A Comparison of Accuracy and Rate of Transcription by Adults With Learning Disabilities Using a Continuous Speech Recognition System and a Traditional Computer Keyboard

Diane C. Millar
University of North Texas

David B. McNaughton and Janice C. Light
The Pennsylvania State University

Abstract

A single-subject, alternating-treatments design was implemented for three adults with learning disabilities to compare the transcription of college-level texts using a speech recognition system and a traditional keyboard. The accuracy and rate of transcribing after editing was calculated for each transcribed passage. The results provide evidence that the accuracy of transcription using the speech recognition system was slightly lower than the accuracy of transcription using the traditional keyboard for all of the participants. The rate of transcription using the speech recognition system was very similar to the keyboard for two of the participants. For the third participant, the transcription rate using the keyboard exceeded the rate using the speech recognition system for all four sessions. Following the fourth and final session, 1 participant indicated that she preferred the speech recognition system, and the others participants stated that they had no preference between the speech recognition system and the keyboard. Educational implications and directions for future research are discussed.

For many individuals frustrated by the mechanics of typing, speech recognition systems may be an attractive alternative to using the traditional computer keyboard for word processing. Speech recognition systems allow individuals to "write" using a word processing program by speaking into a microphone attached to a computer, bypassing the need for a traditional keyboard. The average adult can write with a pen at a rate of approximately 20–29 words per minute with a standard deviation of 3.3 (Birren & Botwinick, 1951), and can type at a slightly higher rate; the average speaking rate for an adult without disabilities is approximately 125–160 words per minute (De La Paz, 1999). Some have suggested (e.g., De La Paz, 1999) that speech recognition systems may allow individuals to compose at rates between their typing and speaking rates.

Speech recognition systems might be especially useful for individuals with learning disabilities. The written compositions of individuals with learning disabilities are typically shorter, less coherent, less organized, and contain more errors in spelling, capitalization, and punctuation compared to the written compositions of individuals without learning disabilities (MacArthur, 2000; Wetzel, 1996). Difficulties with the lower-level skills of transcribing language on to paper may negatively affect higher-order writing skills (e.g., planning and generating content), as college-level students with learning disabilities face significant challenges in spelling and proofreading activities (McNaughton, Hughes, & Clark, 1997). If lower-level skills require significant cognitive effort, working memory required for higher-order writing skills may be depleted (De La Paz, 1999; MacArthur, 2000). Individuals who have difficulty with the lower-level skills of transcribing often become frustrated with the overall process of writing and tend to avoid writing as much as possible (MacArthur, 2000).

One method of alleviating problems associated with transcribing language to paper for individuals with learning disabilities is the use of dictation (i.e., students dictate their compositions to a teacher or to a tape recorder). MacArthur and Graham (1987) demonstrated that students with learning disabilities in the 5th and 6th grades composed longer stories with vocabulary that is more sophisticated and fewer grammatical errors by using dictation compared with handwriting or word processing. Dictation may alleviate some of the mechanical problems that interfere with transcribing ideas to text (MacArthur & Graham, 1987). With the use of dictation, students with learning disabilities also may be less likely to restrict their vocabulary to words that are simple to spell, in an attempt to reduce spelling errors (Hughes, Clark, & McNaughton, 1993).

One alternative to dictation is speech recognition technology. Similar to dictation, speech recognition systems may allow individuals with learning disabilities to bypass problems associated with the mechanics of writing, including the need to spell each word letter-by-letter. Speech recognition systems may offer individuals with learning disabilities potential advantages over dictation, by allowing students to dictate compositions independently, thereby eliminating the need for a teacher or aide to transcribe for the students. Furthermore, speech recognition systems enable students to view and edit the text as it is composed (MacArthur, 2000).

The benefits of speech recognition systems will not be realized unless they are accurate in recognizing the user's speech and efficient in inputting and editing text. Speech recognition technology has improved significantly in the last few years; however, there have been very few research studies investigating the benefits of recent technological changes (MacArthur, 2000). Higgins and Raskind (1995) examined the impact of a speech recognition system on writing with 26 college students with learning disabilities. Each student wrote essays under three conditions: writing without assistance, using dictation with a transcriber, and using a speech recognition system that recognized discrete speech (i.e., single words separated by a slight pause). Papers were scored holistically on a scale of 1–6. The results were mixed: 13 students had higher scores using the speech recognition system compared to writing without assistance, 5 students had higher scores writing without assistance compared to the speech recognition system, and 8 students had the same score for both conditions.

