COMPOSING VIA DICTATION AND SPEECH RECOGNITION SYSTEMS: COMPENSATORY TECHNOLOGY FOR STUDENTS WITH LEARNING DISABILITIES

Susan De La Paz

Abstract. This article provides a rationale for using an oral mode of production as a means of composing. For individuals with learning disabilities (LD), composing orally may allow them to circumvent transcription or text production problems (e.g., handwriting, spelling, punctuation), which in turn may allow greater focus on higher-order concerns such as planning and content generation. Support for this position comes from research on the use of dictation as well as studies of both simulated and existing speech recognition systems involving individuals with LD. Current characteristics of speech recognition systems are described, with an emphasis on what these applications require of teachers and students. Finally, recommendations are offered for teachers who plan to use dictation or speech recognition with individuals with LD.

SUSAN DE LA PAZ, Ph.D., is assistant professor, Vanderbilt University.

People compose by several methods. These include writing by hand, typing either on a typewriter or with a word processor, and dictating to another person or a machine (De La Paz & Graham, 1995). The decision to use any of these methods may, in part, be based on the writer’s ease and speed in composing. According to at least one estimate, most literate adults are capable of writing 15 to 25 words per minute by pen and typing at a slightly faster rate (Wetzel, 1991). In contrast, we speak at a rate of 125 to 160 words per minute. Although existing speech-to-text transcription products such as Dragon NaturallySpeaking (Dragon Systems, Inc., 1997) do not take dictation at this speed, they are able to transcribe continuous speech with relatively high accuracy, approximately four to six times faster than adults typically write. As this technology improves further, it has the potential to revolutionize the composing process, and it may fulfill for many the belief that composing orally “permits thinking to unfold in a natural and unimpeded way... [and] that we are more likely to capture fleeting thoughts and sequences of thoughts than if we had been writing by hand or a typewriter—thus avoiding the uncomfortable feeling that some brilliant idea was irretrievably lost while we were typing” (Gardner, 1980, p. 16).

While advances in oral composition may radically change how people compose and how writing is taught in the near future, it also has a long and distinguished history, ranging from Tiro taking shorthand for Cicero (Halpern & Liggett, 1984), Henry James dictating notes for his writing (Plimpton, 1989), and Alan Gardner composing articles and essays using a tape recorder (Gardner, 1980). In the workplace, dictation has been used for many decades (Buckley, 1940) as letters, records, and reports were transcribed by a stenographer who then used a typewriter (or, in recent years, a word processor) to generate a written copy, providing the writer with a neatly displayed product that was easily reviewed or modified.
By 1972, the first dictation and word processing systems were combined into discrete speech recognition (SR) systems, requiring the writer to speak one word at a time (Lange, 1993; Meisel, 1993). Audio signals were processed by hardware and software in the computer and compared with templates (acoustic word models) already resident in the system's active vocabulary (Milheim, 1993; Williams, 1990). Statistical language models evolved further by using complex matching algorithms in which linguistic and phonetic information was added to the matching process, simultaneously allowing the system's active vocabulary to be updated. Current technology allows SR systems to accept continuous speech by converting the writer's message into digitized signals, which are then transformed by the computer into word output (Milheim, 1993). For the past 25 years, researchers struggled to solve the problem of developing computer programs that could identify acoustic forms, while not comprehending the message that was decoded (Gasser, 1994). This goal has been at least partially realized by using probabilistic models that deal with uncertain or incomplete information, and by applying neural networks that allow systems to "learn" to recognize speech (Lange, 1993).

Despite the impressive gains made in recent years, current continuous SR systems are not capable of truly natural human-computer speech interaction such as the interactive computer HAL envisioned in the movie 2001: A Space Odyssey, 30 years ago (Lange, 1993). Current systems merely translate an utterance or execute oral commands (e.g., "new line" or "caps off"); they do not understand speech and have limited capabilities for interpreting a speaker's message (Milheim, 1993). In anticipation of improvements in SR systems, a small but growing group of researchers has conducted research during the past 10 years to determine how this technology might best be used with persons with learning and writing problems. This article reviews this research and provides suggestions for using dictation and SR systems with individuals with LD as a means for circumventing transcription or text production problems.

