SESSION II

NONTRADITIONAL HUMAN-COMPUTER INTERFACES

C. Michael Levy, *President*

Reading instruction and remediation with the aid of computer speech

RICHARD OLSON, GREGORY FOLTZ, and BARBARA WISE
University of Colorado, Boulder, Colorado

In this paper, we describe the application of new computer and speech synthesis technologies for reading instruction. Stories are presented on the computer screen, and readers may designate words or parts of words that they cannot read for immediate speech feedback. The important contingency between speech sounds and their corresponding letter patterns is emphasized by displaying the letter patterns in reverse video as they are spoken. Speech feedback is provided by an advanced text-to-speech synthesizer (DECTalk). Intelligibility data are presented, showing that DECTalk can be understood almost as well as natural human speech by both normal adults and reading disabled children. Preliminary data from 26 disabled readers indicate that there are significant benefits of speech feedback for reading comprehension and word recognition, and that children enjoy reading with the system.

Recent advances in speech production for microcomputers have opened up some new and exciting possibilities in reading instruction. The system we have developed involves the reading of stories from the CRT screen of an Apple II+ or an IBM PC. When children encounter a word that they are unsure how to read, they can target the word with a mouse or a light pen. The targeted word is then highlighted in reverse video and spoken by the computer.

In addition to giving whole-word feedback, the computer can orthographically segment and speak targeted words at syllable and subsyllable levels. The simultaneous highlighting of the orthographic segment and presentation of its corresponding speech sound provides a powerful way of emphasizing the relations between letter patterns and their sounds during reading of natural text. Thus, the computer-based reading and speech-feedback system is designed to support the long-term improvement

of disabled and beginning readers' phonological decoding skills, in addition to providing on-line feedback for learning how to read specific words in the text.

Our long-term research goal is to determine the optimal levels of feedback segmentation for individual disabled and beginning readers, depending on their unique profiles of reading and language skills. The present studies were designed to answer some preliminary questions concerning the intelligibility of synthesized speech, the optimal methods of targeting words on the screen, and the short-term influences of the different feedback conditions on disabled readers' comprehension and word recognition in short stories.

We will first summarize some results from our earlier research with disabled readers that motivated our development of a computer-based reading and speech-feedback system. Then we will present two experiments on the intelligibility of synthesized speech for normal and disabled readers. Finally, we will present a preliminary study of 26 disabled readers' responses to speech feedback in reading that demonstrates substantial short-term benefits in their comprehension and word recognition.

BACKGROUND: INSTRUCTIONAL NEEDS OF DISABLED READERS

Our earlier research with disabled readers has shown that, as a group, they may be able to read a number of

This research was supported by NICHD Grant No. 2 P01 HD 11681-07; by Grant No. RR07013-85 from the Biomedical Research Support Grant Program, Division of Research Resources, NIH; and by a Digital Equipment Corporate Contributions Grant. We thank the teachers and students from University Hill Elementary School and from Alameda Senior High School for their participation in the research. Vicki Steele edited the phonetic feedback files for the stories, and Mark Nash developed the feedback programs for the IBM PC. Jan Keenan made helpful comments on an earlier version of the manuscript. Requests for reprints should be addressed to the first author at the Department of Psychology, University of Colorado, Boulder, CO 80309.

familiar printed words, primarily through orthographic coding or "sight reading," but that most disabled readers are uniquely deficient in phonetically decoding or "sound-ing out" unfamiliar printed words and in related analytic language skills (Olson, 1985; Olson, Kliegl, Davidson, & Foltz, 1985).¹ Consequently, the development of their reading vocabularies lags substantially behind that of their oral vocabularies.

Longitudinal analyses have shown that most disabled readers who were identified in the early grades retain their reading deficit into the later grades and adulthood, relative to the performance of normal controls (Olson, 1985). However, there are some significant and encouraging exceptions to this group trend. A few children who demonstrate reading difficulties in the early grades eventually achieve normal reading levels.

