2013

Using Text-to-Speech Reading Support for an Adult With Mild Aphasia and Cognitive Impairment

Judy Harvey
University of Nebraska–Lincoln, harvey.judy@gmail.com

Karen Hux
University of Nebraska–Lincoln, khux1@unl.edu

Jeffry Snell
Quality Living, Omaha, NE

Published in *Communication Disorders Quarterly* 35:1 (2013), pp. 39-43; doi: 10.1177/1525740113496843

Published online July 25, 2013, at http://cdq.sagepub.com/content/35/1/39

Using Text-to-Speech Reading Support for an Adult With Mild Aphasia and Cognitive Impairment

Judy Harvey,[1] Karen Hux,[1] and Jeffry Snell[2]

1. University of Nebraska–Lincoln, USA
2. Quality Living, Omaha, NE, USA

Corresponding Author – Judy Harvey, University of Nebraska–Lincoln, 350 Barkley Memorial Center, Lincoln, NE 68583-0738
email harvey.judy@gmail.com

Abstract

This single case study served to examine text-to-speech (TTS) effects on reading rate and comprehension in an individual with mild aphasia and cognitive impairment. Findings showed faster reading, given TTS presented at a normal speaking rate, but no significant comprehension changes. TTS may support reading in people with aphasia when time limitations exist.

Keywords: literacy, reading, case studies, methodology, research, stroke, technology

Language impairments caused by aphasia often include decreased reading efficiency and comprehension (Devlin & Unthank, 2006; Lasker, 2011). Even mild aphasia may cause struggle with decoding words and comprehending sentences common in everyday text. These challenges may restrict an individual's participation in leisure and social activities involving written text comprehension (Dietz, Hux, McKelvey, Beukelman, & Weissling, 2009), and limit self-esteem, cognitive stimulation, and general enjoyment associated with reading (Triandafilou, 2003). Reduced reading comprehension leaves individuals with aphasia at a distinct disadvantage compared with people without reading challenges given today's literacy-dependent world.

Aphasiologists and reading specialists believe that providing multimodality supports may benefit people with reading disabilities (Elkind, Black, & Murray, 1996; Hale et al., 2005; Hecker, Burns, Elkind, Elkind, & Katz, 2002; Higgins & Raskind, 2004; Hux, Weissling, & Wallace, 2008; Lasker, Hux, Garrett, Moncrief, & Eischeid, 1997; Lindstrom, 2007). One method of providing multimodality supports within the reading domain is to use audio recordings as supplements to written texts. Given that computers and electronic readers (i.e., e-readers) routinely include text-to-speech (TTS) features, the provision of audio supports is now readily available (Dietz, Ball, & Griffith, 2011; Lasker, 2011).

Despite easy access to TTS technology, questions remain about whether incorporation of this type of support is truly beneficial to people with aphasia. Multiple variables—such as the quality and rate of digital speech production and the relative benefit of having single or multiple presentation modalities—remain unexplored. For the most part, researchers and practitioners do not yet know whether TTS technology supports or hinders the reading efficiency and comprehension of people with aphasia. Therefore, the research questions explored through this pilot study were as follows: How do (a) reading rate (measured in words per minute [WPM]) and (b) comprehension accuracy (measured in proportion of correct question responses) for an individual with mild aphasia differ across three conditions—no TTS support, TTS support presented at an average listening rate for neurotypical adults (i.e., regular TTS), and TTS support presented at a speed comparable with the silent reading rate in an individual with aphasia (i.e., slow TTS)?

Method

Participant

Kay—a 44-year-old female with cognitive impairments and aphasia—underwent surgical resection for a left hemisphere meningioma in 2002. The meningioma recurred in fall 2007, at which time Kay underwent a debulking procedure followed by radiation. Subsequently, she demonstrated expressive and receptive aphasia for which she received 1 week of inpatient treatment. Kay made adequate progress to return home; however, she required 24-hr support for safety due to persistent mild aphasia, cognitive impairments, chronic headaches, and seizures. Failing to thrive in the home setting, Kay returned to rehabilitation in December 2007 and was later discharged to an assisted living facility where she resided at the time of study participation.