First, I examine the rationale for using an oral mode of production when composing. The second section of this article summarizes research involving dictation and students with LD. Third, experiments investigating SR systems using discrete speech as well as studies that simulate SR systems (to provide better speech/word processing capabilities than are currently available) are reviewed. Fourth, I examine the characteristics of current SR systems (and the use of a related application, speech synthesis) with an emphasis on what these applications require of teachers and students. Finally, recommendations are proposed for teachers who plan to use SR technology for individuals with LD.

**Why Compose Orally?**

While some individuals may be hesitant to use an oral mode of production because it is different from the way they are used to composing (Halpern & Liggett, 1984), dictation has been successfully used by many individuals from a variety of occupations. Howard Gardner, noted for his theory of multiple intelligences (Gardner, 1983), provided an eloquent description of the potential advantages of composing by dictation rather than writing by hand or typing. Based on his own experiences, Gardner observed that dictation can not only increase an author's output, it can also improve the quality of what is produced by compelling the author to plan in advance. Other professional writers, such as novelist Sidney Sheldon and TV personality Steve Allen, have been equally positive about the benefits of dictation (Haggblade, 1990). They indicated that dictation allowed them to increase their productivity, making it possible for them to work on several projects simultaneously. They also echoed Howard Gardner's sentiments that dictation permitted them to compose at rates closer to their speed of thought, thereby allowing them to capture ideas before they were forgotten.

This approach to composing may be especially advantageous for individuals with LD, as it allows them to focus on high-level concerns, such as planning and content generation, rather than on the mechanics of writing. This may provide a powerful benefit, since both children and adults with LD often have difficulty with the mechanical demands for producing text. For example, researchers have consistently found that students with LD make considerably more spelling, capitalization, and punctuation errors than their normally achieving classmates (MacArthur, Graham, & Schwartz, 1991), and their handwriting is both quite slow and legible (Graham & Weintraub, 1996; MacArthur, this issue).

**Text production difficulties.** The mechanical requirements for producing text can interfere with writing in several important ways (Graham, 1990; Scardamalia & Bereiter, 1986). First, having to attend to lower-level skills of getting language onto paper may interfere with higher-order skills such as planning and content generation. To illustrate, when stopping to consider how to spell a word, a writer may forget related ideas or experience disruption in generating additional ideas. According to Scardamalia, Bereiter, and Goelman (1982), having to switch attention to mechanical demands may lead the writer to forget already developed intentions and meanings.
The mechanical requirements of text production may also interfere with an individual’s overall rate of production. Thus, writers often forget ideas because they cannot write their thoughts down fast enough (De La Paz & Graham, 1995). Slow writing, therefore, may further interfere with content generation and remembering ideas or text already planned and held in working memory. A study by Graham (1990) also indicated that children with LD have difficulty with text production skills, as the mechanics of writing interfere with both the quantity and quality of writing produced by fourth- and sixth-grade students with LD.

Difficulties with the mechanics of writing may also undermine a writer’s motivation and persistence during composing. It has been my observation, in talking with students with LD, that text production difficulties can negatively influence writers’ attitudes toward composing and the degree to which they attempt writing tasks in school. In addition, writers may restrict their vocabulary or syntax to avoid words they cannot spell or complex sentences that might become confused during text production (De La Paz & Graham, 1995). Further, if a study by Gerber and his colleagues (1990) is representative, the mechanical difficulties encountered by children with LD persist into adulthood. Their data indicated that 45% of the surveyed adults with LD believed their spelling and writing difficulties continued throughout their adult lives and another 40% believed these difficulties got worse over time.

Summary. Text production difficulties can interfere with composing in several important ways. Attending to spelling and handwriting skills or having a slow rate of production can hinder a writer’s ability to generate ideas during planning. In addition, difficulties with the mechanics of writing can have adverse effects on a writer’s persistence and motivation during composing. Dictation appears to have the potential for increasing how much and how quickly writers produce text. For persons who have difficulties with the mechanics of text production, dictation may allow them to bypass their low-level difficulties, and it may further affect the quality of what is produced. In order to examine the advantages and disadvantages of using dictation with persons with LD more directly, its applications with both children and adults are now reviewed.

Is Dictation Effective for Persons with LD?

