

EFFICACY OF COMPUTERIZED APHASIA TREATMENT: PAST RESEARCH AND FUTURE DIRECTIONS

Richard C. Katz^{1,2}

¹ Phoenix VA Health Care System, Phoenix, Arizona, USA

² Arizona State University, Tempe, Arizona, USA

Key words: *aphasia, therapy, efficacy, computer, computerized therapy*

SUMMARY

Background. *Computerized treatment has evolved considerably over the past 30 years, reflecting changes in treatment strategies as well as advances in technology. The purpose of this report is to review past outcome research and provide direction for future computerized treatment for aphasia.*

Methods. *Relevant literature is described using the five phase treatment outcome model (Robey & Shultz, 1998) for evaluating research. In addition, the application of the Internet and related technology are also considered.*

Conclusions. *Computerized aphasia treatment has been demonstrated to be efficacious in one Phase 3 RCT study (Katz & Wertz, 1997) and in reports of meta-analyses (e.g., Robey, 1998). The Internet offers the opportunity to provide automated computerized treatment to patients in their homes and other locations, outside the traditional confines of the clinic.*

INTRODUCTION

Computers are used in aphasia rehabilitation in three different manners: to assist in communication (e.g., Dynavox, Lingraphica), to assist clinician-provided treatment by presenting stimuli and storing performance, and to provide treatment activities for patients working alone in the absence of a clinician or other. It is this third condition, computer-only treatment, that we are describing in this document.

Computer-only treatment requires that the computer and software be easily accessible and understandable to the patient without assistance from ano-

ther person. The interface must be familiar or at least intuitive, so that use of the treatment software is not impeded by factors unrelated to the program's purpose or content. Error-handling and feedback should reflect the patient's needs as well as the programmer's. Movement through the program should be automatic and incorporate complex branching algorithms to present a variety of appropriate cues and interventions. Levels of difficulty should be hierarchically arranged. In most cases, stimuli should be presented:

- rapidly to improve efficiency, and
- randomly to minimize familiarity with the order of stimulus presentation.

Computer-only treatment can accommodate a number of different treatment models and combinations of models, but the particular attributes of computer-only treatment seem best suited for the stimulation-facilitation model. (Schuell, Jenkins & Jiménez-Pabón, 1964).

According to Fredrick L. Darley (1972), rather than asking the nonspecific question, "Does therapy work?", we should be more specific and ask: "For whom does it work? What kind of treatment works? When do you treat? How much treatment do you provide? How often do you treat? Under what conditions do you treat?" To answer these questions in a comprehensive and standardized way, Robey (1998) and Robey & Shultz (1998) adapted the Five Phase Outcome Research Model (World Health Organization, 2001) for aphasia research (cf: Wertz & Irwin, 2001; Wertz & Katz, 2004).

Phase 1 studies are single subject and small group studies designed initially to detect if treatment is active by measuring outcome at two or more points in time. Phase 2 studies replicate Phase 1 studies with modifications to refine the research question. The purpose of Phase 3 studies is to measure efficacy under optimal conditions by utilizing large subject samples randomly assigned to treatment and no-treatment groups, incorporating powerful inferential statistics, and usually conducted at several different sites to minimize institutional bias. Phase 4 studies measure effectiveness, that is, outcomes under real world conditions, such as performance in actual clinics, rather than optimal conditions. Phase 5 studies measure efficiency by comparing outcomes resulting from two different treatment protocols. The treatment examined in both Phase 4 and 5 studies should have been previously

Table 1. Five-phase research study design by Robey & Shultz (1998), after World Health Organization (2001)

Phase	Description
Phase 1	Is Treatment Active?
Phase 2	Refine Phase I Question
Phase 3	Efficacy
Phase 4	Effectiveness
Phase 5	Efficiency

shown to be efficacious in an earlier Phase 3 study. A similar distinction among study designs is demonstrated in the Levels of Evidence proposed by the American Academy of Neurology (2001). Class 1 studies are randomized control studies and meta-analyses and are similar to Phase 3 studies. Class 2 are observational studies and include case studies and cohort studies. Class 2 studies are similar to Phase 1 and Phase 2 studies. The Class 3 level of evidence does not describe experimental studies at all, but includes expert opinion, nonrandomized historical reports, and professional consensus, a valid, although weaker, form of evidence. Describing levels of evidence in this manner allows for a systematic evaluation of existing support for current and proposed treatment procedures when constructing evidence-based clinical practice guidelines (e.g., Duncan, Zorowitz, Bates, Choi, Glasberg, Graham, Katz, Lamberty & Reker, 2005; Frattali, Bayles, Beeson, Kennedy, Wambaugh, & Yorkston, 2003).