Kay independently managed basic activities of daily living at the assisted living facility. She received help with executive functioning for activities such as planning future events and managing finances and medical care. From the variety of activities offered each day, Kay consistently opted to join ones that were literary in nature: She was active in a book club, she checked out novels and audio books from the library, she had a list of favorite writers, she used the Internet, and she had frequent meetings with friends to read celebrity magazines.

Despite her fondness for books and reading, Kay's cognitive and language profile indicated that performance with text activities would not be strong. On the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), Kay scored at the first percentile because of poor orientation, attention, immediate memory, and delayed recall. Her *t* scores of 23 and 2 on Trails A and B, respectively, of the Trail Making Test (Reitan & Wolfson, 1993) and her inability to perform the Wisconsin Card Sorting Test (Grant & Berg, 1993) confirmed poor problem solving and cognitive flexibility. Her aphasia quotient score of 93 on the Western Aphasia Battery–Revised (Kertesz, 2007) indicated mild anomic aphasia. She spoke in short, syntactically accurate phrases and sentences, but she demonstrated substantial delays when performing confrontation naming and word fluency tasks. Single-word reading was a relative strength as demonstrated by Kay's attainment of 80% to 90% accuracy on the single-word subtests of Reading Comprehension Battery for Adults–2 (LaPointe & Horner, 1998). She also scored 80% on factual paragraph comprehension with and without picture support; however, she achieved only 50% accuracy responding to inferential questions about written paragraphs and questions about morphosyntactic sentence elements. Her average silent reading rate of 102 WPM was slow compared with average rates for neurotypical individuals (i.e., 250–350 WPM; Brown, Fishco, & Hanna, 1993; Hiebert, Samuels, & Rasinski, 2012). Wearing glasses, Kay reported that she could read hardcover and paperback novels, most of which are printed in 11-point font. She passed a visual screening for the experimental task by accurately reading 18-point font words presented on a computer screen.

Materials

The researchers selected 36 reading passages from the American Council on Education (ACE) practice tests (Cengage Learning, 2012) for the experimental stimuli. The passages ranged from 117 to 351 words (M = 225.11, SD = 60.39) and from grade level 5.16 to 15.97 (M = 9.31, SD = 2.26; Readability Formulas, 2012). Five multiple-choice questions accompanying each passage—with five response options each—addressed a mixture of reading concepts relating to comprehending the main idea, comprehending details, recognizing organizational structure, identifying the author's perspective, making inferences, or understanding specific word usage.

The researchers typed the passages into separate Microsoft Word documents on a MacBook Pro computer for later presentation either in print or combined print and TTS modalities. Hard copies of the five multiple-choice comprehension questions for each passage appeared on two separate sheets of 8.5 × 11 inch paper.

The researchers assigned each passage two standardized z scores—one for length and one for grade level. Pearson product- moment correlation computation between length and grade level established that longer passages were not inherently more difficult than shorter ones (r = −.270, p = .111). By summing the two z scores per passage, the researchers obtained a composite score representative of the combined variables of passage length and difficulty. They then sorted passages into three sets of 12 with comparable composite score distributions. Comparable passages were ones that differed by less than .20 on composite scores. The three resulting stimulus sets comprised the passages used for the no TTS, regular TTS, and slow TTS conditions.

The researchers presented the reading passages on a MacBook Pro computer with "Alex" selected as the voice output both for the regular TTS and slow TTS conditions. The stimuli were presented at a comfortable listening level through noise-canceling headphones in both TTS conditions. Using system preferences, the researchers set the speaking rate at "normal" for regular TTS and halfway between "slow" and "normal" for slow TTS. The resulting average presentation rate of 155.44 WPM (SD = 15.53, range = 131.54–183.33 WPM) in the regular TTS condition was comparable with the recommended 150 to 160 WPM rate of presentation for audio books (Williams, 1998). For the slow TTS condition, the selected presentation rate resulted in speech output at 97.59 WPM (SD = 7.21, range = 80.49–111.74 WPM). This rate was comparable with Kay's habitual reading rate in the no TTS condition (M = 102 WPM, SD = 28.06, range = 68.11–147.40 WPM).