Currently, there is not much known about the effects of dictation for individuals with LD. Five studies were located for this review, including fourth- and seventh-grade students with LD (De La Paz & Graham, 1997a; Graham, 1990; MacArthur & Graham, 1987; Reece, 1992, Experiment 3) and postsecondary students (Higgins & Raskind, 1995). In each of these studies, the effects of dictation were determined by comparing essays composed via dictation to essays written using alternative modes of production.

The first of these investigations compared three different methods of text production: dictation, handwriting, and word processing (MacArthur & Graham, 1987). These investigators asked fifth and sixth graders with LD to compose stories in response to pictures. Each of the participating students composed one story using each mode of composing (order of pictures and mode of composing were counterbalanced). While there were no discernable differences in the stories written with paper and pencil or on the word processor, students’ dictated stories were longer, contained fewer grammatical errors, and were of higher quality than stories produced in either of the other two composing conditions. Two factors, slow rate of production and difficulties with mechanics, appeared to account for these differences. Having to attend to the lower-level skills of text production may have caused these writers to forget plans already made or disrupted the processes involved in planning the next unit of text.

In a follow-up study to the MacArthur and Graham (1987) investigation, Graham (1990) adapted experimental procedures devised by Scardamalia, Bereiter, and Goelman (1982) to separate the effects of rate from mechanical interference on the writing of students with LD. In addition to writing and dictation, students composed a paper under a third condition entitled “slow dictation.” During slow dictation, the composition was dictated to an examiner who transcribed it at the same speed that an earlier paper was written. By comparing students’ performance during dictation to slow dictation, the effects of mechanics were removed, isolating the effects due to rate. Correspondingly, the effects of rate were controlled by comparing handwriting and slow dictation, isolating the effects due to mechanics.

In contrast to the earlier MacArthur and Graham (1987) investigation, the composition task used in the follow-up study involved writing an opinion essay. Each student produced three compositions (one each during normal dictation, slow dictation, and handwriting). Results from this study indicated that the fourth- and sixth-grade students’ normally dictated essays were of higher quality than their handwritten essays and were generated much more rapidly (seven times faster for fourth graders and five times faster for sixth graders). Contrary to expectations, there were no differences in the length of the two sets of essays. It was hypothesized that differences between the two studies regarding length may be a direct consequence of differences in the genres studied. That is, generating expository content may be more difficult than producing narrative material for students with LD.
Most importantly, however, the mechanical demands of writing were disruptive for the students participating in this study. In comparison to their written text, essays produced during slow dictation were longer and of higher quality and, at least for the sixth graders, slow dictation resulted in more coherent compositions. In contrast, a slower rate of production did not interfere with the writing performance of the participating students. There was no significant difference in the quality of essays produced during normal and slow dictation, and students actually produced more text when composing during the slow dictation condition. While a slower rate of production may have resulted in some forgetting of intentions and plans, these effects were either minimal or offset by the increased generation of content during slow dictation.

Reece (1992, Experiment 3) compared dictation to a tape recorder and handwriting with fifth and sixth graders who were poor writers (students with LD, low general intelligence, or poor motor coordination). Students generated three compositions (persuasive, personal narrative, and expository) using each mode of production. While there were no significant differences in holistic scores, dictated essays received higher mean ratings than handwritten texts. In contrast, dictated essays were significantly longer than those written by hand, and students spent much less time dictating their essays compared with writing by hand. These findings were in contrast to results from a similar study with average writers (Reece 1992, Experiment 2), in which the benefits of dictation were outweighed by difficulties associated with the writers’ inability to see text as it was produced. Apparently, for poor writers, the constraining effects of mechanics were more problematic than being unable to see one’s text as it was composed.

In a fourth study, Higgins and Raskind (1995) asked postsecondary students with LD at California State University, Northridge, to compose essays by dictating to a human transcriber, and by writing “without assistance” (by hand or via a word processor without spell-checking function). Essays were written in response to a writing task (previous test questions) from the university’s writing proficiency exam, which students were required to pass to graduate. Participants in the study were given instructions identical to those on the actual exam, and compositions were scored by two readers with experience in reading the exams. Readers were blind to the condition within which the essays were composed.

There were no significant differences in holistic quality ratings between dictated essays and those written without assistance. In examining these findings, the authors cited disadvantages in the use of dictation, which may have interfered with its success. First, they hypothesized that students who were dictating did not plan and organize their ideas adequately before beginning to write because they felt that it was not appropriate to keep their transcriber waiting for a long time. Students were also hesitant to ask their transcriber to reread portions of their text, especially when a certain passage needed to be read more than once or twice. Third, some students had trouble reading the transcriber’s handwriting, and their overall reluctance to read as they composed may have reduced their ability to benefit from reading previous portions of their papers.