One factor contributing to the ultimate success of some disabled readers is their continued instructional support in reading through the later grades. Unfortunately, most schools do not have the resources to provide the intense individualized instruction that is needed to motivate and train disabled readers. Many disabled readers become discouraged by their early failures and avoid reading. Others are motivated to continue reading; they show greater improvement, but they are often hampered by weak phonetic coding skills. Disabled readers may correctly guess unfamiliar words in context in association with their limited alphabetic skills, but this approach often leads to decoding errors, because context is not a very reliable predictor for words in continuous text (Gough, 1983). In the absence of feedback for their frequent decoding errors during silent reading, the development of better decoding skills is actively retarded by incorrect learning trials (Jorm & Share, 1983), comprehension suffers, and the pleasures of reading are diminished.

It seemed to us that computer-based reading and speech feedback for decoding difficulties could be very helpful for disabled readers, who could read text for meaning and pleasure while conveniently and privately obtaining the decoding help they need. Reading materials could be selected to reflect the interests and general reading level of each child, and the segmental level of the orthographic and speech feedback could be tailored to match each child's unique profile of reading and language skills. However, there were preliminary questions about the intelligibility of synthesized speech for disabled readers that needed to be answered. These questions are addressed in the following section. Then we will turn to the results of a pilot study that provided synthesized speech feedback to disabled readers in natural text.

INTELLIGIBILITY EXPERIMENTS: NATURAL VERSUS SYNTHESIZED SPEECH

There are two basically different ways of producing speech from a computer: natural human speech can be digitized, stored in memory, and then played back through a digital-to-analog (D/A) converter, or artificial speech

can be synthesized by the computer from text (Witten, 1982). The advantage of synthesized text-to-speech systems is that additional memory demands on the computer are limited to the storage of ASCII code for the text to be read and spoken, and any word in the language can be spoken. The disadvantage of speech synthesizers for small computers has been their poor intelligibility. In contrast, the intelligibility of digitized speech can be equal to the human speech input, but the demands on computer memory are high. McConkie and Zola (1985) partly resolved this memory problem by storing the analog signal for about 3,000 words on a floppy disk. When their subjects target a word for speech feedback, the appropriate analog signal is located on the disk and played back.

Three factors led us to develop our speech-feedback system around one of the more recent text-to-speech synthesizers. First, as shown in the following experiments, speech synthesis for microcomputers has approached the intelligibility levels of natural human speech. Second, the range of text materials in our experiments requires a much larger vocabulary than the 3,000 words used by McConkie and Zola (1985). Third, the speech synthesizer allows us to more conveniently manipulate the level of segmentation and timing for syllable and subsyllable feedback conditions.

The text-to-speech synthesizer used in the present experiments is DECTalk, made by Digital Equipment Corporation. DECTalk may be interfaced with any computer or terminal. It responds to text in ASCII format, which is converted to phonemic code for further processing. A large internal dictionary and rules of English pronunciation are used to do the conversion. Subsequently, the phonemic code is converted into synthesizer control commands. Finally, the speech synthesizer uses the control commands to generate speech. Actually, frames of speech are produced every 100 μ sec. These successive frames sound like a continuous speech pattern. The speech synthesizer used in DECTalk is described by Klatt (1980). Other details on the internal structure of DECTalk are described by Bruckert, Minow, and Tetschner (1983).

In order to justify the use of synthesized speech, it was necessary to show that DECTalk's intelligibility was adequate. Greene and Pisoni (in press) reviewed a number of studies on the intelligibility of several text-to-speech systems compared to natural human speech, and concluded that DECTalk's "perfect Paul" male voice provided the most intelligible synthetic speech. Although statistically significant differences in intelligibility were typically found between DECTalk and natural speech, these differences were usually quite small, and the magnitude of the differences depended on the particular test used.