Wetzel (1996) conducted a case study of a 6th-grade student with a learning disability. After fourteen 30-minute sessions over a 10-week period, the student reached an accuracy level of 74% in the composition of short personal narratives with a speech recognition system that recognized discrete speech. Wetzel (1996) concluded that although the promotional literature suggested adults could attain an accuracy of over 90% using discrete speech recognition, Wetzel observed the student experiencing frustration with errors by the system in recognizing his voice and in producing editing mistakes. The low accuracy rate combined with the additional time required for editing text far outweighed the benefits of using the technology for the student in his study. Wetzel (1996) recommended that research into the potential benefits of speech recognition systems continue with students with learning disabilities once improvements were made in the technology.

More recently developed speech recognition systems such as Dragon NaturallySpeaking recognize continuous speech.[1] The recognition of continuous speech allows individuals to dictate full sentences using natural pauses and intonation. The systems are designed to recognize not only the phonemes spoken by the individual at the word level, but also the syntax of the sentence in order to improve the accuracy of the output.

Hux, Rankin-Erickson, Manasse, and Lauritzen (2000) investigated the accuracy of 3 speech recognition systems, including Dragon NaturallySpeaking, with 1 adult with dysarthria and a second adult without a disability. The participants were asked to read preselected sentences and compose novel sentences using the systems in 5 sessions. The 2 adults were more accurate with Dragon NaturallySpeaking than the other 2 systems. The adult with dysarthria achieved levels of accuracy in the range of 64% to 70% for the preselected passages and 54% to 69% for novel sentences. The speaker without dysarthria achieved accuracy levels in the range of 88% to 97% for preselected passages and 84% to 94% for novel passages.

To date, there is a lack of research available to guide teachers in their decisions about introducing speech recognition systems for students with learning disabilities. The current research study was designed to compare the accuracy and rate of transcription of 3 college-aged individuals with learning disabilities using a speech recognition system that recognized continuous speech and a traditional computer keyboard. The 2 methods of transcription were compared across four sessions by measuring the rate of transcription and the accuracy of transcribing passages of college-level texts after editing.

Method

Research Design

A single-subject, alternating-treatments design was used to compare rate and accuracy of transcription across 2 systems: a continuous speech recognition system and a traditional computer keyboard. The treatments were counterbalanced across sessions (e.g., AB in session 1, BA in session 2) to control for possible order effects (Barlow & Hersen, 1984).

A single-subject, alternating-treatments design was selected for several reasons. First, single-subject designs are ideal for evaluating new interventions because data are collected for individual participants across time (Barlow & Hersen, 1984); this allows for observation of learning curves for each participant. Second, an alternating-treatments design enables researchers to study the relative effectiveness of two or more treatments or conditions without constructing a large between-groups comparison with potential difficulties with intersubject variability (Barlow & Hersen, 1984). Finally, unlike other single-subject designs, participants begin the two interventions immediately, without needing to establish a baseline or consistency in one treatment before introducing the second.

Participants

Participants were recruited by contacting the Office for Disability Services at a large northeastern university and by posting a general announcement about the study via e-mail to undergraduates who had registered to receive services because of a learning disability. Three adults volunteered to participate in the research study. Each met the following criteria: (a) were between 19 and 25 years of age; (b) had been enrolled in a post-secondary class within the past 24 months; (c) had hearing and vision, with or without correction, within normal limits; (d) had a documented learning disability, specifically in writing; (e) was a native speaker of English; and, (f) had no previous experience with speech recognition systems.

Maddie was a 24-year-old senior undergraduate student pursuing a degree in communication disorders. She reported that she had normal hearing and vision with correction. Maddie was first diagnosed with a learning disability (i.e., specifically in written expression) and attention-deficit hyperactivity disorder in her sophomore year of college, two years before the study. She had reading comprehension scores within normal limits and written expression scores at an 8th-grade level. She had special accommodations for taking tests in her undergraduate classes (i.e., she was given tests outside of the classroom with additional time).

Josie was a 22-year-old woman and was not attending college at the time of the study. She reported that she had normal hearing and vision. She had previously completed one year of college toward a degree in leisure studies at the age of 20. She was first diagnosed with a learning disability in the 6th grade. At that time, she was reported to have reading comprehension scores within normal limits and written expression scores below a 3rd-grade level. She reported having special accommodations in high school for taking tests (i.e., she was given additional time) and was provided with access to learning support services. Josie had not been tested since her original diagnosis at the age of 12.

Jane was the identical twin sister of Josie. Jane was pursuing an undergraduate degree in information technology. She reported that her hearing and vision were within normal limits. She was diagnosed with a learning disability in the 6th grade, at the same time as her sister. At the time of testing, she had reading comprehension scores within normal limits and written expression scores below a 3rd-grade level. She reported having special accommodations in school for taking tests (i.e., she was provided with additional time). She also was given access to learning support services in school. Jane had not been tested after her original diagnosis.

