



Published in final edited form as:

Am J Speech Lang Pathol. 2014 May ; 23(2): S343–S360. doi:10.1044/2014_AJSLP-13-0097.

Acquisition and Maintenance of Scripts in Aphasia: A Comparison of Two Cuing Conditions

Leora R. Cherney^{1,2}, Rosalind C. Kaye¹, and Sarel van Vuuren^{3,4}

¹Rehabilitation Institute of Chicago, Center for Aphasia Research and Treatment, Chicago, IL

²Northwestern University, Feinberg School of Medicine, Department of Physical Medicine and Rehabilitation, Chicago, IL

³University of Colorado Boulder, Institute of Cognitive Science, Boulder, CO

⁴University of Colorado Anschutz Medical Campus, School of Medicine, Department of Physical Medicine and Rehabilitation, Denver, CO

Abstract

Purpose—This study was designed to compare acquisition and maintenance of scripts under two conditions: High Cue which provided numerous multimodality cues designed to minimize errors, and Low Cue which provided minimal cues.

Methods—In a randomized controlled cross-over study, eight individuals with chronic aphasia received intensive computer-based script training under two cuing conditions. Each condition lasted three weeks, with a three-week washout period. Trained and untrained scripts were probed for accuracy and rate at baseline, during treatment, immediately post-treatment, and at three and six weeks post-treatment. Significance testing was conducted on gain scores and effect sizes were calculated.

Results—Training resulted in significant gains in script acquisition with maintenance of skills at three and six weeks post-treatment. Differences between cuing conditions were not significant. When severity of aphasia was considered, there also were no significant differences between conditions, although magnitude of change was greater in the High Cue condition versus the Low Cue condition for those with more severe aphasia.

Conclusions—Both cuing conditions were effective in acquisition and maintenance of scripts. The High Cue condition may be advantageous for those with more severe aphasia. Findings support the clinical use of script training and importance of considering aphasia severity.

Cuing is recognized as an integral routine part of aphasia treatment, but the amount and pattern of cuing varies widely, and its application in clinical settings is not always done in a

Corresponding author: Leora R. Cherney, PhD, Center for Aphasia Research and Treatment, Rehabilitation Institute of Chicago, 345 East Superior Street, Suite 1353, Chicago, IL 60611, Phone. 312-238-1117; Lcherney@ric.org.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Deafness and Other Communication Disorders or the National Institutes of Health. AphasiaScripts is a trademark of the Rehabilitation Institute of Chicago.

None of the authors have a financial interest in the software product.

systematic way. Cuing can promote or decrease error production in persons with aphasia (Abel, Schulz, Radermacher, Willmes & Huber, 2005; Conroy, Sage, & Lambon-Ralph, 2009a). Variables affecting error production include the amount of cues and the timing of cue presentation in relation to the required response. If participants receive sufficient cues before attempting a response, the response may be errorless (or error-reduced since ensuring the total elimination of errors is not possible). If participants attempt a response without prior cues or with only minimal cues, there is greater potential for error production. The use of cues and their relation to errorful and error-reducing treatment conditions is a relatively new area of investigation.

Most studies of errorless and errorful learning in aphasia have focused on naming. In errorless learning, the correct spoken or written name is presented along with the target object or concept, so that the person with aphasia is required only to repeat or copy it. In this way, naming errors are avoided (Conroy, Sage, & Ralph, 2009b; Fillingham, Sage, & Lambon Ralph, 2005, 2006; Mckissock & Ward, 2007; Raymer, Strobel, Prokup, Thomason, & Reff, 2010). According to this viewpoint, the pairing of a word with its target strengthens their association. If the correct association between the word and its target is repeated, then correct learning will occur. If an incorrect pairing is made, then that incorrect pairing also will be strengthened and there is an increased likelihood that the error response will be learned. Therefore errors should be avoided to prevent error learning. This type of associative training bypasses the need to retrieve knowledge from long-term memory.

In contrast, a large body of research suggests that learning lasts longest when long-term memory retrieval occurs. The act of retrieval (retrieval practice) strengthens learning, even if errors result. For example, in non-clinical populations, a testing effect has been demonstrated, i.e., after initial studying of target material, learning is boosted by testing rather than additional studying (Karpicke & Roediger, 2008; Roediger & Karpicke, 2006; Tulving, 1967). With “testing”, the individual is required to attempt retrieval from memory, a process that is more effective in the long term than studying, where no retrieval practice is involved. This distinction may be especially relevant to aphasia rehabilitation, where what must be learned is often not new knowledge, but the retrieval of linguistic knowledge that may still reside in long-term memory.

Relatively few studies have directly compared errorless and errorful learning in aphasia to determine the superiority of one condition over the other. In a review of a series of anomia studies, Fillingham and colleagues (Fillingham, Hodgson, Sage, & Lambon Ralph, 2003; Fillingham et al., 2006) found that rate of success was equivalent for errorless and errorful conditions; the number of therapies using errorful techniques outweighed those based on errorless learning; errorless approaches were likely to achieve positive immediate effects, but many did not report long term effects and generalization; and there was a lack of studies reporting a direct comparison of errorful and errorless learning. Although there was some indication that errorful training was more likely to promote long-term retention than errorless training, more recent work has continued to support the findings that there is no essential difference between errorful and errorless learning approaches (Middleton & Schwartz, 2012).

Results are less conclusive when increasing and decreasing cue therapies have been directly compared. Conroy et al., (2009a) found that both therapies were beneficial with no difference between increasing and decreasing cues for accuracy or speed of naming of nouns or verbs. However, participants preferred the decreasing cue therapy which provided them with more opportunities for successful naming and fewer instances of failure. Conversely, Abel et al (2005) demonstrated positive effects with increasing cues alone, or with a combination of increasing and decreasing cues, but not with decreasing cues alone. Furthermore, those with severe naming disorders benefitted more from cuing therapy in general than those with moderate naming disorders, but degree of impairment had no impact on the overall clear finding that there was an advantage to the increasing cues condition.

No studies have yet investigated and contrasted the errorless/errorful learning paradigm in the production of phrases and sentences, which is more representative of real-life functional communication. In script training, functional phrases and sentences are repeatedly practiced within the context of a dialogue or monologue. A growing body of evidence supports the use of script training in aphasia (Bilda, 2011; Cherney, Halper, Holland, & Cole, 2008; Cherney, Halper, & Kaye, 2011; Goldberg, Haley, & Jacks, 2012; Lee, Kaye, & Cherney, 2009; Manheim, Halper, & Cherney, 2009; Nobis-Bosch, Springer, Rademacher, & Huber, 2011; Youmans, Holland, Munoz, & Bourgeois, 2005; Youmans, Youmans, & Hancock, 2011). In most studies, cues and supports have been provided consistently before every response so that script practice may be characterized as potentially error-reducing; use of cues and supports promotes rapid script acquisition, decreases learning of persistent errors, and reduces client frustration (Youmans et al., 2005, 2011). Still, it is not known whether this cued condition is the preferred condition for promoting long-term script learning.