In 1997 Steve Graham and I also compared dictation and handwriting for students with LD, but we extended our investigation to explore the singular and combined effects of dictation and explicit planning instruction. Based on previous work by Graham (1990) and MacArthur and Graham (1987) with students with LD, we expected that dictation would have a positive effect on students’ essay production. In addition, because the planning strategies employed by students with LD during writing are generally considered less mature than those used by their normally achieving peers (De La Paz & Graham, 1995), we hypothesized that planning instruction would be beneficial for students with LD who used either handwriting or dictation as a mode of production. Most important, we predicted that the effect of dictation would be especially pronounced when combined with instruction in planning.

In our study, fifth-, sixth-, and seventh-grade students with LD were randomly assigned to four instructional conditions: (a) planning and dictation, (b) planning and handwriting, (c) essay structure and dictation, and (d) essay structure and handwriting. Students in the two planning conditions were taught a specific strategy for planning opinion essays in advance of writing (see De La Paz & Graham, 1997a, 1997b, for detailed instructional procedures). The advanced planning strategy was taught via the self-regulated strategy development model (Harris & Graham, 1996). Students assigned to the essay structure condition learned about the characteristics of good essays, read and revised sample essays, and composed and shared their own essays with peers. In both conditions, half the students dictated their essays during testing and instruction; the other half wrote their essays by hand.

Consistent with our expectations, students who received instruction in planning and dictation outperformed students in the essay structure condition who wrote by hand. Immediately following instruction, students in the planning and dictation condition produced essays that were longer, more complete, more cohesive, and qualitatively better than those generated by students in the essay structure condition.
maintenACE probe administered two weeks later, their essays were again more complete and qualitatively better. contrary to predictions, however, instruction in planning generally did not benefit dictation more than writing. further, neither dictation nor planning instruction in isolation produced the same level of improvement in the children’s essays as the combination of planning and dictation, suggesting that both factors were responsible for differences between the planning condition and the essay structure condition.

In summary, the results from these five studies were inconsistent. MacArthur and Graham (1987) and Graham (1990) both found that dictation had a beneficial effect on the quality of the compositions produced by students with LD; however, an increase in written output was found only in MacArthur and Graham’s 1987 investigation and in Reece (1992, experiment 3). Dictation was effective in increasing the rate at which students with LD composed in four of the five studies reviewed (De La Paz & Graham, 1997a; Graham, 1990; MacArthur & Graham, 1987; Reece, 1992). These findings may underestimate the effects of dictation for postsecondary students with LD, as Higgins and Raskind’s (1995) study did not assess outcomes for important variables such as rate and length.

Speech Recognition Systems

One potential explanation for why dictation has not consistently improved the quality of students’ papers in these studies is that writers were not typically able to see their text as they composed. This limitation may have obscured benefits gained from a reduction in text production difficulty. Results from Reece’s (1992) and De La Paz and Graham’s (1997a) investigations indicate that planning a paper prior to composing may counter the negative effects of dictating without seeing text as it is produced. The critical nature of advanced planning for the process of dictation is also supported by investigations involving both expert and novice dictators. Interviews and observations with expert dictators, workers who routinely used dictation, showed that they typically planned what to say in advance, jotting down ideas, sketching outlines, or making mental notes (Gould, 1978; Halpern & Liggett, 1984). As Halpern and Liggett (1984) noted, the most compelling feature of their study of expert dictators was “the prominent position of advanced planning” (p. 29). Correspondingly, research with novices, elementary and college-age students with little prior dictating experience, demonstrated that instruction in advanced planning improved the quality of written and dictated text, but that dictation benefited more than writing from such instruction (Reece, 1992, Experiments 1, 4, and 5).

Another solution may be the use of discrete or continuous SR systems. With SR systems, students can read the emerging text while composing. In addition, composing to a machine may be beneficial for students with LD since human transcribers tend to prompt students rather than wait for them to express their ideas. I was only able to locate three studies that examined the use of SR systems (simulated or software) as a means of composing a paper with students who were poor writers or LD (Higgins & Raskind, 1995; Reece, 1992; Wetzel, 1996).