The test used in our intelligibility experiments was designed to compare how well subjects in the feedback studies could recognize the range of words in their reading material when the words were presented by DECTalk or in natural speech. Two experiments were carried out: one with college-age normal readers and the other with reading-disabled children. The stimuli for both experi-

ments were 120 different words from an eighth-grade-level story. These words provided an ecologically valid range of word lengths, word frequencies, and grammatical forms that are encountered in text at the eighth-grade level. The words in the story were randomized and presented at the rate of one word every 5 sec by DECTalk's "perfect Paul" male voice. The series of computer-generated words was recorded on tape. The same words in a different random order were recorded in natural speech at 5-sec intervals.

Eight normal readers from an introductory psychology class participated in the first experiment as part of course requirements. After listening to a 500-word story that was read by DECTalk to familiarize students with the dialect, they listened to each word on the two lists and repeated what they thought they had heard.² The presentation order for the two lists was counterbalanced across subjects. Oral responses were recorded on tape and later transcribed for accuracy analyses. Any significant phonemic deviation from the correct pronunciation for a word was scored as an error.

The mean percent correct recognition rates for DECTalk (95.8%) and natural speech (96.1%) were not significantly different for the psychology students [$t(7) = -.12$]. Most of the errors in both synthetic and natural speech conditions were due to single phoneme substitutions or additions (e.g., "free" for *three*, "black" for *lack*, "skin" for *skim*, "great" for *rate* [target words in italic]).

The high recognition rate of the psychology students for isolated words produced by DECTalk suggested that synthetic speech has reached intelligibility levels that are useful for speech feedback in reading. However, Greene and Pisoni (in press) suggested that younger children and subjects with language disabilities may have more serious problems understanding synthesized speech. Since our subjects were disabled and beginning readers, it is important to demonstrate that their recognition rates with DECTalk are also high.

Nine disabled readers between 8 and 12 years of age, who also participated in the pilot study on speech feedback, were given the intelligibility tests for DECTalk and natural speech. The only difference in procedure from the study with the psychology students was that the discrimination tests were presented at the end of the feedback training session. During this session, subjects listened to about 500 words of oral instruction produced by DECTalk, in addition to the words they targeted. Thus, the amount of exposure to DECTalk speech prior to the intelligibility tests was roughly similar to that for the psychology students.

The mean correct recognition percentages for DECTalk (94.5%) and natural speech (98.4%) were slightly, but significantly, different in this sample [$t(8) = -2.90$, $p < .05$]. However, the most important result is that DECTalk recognition for the disabled readers (94.5%) was nearly as high as for the psychology students (95.8%) [$t(15) = .6$]. It is clear that disabled readers were also able to recognize DECTalk's synthesized speech at a high

level. Their recognition rate would possibly be even higher if subjects heard words while reading text because the context would constrain the alternative responses. This prediction was confirmed in the feedback study with disabled readers, as described in the following section.

Future advances in speech synthesis technology should continue to reduce the small differences in intelligibility between synthesized and natural speech. In addition to DECTalk, high-quality synthesized text-to-speech systems are becoming available from other companies. The most recent 3.0 version of the Prose 2000 produced by Speech Plus has intelligibility levels very close to those of DECTalk (Greene & Pisoni, in press). Greene and Pisoni have also reviewed several less expensive systems, such as the Echo II for the Apple computer, but these systems have intelligibility levels that are substantially lower. Our evaluations of the Echo II synthesizer indicated that it was not suitable for the feedback system. It seems likely that the increasing availability and decreasing cost of high-quality synthesized speech for microcomputers will lead to its wider use in computer-based instruction.

DISABLED READERS' RESPONSES TO SPEECH FEEDBACK IN READING

A pilot study with disabled readers was designed to evaluate their responses to the computer-based reading and speech-feedback system. The primary questions addressed by the study were the following: (1) whether subjects would be sensitive to their decoding errors and target the problem words during the normal course of reading stories; (2) whether the targeted words would be subsequently recognized in posttraining tests; (3) whether speech feedback would improve text comprehension. Additional questions concerned the subjects' abilities to use a mouse and light pen as targeting devices, their responses to reading text at different difficulty levels, and the relative benefits of segmenting the orthographic and speech feedback at different levels. Finally, we wished to determine whether disabled readers would enjoy reading with the system and whether they would be motivated to continue using it in a long-term study of reading with the different segmentation conditions.