Materials

The continuous speech recognition system used in the study was Dragon NaturallySpeaking, Version 5.0, by Dragon Systems, Inc. When using this software, individuals speak sentences with natural intonation. A traditional computer keyboard with a QWERTY layout was used for the transcription of the passages via the keyboard. The word processing program used for the transcription of the passages was Dragon Pad, a standard word processing program that was part of the Dragon NaturallySpeaking software package. This software was used in both treatment conditions.

The stimuli used in each session were selected based on the recommendations of a learning disability specialist from the university. She identified two general education classes (i.e., introduction to psychology and elementary astronomy) that were typically identified as challenging for students with learning disabilities. The passages used as stimuli were from textbooks used in the identified classes. These college-level texts were representative of the type of material used by the participants in school. Ten passages were randomly selected from each of the textbooks: 10 psychology and 10 astronomy passages. Each of the passages was 100–115 words in length, had a reading level of grades 10–11, and was randomly assigned to a treatment condition, use of a keyboard or speech recognition, within a session. None of the passages was used more than once for each participant. The passages were randomized so that the participants did not have the same passages in the same order. During each session, the participants completed 4 passages (2 novel passages for each treatment condition): 1 astronomy passage and 1 psychology passage for the keyboard condition and 1 astronomy passage and 1 psychology passage for the speech recognition condition.

Measures

The two dependent measures were: (a) accuracy of transcription after editing and (b) rate of transcription in words per minute (i.e., the number of passage words correct/total minutes of transcription).). The accuracy of transcription after editing was calculated by comparing the textbook passage to the transcribed passage. The number of correctly transcribed words and the number of transcribed words that were incorrect or omitted were totaled. The accuracy of transcription was calculated as the percentage of the number of words correct/ number of passage words correct omitted.

A stopwatch was used in the calculation of the rate of transcription. Timing began at the first keystroke or spoken word and ended when the subject said "done" for the keyboard condition, "microphone off" for the speech recognition condition, or when 10 minutes had elapsed. The rate of transcription was calculated in words per minute.

As a social validation measure, participants were asked to provide feedback regarding the two methods of transcription during a structured interview, with the first author, following each session. See Figure 1 for the social validation form used as a guide for the interview. The feedback was used to determine which method of transcription the participants preferred and whether their opinions changed over time; preferences for each participant across sessions were analyzed to determine any patterns.

Reliability of the measures. Inter-rater reliability was calculated for 25% of the data. Four passages were randomly selected from the 16 total passages for each participant. A second researcher viewed videotapes and calculated the accuracy and rate of transcription for each passage. The rates were calculated by watching the videotapes and timing the transcriptions. For each measure, the inter-rater reliability was calculated as a percentage of the number of agreements divided by the total number of agreements and disagreements. The mean reliability was 99.6%, with a range of 98%–100%.

Procedures

Participants followed the instructions for using the speech recognition system that accompanied the Dragon NaturallySpeaking software. According to the manual, minimal time is required to learn to use the system and have accurate recognition of individuals' voices. The instructions describe how to connect the microphone to the computer, position the headset microphone, and train the system to recognize individuals' speech patterns. The manual states that the system requires 3 to 5 minutes to learn speakers' voices, and it suggests that individuals complete a short tutorial at the start of their first experience using the speech recognition system. The manual suggests that if individuals have difficulty with dictation, they periodically update their speech files by reading short passages; no additional instructional procedures were provided. The goal of the study was to investigate the effectiveness of transcribing text using the speech recognition system in comparison to the keyboard, according to the instructions provided with the system; therefore, participants received no additional instruction, and their performance was followed across 4 sessions (2 sessions each week for two weeks). The order of treatments was counterbalanced across sessions to avoid order effects. The time for transcribing each passage was limited to 10 minutes in the event that the participants became frustrated and were unable to finish the passages in a reasonable time. Each session was videotaped to allow for accuracy in the collection of data and reliability.

Keyboard. During each session, participants were instructed to copy the passages into a word processing document using the keyboard and to edit mistakes after completing each sentence.

Speech recognition system. At the first session, participants completed a training task, provided by the speech recognition system to allow the system to "learn" their speech patterns, and a tutorial so participants could learn the editing commands. Participants completed the task in approximately 20 minutes. During each subsequent session, participants were instructed to first complete a shortened training to improve the system's recognition of the participants' speech patterns as they read. Participants were asked to dictate astronomy and psychology passages using the speech recognition system and to edit their mistakes after each sentence. The speech profiles for the participants were saved at the end of each session.

Figure 1. Feedback form for interviewing participants following each session.