A study by Reece (1992, Experiment 3) compared the use of a simulated continuous SR system to dictation to a tape recorder and writing by hand. Students composed three compositions using each mode of production. To accomplish the simulation, students dictated into a microphone that was attached to a computer monitor while a typist sat unseen behind a large cardboard screen. The typist attempted to transcribe the dictated material as quickly as it was spoken, enabling students to watch the emerging text. Results indicated that texts produced using the simulated SR system were qualitatively better than handwritten texts. Similarly, the simulated SR system allowed poor writers to produce significantly more mature text features (a measure based on rhetorical and grammatical structure) than handwritten texts. Surprisingly, the study found there was no statistically significant difference in the overall quality or maturity of text features of texts produced using the two speech-based production methods. One possible reason why no difference was found in quality or text maturity between the simulated SR system and dictation may be a consequence of how students composed. In comparison to average writers (Reece, 1992, Experiment 2), the poor writers in the present study showed less planning when writing. Thus, these students may have used an approach to composing that minimized the need for them to see text as it was composed.

Higgins and Raskind (1995) also compared a SR system to two other modes of production by evaluating essays composed by postsecondary students with LD. Students wrote essays on established writing proficiency exam topics. Each participating student used a discrete SR system, dictated an essay to a human transcriber, and wrote “without assistance” (by hand or via a word processor without spell-checking function).

The SR system used in this investigation (Dragon Dictate) required individuals to dictate one word at a time. In addition, students learned to use both voice commands and the keyboard to correct inaccurate “guesses” made by the program. On average, students spent close to six hours learning to use the SR system. An important component of the training procedures.
required each individual to continue using the equipment until the system could successfully recognize 75-80% of the speaker's words. Although many students initially required prompting from an examiner not to allow the system to pass over wrong word choices, all were able to correct 90% of the system's errors by the end of the training period.

Essays produced with SR were of significantly higher quality than those written by hand. However, there were no discernable differences between essays dictated to human transcribers and those created with the discrete SR system. In an effort to learn more about the two speech-based modes of production, the authors compared the participating students' compositions to those previously written by students in the university's general population. Results indicated that when students with LD used the SR system, their essays received “passing” scores as often as those of students in the general population. This was not the case when students with LD wrote by hand or dictated.

Finally, when syntactic complexity and fluency measures were correlated with holistic scores, the number of words containing at least seven letters accounted for a significant portion of the quality ratings. One possible explanation for this outcome was that the SR allowed students the freedom to use their more extensively developed oral vocabularies since they did not have to check the spelling of words. It also appeared that the SR condition may have encouraged long words because the program was better at making correct guesses for longer words than for short, single-syllable ones.

Wetzel (1996) explored the use of a discrete SR system in a case study involving a sixth-grade male with LD. In this investigation, the child learned to use a discrete SR system over a period of 10 weeks. Although anecdotal, the observations reported in this case study may be useful for individuals considering this technology, especially because the limitations cited in the study remain inherent to existing discrete SR systems. In addition to requiring discrete speech, the system required each user to use an error-correction procedure that was often cumbersome to follow. As soon as the system “heard” a word, it placed its best guess on the screen with a box appearing below giving the selected word and alternatives. If the selected word was correct, the user continued by dictating the next word. If the selected word was in error, the user selected the correct item from a menu of choices, or if it did not appear on the screen, the writer had to type it in. The target student used this system for a total of 14 sessions, and at the end of this period the system reached recognition levels of up to 74%, which was far below the 90% average accuracy level reported in the promotional literature for adults.

Some of the problems encountered by this student were that incidental sounds (such as when breathing deeply or coughing) were recognized by the system as extra nonsensical words. In addition, due to his poor spelling skills, the student was unable to correct many errors that required his typing them into the system. Apparently, his problems were compounded when he began sounding out words while the microphone was on, which resulted in additional unwanted words appearing on the screen. Understandably, the child became frustrated at times, and more importantly, these problems slowed the correction procedure and interfered with his dictation. Interestingly, the system requirements proved so complex that the researcher decided to prompt the child to plan before he began composing. This was necessary because the student was required to tell commands, such as “Capitalize the next word,” as he composed. To overcome memory limitations and speak fluently, the student was taught to first think aloud about the composition he planned to tell, and then to write down key words on a separate paper to serve as a planning guide for recalling both the content and the order of ideas as he dictated to the computer.