Method

Subjects. Fifteen reading-disabled children between 8 and 12 years of age were obtained from elementary schools in Boulder, Colorado. An additional group of 11 disabled readers between 15 and 18 years of age was obtained from a local high school. The subjects were referred to the study by their reading teachers based on their normal-range IQ and on reading performance that was substantially below what was expected for their grade level. Most of the children met the strict criteria for reading disability used in previous research (Olson et al., 1985). Within these criteria, the children varied widely both in their absolute reading levels and in the severity of their reading disabilities. We wished to determine the

usefulness of the system for disabled readers at different levels of reading proficiency and age.

Apparatus and software. We used two different feedback systems. First, a prototype system was developed for an Apple II+ that used a mouse as a targeting device. A Sider hard disk was interfaced to the Apple for storing text displays, speech instructions for DECTalk, and response data. Apple Pascal was the primary programming language, and a number of low-level assembly language routines were incorporated into the host program to handle the mouse, provide rapid disk I/O, display screens, highlight selected words or parts of words in the text, control timing functions, and gather subject responses from the mouse and keyboard. Text displays were white on black, upper- and lowercase, in 40-column format. Words or parts of words were designated for feedback by placing the mouse cursor beneath a word and pressing the mouse button. Speech instructions were then sent over a serial line to DECTalk, and the targeted word or word segment was highlighted in reverse video while the speech was being produced.

A second system was developed for an IBM PC that employed a light pen as a targeting device. A RAM disk was used in the PC to store and present displays and their related speech instructions for each word so that text pages could be rapidly turned while the subjects read stories. An HEI light pen, when pressed against the glass, signaled the character location. As with the Apple, the appropriate speech code was sent to DECTalk, and the targeted word or segment was displayed in reverse video during the speech signal. Text displays on the PC were black on white in 40-column format with up to 12 double-spaced lines.

A voice key determined the latency between sending an ASCII string to DECTalk and the beginning of the pronunciation. When the character string was sent in DECTalk's phonemic instruction mode, the delay to speech onset ranged from 140 to 200 msec. Highlighting of a word or word segment was nearly instantaneous with the press of the mouse button or light pen. Therefore, there was a delay of approximately 140–200 msec between the highlighting of a word or word segment on the screen and the beginning of the speech signal.

Text materials and segmentation conditions. Reading material and comprehension questions for the study were drawn from the Spache Diagnostic Reading Scales (Spache, 1963). There were four different stories at Grade Levels 1, 2, and 3 that were approximately 100 words in length. There were two stories for each of Grade Levels 4 through 8 that were each about 200 words in length. Four comprehension questions were selected from the Spache scales for each story.

Three different segmentation conditions were used for the orthographic and speech feedback. When a word was targeted by the subject, feedback could be presented at the whole-word, syllable, or subsyllable levels. In the syllable condition, words were segmented into syllables generally according to the American College Dictionary.³

When a word was targeted, the first syllable was immediately highlighted in reverse video and the speech followed within 140 to 200 msec. Highlighting was maintained for 600 msec for each syllable, and then it was returned to normal contrast as the next syllable was highlighted and spoken.

In the subsyllable condition, syllables were segmented into an "onset" consisting of a consonant or consonant cluster (if present in the syllable), and a "rime" consisting of a vowel, which may or may not have been followed by up to four consonants (Treiman, 1985). Each subsyllable segment was highlighted and spoken within a 400-msec time frame. On average, there were about 1.5 subsyllable segments per syllable, so that the mean highlighting time for words in the syllable and subsyllable conditions was equivalent. The total highlighting time in the whole-word condition was based on 600 msec per syllable to equate highlighting time per word in each of the three conditions.