Feedback

Subject Number: _________

Date: ___________________ Session: _________________

1. I found the accuracy of the speech recognition system in recognizing my speech:

   Better than the last session ________

   The same as the last session ________

   Worse than the last session ________

Comments:

2. I found the editing process with the speech recognition system:

   Easier than the last session ________

   As difficult as the last session ________

   More difficult than the last session ________

Comments:

3. Comparing writing using the keyboard and the speech recognition system:

   I prefer to use the keyboard ________

   I prefer to use the speech recognition system ________

   I would choose either the keyboard or the speech recognition system ________

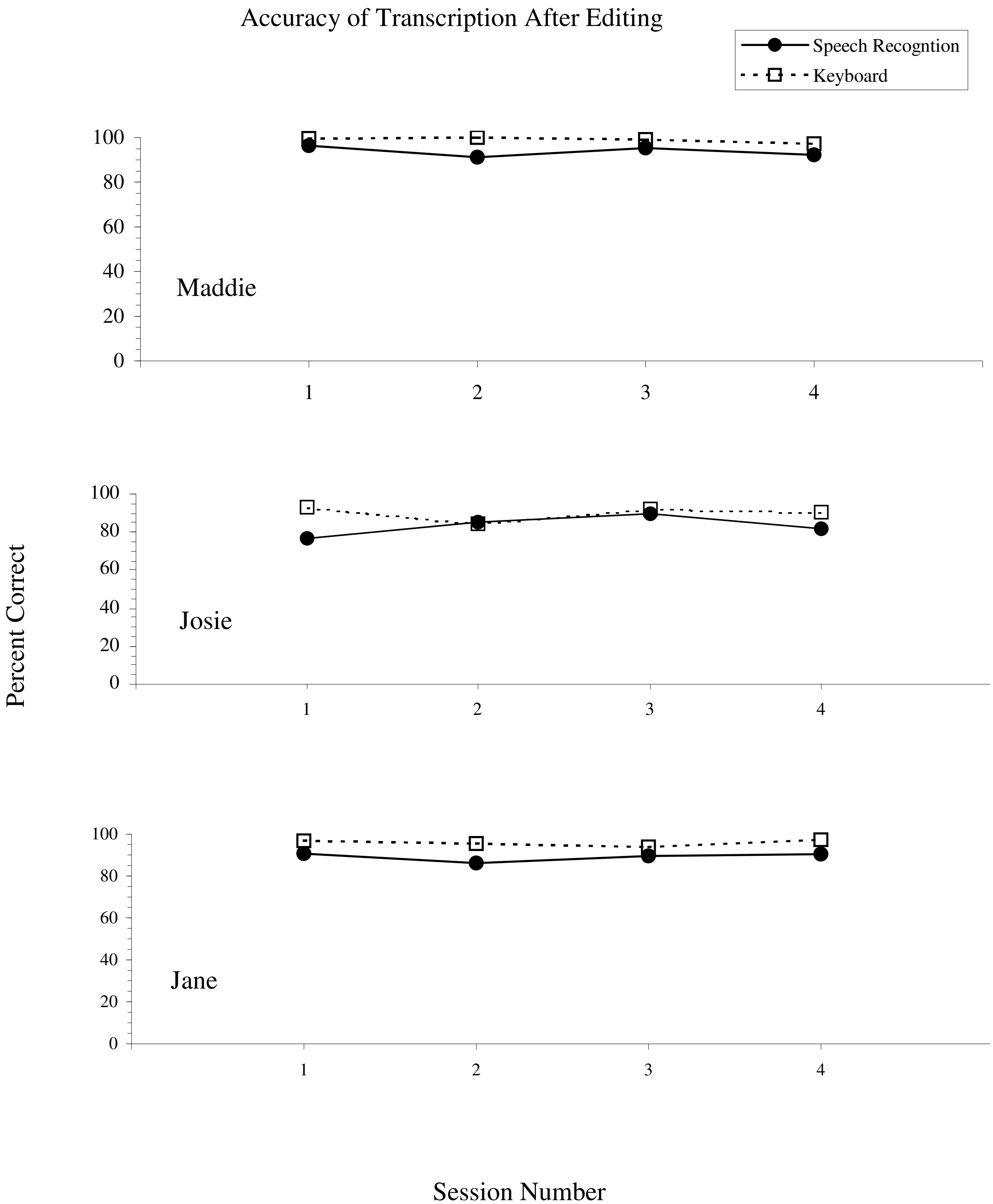
Comments:

Procedural reliability. Procedural reliability was determined for 25% of the data to ensure the integrity of the procedures. A trained researcher then viewed videotapes of the sessions and indicated whether the instructor had correctly followed each of the instructional steps for each condition. Reliability was calculated as a percentage of the number of correct instructional steps divided by the total number of correct, incorrect, or omitted sentences. The mean procedural reliability was 100%.