PAST RESEARCH

One of the earliest reports of a computerized aphasia treatment outcome study described a program that incorporated complex branching algorithms to provide multilevel intervention to improve spelling graphically (typing). Seron, Delouche, Moulard & Rouselle (1980) described a computer-assisted program that provided cuing for aphasic patients learning to type words from dictation. The clinician said the target word out-loud and the subject typed a response on the computer keyboard. (The clinician had to know in advance the order of the stimuli programmed in the computer.) Intervention consisted of three levels of feedback: the number of letters in the target word; whether the letter typed was in the word; and when the correct letter was typed, whether that letter was in the correct position. The five subjects completed the program in 7 to 30 sessions. Pre- and post-treatment tests required the subjects to write a generalization set of single words to dictation. A decrease ($p < 0.05$) in the number of misspelled words and in the total number of errors made on the post-treatment test suggested that the computer program had improved spelling of words written by hand. Four of the five subjects maintained improved performance on a second post-treatment test administered 6 weeks later.

Building from the results reported by Seron et al. (1980), Katz and Nagy (1984) developed a Phase 1 Class 2 study to test a program that used complex branching steps to evaluate responses and provide patients with specific feedback in a computer-only task that involved confrontation naming, spelling, typing and handwriting. A stimulus was randomly selected by the program, then a graphic representation of the stimulus was presented on the monitor. The subject responded by typing on the keyboard. Feedback consisted of auditory sounds (for correct and incorrect) and text printed on the screen. Single and multiple cues from a hierarchy of six (Table 1) were selected by the program in response to the number of errors made for each of seven stimuli. A seven-point scoring system was used to describe perform-

ance and track the effectiveness of the various cues. Additional feedback included repetition of the successful and most recently failed cues. At the end of the computer session, pencil-and-paper copying assignments automatically generated via the computer printer were completed by the subject. Pre- and post-writing tests revealed improved spelling of the target words for seven of the eight aphasic subjects ($p < 0.01$).

Katz and Wertz (1997) conducted a Phase 3 Class 1 longitudinal randomized control treatment group study to investigate the effects of computerized language activities and computer stimulation on language test scores for chronic aphasic adults. Fifty-five chronic aphasic subjects who were no longer receiving speech-language therapy were randomly assigned to one of three conditions:

- 78 hours of Computer Reading Treatment;
- 78 hours of Computer Stimulation ("non-language" activities);
- No treatment.

The Computer Reading Treatment software consisted of 29 activities, each containing eight levels of difficulty, totaling 232 different tasks. Treatment tasks required visual-matching and reading comprehension skills, displayed only text (no pictures), and used a standard, match-to-sample format with two to five multiple choices. Treatment software automatically adjusted task difficulty in response to subject performance by incorporating traditional treatment procedures, such as hierarchically arranged tasks and measurement of performance on baseline and generalization stimulus sets, in conjunction with complex branching algorithms. Software used in the Computer Stimulation condition was a combination of cognitive rehabilitation software and computer games that used movement, shape, and/or color to focus on reaction time, attention span, memory, and other skills that did not overtly require language or other communication abilities. Subjects in the two computer conditions worked on the computer for 3 hours per week for 26 weeks. Clinician interaction during the two computer conditions was minimal. Subjects from all three conditions were tested using the Porch Index of Communicative Ability (PICA) (Porch, 1981) and Western Aphasia Battery

Table 2. Scores and corresponding hierarchical cues

Attempt	Score if correct	Intervention if error
1 ST	7	Repeat stimulus
2 ND	6	Anagram without feedback
3 RD	5	Delayed response after model
4 TH	4	Anagram with feedback
5 TH	3	Multiple choice cue
6 TH	2	Immediate response with model
7 TH	1	Flag stimulus for clinician/Score = 0

(WAB) (Kertez, 1982) at baseline, three months, and six months. Significant improvement over the 26 weeks occurred on five language measures for the computer reading treatment group, on one language measure for the computer stimulation group, and on none of the language measures for the no-treatment group. The computer reading treatment group displayed significantly more improvement on the PICA Overall and Verbal modality percentiles, and on the WAB Aphasia Quotient and Repetition subtest, than the other two groups. The results suggest that:

- computerized reading treatment can be administered with minimal assistance from a clinician;
- improvement on the computerized reading treatment tasks generalized to non-computer language performance;
- improvement resulted from the language content of the software and not stimulation provided by a computer;
- the computerized reading treatment we provided to chronic aphasic patients was efficacious.