The rationale for script-training methodology is derived from the instance theory of automatization (Logan, 1988). The theory suggests that automaticity of skills is achieved by retrieving memories of complete, context-bound, skilled performances. These memories are formed with repeated exposures to and practice on the same task. The focus of instance theory and most other theories of automaticity is on the skill acquisition phase. However, long term retention (maintenance) may be a better index of learning. Although previous studies show that script training promotes acquisition, long term retention of the acquired scripts has not been systematically evaluated.

The purpose of this study was to evaluate script training under two cuing conditions that we have called High Cue and Low Cue. The High Cue condition has the intent of providing an error-reducing learning environment because participants are provided with numerous multimodality cues before responding such that errors are minimized. In contrast, participants in the Low Cue condition attempt a response with reduced or no cues, thereby creating a learning environment where both errors and retrieval practice are more likely. The impact of the High and Low Cue conditions on both acquisition and maintenance of script training was evaluated. Specifically, we addressed the following questions:

1. Does computer-based script training (via *AphasiaScripts*TM) result in the acquisition and maintenance of conversational scripts?

2. Do the High Cue and Low Cue training conditions differentially affect the acquisition and maintenance of conversational scripts?
3. Does the severity of aphasia, as measured by the Western Aphasia Battery-Revised (Kertesz, 2007) Aphasia Quotient (WAB AQ), impact the relationship between High and Low Cue training conditions and their effect on the acquisition and maintenance of conversational scripts?
4. Do persons with aphasia have a preference for one cue condition over the other?

We hypothesized that computer-based script training would result in the acquisition and maintenance of conversational script production. We also hypothesized a differential effect of the training conditions, with no difference between conditions for script acquisition, but an advantage of the Low Cue condition for script maintenance because of the importance of retrieval practice for long term learning. There was insufficient evidence to hypothesize whether severity would impact the relationship between High and Low Cue conditions, but we hypothesized that participants, regardless of severity, would prefer the High Cue condition since it would be less frustrating.

Method

Experimental Design

A randomized, controlled, crossover design was used to examine the effects of High Cue vs. Low Cue training conditions on the acquisition and maintenance of conversational script production. Each participant practiced a script under one training condition for three weeks, followed by a three-week washout period; then a second script was practiced for three weeks under the other training condition. For each participant, scripts were matched for length and grammatical complexity. However, they were adjusted across participants according to severity. Script production was probed for accuracy and rate. For each training condition, script performance was measured with three baseline probes, six treatment probes during the three weeks of practice, one post-treatment probe, and two maintenance probes at three and six weeks after the conclusion of treatment. The study was approved by the Institutional Review Board of Northwestern University.

Participants

Participants were eight individuals (6M; 2F) with chronic aphasia due to a single left-hemisphere stroke. Participants ranged in age from 25 through 66 years (mean=52.0, SD=14.0). Time post-onset was between 8 and 59 months (mean=26.4, SD=19.2), and education ranged from 11 to 18 years (mean=14.3, SD=2.3). Table 1 shows the demographic data and aphasia severity based on the Western Aphasia Battery-Revised Aphasia Quotient (WAB-R AQ; (Kertesz, 2007) of each participant. WAB-R AQ ranged from 28.1 to 80.1 (mean=58.0, SD=18.5), with all but one participant presenting with a nonfluent aphasia. To participate in the study, participants were required to be native speakers of English, pass a pure-tone hearing screening at 25 dB in the better ear, and have no history of active substance abuse, significant psychological problems, or neurological conditions other than stroke. None of the participants received other speech/language treatment for at least one month prior to the study or during the study.

Randomization—A WAB-R AQ cut-off of 60 was used to differentiate more severe and less severe aphasia groups. The sequence of the conditions (i.e., High Cue - Low Cue vs. Low Cue - High Cue) was randomized within each severity group. Participants were assigned to a severity group, and then a treatment condition sequence based on the order in which they were enrolled. This allowed each of the training condition sequences to be matched for aphasia severity as shown in Table 2.

Script Development—We developed six personalized scripts of equal length and grammatical complexity for each participant-- three for the High Cue condition and three for the Low Cue condition. In each condition, one script served as the trained script, one as the untrained script, and one as a generalization script. All scripts were dialogues, with the digital therapist initiating the conversation and the participant with aphasia responding.

Personalized scripts were developed from templates that we had written in advance. For each topic, there was a template for more severe participants and a similar but more complex one for less severe participants. Scripts were personalized with the participants' own choices, such as favorite restaurant names and food items, determined in an interview with the participant. Then length and complexity were adjusted to ensure that the six scripts for each participant were comparable in terms of number of words, sentences, syllables, verbs, morphemes, average words per sentence, average syllables per word, grammatical complexity (ratio of morphemes to words) and the Flesch Reading Ease score, a formula combining words per sentence and syllables per word. As an example, Table 3 shows the counts for each of the six scripts written for WELED, the subject with the highest AQ (80.1) and for SMIDA, the subject with the lowest AQ (28.1). Scripts always consisted of 10 turns.

Treatment Intervention

Treatment was provided using a computer program, *AphasiaScripts*[™]. Script training was delivered by an anthropomorphically accurate “digital” therapist capable of visually modeling speech and interactively guiding treatment. By removing clinician-related variables (e.g., clinician expertise, personality factors) that potentially influence treatment outcomes, treatment fidelity was ensured. The treatment software has experimental support (Cherney, Halper, Holland, & Cole, 2008; Lee, Kaye, & Cherney, 2009; Manheim, Halper, & Cherney, 2009). Furthermore, it accommodates the manipulation of variables such as cueing, allowing comparison of different conditions as described below.

AphasiaScripts[™] treatment has three parts. First, individuals with aphasia listen silently to the whole conversation while it appears on the screen. Then, participants repeatedly practice each turn of the conversational script, first in unison with the digital therapist and then independently. Third, the entire conversation is rehearsed while taking turns with the digital therapist. Fig. 1 is a screen shot of the digital therapist with a customized script during conversational practice. The digital therapist “speaks” the words with mouth movements similar to that of a real person and the word is highlighted as it is spoken. As the person with aphasia masters the script, cues are removed one by one by the participant, so that eventually practice is accomplished with the digital therapist in a situation that simulates real conversation. Cues that can be used or removed by the participant during conversation

practice include highlighting of the written word, the entire written sentence, the digital therapist's mouth movements, and auditory cues of the digital therapist's voice for choral production of the sentence.