Data Analysis

Data on rate and accuracy were presented separately in graphic form to facilitate visual inspection of divergence of treatments and changes in level and slope across the two systems, as recommended for alternating treatment designs (Barlow & Hersen, 1984).

Figure 2. Accuracy of transcription after editing for the three participants using the speech recognition system and the traditional keyboard.



Results

Accuracy

Accuracy of transcription was used as a measure of effectiveness. Figure 2 presents the accuracy of transcription after editing for Maddie, Josie, and Jane. In all 4 sessions, Maddie had a higher mean accuracy of transcription using the traditional keyboard of 98.9%, with a range of 97.2%–100% compared to the speech recognition system, 93.8%, with a range of 91.2%–96.4%. Josie had a higher mean accuracy of transcription using the traditional keyboard (90.1%, with a range of 84.1%–93.4%) compared to the speech recognition system (83.4%, with a range of 76.5% - 89.9%). There was a slight overlap in levels of accuracy using the keyboard and speech recognition system in the second session. The mean levels of accuracy were 81.2% for the keyboard and 85.5% for the speech recognition system). In all 4 sessions, Jane had a higher mean accuracy of transcription using the traditional keyboard (95.8%, with a range of 93.7%–97.3%) compared to the speech recognition system (89.2%, with a range of 86.2%–90.7%).

Maddie achieved high levels of accuracy of transcription using the speech recognition system from the start. The mean accuracy was 96.4% for session 1; there was no learning curve across the 4 sessions. Josie demonstrated an increase in the mean accuracy of transcription using the speech recognition system from the first session to the second session (76.5%–85.5%). She had very little change in the levels of accuracy of transcription for the remaining sessions. Jane achieved similar levels of accuracy of transcription using the speech recognition system across all 4 sessions. For sessions 1 – 4, the means were 90.7%, 86.2%, 89.6%, and 90.4%. Therefore, similar to Maddie, there was no learning curve evident for Jane.

Overall, the participants had a higher level of accuracy of transcription using the traditional keyboard compared to the speech recognition system across the 4 sessions; however differences between the conditions were very small, a mean gain of five percentage points in accuracy for Maddie, seven percentage points for Josie, and six percentage points for Jane.

Rate of transcription

Figure 3 presents the data for the rate of transcription (wpm) for Maddie, Josie, and Jane using the speech recognition system and the traditional keyboard. In general, Maddie had a faster rate of transcription using the traditional keyboard— a mean of 20.4 wpm, with a range of 17.3–22.5 wpm compared to the speech recognition system (a mean of 15.9 wpm, with a range of 8.2 – 21.7 wpm). By the fourth session, Maddie had a rate of transcription of 20.7 wpm using the traditional keyboard and a rate of transcription of 21.7 wpm using the speech recognition system.

In the first 3 sessions, Josie had a higher rate of transcription using the traditional keyboard, a mean of 15.6 wpm, with a range of 14.4–17.6 wpm, compared to the speech recognition system, which had mean of 10.2 wpm, with a range of 5.4 – 15.1 wpm. By the fourth session, however, Josie demonstrated an overlap in the rates of transcription in the fourth session (14.2 wpm using the keyboard and 15.8 wpm using the speech recognition system). She demonstrated a learning curve in the rate of transcription using the speech recognition system. The rate consistently increased across the 4 sessions using the speech recognition system.

Jane had a higher rate of transcription using the traditional keyboard compared to the speech recognition system for all 4 sessions. She had a mean rate of transcription of 17.0 wpm using the traditional keyboard, with a range of 16.218.7 wpm, and a mean rate of transcription of 9.7 wpm using the speech recognition system, with a range of 8.9–10.7 wpm. There was very little change in the rate of transcription using the speech recognition system across the 4 sessions for Jane. The mean rate of transcription using the speech recognition system for the first session was 8.9 wpm and 10.7 for the last session.

Social validation

Table 1 presents the results of the participants' feedback following each of the 4 sessions. The participants were asked whether they preferred to use the traditional keyboard, the speech recognition system, or to choose either method for writing. Maddie indicated that she had no preference between the speech recognition system and the traditional keyboard for the first 3 sessions. Following the 4th session, she stated, "I find it easier to use the speech system over the keyboard. It is less time consuming and I have noticed that I am less anxious." Both Josie and Jane initially preferred the keyboard to the speech recognition system. However, by the end of the 4th session, both Josie and Jane indicated that they viewed the traditional keyboard and the speech recognition system as equivalent. Josie stated, "After getting used to the speech recognition system, it became easier to use, about the same as the keyboard." Discussing the speech recognition system, Jane said, "If you worked with it every day, it would be very helpful. The commands are difficult."