Over a decade earlier, Vaughn (1980) and her associates (Vaughn, Amster, Bess, Gilbert, Kearns, Rudd, Tidwell, & Ozley, 1987) adapted computer technology in a different direction, using small, microprocessor-driven devices connected to minicomputers, like the PDP-11, via telephone modems, to present reading, writing, listening and speaking tasks to aphasic and dysarthric patients. REMATE (Remote Machine-Assisted Treatment and Evaluation) was designed to bring diagnosis and treatment to patients with communication disorders who were un-served, underserved, or living in remote areas. Although their publications focused more on feasibility than on outcome, they developed and established the infrastructure for remote treatment that leads up to our present efforts.

FUNCTI ONS

Find the word below that goes best with the word in the box.

Restaur ant

1. Si t
2. Buy
3. Eat
4. Drive
5. Sleep
6. Put

Select the key (1-6) that goes best with the word: __

Figure 1. Example computer display for Katz & Wertz (1997)

FUTURE DIRECTIONS

Katz & Ross (2004) proposed a study designed to create and evaluate an automated, multi-level, comprehensive, computerized auditory stimulation program provided via the Internet for treating a variety of aphasic communication impairments. The program will present an interactive series of language activities to aphasic patients. To optimize learning, the program will utilize complex branching algorithms (Katz & Wertz, 1997; Katz & Nagy, 1984) that monitor accuracy and adjust task difficulty and, when indicated, provide cuing and other interventions, according to each patient's pattern of response. Use of the internet will allow patients to access the program from various locations, and clinicians to easily monitor progress and modify the treatment material accordingly.

A prototype of the program, called SPARK (Stimulation Program for Assisting Rehabilitation by Katz & Karow), developed with Dr. Colleen Karow from the University of Tennessee, is being tested as a Phase 1 study, designed to determine if the treatment is "active." Ten aphasic subjects will each receive 50 hours of computerized treatment, five one-hour sessions per week for ten weeks (Caspari & Katz, 2001). Initially, subjects will be trained for up to three additional sessions to use the computer and treatment program interface (Katz & Wertz, 1992; Petheram, 1988). All treatment will be delivered in the clinic, with a clinician present only to provide initial orientation, to monitor progress, and to insure compliance. Any assistance required from the clinician (e.g., patient unable to complete a task after repeated attempts) will be carefully documented and incorporated into refinement of the treatment protocol (Phase 2).

Outcome measures will be administered to all study patients twice, once upon entry into the study and again at the conclusion of subject participation. At each point, measurements will be administered in random order within a seven day period. Measures will be administered by a trained speech-language pathologist, licensed by the state and certified by the American Speech-Language-Hearing Association (ASHA). Tests will be administered according to each test's standardized (or recommended) protocol. Scores derived from at least 25% of the tests administered will be reviewed for accuracy by a second speech-language pathologist who has no knowledge of subject task performance. Scoring discrepancies will be resolved by consensus or by a third speech-language pathologist.

The outcome measures are:

- the WAB, to classify aphasia by type and measure language impairment;
- the *Revised Token Test* (McNeil & Prescott, 1978), to measure auditory comprehension ability.

Effect size (Robey, Schultz, Crawford & Sinner, 1999) for each test will be calculated, e.g., an improvement of 5 or more AQ points on the WAB (Katz & Wertz, 1992; 1997).

To meet the needs of a heterogeneous aphasic population, the treatment material developed is suitable for aphasic patients demonstrating a range of levels of impairment. Presently, the treatment hierarchy consists of 112 treatment tasks arranged in order of presumed difficulty by systematic manipulation of five parameters. The two stimulus parameters are

- frequency of occurrence of stimulus word (frequent and infrequent);
- semantic confusion, i.e., semantic distance between stimulus word and multiple choice foils (different and related).

The three task (presentation) parameters are:

- number of multiple choices (2-6);
- number of critical elements (1-5);
- duration of forced delay of response in seconds (0, 5).

The five parameters have a total of 16 variables, and are combined in a sequence to create 112 distinct tasks. Each task has 10 stimuli, totaling 1120 stimulus item presentations for each patient completing the hierarchy.

