance the written performance of students with mild disabilities. ERIC Document Reproduction Service No. ED463125.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology, Research, and Development, 42(2)*, 21-29.

De La Paz, S. (1999). Composing via dictation and speech recognition systems: compensatory technology for students with learning disabilities. *Learning Disabilities Quarterly, 22,* 173-182.

Edyburn, D. L. (2005). Assistive technology and students with mild disabilities: From consideration to outcome measurement. In D.L. Edyburn, K. Higgins, & R. Boone (Eds.), *Handbook of special education technology research and practice* (pp. 239-270). Whitefish Bay, WI: Knowledge by Design, Inc.

Edyburn, D. L. (2001). Critical issues in special education technology research: What do we know? What do we need to know? In M. Mastropieri, & T. Scruggs, (Eds.), *Advances in learning and behavioral disabilities, Vol. 15,* NY:JAI Press, pp. 95-118.

Handley-More, D. (2003). Facilitating written word using computer word processing and word prediction. *American Journal of Occupational Therapy, 57(2)*, 139-151.

Higgins, E. L. & Raskind, M. H. (2000). Speaking to read: The effects of continuous vs. discrete speech recognition systems on the reading and spelling of children with learning disabilities. *Journal of Special Education Technology*, *15(1)*, 19-30.

Higgins, E. L., & Raskind, M. H. (2004). Speech recognition-based and automaticity programs to help students with severe reading and spelling problems. *Annals of Dyslexia, 54(2),* 173-177.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*, 165-179.

Kennedy, C. H. (2005). *Single-case designs for education research*. Boston, MA: Allyn and Bacon

Lewis, R. (1998). Assistive technology and learning disabilities: Today's realities and tomorrow's promises. *Journal of Learning Disabilities, 31*, 16-26.

Lewis, R. B., Graves, A. W., Ashton, T. M., & Kieley, C. L. (1998). Word processing tools for students with learning disabilities: A comparison of strategies to increase text entry speed. *Learning Disabilities Research & Practice*, *13*, 95-108.

MacArthur, C. (1998). Word processing with speech synthesis and word prediction: Effects on the dialogue journal writing of Students with learning disabilities. *Learning Disabilities Quarterly, 21*, 151-166.

MacArthur, C. A., & Cavalier, A. R. (2004). Dictation and speech recognition technology as test accommodations. *Exceptional Children*, *71*(1), 43-58.

MacArthur, C., Ferretti, R., Okolo, C., & Cavalier, A. (2001). Technology applications for Students with literacy problems: A critical review. *The Elementary School Journal*, *101(3)*, 273-301.

MacArthur, C., & Graham, S. (1987). Learning disabled students' composing under three methods of text production: Handwriting, word processing, and dictation. *Journal of Special Education*, *21*, 22-42.

MacArthur, C. A., Graham, S., Haynes, J. A., & De La Paz, S. (1996). Spelling checkers and students with learning disabilities: Performance comparisons and impact on spelling. *Journal of Special Education*, *30,* 35-57.

MacArthur, C., & Schwartz, S. S. (1990). An integrated approach to writing instruction: The Computers and Writing Instruction Project. *LD Forum*, *16(1),* 35-41.

McNaughton, D., Hughes, C., & Ofiesh, N. (1997). Proofreading for students with learning disabilities: Integrating computer and strategy use. *Learning Disabilities Research & Practice*, *12*, 16-28.

Montgomery, D. J., Karlan, G. R., & Coutinho, M. (2001). The effectiveness of word processor spell checker programs to produce target words for misspellings generated by students with learning disabilities. *Journal of Special Education Technology*, *16(2)*, 27-40.

Parette, H. P., Wojcik, B. W., Peterson-Karlan, G., & Hourcade. J. J. (2005). Assistive technology for Students with mild disabilities: What's cool and what's not. *Education and Training in Developmental Disabilities*, *40*, 320-331.

Quinlan, T. (2004). Speech recognition technology and students with writing difficulties: Improving fluency. *Journal of Educational Psychology*, *96*, 337-346.

Raskind, M. H., & Higgins, E. H. (1995). The effects of speech synthesis on proofreading efficiency of postsecondary students with learning disabilities. *Learning Disability Quarterly*, *18*, 141-158.

Raskind, M. H. & Higgins, E. L. (1999). Speaking to read: The effects of speech recognition technology on the reading and spelling performance of children with learning disabilities. *Annals of Dyslexia*, *49*, 251-281.

Reagan, K. S., Mastropieri, M. A., & Scruggs, T. E. (2005). Promoting expressive writing among students with emotional and behavioral disturbances via dialogue journals. *Behavioral Disorders*, *31*, 35-52.