Figure 3. Rate of transcription using the speech recognition system and the traditional keyboard for Maddie, Josie, and Jane.

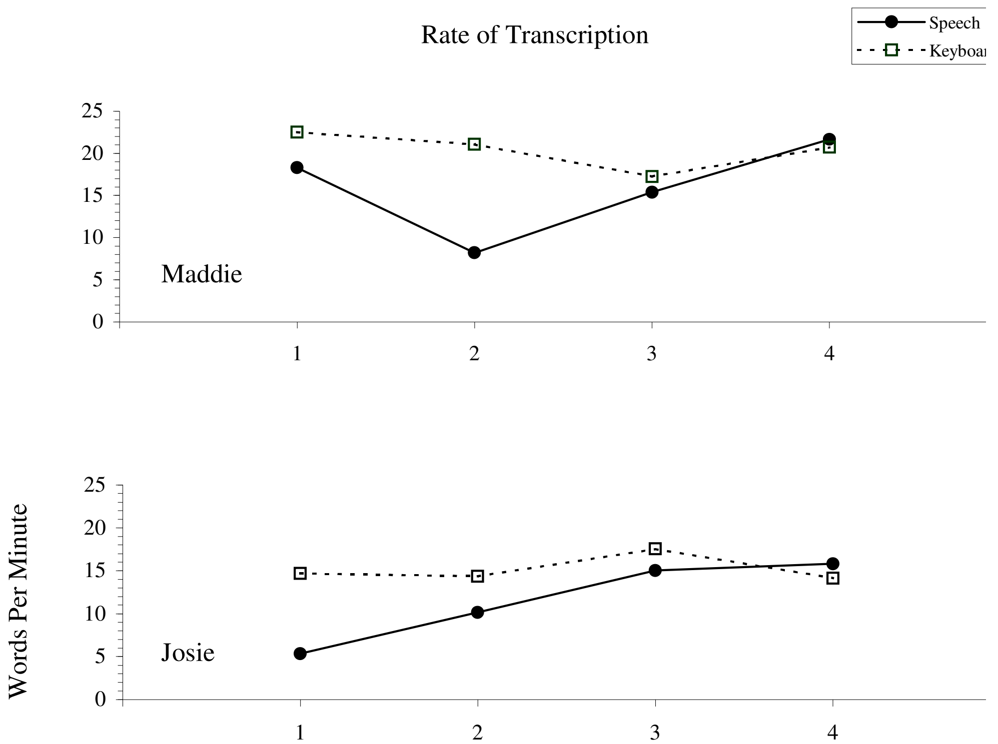
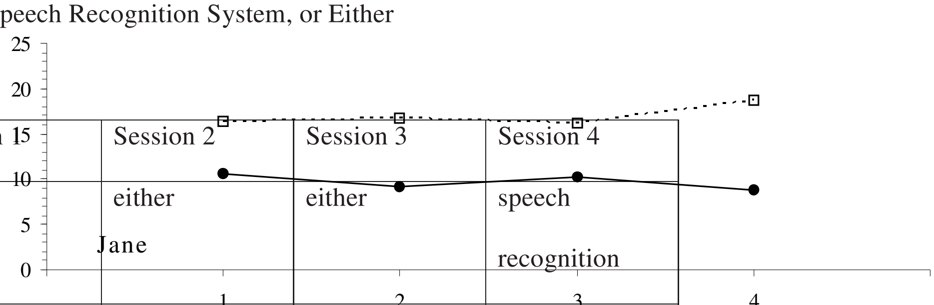


Table 1. Maddie's, Josie's, and Jane's Preferences Following Each Session for the Traditional Keyboard, the Speech Recognition System, or Either



	Session 1	Session 2	Session 3	Session 4
Maddie	either	either	either	speech recognition
Josie	keyboard	keyboard	keyboard	either
Jane	keyboard	keyboard	either	either

Discussion

Like any assistive technology, speech recognition can be evaluated with respect to its effectiveness, efficiency, and acceptability (McNaughton, Hughes, & Clark, 1997; Wetzel, 1996). In this study, effectiveness of the method of transcription was calculated by determining the accuracy of the system in transcribing the passage. Efficiency of transcription was measured using the rate (words per minute) information for each technique. Finally, acceptability was measured by asking participants to identify system preferences (i.e., traditional keyboard and/or a speech recognition system) following each session.