Task structure and response requirements are consistent among tasks. All tasks use a standard, match-to-sample format that displays two to six response choices. Stimuli consist of spoken (digitized) speech presented by the computer through high quality speakers, positioned on either side of the computer monitor. Response choices are displayed in one or two rows centered on the screen. Subjects respond by using the mouse to move a large cursor (an arrow) to point to the desired response item(s) and clicking the button on the mouse to indicate the selection of the response item(s). For multiple critical elements, responses do not have to be in the correct order to be considered correct. Activities and tasks are arranged sequentially in a hierarchy of assumed difficulty, determined by the stimulus, task, and the number and type of response foils.



Figure 2. Example of stimulus screen with multiple foils for SPARK

In the first task in the hierarchy (Task #1), the computer produces (via digitized speech) one frequent word, while simultaneously displaying on the screen a picture representing that same word and another picture (the foil) representing an unrelated frequent word. Half-way through the hierarchy, during Task #56, the computer produces (via digitized speech) four infrequent words while simultaneously displaying on the screen six pictures representing the four spoken words plus two other, semantically related, words. In the next task (Task #57), the computer produces (via digitized speech) one frequent word, and after a five-second delay, displays on the screen a picture representing that same word and another picture (the foil) representing an unrelated frequent word. In the last task in the current hierarchy (Task # 112), the computer produces (via digitized speech) four infrequent, related words and, after a delay of 5 seconds, displays on the screen pictures representing those same four words and two other words, also semantically related and infrequently used.

The program automatically provides feedback (correct or incorrect) each time a patient responds to a stimulus and moves on to the next item. If the response is correct, a message is provided indicating the correct response was selected and, after a delay, the program presents the next stimulus (and response array) in the task. If the response is incorrect, a message is provided indicating the wrong response was selected, the correct response is indicated, and, after a delay, the program presents the next stimulus (and response array) in the task. Feedback is limited to correct or incorrect only; corrective feedback and second chances are not provided, as the objective of the stimulation-facilitation approach is to facilitate and maximize reorganization and recovery of language by providing the aphasic patient with the opportunity to respond quickly and frequently to numerous and intense auditory stimuli. When the patient has responded to the last (10th) stimulus in a task, a summary of the task performance is presented (task accuracy, list of correct items, list of items on which errors were made, etc.). When a patient accurately responds to at least 80% of the stimuli within a task, the program automatically moves up the hierarchy to the next task.

If the patient does not achieve 80% or greater after a task is attempted five times, the program automatically moves down the hierarchy to the last task successfully completed, and alerts the clinician to monitor the patient's performance should criterion not be reached and clinician intervention be required. A patient's participation in the program will be discontinued after:

- clinician intervention is required for three consecutive tasks;
- 50 hours of treatment have been completed;
- the patient completes the hierarchy;
- the patient requests termination of treatment.

Because the intent of this study is to test the effect of auditory stimulation software, the software and protocols are designed to require minimal involvement from a clinician while the subject uses the computer. The research clinician, however, has many responsibilities in the course of this study. The cli-

nician verifies attendance, instructs the subject how to access and operate the software, and monitors performance. If the subject fails to achieve criterion on a particular task, the clinician observes subject performance more closely to determine what intervention is necessary in case the subject requires help, recording in a logbook the occurrence and content of any assistance provided. Unless performing one of these responsibilities, the clinician remains close, but out of direct line of sight while the subject runs the program.

DISCUSSION

This paper has briefly described various efforts to develop and measure the effect of automated computerized treatment programs used independently by aphasic adults. These computer-only treatment studies utilized programs that presented stimuli automatically and rapidly and incorporated complex branching algorithms that presented hierarchically arranged cues. Although a variety of treatment models can be used in computer-only treatment programs, stimulation-facilitation appears to be the treatment model best suited for independent computer use. The research described includes small Phase 1 and Phase 2 studies on single computers and a large multi-center randomized control Phase 3 study. Our current project to develop remote, interactive, computerized treatment over the internet will permit patients with aphasia to receive treatment from their homes or other convenient and familiar locations.

ACKNOWLEDGEMENT

An earlier form of this paper was presented at the Wernicke-Alzheimer Memorial Conference held in Wroclaw, Poland, December 7-10, 2005.