Procedures

Kay participated in nine sessions, reading and responding to questions relating to four passages during each session. Each of the first six sessions included passages presented in the no TTS and regular TTS conditions; the last three sessions included passages presented in the slow TTS condition only. In the no TTS condition, Kay read silently to herself while the researcher measured the elapsed time to determine her habitual reading rate in WPM. In the regular and slow TTS conditions, Kay read silently to herself while simultaneously listening to the TTS feature of the computer. After each passage, the researcher presented the five associated multiple-choice comprehension questions and response choices. To prevent reading comprehension difficulties from interfering with Kay's understanding of the passage questions, the researcher read the questions and response options aloud while simultaneously showing them to Kay. Kay verbalized her answers, and the researcher recorded them on a response form.

Data Analysis

The dependent variables were Kay's reading rates and comprehension accuracy in the three experimental conditions. The researchers used repeated measure analyses of variance (ANOVAs) and Fisher's least significant difference (LSD) testing to determine the presence of any significant differences across conditions. The researchers also asked Kay for her preferences and self-evaluation of her reading comprehension accuracy given the three experimental conditions.

Results

Kay's habitual reading rate in the no TTS condition ranged from 68.11 to 147.40 WPM (M = 102 WPM, SD = 28.06). As expected, a significant difference in the reading rate existed across the three conditions, $F(2, 35) = 29.94$, $p < .000$. Computation of Fisher's LSD confirmed the expected finding of no significant difference between slow TTS and no TTS conditions, $t(1, 11) = .156$, $p = .879$. Computation of Fisher's LSD also confirmed the expected significant difference between regular TTS and slow TTS conditions, $t(1, 11) = -11.480$, $p = .000$, and regular TTS and no TTS conditions, $t(1, 11) = 6.282$, $p = .000$.

Kay's accurate responses to the five comprehension questions associated with a given passage ranged from 0 to 4 across the 36 trials. Her overall comprehension accuracy score across passages in the no TTS condition was 13 of 60 (per passage range = 0–3, M = 1.6, SD = 0.98); in the regular TTS condition, her comprehension accuracy score was 20 of 60 (per passage range = 0–4, M = 1.08, SD = 1.24); and in the slow TTS condition, she scored 18 of 60 (per passage range = 0–3, M = 1.5, SD = 0.90). Computation of a repeated measure ANOVA revealed no significant accuracy difference across conditions, $F(2, 35) = 0.977$, $p = .387$.

Kay reported a preference for both TTS conditions to the no TTS condition and stated she would choose to use TTS in the future given the opportunity. She erroneously reported believing her comprehension was better in both TTS conditions than in the no TTS condition.

Discussion

The major finding in this pilot study was that regular TTS increased Kay's reading rate while neither positively nor negatively affecting her comprehension. Kay's habitual silent reading rate was significantly slower than that of the average neurotypical reader. For people with slower-than-average reading rates but who could comprehend connected speech presentation at average or above-average rates, TTS may make reading more efficient when performing time-sensitive activities such as finishing a book to be discussed at a predetermined time, reviewing forms to be signed at doctor appointments, performing work assignments, and meeting deadlines.