In the syllable and subsyllable segmentation conditions used in the present study, all of the segments were presented automatically and sequentially from left to right across the word, regardless of which character was targeted in the word. Two additional segmentation conditions have been designed to allow for the isolated presentation of the specific syllable or subsyllable segments that are targeted. These segmentation conditions will be evaluated in a subsequent study.

Design and procedure. Subjects were given a word recognition test at the beginning of the session, and their scores on this test were used to select reading material at the appropriate difficulty level. Then they sat in front of the computer and listened to DECTalk present the instructions for reading and targeting words in the stories. These instructions included a request that they raise their hands when DECTalk spoke a word that they did not recognize. Subjects participated in a practice session with the targeting device wherein subjects targeted 20 numbers that successively appeared in random positions on the screen. In the final instructions, the subjects were told to read the stories for meaning and to target each word that they did not recognize. They were told that they would later be tested on the words they targeted, along with other difficult words in the text, and that they would be asked some comprehension questions at the end of each story. The instructions emphasized that the subjects should attempt to decode difficult words before targeting them for feedback. In the oral reading conditions, they had to attempt an oral reading of each word before targeting it. This allowed us to evaluate the benefits of targeting for words that were read incorrectly prior to targeting. Very few correct words were targeted by the subjects.

Six stories were read during the test session. In the first two stories, presented at a difficulty level appropriate for a subject's reading grade level, the subject was expected to make relatively few decoding errors. Difficulty levels for the next two stories were one grade level above the subject's reading grade, and the last two stories were two

grade levels above the subject's reading grade. After each story, the subject was asked four comprehension questions by the experimenter. Then the computer presented a word recognition test that included a randomized list of all of the words that were targeted plus six of the most difficult words in the text (determined from a previous study). The subject's oral response to each word in the recognition test was recorded as correct or incorrect by the experimenter, and then the next word was presented. No feedback was given about the accuracy of responses. A final recognition test was presented after all of the stories had been read. This test included all of the words that had been targeted in the six stories.

The first two stories in the session were read aloud with speech feedback. The second pair of stories at the next higher grade level included one that was read aloud with speech feedback and one that was read aloud without speech feedback in counterbalanced order across subjects. In the control condition without feedback, the synthesizer was simply turned off and the subject was asked to continue targeting words he/she was uncertain how to read. The highlighting pattern that was appropriate for the subject's assigned segmentation condition was continued for each targeting response in the no-feedback condition. The subject was asked to study and to try to decode the targeted words while they were highlighted on the screen. The last two stories at the highest difficulty level were both read with speech feedback, but one was read aloud and the other silently in counterbalanced order. All of the oral reading was tape-recorded for later analysis of reading errors. Most subjects completed the entire session within 1½ h.

Subjects were alternately assigned to the Apple-mouse and IBM PC-light-pen systems. Within the targeting conditions, we had hoped to balance the assignment of subjects to the three segmentation conditions, but software problems with the subsyllable and whole-word conditions led to a larger number of subjects in the syllable segmentation group in this pilot study. There were 6 subjects in the whole-word condition, 14 subjects in the syllable condition, and 6 subjects in the subsyllable condition.

Results and Discussion

Dependent variables included the percentage of content words in the text that were read incorrectly (from oral stories only), the percentage of content word errors that were targeted, the percentage of targeted oral errors that were read correctly on the recognition test following each story, the percentage of targeted oral errors from each story that were correctly recognized in the final test at the end of the session, and percent comprehension (see Table 1). The results for these variables will be discussed, following a brief consideration of some small differences in performance between subjects who used the mouse and light-pen targeting devices.