Scherer, M. J. (2005). Living in the state of stuck: How assistive technology impacts the lives of people with disabilities, (4th ed.). Brookline, MA: Brookline Books.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. Remedial *and Special Education*, *8(2)*, 24-33.

Sitko, M. C., Laine, C. J., & Sitko, C. J. (2005). Writing tools: Technology and strategies for struggling writers. In D.L. Edyburn, K. Higgins, & R. Boone (Eds.), *Handbook of special education technology research and practice* (pp. 571-598). Whitefish Bay, WI: Knowledge by Design, Inc.

Sturm, J. M., & Rankin-Erickson, J. L. (2002). Effects of hand-drawn and computer-generated concept mapping on the expository writing of middle school students with learning disabilities. *Learning Disabilities Research & Practice*, *17*, 124-139.

Todman, J., & Dugard, P. (2001). *Single-case and small-n experimental designs: A practical guide to randomization tests.* Mahwah, NJ: Lawrence Erlbaum Associates.

Tumlin, J., & Heller, K. (2004). Using word prediction software to increase typing fluency with students with physical disabilities. *Journal of Special Education Technology*, *19(3).* Retrieved September 24th, 2006 from http://jset.unlv.edu/19.3/tumlin/first.html

Weller, H. G. (1996). Assessing the impact of computer-based learning in science. *Journal of Research on Computing in Education*, *28,* 461-485.

Williams, S. (2002). How speech-feedback and word prediction software can help Students write. *TEACHING Exceptional Children*, *34*, 72-78.

Zhang, Y. (2000). Technology and the writing skills of students with learning disabilities. *Journal of Research on Computing in Education*, *32*, 467-478.

Table 1

Randomly Assigned Order for Different Word Prediction Programs' Implementation across Six Students with Writing Difficulties

Students	Week 1 Baseline	Week 2 Treatment 1	Week 3 Treatment 2	Week 4 Treatment 3
Student 1	Word Processing	Word Prediction 1	Word Prediction 2	Word Prediction 3
Student 2	Word Processing	Word Prediction 2	Word Prediction 3	Word Prediction 1
Student 3	Word Processing	Word Prediction 3	Word Prediction 1	Word Prediction 2
Student 4	Word Processing	Word Prediction 1	Word Prediction 2	Word Prediction 3
Student 5	Word Processing	Word Prediction 2	Word Prediction 3	Word Prediction 1
Student 6	Word Processing	Word Prediction 3	Word Prediction 1	Word Prediction 2