REFERENCES

- American Academy of Neurology Practice Guidelines - Practice Statement Definitions (2001).
Caspari, I. & Katz, R. (2001, June). Comparative effects of ten or twenty hours of treatment on treatment effectiveness in aphasia. Poster presented at the Clinical Aphasiology Conference, Santa Fe, New Mexico, USA.
Darley, F. L. (1972). The efficacy of language rehabilitation in aphasia. *Journal of Speech and Hearing Research*, 37, 3-21.
Duncan, P. W., Zorowitz, R., Bates, B., Choi, J. Y., Glasberg, J. J., Graham, G. D., Katz, R. C., Lamberty, K. & Reker, D. (2005). Management of adults stroke rehabilitation care: A clinical practice guideline. *Stroke*, 36, 100-143.
Frattali, C., Bayles, K., Beeson, P., Kennedy, M. R. T., Wambaugh, J., & Yorkston, K. M. (2003). Development of evidence-based practice guidelines: Committee update. *Journal of Medical Speech-Language Pathology*, 3(11), ix-xviii.
Katz, R. C. & Ross, K. B. (2004). An internet-based automated auditory stimulation treatment for aphasic adults. *Acta Neuropsychologica*, 2(1), 25-34.
Katz, R. C., & Nagy, V. T. (1984). An intelligent computer-based task for chronic aphasic patients. In R. H. Brookshire (Ed.), *Clinical aphasiology: 1984 conference proceedings* (pp. 159-165). Minneapolis, MN: BRK.

- Katz, R. C., & Wertz, R. T. (1992). Computerized hierarchical reading treatment in aphasia. *Aphasiology*, 6, 165-177.
- Katz, R. C., & Wertz, R. T. (1997). The efficacy of computer-provided reading treatment for chronic aphasic adults. *Journal of Speech, Language and Hearing Research*, 40(3), 493-507.
- Kertesz, A. (1982). *Western aphasia battery*. New York: Grune & Stratton.
- McNeil, M.R. & Prescott, T.E. (1978). *Revised Token Test*. Austin, Texas: Pro-Ed.
- Petheram, B. (1988). Enabling stroke victims to interact with a minicomputer -- comparison of input devices. *International Disabilities Studies*, 10(2), 73-80.
- Porch, B. E. (1981). *Porch Index of Communicative Ability*, Vol. 1: Administration, scoring and interpretation (3rd ed.). Palo Alto, CA,: Consulting Psychologists Press.
- Robey, R.R. (1998). A meta-analysis of clinical outcomes in the treatment of aphasia. *Journal of Speech, Language, and Hearing Research*, 41, 172-187.
- Robey, R.R., & Schultz, M.C. (1998). A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology*, 12, 787-810.
- Robey, R.R., Schultz, M.C., Crawford, A.B. & Sinner, C.A. (1999). Single-subject clinical-outcome research: designs, data, effect sizes, and analyses. *Aphasiology*, 13, 445-473.
- Schuell, H., Jenkins, J. J., & Jiménez-Pabón, E. (1964). *Aphasia in adults*. New York: Harper & Row.
- Seron, X., Deloche, G., Moulard, G., & Rouselle, M. (1980). A computer-based therapy for the treatment of aphasic subjects with writing disorders. *Journal of Speech and Hearing Disorders*, 45, 45-58.
- Vaughn, G. R. (1980, August). REMATE (Remote Machine Assisted Treatment and Evaluation): Communication outreach innovative health care delivery system for persons with communicative disorders workshop. Birmingham: Veterans Administration Medical Center.
- Vaughn, G.R., Amster, W.W., Bess, J.C., Gilbert, D.J., Kearns, K.P., Rudd, A.K., Tidwell, A.A. & Ozley, C.F. (1987). Efficacy of remote treatment of aphasia by TEL-Communicology. *Journal of Rehabilitative Research and Development*, 25(1), 446-447.
- Wertz, R. T. & Katz, R. C. (2004). Outcomes for computer-provided treatment for aphasia. *Aphasiology*, 18(3), 229-244.
- Wertz, R.T. & Irwin, W.H. (2001). Darley and the efficacy of language rehabilitation in aphasia. *Aphasiology*, 15, 231-247.
- World Health Organization. (2001). *International classification of functioning, disability, and health (ICF)*. Geneva: World Health Organization. [Online]. Available: <http://www3.who.int/icf/icftemplate.cfm> [2004, March 18].

Address for correspondence:

Richard C. Katz, Ph.D.
Phoenix VA Health Care System
Phoenix, Arizona, USA
e-mail: richard.katz@va.gov

Received: 21 April 2007

Accepted: 28 March 2008