In the practice session, subjects were faster in targeting numbers on the screen with the light pen than they were with the mouse. However, this difference did not

result in a significant difference in the frequency with which subjects targeted words in the stories. Mouse subjects averaged 53.7% targeting responses for words that were initially read incorrectly, whereas light-pen subjects averaged 58.6% targeting responses. Although it appeared that the light pen was easier to use, and a larger sample might reveal a statistically significant advantage in targeting for the light pen, it was clear that subjects could effectively target words with the mouse. The general pattern of results for all of the dependent variables was very similar for subjects who used the mouse and light pen. Therefore, the data from the two conditions were combined for the following analyses.

The means in Table 1 are based on 9 subjects who had at least 4% but not more than 24% oral reading errors and who targeted at least 47% of their oral errors. Although some useful data was obtained from all of the subjects, some subjects made too many or too few oral reading errors due to the difficulty level of the text materials. Two of the younger and most severely affected disabled readers made over 50% oral reading errors because our first-, second-, and third-grade stories were too difficult for them. These subjects said they liked "reading" with the system. It markedly aided their comprehension, and they recognized an average of 43% of targeted oral errors on the first test and about 20% on the second test. Their recognition performance was not as good as for the other subjects, because they were overwhelmed by the number of unfamiliar words they had to learn and remember for the recognition tests. Some of the older high school subjects made less than 3% oral reading errors and their individual story data were too unstable for the following analyses. Finally, several subjects, who targeted a very low percentage of their oral reading errors, were excluded. These disabled readers had been told by their teachers to rely heavily on context in their reading of unfamiliar words, and this may have lowered their sensitivity to decoding errors. In post-experimental training with these subjects, sensitivity to decoding errors could be improved by pointing the er-

Table 1
Mean Percentages for Selected Disabled Readers

	All Oral Feedback Stories (4)	Feedback Versus No Feedback		Oral Versus Silent	
		F	NF	Oral	Silent
% Oral Errors	15.5	16.5	18.5	22.9	-
% Targeted Oral Errors	65.7	58.9	29.0*	68.1	(18.8% Words Targeted)
% Recognized Test 1	74.7	89.6	15.9†	58.7	55.7
% Recognized Test 2	55.8	67.5	26.6*	52.4	46.6
% Correct Comprehension	-	88.9	63.0*	75.0	65.8

Note - The means and *p* values for the recognition tests in the feedback versus no-feedback conditions are based on 6 subjects who targeted words in the no-feedback condition. All others based on 9 subjects. **p* < .05; †*p* < .01.

rors out as subjects read. It is clear from these results that pretraining in sensitivity to decoding errors will be needed for some children to take full advantage of the feedback.

Column 1 in Table 1 contains the mean results from the 9 selected subjects for the four stories that were read aloud with feedback (two stories at reading grade level and one each from the higher grade levels). Over the four stories, the subjects averaged 15.5% oral errors, of which 65.7% were targeted for feedback. This demonstrates a reasonable amount of sensitivity to their decoding difficulties. Of the words that were initially read incorrectly in the oral stories and targeted, 74.7% were recognized in the tests that followed the comprehension questions after each story, and 55.8% were correctly recognized in the final recognition test for all targeted words at the end of the session. Both tests showed an impressive gain in word recognition over the duration of the test session. It should be noted that the recognition scores were for words taken out of context and presented in isolation. It is likely that recognition scores would be even higher if the words were read again in context.

The results presented in Columns 2 and 3 of Table 1 are for the speech-feedback and no-speech-feedback conditions presented in the third and fourth stories of the test session. There was a significantly lower rate of targeting in the no-feedback condition [$t(8) = 2.43, p < .05$], although the subjects were explicitly instructed to continue targeting words they were uncertain how to read. In fact, 3 of the subjects did not target any words in the no-feedback condition, although they missed some of the more difficult words that were automatically included in the recognition test. They may have realized that targeted words would show up on the recognition test and that they would have problems reading them without feedback.