Practitioners in communication disorders as well as researchers examining TTS have suggested that multimodality stimulus presentation may help reading comprehension, fluency, speed, and concentration (Elkind et al., 1996; Hale et al., 2005; Hecker et al., 2002; Higgins & Raskind, 2004; Hux et al., 2008; Lasker et al., 1997; Lindstrom, 2007). In this single-case pilot study, the addition of auditory support to a reading task did not result in improved comprehension, thus calling into question the benefits of TTS as a means of providing multimodality reading support for individuals with aphasia. However, other factors may have influenced the performance of the individual case presented herein. In particular, the participant's slow reading rate may have created a processing speed mismatch between the auditory and written language modalities in the regular TTS condition. This possibility prompted the researchers to slow the TTS presentation rate, because they hypothesized that synchronizing the TTS rate to match more closely the participant's habitual reading rate would result in improved comprehension. The results from the slow TTS condition, however, did not confirm this hypothesis, and, as such, other explanations for the lack of significant reading comprehension improvement warrant consideration.

One possibility is that the TTS rate selected by the researchers for the slow TTS condition was not optimal. Given the opportunity to self-select a rate, people with reading challenges associated with aphasia may choose a presentation rate that differs from their silent reading rate. Although selecting a person's habitual silent reading rate as a target TTS rate seems logical, empirical evidence supporting this assumption does not exist. In fact, additional research may reveal that an even slower rate of TTS presentation is most beneficial. Alternately, future research may show that the manner of slowing the auditory presentation of written material is of greater importance than the actual WPM decrease. Other research suggests that inserting pauses between key words, phrases, or sentences may have a greater positive effect on reading comprehension than an across-the-board slowing of rate (Aarons, 1994).

Another possibility is that the researchers did not manipulate the most salient TTS variable. Their decision to decrease the rate of TTS presentation was only one of many possible variables available for investigation. Previous researchers have identified other factors that could potentially affect reading comprehension: (a) the quality of the auditory TTS signal (e.g., intelligibility and naturalness of prosody; Kimelman, 1999; Olson, Foltz, & Wise, 1986); (b) the inclusion of other multimodal supports such as visual graphics (Dietz et al., 2009); and (c) the incorporation of aphasia-friendly features in the written material (e.g., font size, amount of white space, and passage simplicity; Worrall, Papanathansiou, & Sherratt, 2013).

Some aspects of digitized speech provided via TTS technology are not easy to modify. In this study, the researchers selected the MacBook Pro computer for TTS presentation, because they judged the "Alex" option as superior to other available e-readers and TTS software applications with respect to speech clarity. Despite this judgment, the output was unmistakably synthesized and less natural than human speech in tone and timbre. As a rule, prosody is an aspect of personal communication missing from synthesized speech. The natural human voice supplies meaning through inflection and suprasegmental cues. Kimelman (1999) found that individuals with aphasia demonstrated improved listening comprehension when speakers stressed unpredictable target words during the presentation of connected speech. Because these meaningful cues are not available via TTS technology, people with aphasia may not benefit from TTS as a multimodality support, but they may find the addition of natural speech advantageous.

This study was piloted to explore the use of TTS as a reading accommodation for individuals with aphasia. Technological advances have increased the amount of communication that takes place through written language now more than ever before. In particular, people routinely text, e-mail, and Google to obtain and transfer information of all types. As such, reading has joined the ranks of dressing and bathing as an activity of daily living. Further investigation of the benefits and challenges associated with comprehending TTS may provide guidelines about supporting people with aphasia to better participate in the world around them.

References

Aarons, B. (1994). Techniques, perceptions, and applications of time-compressed speech. Proceedings of the 6th ACM International Conference on Multimedia, 97–103.

Brown, J., Fishco, V., & Hanna, G. (1993). *The Nelson-Denny Reading Test*. Geneva, IL: Riverside.

Cengage Learning. (2012). *Student Resource Center. Reading: ACE practice tests*. Retrieved from http://college.cengage.com/devenglish/resources/reading_ace/students/index.html

Devlin, S., & Unthank, G. (2006). Helping aphasic people process online information. *Proceedings from ASSETS 2006: The Eighth International ACM SIGACCESS Conference on Computers and Accessibility*, 225–226.