In summary, the results from these investigations provide modest support for the use of SR systems for persons with LD. In both empirical studies involving the use of SR as a mode of production, students’ compositions were of higher quality than those written by hand (Higgins & Raskind, 1995; Reece, 1992, Experiment 3). In contrast, and for reasons not readily apparent, dictation to a simulated continuous or existing discrete SR system did not significantly improve compositions when compared to dictating to a tape recorder or transcriber. At the time this article was written, it remains an empirical question whether newly developed continuous SR systems (such as Dragon NaturallySpeaking) offer advantages not seen to date in studies involving discrete SR systems.

Results from several studies involving both dictation and SR systems (De La Paz & Graham, 1997a; Reece, 1992; Wetzel, 1996) underscore the importance of planning before composing when writers use dictation. For students in these studies, overall writing skills improved only when planning (or instruction in planning) was combined with composing orally. Finally, it will be obvious to most readers that writing instruction should focus on improving the text production skills of individuals with LD, in addition to finding ways to circumvent these problems. Individuals in each of the reviewed studies struggled with text production difficulties such as handwriting, spelling, punctuation, and rate of production. Compensatory strategies such as the use of dictation or SR systems should be best viewed as part of a multifaceted writing program.
Characteristics of Speech Input Technology

When using speech recognition systems, writers typically speak into a microphone, and the computer converts the acoustic waveform of the vocalization into a digitized version (Cavalier & Ferretti, 1996). The computer then extracts spectral data from the digitized signals and compares them to templates representing vocabulary items that have previously been entered and stored in the system. Recognition occurs when the analysis turns up a match (Cavalier & Ferretti, 1996). To provide readers with a more thorough description of how SR systems work, this section describes characteristics of most SR software and the cognitive demands that SR systems impose on writers. I will then briefly describe how a related word processing tool (speech synthesis) may be combined with SR technology to maximize its usefulness.

While specific speech recognition systems can vary significantly, most may be described along at least five dimensions, including training requirements, the ability to handle continuous speech, size of the system’s active vocabulary, error-correction procedures and overall accuracy (Cavalier & Ferretti, 1996; Milheim, 1993). Distinctions between these features are somewhat arbitrary, however, as they overlap with each other (e.g., accuracy and training or vocabulary size).

Training requirements. Speech recognition systems differ by the degree to which they require an enrollment procedure before using the program. Speaker-dependent systems require every user to build a recognition template in the system. This training can require up to several hours to complete, although the processes used differ according to the system. For some systems, a given lexicon has to be repeated several times, word after word, whereas others end the training as soon as the recognition rate is satisfactory, and still others adapt themselves progressively to each speaker during the first hours or days of use (Gasser, 1994). In contrast, speaker-independent systems use previously produced templates provided by system manufacturers, which are generated from a statistical sample of potential users (Milheim, 1993). These systems are especially important in situations in which a large number of people need to interact with the same system (Cavalier & Ferretti, 1996). Finally, some SR products (e.g., Naturally Speaking and ViaVoice) work out of the box but their performance improves with training.

Speech continuity. Until recently, most systems used isolated word recognition, which required the user to segment speech into discrete units, pausing briefly between individual words or utterances. Affordable continuous speech recognition systems are now available, allowing users to speak continuously and more naturally, although they also are not completely natural because accuracy is dependent on the user having consistent patterns of pronunciation. Each system requires the user to learn and use a series of commands for routine use, such as saying, “Delete that” followed by a word or phrase that the speaker wants to cut, or “Cap word” to capitalize the first letter of a proper noun in the middle of a sentence. Dragon NaturallySpeaking allows users to choose between using speech or a mouse to format text, and to use commands to control the microphone, move around a document, edit and format text, and play back recorded speech.

Vocabulary. Although systems with small vocabularies are still available, active vocabularies can range from 20,000 to 55,000 immediately recognizable words, depending on the system’s cost, and less common words are retrieved from a larger, back-up dictionary. When first starting systems such as Dragon NaturallySpeaking, users receive a standard 30,000-word general vocabulary file containing available words and a language model. Users then customize the active vocabulary and its language model by using a “vocabulary builder,” so it more closely matches the content they dictate. In addition, the more consistently a user dictates vocabulary items, the better the system’s recognition rate.