The recognition test means and significance tests between the feedback and no-feedback conditions were based on the 6 subjects who targeted words in the no-feedback condition. For both the recognition test that followed each story and the final recognition test, the percentage of targeted oral errors was significantly higher in the speech-feedback condition [$t(5) = 5.16, p < .01$; and $t(5) = 2.61, p < .05$, respectively]. There were a few obvious cases in which the separate highlighting of syllables helped subjects decode words without speech feedback, but it was apparent from the results that speech feedback was the primary factor that improved word recognition. The bottom row in Columns 2 and 3 of Table 1 shows that the benefits of speech feedback were also reflected in the subjects' comprehension scores [$t(8) = 2.53, p < .05$].

Columns 4 and 5 in Table 1 present means for the oral and silent reading conditions in the last two stories of the series. The targeting frequencies, performance in the recognition tests, and comprehension scores were not significantly different in the two conditions, although there was a trend toward better performance in the oral reading condition. The silent condition was included because we plan to have children read silently in a long-term study with feedback systems set up in the schools. It was neces-

sary to show that children would also target words and benefit from speech feedback in silent reading.

One issue of concern was the effect of different test difficulty levels on the recognition rates for targeted words. The percentage of oral errors is higher and the recognition rates are lower for the most difficult stories, represented in Column 4, compared to rates for feedback stories that were one grade level lower, shown in Column 2, and the differences were greater for stories at the lowest difficulty level. The declining trend in percent recognition across difficulty level is confounded by presentation order in the series of stories, but a similar trend was observed across subjects when those with many oral errors were compared with those subjects with fewer errors. It seems that when the error rates get too high because of text difficulty, subjects become overwhelmed and show less benefit from their feedback. The natural flow of reading is seriously disrupted when there is a high frequency of decoding errors and targeting. An optimal range of oral errors seemed to be between 5% and 15% for attaining high recognition rates for targeted words, for adequate comprehension, and for maintaining the natural flow of reading.

The final issue addressed by the study was the effects of the whole-word, syllable, and subsyllable feedback conditions. Of the 9 subjects whose data were included in Table 1, 2 had subsyllable feedback, 5 had syllable feedback, and 2 had whole-word feedback. The mean targeted word recognition scores for the immediate and final tests were 70.5% and 35.8% for subsyllable, 76.5% and 54% for syllable, and 77.6% and 61.5% for whole-word conditions. Of course the samples in the feedback conditions are too small for any of these differences to be significant. The lower recognition rates for the 2 subjects in the subsyllable condition may have been due to their higher rates of oral reading errors and targeting in the stories, rather than the type of feedback.

One severely affected disabled reader assigned to the subsyllable feedback condition did not finish the session. This child had great difficulty blending the subsyllable segments and failed to recognize many of the words after speech feedback. The child was better able to recognize multisyllabic words from feedback when they were segmented syllabically. The 2 subsyllable subjects included in Table 1 were unable to successfully blend and recognize a few of the feedback words. Subjects who received syllable feedback nearly always recognized the words if they were in their oral vocabularies. Thus, the preliminary results from these few subjects suggest that subsyllable segmentation may not be an effective form of feedback to aid in word recognition for some severely disabled readers, at least over the brief training period of this study.

Of course it is the long-term consequences of extended training with the different feedback conditions that are most important for the remediation of reading disabilities and for teaching beginning readers. The type of feedback that provides the greatest short-term gains in recognizing targeted words may or may not be the best way

to build a child's phonological decoding skills, which can be applied to successfully decode new words without feedback. At least for normal beginning readers, a phonics or decoding approach is viewed by most researchers as an important component in reading instruction (see Perfetti, 1985, for a review). Our subsyllable segmentation condition bears the greatest similarity to a phonics approach. However, the best instruction for the normal reader is not necessarily the best instruction for the disabled reader. Even within the disabled group, there is a wide range of analytic language and component reading skills (Olson, 1985). Individual differences in these skills may be related to the optimal level of segmentation in speech feedback. This issue will be addressed in a planned long-term training study with disabled and beginning readers.