Training Conditions—Appendix A shows the step-by-step procedures for the High Cue and Low Cue conditions. In the High Cue condition, *AphasiaScripts*[™] was administered in its current form. The participant was given all the available cues during sentence and conversation practice prior to any independent attempts at production. In the Low Cue condition, *AphasiaScripts*[™] was modified so that during sentence and conversation practice, only written sentences were provided, without any auditory or oral-motor cues from the digital therapist. The participant was required to produce each turn independently for three trials. Only then did the digital therapist provide a correct model, allowing the participant to listen to the correct production and watch the oral motor movements during a choral production of the turn. Training in each condition continued for three weeks, with a three-week washout period between conditions. Participants practiced six days a week for 90 minutes a day at home on a loaned laptop. The 90 minutes of daily practice were divided into three 30-minute sessions. Participants consistently received the same amount and type of cuing as designated by the condition that they were assigned and as illustrated in Appendix A. Within that condition, they could choose how many times to practice a specific sentence of the script. They could also select how many times to complete the sentence practice versus the conversation practice during their scheduled 30 minute treatment session. More severe participants were instructed to do sentence practice only during the first and second weeks, and to add conversation practice in the third week. Less severe participants added the conversation practice to the sentence practice in the second week. This was true in both the High and Low Cue conditions.

To ensure compliance, the computer program captured log on and log off times, every key stroke made by the participant during the daily treatment sessions, and audio recordings of the last two attempts at sentence production (one spoken independently and one spoken chorally) during the sentence practice sequence. Participants were also provided with a paper-and-pencil log to complete with the start and finish times of their daily treatment sessions. Participants visited the clinic once a week. During these visits, the research speech-language pathologist (SLP) downloaded the computer logs and checked them against the paper-and-pencil logs to ensure that the participant had practiced the designated amount of time during the week. She also observed the participant do one 30-minute practice session to ensure that he or she was attempting each part of the treatment sequence.

Data collection and analysis

Probes—Treatment and probe sessions were set up by the research SLP via a calendar function integrated into the computer software. This function ensured that probe sessions were conducted at the beginning of the day always prior to any daily treatment sessions. Trained and untrained scripts were probed the same number of times during baseline, treatment and post-treatment whereas generalization scripts were probed only during baseline and post-treatment. The primary outcome measures were the accuracy and rate of script production during oral reading of the script. Each turn (i.e., both the digital therapist's

and the participant's response) appeared on the screen, the digital therapist read her part, and then the participant read his or her part aloud without any cues from the digital therapist. The participant pressed the space bar at the end of each turn to indicate that he or she had finished speaking and to bring the next turn up on the screen.

Audio-recordings of the probes were captured by the computer software. For accuracy, each script-related word was transcribed and scored on a 6-point scale, the NORLA-6 (Naming and Oral Reading for Language in Aphasia 6-Point Scale), a standardized rule-based scoring system (Gingrich, Hurwitz, Lee, Carpenter, & Cherney, 2013). The scale ranges from 0 (no response) and 1 (unintelligible or unrelated response) to 4 (accurate but delayed or self-corrected response) and 5 (accurate and immediate response). Semantic or phonological paraphasias are scored as 2, while appropriate and intelligible responses with minor errors such as the omission of a grammatical morpheme are scored as 3. The NORLA-6 has previously demonstrated evidence supporting its validity and reliability (Gingrich et al., 2013). NORLA-6 scores were shown to be significantly correlated ($p < .001$) with accuracy on the Gray Oral Reading Test (Wiederholt & Bryant, 2001) ($r_s = .84$). Inter-rater and test-retest reliability were assessed using intraclass correlations, which were high ($ICC > .909$) (Gingrich et al., 2013). The duration of each production (i.e., from the initiation of the participant speaking to the pressing of the space bar to indicate the end of the turn) was automatically captured and calculated by the computer software.

Inter-rater Reliability—In addition to the previously demonstrated reliability of the NORLA-6, 10% of the script probes, including trained, untrained and generalization probes, were randomly selected for scoring by a second rater. A score for each sentence of the script was obtained by summing the word scores for that sentence. Inter-rater reliability (Pearson's r) for accuracy of script sentences was 94%. Since the length of time taken for each turn was automatically calculated by the computer, it was deemed to be reliable after statistical and manual verification to detect and correct outliers. Most recordings were properly segmented and we had to manually adjust only 0.52% of the recordings (i.e., 27 out of 5241 probe sentences).

Analyses—Percent accuracy was the total score for each sentence divided by the maximum score that could be achieved (5 points per word multiplied by the number of words in the turn). Rate was calculated as the number of script-related words (defined as having scores of 3 to 5) produced per minute (WPM). The accuracy and rate scores for any probe session were averaged over all ten turns of the probed script to yield the final score. Baseline, treatment and maintenance probe scores for each script were plotted for each participant and the graphs were visually analyzed. For the purposes of this paper, we present data from the trained and untrained probes.

Significance testing was conducted using the group gain scores from the baseline mean to post-treatment, and from baseline to the mean of the three - and six-week maintenance times. For each question, a Bonferroni adjustment for the p-values was used to determine significance.

Effect sizes (ES), a measure of the magnitude of change, were also calculated for acquisition (using the group gain scores from baseline to post-treatment) and maintenance (using the group gain scores from baseline to the mean of the three- and six-week maintenance probes). Effect sizes for both accuracy and rate were calculated using Cohen's (1988) *d* statistic. Since we were interested in the difference in gain between trained and untrained script probes, ES was calculated by subtracting the mean gain for the untrained script probes from the mean gain for the trained script probes, and then dividing by the pooled standard deviation of these trained and untrained gain scores. This is a conservative approach that controls for practice effects from repeated probes by removing the gain associated with the untrained script from that of the trained script. Effect sizes were interpreted using benchmarks for interpreting Cohen's *d*, whereby 0.2 equates to a small effect, 0.5 equates to a medium effect, and 0.8 equates to large effects (Cohen, 1988).

Assessing Participant Preference—Following completion of the second treatment condition (i.e. at the end of 9 weeks), each participant was interviewed about his or her experience participating in the treatment study. Appendix B includes a copy of the questionnaires and rating scales that were used. The research SLP clearly identified the training condition that had occurred first before asking the participant three questions about the first training condition: How much did you like the treatment? How hard did you work? How much did it help you? Questions were presented one at a time, written as well as spoken aloud, with a response rating scale where 1 represented “not at all”, 3 represented “somewhat”, and 5 represented “very much”. Then, the SLP identified the training condition that had occurred most recently and asked the same three questions in the same way with the similar rating scale. Finally, she asked the participant to compare the two training conditions. Questions were written, the SLP read each question aloud to the participant making sure that they were understood, and the participant was required to point to the written choice of Treatment 1 or Treatment 2. Participant responses to the comparative questions are reported in the results.

Results

The Appendices include graphs of each participant's probe performance on the trained and untrained scripts for each of the conditions, High Cue and Low Cue. Appendices C and D show percent accuracy for the participants with less severe (Appendix C) and more severe aphasia (Appendix D). Appendices E and F show words per minute (WPM) for the participants with less severe (Appendix E) and more severe aphasia (Appendix F). For ease of comparison, the High Cue and Low Cue conditions are superimposed over each other rather than being shown consecutively with the 3-week washout between conditions. The horizontal axis represents the weeks of each training condition, and not the weeks of the entire study. Aggregate data are used in the following sections to answer each of the research questions.