Figure 1: Total Number of Words in Journal Entries across Programs and Students

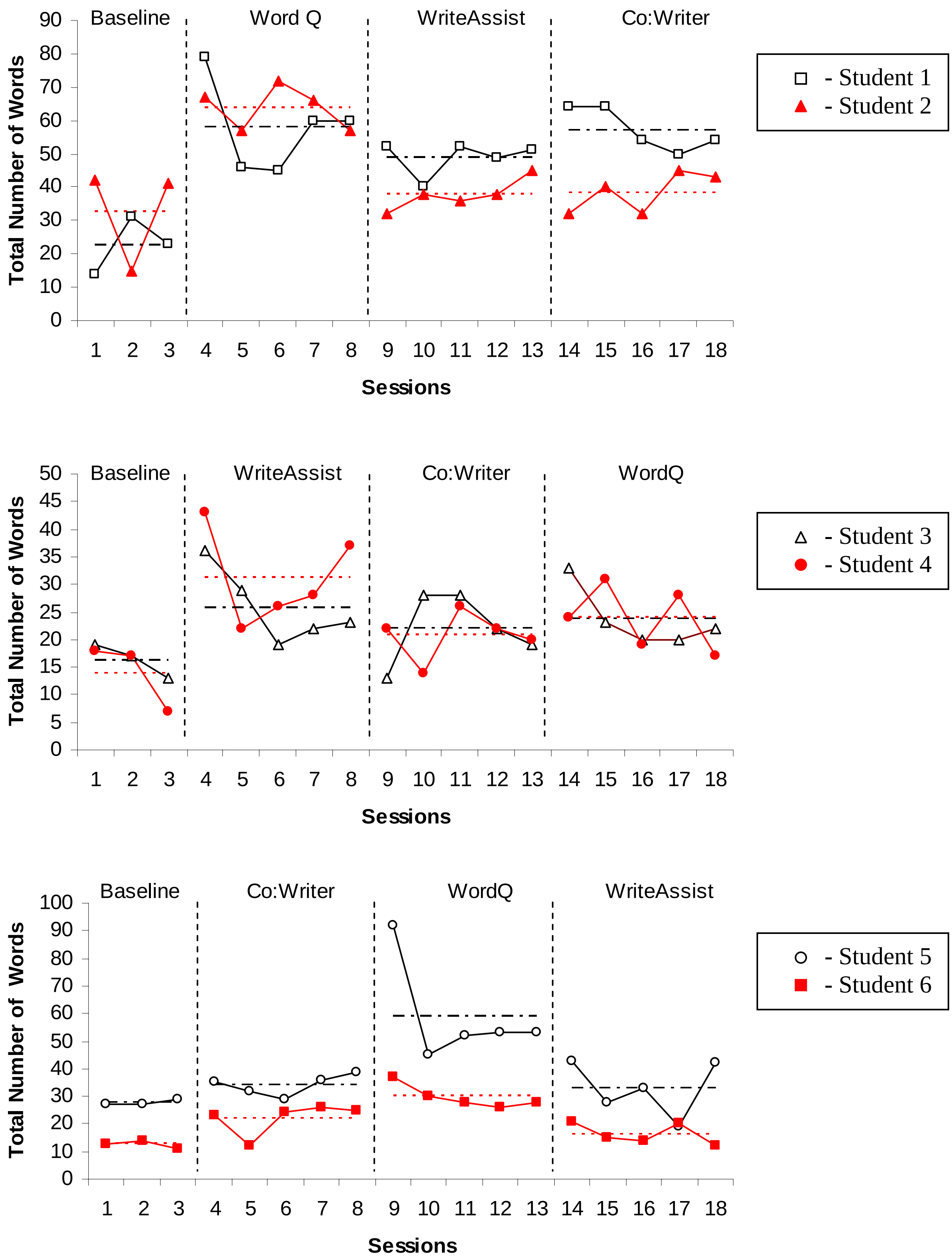

Figure 2: Word Fluency in Journal Entries across Programs and Students