Error-correction procedures. As noted earlier, error-correction procedures can be cumbersome for users’ to learn and follow. In addition, continuous SR systems work differently from discrete systems. The most apparent difference is that continuous systems do not adapt continually to the user’s speech whereas discrete systems do. The result is that when using discrete SR systems, users must make corrections as errors occur, while continuous systems allow users to make corrections at any time. Recognition failures in discrete systems typically result in the system providing a list of several words for the user to search for the target word. In contrast, many continuous systems allow speakers to repeat an unknown word rather than guess the correct word. Both discrete and continuous systems require users to spell (orally or by keyboard) the correct word when the system remains unable to recognize the target item; however, this process is easier to execute in continuous systems. When using Dragon Naturally Speaking, for example, and the user says, “Spell that” followed by four to six letters of the correct word, a corrections dialog box (incorporating a word prediction feature) also opens, allowing the user to choose the correct text or continue spelling. Finally, once a user pronounces or spells an unknown word once, the system “learns” the item and automatically adds it to the user’s active vocabulary.

Accuracy. An analogy can be made between handwriting and spelling accuracy in texts produced when...
writing by hand and in the overall accuracy of SR systems in taking a user's dictation. That is, users are not likely to use SR systems which fail to reach a high level of accuracy in text translation, just as poor spelling and handwriting interfere with the overall quality of a composition. Multiple factors influence accuracy including the similarity of words, the fact that connected speech involves co-articulation (i.e., overlapping phonemes), variability in a user’s speech (e.g., having a cold, being fatigued, and saying mispronunciations) or in physical environment (e.g., inconsistent placement of the microphone). Using words that are spelled differently from the way they are pronounced also leads to system errors. Although Dragon NaturallySpeaking is advertised as taking dictation with high rates of recognition accuracy in the general population, it remains to be seen whether persons, and particularly students with LD, can achieve the claims made in current promotional literature.

What kinds of difficulties may be anticipated by SR users who have LD? First, it is apparent that SR imposes a substantial cognitive load on individuals wishing to compose using an oral mode of production. At this point, it frees users from worrying about spelling and handwriting, but it imposes new burdens—careful speech, vocabulary building, explicit punctuation, error correction, and playback and editing procedures, not to mention the initial training requirements. Second, as noted by researchers (Higgins & Raskind, 1995; Wetzal, 1996), individuals with LD may need guidance as they initially learn to use speech commands, the keyboard, and specific error correction procedures that are necessary to operate the system. Further, students must realize that incidental sounds (such as breathing deeply or coughing) will be recognized by the system as extra nonsensical words.

Equally important for many individuals with learning problems, re-reading text will be problematic, due to poor decoding skills. Systems with built-in speech synthesis may help provide a solution as this feature allows the user to hear what was written by translating text into speech. This “playback” feature may provide additional, unique advantages to persons with LD who use speech recognition systems. Persons with more proficient oral language skills may be able to use these skills to monitor the adequacy of their writing (Rosegrant, 1988). In addition, the auditory input may reduce the user’s burden of reading while editing (Burenstein, 1995). Students may learn to identify more errors of punctuation, grammar and syntax when listening to previously written passages (Espin & Sindelar, 1988; Raskind & Higgins, 1995) which in turn may result in more substantive revisions to their papers (MacArthur, 1996). Speech synthesis can also be used to allow individuals to check their own spelling (MacArthur, this issue).

RECOMMENDATIONS

While our understanding of the impact of dictation and SR systems on the composing of persons with LD remains unclear, the latter may allow individuals to transcribe at rates closer to the speed of speech, a development that may result in improved writing for some. Luckily, educators in special education can use the existing knowledge regarding the writing of individuals with LD to inform their efforts to accomplish this objective.

First, converging evidence points to the importance of teaching individuals with LD to plan before composing via an oral mode of production. It appears that neither dictation nor SR is sufficient by itself to offset the difficulties that persons with LD have in composing (De La Paz & Graham, 1997a; Reece, 1992) and, as is the case for assistive technology in general, these compensatory strategies must be combined with instruction to maximize their effectiveness (MacArthur, 1996). In addition, because planning is central to effective dictation (Halpern & Liggett, 1984; Reece, 1992; Wetzal, 1996) and persons with LD often use immature or ineffective planning strategies (De La Paz & Graham, 1997a, 1997b), planning must be a central element of any dictation program. As with other modes of composing, good writing is dependent upon thoughtful planning and reflection.