Effectiveness

The results of the present study indicate that the accuracy of transcription using the traditional keyboard was slightly higher for all 3 participants (Maddie [5%], Josie [7%], and Jane [6%]) compared to the speech recognition system for the final session. The accuracy levels may have been slightly better for the traditional keyboard compared to the speech recognition system for several reasons. One possible explanation is that the participants were more familiar with the processes required for transcribing and editing using the keyboard than the speech recognition system, and none had any experience with a speech recognition system before the study. Participants had to learn spoken commands to edit their passages using the speech recognition system.

A second possible explanation for the higher accuracy levels with the keyboard compared to the speech recognition system may be related to differences in the transcription methods using the systems. Using the speech recognition system, participants read entire sentences without pausing. At the end of each dictated sentence, participants were required to go back through the passage and check for incorrect, additional, or omitted words. Although the speech recognition system did not misspell words, the system might not have recognized the words correctly. This may have resulted in the transcription of incorrect words that sound similar to the target words (e.g., computer/comforter). The errors produced using speech recognition systems may be especially difficult to detect because they are correctly spelled words. Participants tended to overlook incorrect words that were visually similar to the target words (e.g., performance/permanence) and words that were omitted. The failure to correct mistakes during editing with the speech recognition system negatively affected accuracy levels.

Overall, participants experienced problems in producing texts with low levels of spelling errors. Error rates for Maddie, Josie, and Jane for speech recognition (6.2%, 16.6%, and 10.8%) and for keyboarding (1.1%, 9.9%, and 4.2%) were, except for Maddie in the keyboard condition, far above those of college students without disabilities. Students without learning disabilities typically have an error rate of 1.1% (McNaughton et al., 1997); ideally, error rates should be within two standard deviations of the performance of students without learning disabilities, that is, lower than 2.9% error rate (McNaughton et al., 1997).

Efficiency

Another important factor to consider in the comparison of the speech recognition system and the traditional keyboard was the rate of transcription. Wetzel (1996) suggested that low accuracy rate combined with the time required for editing text outweighed the benefits of using the speech recognition system at the time of his study. The results of the present study indicated that by the 4th session, two participants transcribed passages from college-level texts at only a slightly higher rate using the speech recognition system compared to the traditional keyboard. Maddie's and Josie's rates of transcription using the speech recognition system during the final session were 21.7 and 15.8 wpm, respectively, and 20.7 and 14.2 wpm using the keyboard. In the 4th session, Jane transcribed at a rate of 8.9 wpm using the speech recognition system and 18.7 wpm using the keyboard.

One of the suggested advantages of speech recognition systems is that individuals can "write" at rates between their typing and speaking rates, however such high rates of text entry were not seen in this study. There are several possible explanations why participants' rates of transcription using the speech recognition system were not significantly higher compared to the keyboard. Participants' familiarity with a traditional keyboard may be one explanation for its comparative advantage over the speech recognition system. All participants were observed to use appropriate keyboarding skills (i.e., they all used 10 fingers to type) but they were not especially rapid typists (Birren & Botwinick, 1951). A second possible explanation is that the time spent rereading and editing mistakes using the speech recognition system may have resulted in a slower rate of transcription than was expected. Participants may have had difficulty recognizing correctly spelled words that were inappropriate in context as errors. Also, participants had to learn spoken commands for editing that were unique to the speech recognition system. The lack of familiarity with the editing commands may have resulted in an increase in the time required for transcription of the passages. Finally, participants were involved in only 4 sessions. For some individuals, increasing the number of sessions may improve familiarity with the editing commands and increase the rate of transcription. A clear increase in rate over time was observed for Josie, a small increase was observed for Maddie, and no change was observed for Jane.

Acceptability

The speech recognition system was not more accurate or significantly faster to use compared to the traditional keyboard; however, 1 participant indicated that she would rather use the speech recognition system rather than the keyboard. The other 2 participants indicated that they viewed the two methods as equivalent for writing by the end of the 4th session. Although the rate of transcription using the speech recognition system was not significantly higher than the rate of transcription using the keyboard, the speech recognition system may be less frustrating than using the traditional keyboard for some individuals with learning disabilities. In this study, Maddie commented that she was "less anxious" when using the speech recognition system. Some individuals with learning disabilities may choose to use a speech recognition system for writing because it reduces the need to spell individual words and allows them to give more attention to the content of their writing.

In the present study, participants transcribed college-level texts and were not required to compose novel passages. It is not clear if their preferences would change when adding higher-order writing demands (e.g., planning, generating content). MacArthur and Graham (1987) proposed that individuals with learning disabilities may write longer compositions using a speech recognition system if they find the task of writing less burdensome.