Question 1: Does computer-based script training (via *AphasiaScripts*[™]) result in the acquisition and maintenance of conversational scripts?

Data were combined for all subjects and for both conditions. Three weeks of computer-based script training resulted in increased accuracy and rate of script production. Mean (SD) baseline performance was 50.0 (26.4) % for accuracy and 23.7(20.6) WPM for rate. At the end of training, it had improved to 77.8 (19.6) % and 60.3 (30.5) WPM for accuracy and rate respectively. Using a 2-tailed t-test and a Bonferroni adjustment, these gains were significant for both accuracy, $t(7) = 5.40$, $p < 0.0125$, and rate, $t(7) = 5.87$, $p < 0.0125$.

Although there was a slight drop in performance noted at both three weeks and six weeks post-treatment, the decreases were small. At three weeks post treatment, the mean (SD) scores for accuracy were 72.2 (22.4) and the mean scores for rate were 55.2 (34.0). By six weeks post treatment, these scores had declined slightly to 68.6 (24.7) for accuracy and 51.4 (35.8) for rate. For purposes of analysis, we used the mean of the probe scores obtained at three and six weeks post-treatment as a single measure of maintenance. Using a 2-tailed t-test and a Bonferroni adjustment, the overall gains from baseline to the mean of three and six weeks maintenance were significant for both accuracy, $t(7) = 3.71$, $p < 0.0125$, and rate, $t(7) = 4.21$, $p < 0.0125$.

Interestingly, there were also small improvements on the untrained scripts, possibly reflecting a practice effect from the repeated probes. Mean (SD) baseline performance on the untrained scripts was 48.2 (28.4) % for accuracy and 22.0 (21.0) WPM for rate. At the end of the treatment phase, mean performance on the untrained script had improved to 55.5 (24.8) % and 29.6 (22.1) WPM for accuracy and rate respectively. A two-tailed t-test with a Bonferroni adjustment for p-values indicated that the gain on the trained scripts was significantly greater than on the untrained scripts for accuracy, $t(7) = 5.46$, $p < 0.0125$ and rate, $t(7) = 5.38$, $p < 0.0125$ at post treatment. Similarly, the gains on the trained script from baseline to the mean of three and six weeks maintenance were significantly greater than those of the untrained script for accuracy, $t(7) = 4.03$, $p < 0.0125$ and rate, $t(7) = 4.25$, $p < 0.0125$. These results indicate that the improvement on the trained scripts was more than merely a function of probe practice, but true acquisition and maintenance of the trained script.

Table 4 shows the effect sizes (Cohen's d) for the gain scores related to acquisition (baseline to post-treatment) and maintenance (baseline to mean of three- and six- weeks post treatment) of the trained versus untrained scripts for all subjects and both conditions combined. These effect sizes are considered to be large using benchmarks presented by Cohen (1988) for interpreting Cohen's d , whereby 0.2 equates to a small effect, 0.5 equates to a medium effect, and effects larger than 0.8 equate to large effects.

Question 2: Do High Cue and Low Cue training conditions differentially affect the acquisition and maintenance of conversational scripts?

To answer this question, data were combined for all eight subjects but separated out by condition. Figure 2 shows the mean probe performance on trained and untrained scripts for all eight subjects under the High Cue and Low Cue conditions. Visual analysis indicates a

similar pattern of acquisition for both conditions. However, performance on trained scripts reached a higher level of accuracy and rate in the High Cue condition. Maintenance at three and six weeks post treatment appeared relatively similar in both learning conditions.

Using a two-tailed t-test, gain scores from baseline to post-treatment on trained scripts showed no significant difference between the two conditions for either accuracy, $t(7) = 1.83$, $p > .05$, or rate, $t(7) = 1.03$, $p > .05$. Similarly, there was no significant difference between the two conditions for gain scores from baseline to the mean of the three and six week maintenance probes for accuracy $t(7) = 1.09$, $p > .05$, or rate $t(7) = 0.72$, $p > .05$.

Effect sizes (Cohen's d) for the gain scores of the trained versus untrained scripts for the High Cue and Low Cue conditions are shown in Table 4 for all eight subjects combined. For each condition, the effect sizes can be benchmarked as large. When the effect sizes are compared across conditions, the effect sizes are larger for the High Cue condition for both accuracy and rate. The differences between conditions were larger during acquisition than they were for maintenance.

Question 3. Does the severity of aphasia, as measured by the Western Aphasia Battery-Revised (Kertesz, 2007) Aphasia Quotient (WAB AQ), impact the relationship between High Cue and Low Cue training conditions and their effect on the acquisition and maintenance of conversational scripts?

Figure 3 shows the mean probe performance on accuracy and rate for the four participants with less severe aphasia and for the four participants with more severe aphasia in each training condition. The patterns of acquisition and maintenance appear similar for the High Cue and Low Cue conditions in both severity groups, and no differences emerged on significance testing for either accuracy or rate.

Table 4 shows the effect sizes for each condition from baseline to post-treatment and baseline to the mean of three- and six-week maintenance times for the participants with less severe and more severe aphasia separately. In all instances, effect sizes are larger for the High Cue condition than for the Low Cue condition. However, these differences were relatively small in the less severe group. For the more severe group, the differences in the effect sizes across conditions were larger, especially for accuracy.

Question 4. Do persons with aphasia have a preference for one cue condition over the other?

At the end of treatment participants were asked for their feedback about the two learning conditions. They were asked to indicate by pointing or marking their answer in a forced choice written format, which condition they liked better and which condition they thought helped them more. Table 5 shows their responses and also indicates which condition the participant received first.

All the participants with more severe aphasia preferred the High Cue condition and felt that it helped them more. The answers were more varied in the less severe participants. WELED, who had the mildest aphasia, found both conditions equal, while AMBDE, the only fluent participant in the study, preferred the Low Cue condition. With further questioning, he

revealed that he had been overwhelmed by the number of cues given in the High Cue condition. Three out of four participants with less severe aphasia felt each condition helped them about equally. One subject, ABEJO, indicated that he preferred the High Cue condition, but felt the Low Cue condition was more helpful because it was more challenging. Neither the condition that the participant received first nor the script appeared to impact the participants' responses or their gains over baseline in each condition of the cross-over study. This was verified statistically. Paired 2-tailed t-tests comparing baseline accuracy and rate (WPM) scores pair-wise across High Cue and Low Cue conditions showed no significant differences between baseline scores (within participants) across training condition for either measure. Paired 2-tailed t-tests comparing baseline accuracy and rate (WPM) scores between trained and untrained scripts also showed no significant differences between baseline scores (within participants), suggesting that the scripts functioned as equivalent measures.