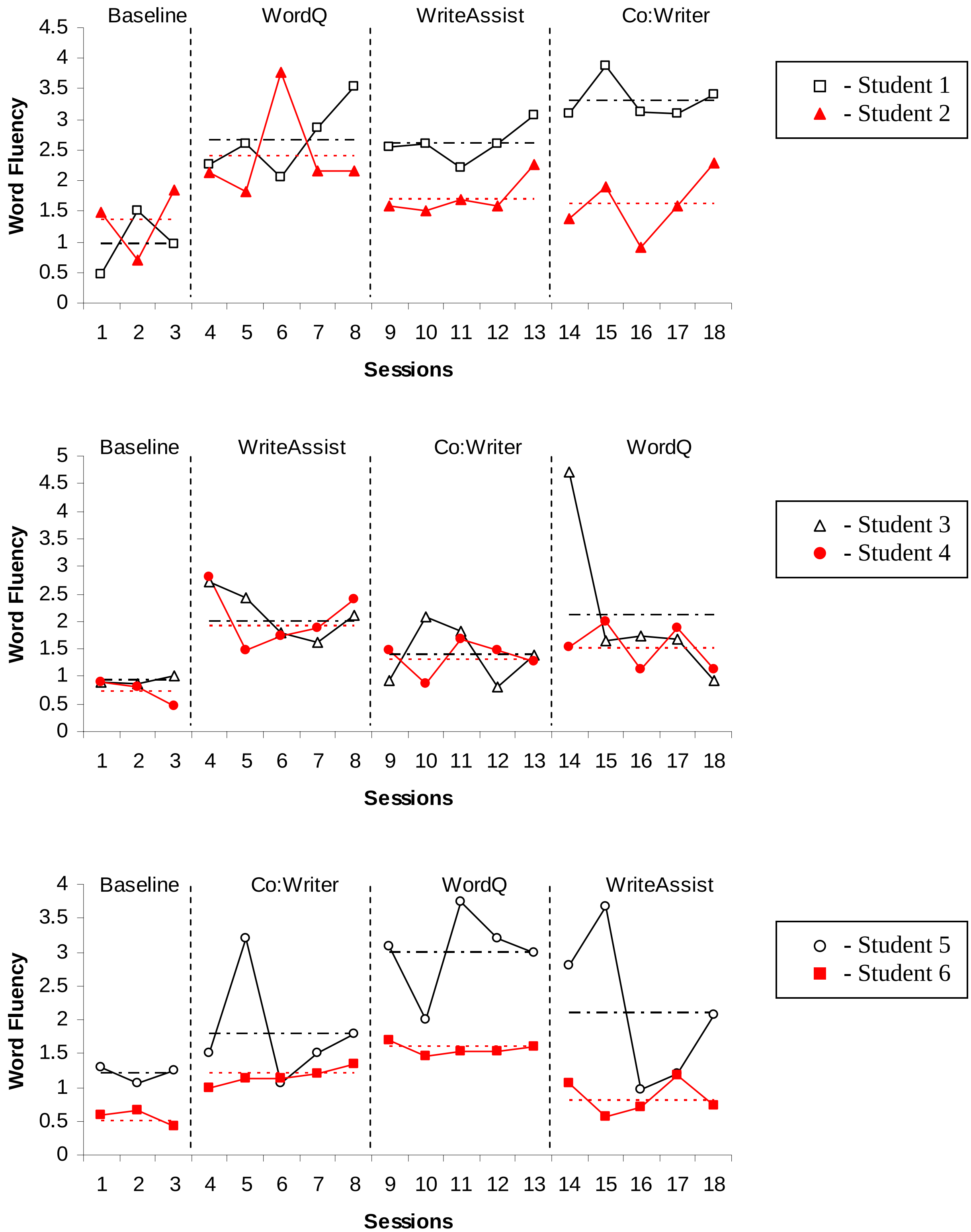

Figure 3: Sentence Fluency in Journal Entries across Programs and Students

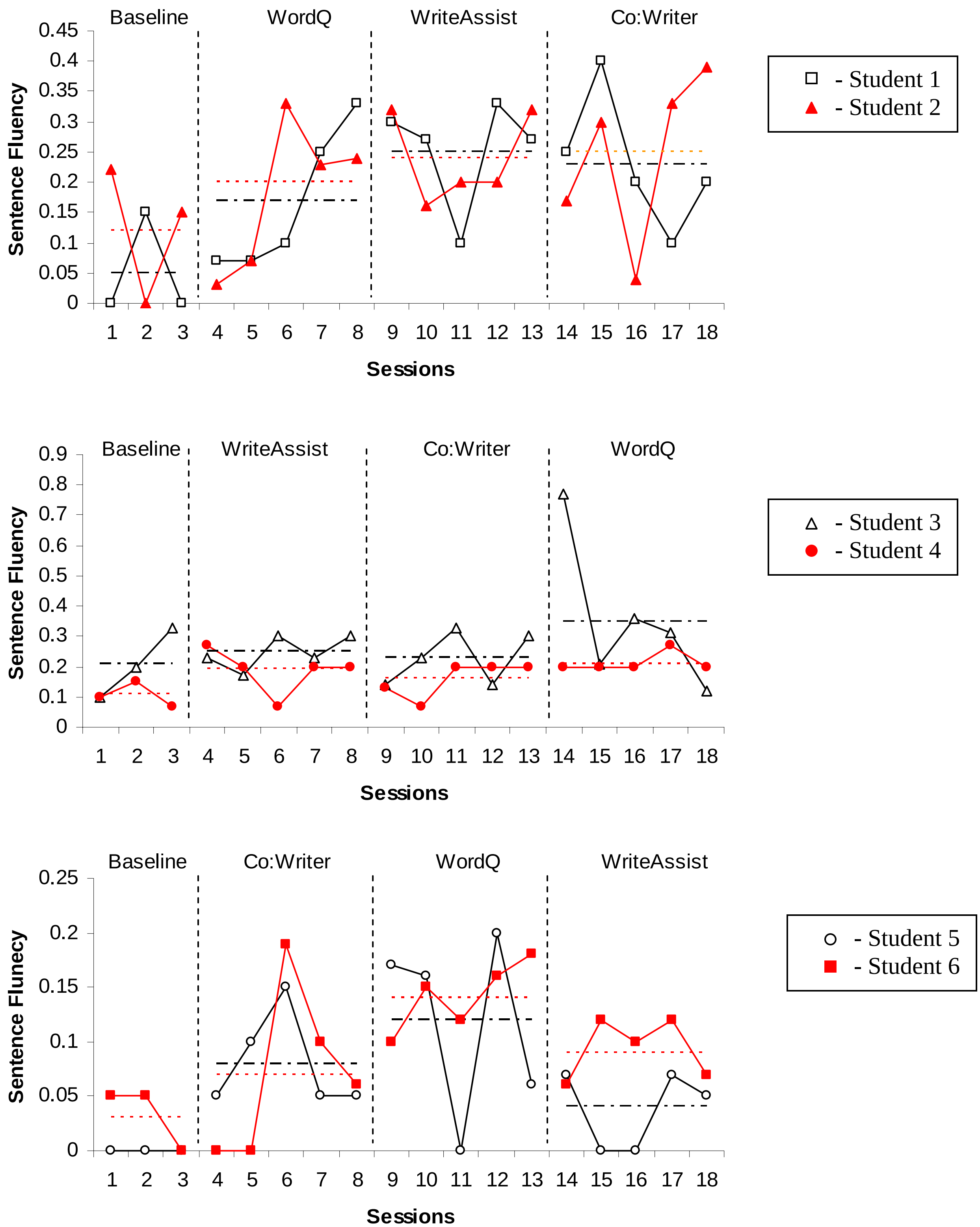

Figure 4: Composition Rate of Journal Entries across Programs and Students