Second, individuals with LD who use SR systems should receive scaffolded guidance from special educators as they learn to use this technology. When providing a model of one's unique speech characteristics, the user should build recognition templates in settings where the system will be used, and he or she should attempt to become as consistent as possible in terms of vocabulary selection and pronunciation (Cavalier & Ferretti, 1996). In addition, users must learn speech commands, mouse and keyboard operations, and error-correction procedures that require them to develop efficient strategies for spelling words that are not initially recognized by the SR system. Even though speech synthesis may facilitate this process, the system does not provide the user with help for words that sound alike but are spelled differently (Burenstein, 1995).

Third, this new technology may be used in combination with speech synthesis to teach students more about revision (Burenstein, 1995; Rosegrant, 1988). Speech output may help users detect errors that are less obvious than if reading along silently without this cueing feature. Moreover, because immature writers need to develop the ability to "hear" whether a text sounds appropriate, the oral feedback provided by speech synthesis may be used
to prompt writers to return to a composition and attempt corrections. When something “sounds” wrong, individuals may make revisions and hear how alternative versions sound, judging repeatedly whether the text adequately conveys their ideas. Writers may even be able to use speech output to develop a better sense of an audience's needs (Rosegrant, 1988) as they may begin to recognize omissions, substitutions and confusions that make it difficult for an audience to gain intended meanings.

Fourth, special educators need to continue their efforts to help individuals with LD learn about written language, including conventions ranging from mechanical to stylistic concerns. Neither dictation nor speech recognition systems alleviate the need for persons to compose using correct grammar, and knowledge of punctuation rules is still essential. When giving businessmen advice on dictating letters more than one half century ago, Earle Buckley (1940) urged writers to be concise rather than rambling and wordy, to check that various thoughts and paragraphs were related, to make text sound like it was coming from a real person, and to have personality. Composing orally may help motivate students with LD to learn more about written language conventions as the current technology frees them from worrying about poor spelling and handwriting. One high school student with LD recently told me that composing his assignments through a SR system did not help him get better grades, yet it did motivate him to finish more assignments because his work was completed more easily.

Fifth, modifications in statewide competency tests often permit students to use dictation as an alternative method of responding, yet special educators may not be aware of common restrictions for the use of such accommodations. Results from a survey of state assessment guidelines by Thurlow and her colleagues (1995) indicated that modifications do not remain consistent from year to year in many states, and that some states reject accommodations that other states accept. However, at the time of Thurlow's review, 15 states allowed the use of a scribe for purposes ranging from “recording answers” to “taking dictation.” In most of these states, students were required to tape or dictate verbatim responses for later verbatim transcription by school personnel. Some states specified that eligible students were to provide all capitalization, punctuation, spelling and paragraphing, so that the only writing skill not performed by the student was handwriting. Most states indicated that the scribe should have experience in working with the student, and more importantly, that it was the state's intent to allow alternative modes of responding only for students who used an oral mode of composition on a routine basis prior to the state assessment. Cases in which students use dictation routinely, either as a prewriting activity or as a means of composition, were more likely to be viewed as engaging in a habitual adaptation rather than a special accommodation (Siskind, 1993). These guidelines serve to reinforce previous recommendations throughout this article, and will most likely be applied in a similar way to the future use of speech recognition systems.

In conclusion, the time for placing greater emphasis on the use of dictation and other oral modes of production has now arrived, as technology has made this a more viable option for students with LD. It is important now to start considering how we can best use such tools, and what teachers and students need to do to make them as effective as possible. In an age where electronic products rapidly become smaller and more sophisticated, we may all soon prefer to talk to our computers instead of struggling with keyboards or handwritten forms of composing.

REFERENCES


Meisel, W.S. (1993). Talk to your computer: Voice technology lets you verbally command your computer or convert speech to text. Byte, 18, 113-120.


AUTHOR NOTES
The author thanks the reviewers for comments on an earlier version of this article.

1 This SR product preceded the same company's recently developed NaturallySpeaking SR system and remains an available option for individuals to consider.

Requests for reprints should be addressed to: Susan De La Paz, delapaz@worldnet.att.net