Discussion

In this randomized cross-over study, eight individuals with chronic aphasia received intensive computer-based script training under two conditions: High Cue in which participants received sufficient cues before responding with the intent of minimizing errors, and Low Cue in which participants attempted a response with the minimum amount of cues, thereby creating an environment where errors were allowed. Each condition lasted three weeks, with a three-week washout period. Performance measures included accuracy and rate of script production. Training resulted in significant gains in script acquisition with maintenance of skills at three- and six- weeks post treatment. Differences between the cuing conditions were not significant for either acquisition or maintenance. This is consistent with our original hypothesis regarding acquisition, but not maintenance where we expected an advantage of the Low Cue condition.

A comparison of effect sizes provides a different interpretation. The High Cue condition was favored over the Low Cue condition, more so during the acquisition phase than during the maintenance phase. However, the primary goal of treatment is to maximize long term learning beyond the therapy session, and performance changes during practice or in the short-term do not always predict long term maintenance (Schmidt & Bjork, 1992). Careful examination of trends during the maintenance period shows that the reduction in performance for both accuracy and rate from post-treatment to 3 weeks and from 3 weeks to 6 weeks was greater in the High Cue condition than the Low Cue condition (see Figure 2). If these trends were to continue over the following weeks, they may have led to an advantage of the Low Cue condition in the longer term, a finding that would have been consistent with our hypothesis.

When severity of aphasia was considered, there were no significant differences between conditions. However, the magnitude of change was larger in the High Cue condition versus the Low Cue condition for both severity groups, with the differences in effect sizes being greater between conditions for those with more severe aphasia. The impact of severity was also reflected in the participants' personal preferences. Participants with more severe aphasia all preferred the High Cue condition and felt that it helped them more. Answers

were more varied for the participants with less severe aphasia, several of whom reported no preference for conditions. Examination of trends during the maintenance period also showed differences for the more and less severe groups. The greater reduction in performance in the High Cue condition discussed previously was evident for those with more severe aphasia and not for those with less severe aphasia (see Figure 3). If these trends would have continued, the Low Cue condition may have been favored in the longer term for those with more severe aphasia.

The finding that training results in script acquisition and maintenance is consistent with previous studies of script training (Bilda, 2011; Cherney et al., 2008, 2011; Goldberg et al., 2012; Lee et al., 2009; Manheim et al., 2009; Youmans et al., 2005, 2011) and serves to reinforce the use of script training as a treatment option for aphasia. Script training has been delivered by speech-language pathologists either in person or by video conferencing (Goldberg et al., 2012; Youmans et al., 2005), and by computer via prerecorded video clips or a digital therapist (Bilda, 2011; Cherney et al., 2008). Script training has even been delivered by a handheld barcode scanner (Nobis-Bosch et al., 2011). Although findings across studies have been positive, further investigation of the optimum procedures and dose of script training is warranted. While specific procedures have varied, common to all studies, including the current one, is the high intensity at which the scripts were practiced. For example, Bilda (2011) provided 10 days of computer-based therapy with three hours of daily training. Although other studies provided only three script-training sessions per week, these were supplemented by daily independent practice for 15–30 minutes (Goldberg et al., 2012; Youmans et al., 2005). Like the current study, scripts were practiced for three weeks in the Goldberg (2012) study, whereas they were practiced until a criterion of 90% accuracy was obtained for two consecutive sessions (a range of 5–11 sessions) in the Youmans (2005) study. Yet, simply providing more opportunities for practice may not be the most cost-effective way of improving outcomes in the long term. Rather, careful consideration of learning theory and the application of learning principles can assist investigators in determining optimum procedures for script training. The current study is a first step in this endeavor and highlights the importance of considering the amount and timing of cuing provided during practice.

For this study, we contrasted two different cuing conditions with the intent of comparing script training in an errorful and an error-reducing learning environment. However, we did not determine the number of resulting errors produced by participants during each treatment condition. As a result, we cannot be sure that errorful and error-reducing training conditions actually occurred and conclusions can only be made about the impact of the cuing conditions. Future studies that investigate errorful and error-reducing training conditions must verify the number of errors produced by participants. Nevertheless, findings from previous studies about errorful and error-reducing learning in aphasia can inform the discussion.

Interestingly, McKissock & Ward (2007) found that both errorless and errorful conditions with feedback were equivocal on a naming task and both were significantly better than errorful learning without feedback. They concluded that production of errors during training is not an important factor; of importance is that a correct response is given as feedback. In

the current study, both script training conditions included presentation of the correct response by the digital therapist, which may have served as a form of feedback. Future exploration of the type, amount, and timing of feedback in combination with different cue conditions may provide additional important information regarding optimum script training procedures.

The current study highlights the importance of considering the impact of participant variables such as severity on performance in different learning conditions. Although not significant, effect size differences between the cue conditions were evident mostly for those with more severe aphasia. In this regard, it is possible that the High and Low Cue conditions were not sufficiently distinct in terms of the opportunity for retrieval practice and the chance to make errors. In the High Cue condition, although repetition (associated with errorless or error-reduced learning) was emphasized, participants were required to attempt the phrase or sentence independently, thereby giving them an opportunity for some retrieval practice. Another opportunity for retrieval practice occurred when scripts were probed twice per week during training. In other words, the amount of retrieval practice in the High Cue condition may have been sufficient to produce similar results across conditions, particularly for those with less severe aphasia. In contrast, for those with more severe aphasia, the degree of difficulty encountered with the independent productions of phrases and sentences was so high in both conditions that participants may have disregarded that part of the training sequence. With the chance for retrieval practice eliminated, the High Cue condition now provided more opportunities for repetition leading to greater mastery of the script than in the Low Cue condition. In sum, it is possible that the High and Low Cue conditions provided the less severe participants with equally sufficient opportunities for retrieval practice; for the more severe participants, the Low Cue condition afforded little opportunity for anything but errors, while the High Cue condition was more successful by providing both errorless and errorful learning opportunities.

Script difficulty is another factor to consider. Even though we attempted to adjust script difficulty to aphasia severity, the participants with less severe aphasia may have responded equally to the High and Low Cue conditions because their scripts were not challenging enough for them to benefit from the extra cues provided in the High Cue condition. Figure 3 illustrates how much better participants with less severe aphasia performed at baseline for rate and accuracy as compared to the participants with more severe aphasia, even before training began. For accuracy, baseline means (SD) were 73.6 (10.2) % for participants with less severe aphasia and 26.3 (5.2) % for those with more severe aphasia; for rate, baseline means (SD) were 40.1 (16.2) WPM for participants with less severe aphasia and only 7.4 (4.1) WPM for those with more severe aphasia. Had the less severe participants' scripts been more complex, the High Cue condition may have been of more benefit to them than the Low Cue condition.