One very encouraging result of this study was the unanimously enthusiastic responses of the subjects to the computer-based reading and speech-feedback system. They would like to have it available at home or school so they could read interesting materials that would otherwise be too difficult. All of the subjects indicated a willingness to participate in future research with the system.

In conclusion, the quality of synthesized speech for microcomputers has reached a level that makes it useful for helping children learn to read. Our efforts were focused on the training of disabled readers with speech feedback because these children have the greatest need for individualized instruction to improve their word decoding skills. The system is also being evaluated as an instructional supplement for normal beginning readers. Our approach was to provide speech feedback for word decoding difficulties in the context of reading stories for meaning. Thus, children are taught how to identify printed words, and at the same time use their language ability while reading. The results of the study provided strong support for this approach, but further research is needed to evaluate the optimal methods of presenting orthographic and speech feedback for individual children over a longer training period.

REFERENCES

- BRUCKERT, E., MINOW, M., & TETSCHNER, W. (1983). Three-tiered software and VLSI aid developmental system to read text aloud. *Electronics*, 56(8), 133-138.
- GOUGH, P. B. (1983). Context, form, and interaction. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 203-211). New York: Academic Press.
- GREENE, B. G., & PISONI, D. B. (in press). Perception of synthetic speech by adults and children: Research on processing voice output from text-to-speech systems. In L. E. Bernstein (Ed.), *The vocally impaired* (Vol. 2). New York: Academic Press.
- JORM, A. F., & SHARE, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics*, 4, 103-147.
- KLATT, D. H. (1980). Software for a cascade/parallelized formant synthesizer. *Journal of the Acoustical Society of America*, 67, 971-995.
- MCCONKIE, G. W., & ZOLA, D. (1985, April). *Computer aided reading: An environment for developmental research*. Paper presented at the meeting of the Society for Research in Child Development, Toronto.
- OLSON, R. K. (1985). Disabled reading processes and cognitive profiles. In D. Gray & J. Kavanagh (Eds.), *Biobehavioral measures of dyslexia* (pp. 215-243). Parkton, MD: York Press.
- OLSON, R. K., KLIEGL, R., DAVIDSON, B. J., & FOLTZ, G. (1985). Individual and developmental differences in reading disabilities. In J. Kavanagh, D. Gray, & D. Pearl (Eds.), *Reading research: Advances in theory and practice* (Vol. 4, pp. 1-64). New York: Academic Press.
- PERFETTI, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- SCHWAB, E. C., NUSBAUM, H. C., & PISONI, D. B. (in press). Effects of training on the perception of synthetic speech. *Human Factors*.
- SPACHE, G. D. (1963). *Diagnostic reading scales*. Monterey: McGraw-Hill.
- TREIMAN, R. (1985). Phonemic analysis in spelling and reading. In T. H. Carr (Ed.), *The development of reading skills. New Directions for Child Development*, No. 27 (pp. 5-18). San Francisco: Jossey-Bass.
- WITTEN, I. H. (1982). *Principles of computer speech*. New York: Academic Press.

NOTES

1. The criteria for reading disability in the present study were a minimum IQ of 90, reading performance at least 2 years below expected grade level, normal educational opportunity and socioeconomic level, and no uncorrected sensory deficiencies. All of the children spoke English as a first language.
2. Schwab, Nusbaum, and Pisoni (in press) have demonstrated that a short period of exposure to synthetic speech can substantially improve subjects' recognition performance. Since subjects in the feedback studies were to have a substantial amount of experience listening to DECTalk, we were interested in recognition performance for words after the subjects had become familiar with DECTalk's dialect.
3. The orthographic segmentation of syllables in most dictionaries does not acknowledge some linguistic complexities. For example, Treiman (1985) has noted the disagreement among linguists on where to segment a two-syllable word such as *open*. Some say the /p/ is part of the first syllable, some say it is part of the second syllable, and some say it is part of both syllables. We are conducting a series of studies in collaboration with Becky Treiman to resolve these issues and determine the best approach to syllable and subsyllable segmentation for the different feedback conditions.