Clinical Implications and Directions for Future Research

This study was an initial investigation of factors associated with the use of speech recognition systems as a writing method for postsecondary students with learning disabilities. Given the small number of subjects and the nature of the writing tasks, caution must be exercised in the interpretation of the findings. The participation of students with a different type of disability, the use of a different writing task, and the provision of a longer period for training might lead to different results. However, this preliminary investigation does provide some initial information for individuals interested in the use of speech recognition technology by individuals with learning disabilities.

First, based on the results of this study, the effectiveness and efficiency of speech recognition technology will vary by individual. Josie clearly showed improvement in rate over the 4 sessions, however Jane did not. While it is unclear if additional training sessions would have improved the performance of Jane, it is important not to underestimate the amount of time needed to gain competence in a speech recognition system before making a decision about its personal usefulness. Based on this study, it is recommended that individuals considering the use of speech recognition technology participate in at least 4 training sessions before making a decision about its potential usefulness. Also, the small number of subjects and the limited documentation of their disability in writing restrict discussion of the generalizability of the results of the study. Future research is needed with a wider range of individuals with learning disabilities.

Second, current speech recognition technology may not be enough to close the gap with the performance of individuals without disabilities. Even with speech recognition technology, the participants in this study produced text with error rates far above those of students without disabilities. Additional training in proofreading strategies may be needed to support effective use of this technology (McNaughton, Hughes, & Ofiesh, 1997). Also, it is important to note that the results of this study of transcription may not generalize to tasks requiring the composition of novel texts by individuals with learning disabilities. However, the objective of this study was to provide preliminary data on the accuracy and speed using the speech recognition system with adults with learning disabilities, without the interference of the additional demands required in composing text. Future studies should investigate the use of the speech recognition system in novel writing.

Third, factors other than accuracy and speed appear to affect the decision about system use. For example, although Jane was never faster with speech recognition technology, she identified speech recognition as comparable in acceptability to the standard keyboard following sessions 3 and 4. While it is unclear what other factors may affect this decision (e.g., ease of use, reduced spelling demands), the decision about the technology to be used should be based both on an individual's observed performance and their stated preferences.

Finally, this is a rapidly changing field. As speech recognition technology becomes more sophisticated, clinicians and individuals with disabilities should revisit this technology to see if software and hardware improvements make the technology more effective, efficient, and more acceptable to the individual with a disability.

Based on the limitations of the present study, future studies should include individuals with a wider range of learning disabilities to enhance the generalizability of the results to a wider population. Future research should also investigate the use of authentic writing tasks to determine the effects of using the speech recognition system in tasks that require higher-level writing skills (e.g., planning, developing ideas). Finally, studies should investigate learning over a larger number of sessions to investigate possible learning curves in the accuracy and rate of transcription or writing. While this technology continues to show promise, additional research is needed to determine how speech recognition technology can best assist individuals with learning disabilities.

Endnote

[1]Dragon NaturallySpeaking®, Version 5.0 is a continuous speech recognition system by Dragon Systems, Inc., a Lernout and Hauspie Company, 320 Nevada Street, Newton, MA 02460. Telephone: 1-617-965-5200. E-mail: info@dragonsys.com. Web site: http://www.dragonsystems.com

Author Note

The authors express their sincere appreciation to the adults who generously volunteered their time to participate in the study. The authors would also like to thank Kathryn Drager and Carol Hammer for their contributions.

References

Barlow, D., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change*. Allyn and Bacon: Boston, MA.

Birren, J., & Botwinick, J. (1951). The relation of writing speed to age and to the senile psychoses. *Journal of Consulting Psychology, 15*, 243–249.

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

Higgins, E., & Raskind, M. (1995). Compensatory effectiveness of speech recognition on the written composition performance of postsecondary students with learning disabilities. *Learning Disability Quarterly, 18*, 159–174.

Hughes, C., Clark, K., & McNaughton, D. (1993). *An analysis of the spelling errors produced by college students with learning disabilities*. Paper presented at the annual meeting of the Learning Disabilities Association, San Francisco.

Hux, K., Rankin-Erickson, J., Manasse, N., & Lauritzen, E. (2000). Accuracy of three speech recognition systems: Case study of dysarthric speech. *Augmentative and Alternative Communication, 16*, 186–196.

MacArthur, C. A. (2000). New tools for writing: Assistive technology for students with writing difficulties. *Topics in Language Disorders, 20*, 85–100.

MacArthur, C. A., & 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.

McNaughton, D., Hughes, C., & Clark, K. (1997). The effect of five proofreading conditions on the spelling performance of college students with learning disabilities. *Journal of Learning Disabilities, 30*, 643–651.

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

Wetzel, K. (1996). Speech-recognizing computers: A written-communication tool for students with learning disabilities? *Journal of Learning Disabilities, 29*, 371–380.