While we considered only aphasia severity as determined by the WAB AQ, other participant characteristics may impact outcomes and require further investigation. Fillingham and colleagues identified cognitive variables such as attention, recall memory, working memory and executive function as important characteristics required by persons with aphasia to benefit from either errorful or error-reducing learning conditions during naming (Fillingham

et al., 2005, 2006; Lambon Ralph, Snell, Fillingham, Conroy, & Sage, 2010). Furthermore, two factors, a cognitive factor (attention, memory and executive function) and a language factor (pre-intervention naming performance) were related to the degree of improvement in an errorful naming task involving accumulating cues (Lambon Ralph et al., 2010). Future research on script training should consider these factors because they may identify participants more likely to benefit from either High Cue or Low Cue training. In particular, reading level may be a critical factor. In our study, only one participant, SMISC, made little change in the Low Cue condition. Although his overall aphasia severity was similar to that of another participant, it was his WAB Reading Subscale score of 9 that differentiated him from all participants. Practice in the Low Cue condition may have been severely limited because it provided only the written sentence for him to read aloud.

In this study, we calculated effect sizes for the gain scores of the trained versus untrained scripts *within* conditions and participant groups, and interpreted the magnitude of change using established benchmarks for Cohen's *d*. In all cases, the effect sizes were large, indicating a meaningful result with respect to the acquisition and maintenance of trained scripts. We then compared the effect sizes *across* conditions and participant groups and found that the effect sizes were always larger in the High Cue condition, particularly when the aphasia was more severe. These findings were in the absence of a statistically significant difference between conditions and participant groups, which raises the issue of whether the study was sufficiently powered. Power was sufficient to find a difference between the acquisition and maintenance of the trained versus untrained scripts, as described in Question 1. While the inclusion of more participants might have resulted in a statistically significant difference between the High Cue and Low Cue conditions, it is also possible that there is no real difference between the conditions regardless of the number of participants, a finding that is consistent with previous studies comparing errorful and error-reducing treatment for naming in aphasia.

We have been reporting group results that reflect mean gains that have the potential to obscure individual performance patterns. Therefore, it is valuable to inspect individual participant data. Individual results were similar to the group results for all participants except for one, STEOM, who demonstrated greater gains in accuracy and rate for both acquisition and maintenance in the Low Cue condition than in the High Cue condition (see Appendices C, D, E and F). The Low Cue condition was the second condition that he received, and it is possible that improvements in the High Cue condition were hindered by slow learning of the computer program and how to navigate through it. Most participants, except STEOM, were either knowledgeable about computers prior to enrollment in the study or had assistance at home for setting up the computer and initiating its use. Interestingly, STEOM reported that he preferred the High Cue condition, worked harder in that condition, and that it helped him more than the Low Cue condition.

This study extends previous work in aphasia by directly comparing High Cue and Low Cue conditions in a script-training task and including assessment of maintenance as well as acquisition. Generalization of results is limited because the sample size was small, and all but one participant presented with chronic nonfluent aphasia. Future studies are needed with larger numbers of participants including those with acute/subacute aphasia of various types,

and with extended maintenance periods to permit the assessment of long term learning. Nevertheless, results support the clinical use of script training and highlight the importance of considering cuing conditions in relation to aphasia severity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by the National Institute on Deafness and Other Communication Disorders, Award Number 1R01DC011754 (to L.R.C. and S.V.V.).

The authors extend their thanks to Julia Carpenter, Rachel Hitch, Rosalind Hurwitz and Jaime Lee who collected and scored data, to Anita Halper and Audrey Holland who served as consultants on the grant, and to Nattawut Ngampatipatpong who modified the AphasiaScripts software for the purposes of this study.

References

- Abel S, Schultz A, Radermacher I, Willmes K, Huber W. Increasing and decreasing cues in naming therapy for aphasia. *Aphasiology*. 1995; 19:831–848.
- Bilda K. Video-based conversational script training for aphasia: a therapy study. *Aphasiology*. 2011; 25:191–201.
- Cherney LR, Halper AS, Holland AL, Cole R. Computerized script training for aphasia: preliminary results. *American Journal of Speech Language Pathology*. 2008; 17:19–34. [PubMed: 18230811]
- Cherney LR, Halper AS, Kaye RC. Computer-based script training for aphasia: emerging themes from post-treatment interviews. *Journal of Communication Disorders*. 2011; 44:493–501. [PubMed: 21612787]
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2. Hillsdale, NJ: Erlbaum; 1988.
- Conroy P, Sage K, Ralph MAL. The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. *Aphasiology*. 2009a; 23:707–730.
- Conroy P, Sage K, Ralph MAL. Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology*. 2009b; 23:1311–1337.
- Fillingham JK, Hodgson C, Sage K, Lambon Ralph MA. The application of errorless learning to aphasic disorders: A review of theory and practice. *Neuropsychological Rehabilitation*. 2003; 13:337–363. [PubMed: 21854317]
- Fillingham JK, Sage K, Lambon Ralph MA. The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation*. 2006; 16:129–154. [PubMed: 16565031]
- Fillingham J, Sage K, Lambon Ralph M. Further explorations and an overview of errorless and errorful therapy for aphasic word-finding difficulties: The number of naming attempts during therapy affects outcome. *Aphasiology*. 2005; 19:597–614.
- Gingrich, L.; Hurwitz, R.; Lee, J.; Carpenter, J.; Cherney, LR. Quantifying Naming & Oral Reading Performance in Aphasia: The NORLA-6 Scale. Presented at the Annual Convention of the American Speech Language Hearing Association; Chicago, IL. 2013 Nov.
- Goldberg S, Haley KL, Jacks A. Script training and generalization for people with aphasia. *American Journal of Speech Language Pathology*. 2012; 21:222–238. [PubMed: 22442283]
- Karpicke JD, Roediger HL. The critical importance of retrieval for learning. *Science*. 2008; 319:966–968. [PubMed: 18276894]
- Kertesz, A. *Western Aphasia Battery-Revised*. San Antonio, TX: PsychCorp; 2007.
- Lambon Ralph MA, Snell C, Fillingham JK, Conroy P, Sage K. Predicting the outcome of anomia therapy for people with aphasia post CVA: both language and cognitive status are key predictors. *Neuropsychological Rehabilitation*. 2010; 20:289–305. [PubMed: 20077315]