Dietz, A., Ball, L., & Griffith, J. (2011). Reading and writing with aphasia in the 21st century: Technological applications of supported reading comprehension and written expression. *Topics in Stroke Rehabilitation*, *16*, 758–769.

Dietz, A., Hux, K., McKelvey, M., Beukelman, D., & Weissling, K. (2009). Reading comprehension by people with chronic aphasia: A comparison of three levels of visuographic contextual support. *Aphasiology*, *23*, 1053–1064.

Elkind, J., Black, M. S., & Murray, C. (1996). Computer-based compensation of adult reading disabilities. *Annals of Dyslexia*, *46*, 159–186.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical of grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Grant, D., & Berg, E. (1993). *Wisconsin Sorting Test*. Luz, FL: Psychological Assessment Resources.

Hale, A. D., Skinner, C. H., Winn, B. D., Oliver, R., Allin, J. D., & Molloy, C. M. (2005). An investigation of listening and listening-while-reading accommodations on reading comprehension levels and rates in students with emotional disorders. *Psychology in the Schools*, *42*, 39–51.

Hecker, L., Burns, L., Elkind, J., Elkind, K., & Katz, L. (2002). Benefits of assistive reading software for students with attention disorders. *Annals of Dyslexia*, *52*, 243–272.

Hiebert, E., Samuels, S., & Rasinski, T. (2012). Comprehension-based silent reading rates: What do we know? What do we need to know? *Literacy Research and Instruction*, *51*, 110– 124.

Higgins, E. L., & Raskind, M. H. (2004). Speech recognition-based and automaticity programs to help students with severe reading and spelling problems. *Annals of Dyslexia*, *54*, 365– 392.

Hux, K., Weissling, K., & Wallace, S. (2008). Communication-based interventions: Augmented and alternative communication for people with aphasia. In R. Chapey (Ed.), *Language intervention strategies in aphasia and neurogenic communication disorders* (5th ed., pp. 814–836). Philadelphia, PA: Lippincott.

Kertesz, A. (2007). *Western Aphasia Battery–Revised*. San Antonio, TX: PsychCorp.

Kimelman, M. D. (1999). Prosody, linguistic demands, and auditory comprehension in aphasia. *Brain and Language*, *69*, 212–221.

LaPointe, L. L., & Horner, J. (1998). *Reading Comprehension Battery for Adults–2*. Austin, TX: ProEd.

Lasker, J. (2011). Assistive technology. In L. LaPointe (Ed.), *Aphasia and related neurogenic language disorders* (4th ed., pp. 185–196). New York, NY: Thieme.

Lasker, J., Hux, K., Garrett, K. L., Moncrief, E. M., & Eischeid, T. J. (1997). Variations on the written choice communication *Harvey et al.* 43 strategies for individuals with severe aphasia. *Augmentative and Alternative Communication*, *13*, 108–116.

Lindstrom, J. H. (2007). Determining appropriate accommodations for postsecondary students with reading and written expression disorders. *Learning Disability Research and Practice*, *22*, 229–236.

Olson, R., Foltz, G., & Wise, B. (1986). Reading instruction and remediation with the aid of computer speech. *Behavior Research Methods, Instruments & Computers*, *18*, 93–99.

Readability Formulas. (2012). *Free readability consensus calculator*. Retrieved from http://www.readabilityformulas.com/ free-readability-formula-tests.php

Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation* (2nd ed.). Tucson, AZ: Neuropsychology Press.

Triandafilou, J. S. (2003). Reading for life: A book club for individuals with aphasia. *ASHA Leader*, *8*, 14.

Williams, J. R. (1998). Guidelines for the use of multimedia in instruction. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, *42*, 1447–1451.

Worrall, L., Papanathansiou, I., & Sherratt, S. (2013). Therapy approaches to aphasia. In I. Papanathansiou, P. Coppens, & C. Potogas (Eds.), *Aphasia and related neurogenic communication disorders* (pp. 93–112). Burlington, MA: Jones and Bartlett.