- Lee JB, Kaye RC, Cherney LR. Conversational script performance in adults with non-fluent aphasia: treatment intensity and aphasia severity. *Aphasiology*. 2009; 23:885–897.
- Logan GD. Toward an instance theory of automatization. *Psychological Review*. 1988; 95:492.
- Manheim LM, Halper AS, Cherney L. Patient-reported changes in communication after computer-based script training for aphasia. *Archives of Physical Medicine and Rehabilitation*. 2009; 90:623–627. [PubMed: 19345778]
- Mckissock S, Ward J. Do errors matter? Errorless and errorful learning in anomic picture naming. *Neuropsychological Rehabilitation*. 2007; 17:355–373. [PubMed: 17474061]
- Middleton EL, Schwartz MF. Errorless learning in cognitive rehabilitation: A critical review. *Neuropsychological Rehabilitation*. 2012; 22:138–168. [PubMed: 22247957]
- Nobis-Bosch R, Springer L, Radermacher I, Huber W. Supervised home training of dialogue skills in chronic aphasia: a randomized parallel group study. *Journal of Speech, Language and Hearing Research*. 2011; 54:1118.
- Raymer A, Strobel J, Prokup T, Thomason B, Reff KL. Errorless versus errorful training of spelling in individuals with acquired dysgraphia. *Neuropsychological Rehabilitation*. 2010; 20(1):1–15. [PubMed: 19449254]
- Roediger HL, Karpicke JD. Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science*. 2006; 17:249–255. [PubMed: 16507066]
- Schmidt RA, Bjork RA. New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*. 1992; 3:207–217.
- Schuele CM, Justice LM. The Importance of Effect Sizes in the Interpretation of Research : Primer on Research: Part 3. *The ASHA Leader*. 2006 Aug 15.
- Tulving E. The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*. 1967; 6:175–184.
- Wiederholt, JL.; Bryant, BR. *Gray Oral Reading Tests -4*. Austin, TX: Pro-Ed; 2001.
- Youmans G, Youmans SR, Hancock AB. Script training treatment for adults with apraxia of speech. *American Journal of Speech Language Pathology*. 2011; 20:23–37. [PubMed: 20739633]
- Youmans G, Holland A, Munoz M, Bourgeois M. Script training and automaticity in two individuals with aphasia. *Aphasiology*. 2005; 19:435–450.

SCRIPT PRACTICE

Waitress: Welcome to Jameson's. How are you tonight?
You: I'm fine. **Thanks**. Is chicken parmesan on the menu tonight?

Waitress: Yes, it is.
You: Sounds good. I'll have that. What comes with it?

Waitress: Either soup or salad.
You: I'm in the mood for a garden salad. May I have that?

Waitress: Sure.
You: And could you put in lots of cucumbers?

Waitress: No problem.
You: Actually, I changed my mind about my order. I'd rather have gnocchi.

Waitress: Anything to drink?
You: I'll have iced tea with dinner.

Waitress: How about dessert?
You: Let me look at the menu a minute. I think I'll have cheesecake.

Waitress: Sorry we ran out.
You: Then I won't get dessert, thanks. But can I have a decaf coffee? With no milk or sugar.

Waitress: Sure.
You: I'm going to a movie in forty-five minutes, so I'm in a hurry. Can you take my credit card, now?

Waitress: Of course.
You: Thanks. And can I have a box for the leftovers?




Figure 1.

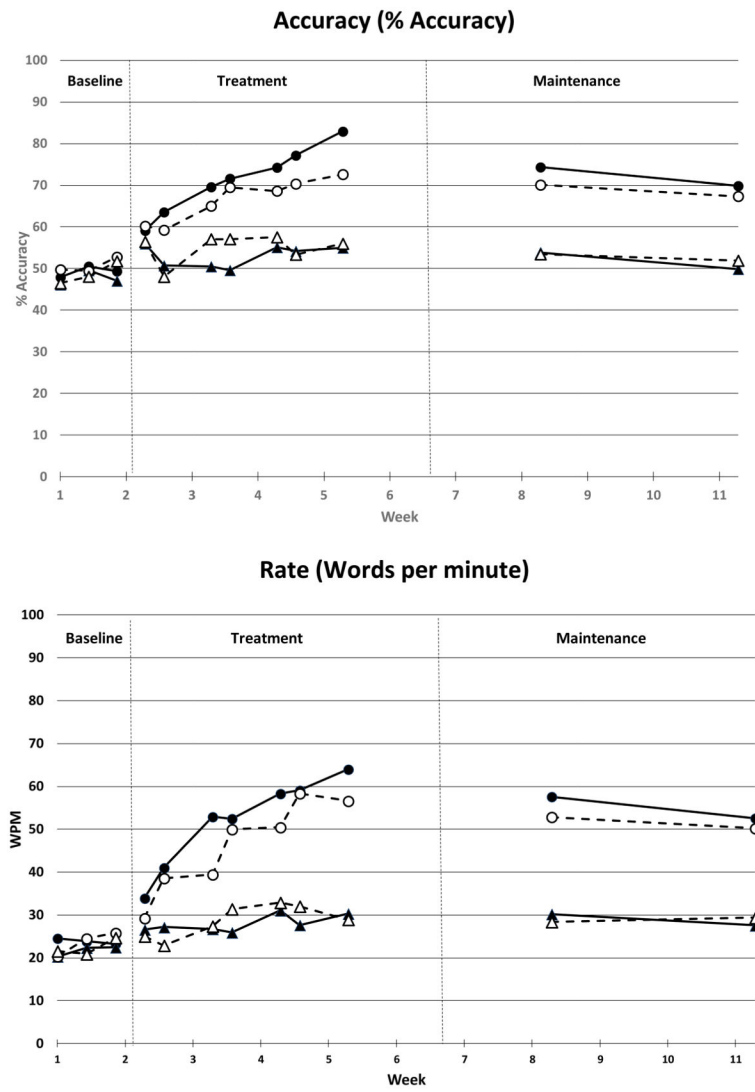


Figure 2.

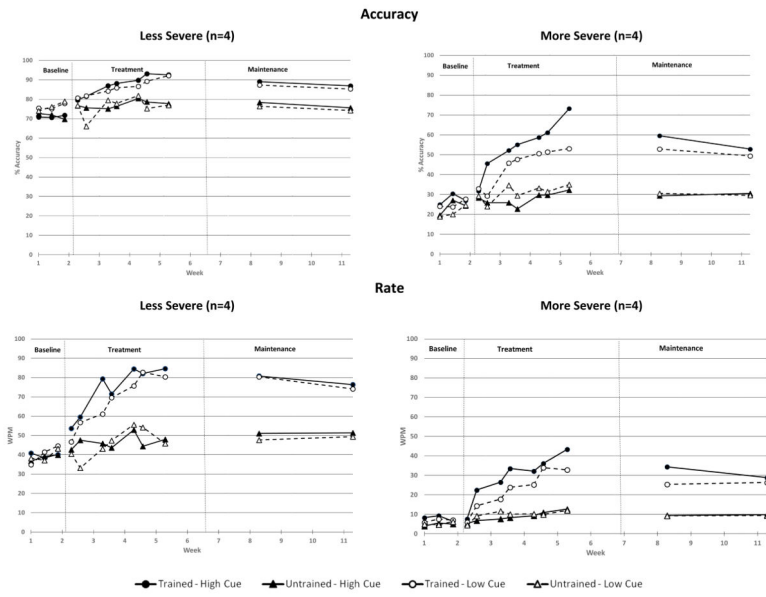


Figure 3.

TABLE 1

Demographic and language characteristics of study participants

Participant	Age (years)	Gender	Handedness	TPO (months)	Education (years)	WAB SS	WAB Comp	WAB Rpt	WAB Nmg	WAB AQ	WAB Reading Sub-Score	Aphasia Type	Treatment Sequence
AQ > 60													
ABEIO	51.8	M	R	48	16	13	9.8	5.3	6.3	68.8	78	NonFluent	Low - High
AMBDE	61.6	M	R	21	11	16	9.15	5.1	7	74.5	74	Fluent	High - Low
PIESH	66.4	F	R	59	14	10	8.9	8.4	6.5	67.6	63	NonFluent	High - Low
WELED	64.5	M	R	37	14	13	9.75	8.4	8.9	80.1	100	NonFluent	Low - High
AQ < 60													
CAVCH	25	F	L	13	13	10	7	3.1	7.1	54.4	62	NonFluent	Low - High
SMIDA	44.5	M	R	8	18	5	6.15	0.9	2	28.1	45	NonFluent	High - Low
SMISC	59.3	M	L	10	16	5	4.3	6.7	1.6	35.2	9	NonFluent	Low - High
STEOM	42.9	M	R	15	12	12	5.6	5.2	4.7	55	48	NonFluent	High - Low

Note. WAB AQ = Western Aphasia Battery-Revised Aphasia Quotient (Kertesz, 2007); SS = Spontaneous speech score; Comp = Auditory verbal comprehension; Rpt = Repetition; Nmg = Naming and word finding; Low = Low Cue condition; High = High Cue condition

TABLE 2

Mean aphasia severity of participants randomized to each treatment sequence

Treatment Sequence	WAB AQ			n
	Mean	SD	Range	
High Cue – Low Cue	56.3	20.5	28.1 – 74.5	4
Low Cue – High Cue	59.6	19.4	35.2 – 80.1	4

Note. WAB AQ = Western Aphasia Battery-Revised Aphasia Quotient (Kertesz, 2007)

TABLE 3

Script counts for WELED (AQ= 80.1) and SMIDA (AQ=28.1)

	Words	Sentences	Syllables	Verbs	Morphemes	ASL	ASW	Flesch Reading Ease	Grammatical Complexity
WELED AQ = 80.1 (less severe aphasia)									
Set A									
Restaurant (T)	122	20	148	32	141	6.10	1.21	98.01	1.2
Grocery (U)	120	19	145	33	152	6.32	1.21	98.20	1.3
Dinner Plans (G)	120	18	142	33	143	6.67	1.18	99.96	1.2
Set B									
Catching Up (T)	124	20	150	32	154	6.2	1.2	98.2	1.2
Meeting Someone New (U)	118	19	139	32	137	6.2	1.2	100.9	1.2
Plans (G)	121	18	144	35	139	6.7	1.2	99.3	1.1
SMIDA AQ = 28.1 (more severe aphasia)									
Set A									
Restaurant (T)	54	12	69	11	53	4.50	1.28	94.17	1.0
Grocery (U)	53	12	67	12	54	4.42	1.26	95.40	1.0
Dinner Plans (G)	54	12	69	11	56	4.50	1.28	94.17	1.0
Set B									
Catching Up (T)	54	12	68	11	58	4.5	1.3	95.7	1.1
Meeting Someone New (U)	51	12	66	11	54	4.3	1.3	93.0	1.1
Plans (G)	54	12	68	12	60	4.5	1.3	95.7	1.1

Note. T=trained script; U = untrained script; G=generalization script; ASL= average sentence length; ASW=average syllables per word; Flesch Reading Ease score: 206.835 – (1.015 x ASL) – (84.6 x ASW); increased score = increased reading ease

Grammatical Complexity = ratio of morphemes to words

Effect sizes for acquisition (baseline to post-treatment) and maintenance (baseline to average of 3 and 6 weeks post treatment) of trained versus untrained scripts.

TABLE 4

Question 1: Does computer-based script training result in the acquisition and maintenance of conversational scripts?

Effect sizes are shown for all subjects and both conditions combined*

Gains	Accuracy (% Accuracy)	Rate (WPM)
Baseline to post-treatment	1.49	1.96
Baseline to 3week/6week mean	1.18	1.51

Question 2. Do the high cue and low cue training conditions differentially affect the acquisition and maintenance of conversational scripts?

Effect sizes are shown for all subjects under each condition (High Cue and Low Cue)**

Gains	Accuracy (% Accuracy)		Rate (WPM)	
	High Cue	Low Cue	High Cue	Low Cue
Baseline to post-treatment	2.00	1.05	2.41	1.57
Baseline to 3week/6week mean	1.48	0.93	1.89	1.19

Question 3. Does the severity of aphasia impact the relationship between High and Low Cue training conditions and their effect on the acquisition and maintenance of conversational scripts?

Effect sizes are shown for participants with less severe and more severe aphasia under each condition***

Gains	Accuracy (% Accuracy)		Rate (WPM)	
	High Cue	Low Cue	High Cue	Low Cue
Less severe aphasia				
Baseline to post-treatment	1.16	1.10	2.71	2.00
Baseline to 3week/6week mean	1.00	0.74	2.30	1.55
More severe aphasia				
Baseline to post-treatment	2.84	1.00	2.11	1.14
Baseline to 3week/6week mean	1.97	1.12	1.49	0.84

* Numerators: Mean of gains from baseline for TRAINED vs. UNTRAINED (n=16, conditions combined); Denominators: Pooled standard deviations of gains from baseline for TRAINED and UNTRAINED (n=16, conditions combined)

** Numerators: Mean of gains from baseline for High/Low Cue TRAINED vs. UNTRAINED (n=8); Denominators: Pooled standard deviation of gains from baseline for High/Low Cue TRAINED vs. UNTRAINED (n=8)

*** Numerators: Mean of gains from baseline for High and Low Cue TRAINED vs. UNTRAINED for less/more severe participants (n=4); Denominators: Pooled standard deviation of gains from baseline for High/Low Cue TRAINED vs. UNTRAINED for less and more severe participants COMBINED (n=8)

TABLE 5

Participants' preferences for High Cue vs. Low Cue training conditions

Less severe aphasia	Which 1st?	Like better?	Helped more?	Worked harder?
AMBDE	High Cue	Low Cue	Same	Same
PIESH	High Cue	High Cue	Same	Low Cue
ABEJO	Low Cue	High Cue	Low Cue	Low Cue
WELED	Low Cue	Same	Same	Same
More severe aphasia	Which 1st?	Like better?	Helped more?	Worked harder?
SMIDA	High Cue	High Cue	High Cue	Same
STEOM*	High Cue	High Cue	High Cue	High Cue
CAVCH	Low Cue	High Cue	High Cue	Low Cue
SMISC	Low Cue	High Cue	High Cue	Low Cue

* STEOM was the only participant who performed better in the Low Cue than High Cue condition