Script Training and Generalization for People with Aphasia

Samantha Goldberg

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in the Department of Allied Health Sciences, Division of Speech and Hearing Sciences

Chapel Hill
2011

Approved by:

Dr. Katarina Haley
Dr. Adam Jacks
Dr. Melody Harrison
Abstract

Samantha Goldberg: Script Training and Generalization for People with Aphasia
(Under the direction of Dr. Katarina Haley)

A multiple baseline study was conducted to test the effect of script training intervention on fluency, rate of speech, grammatical complexity, and content production of target scripts for two participants with non-fluent aphasia. Script training occurred in person and via Skype. Both participants showed improvement on all variables for both scripts, with the greatest effects for rate of speech for both participants, as well as use of script-related words for Participant A and number of disfluencies Participant B. During generalization probes, both participants demonstrated functional use of their learned scripts in a less structured communicative context. Factors that influenced accuracy and rate of script learning included self-cueing skills, motivation, and level of impairment.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................ vi

LIST OF FIGURES ...................................................................................................... vii

Chapter

I. INTRODUCTION ..................................................................................................  1

II. SCRIPT TRAINING .............................................................................................  2

III. PREVIOUS OUTCOMES ....................................................................................  4
    Candidacy ...........................................................................................................  5
    Script Development .........................................................................................  6
    Generalization Effects .....................................................................................  7
    Summary of Previous Research ......................................................................  9
    Rationale for Further Research .................................................................... 10

IV. PURPOSE ............................................................................................................ 12

V. METHODS .......................................................................................................... 14
    Participants ...................................................................................................... 14
    Script Development ......................................................................................... 14
    Baseline Conversation .................................................................................... 15
    Research Design ............................................................................................. 18
    Script Training Intervention ......................................................................... 19
    Frequency and Intensity of Intervention ......................................................... 22
    Generalization Probe ..................................................................................... 22
LIST OF TABLES

Table

1. Effect Sizes for Participant A................................................................. 34
2. Effect Sizes for Participant B................................................................. 34
LIST OF FIGURES

Figure

1. Participant A, Percent Script Related Words ................................. 35
2. Participant B, Percent Script Related Words ................................. 35
3. Participant A, Rate of Speech ...................................................... 36
4. Participant B, Rate of Speech ...................................................... 37
5. Participant A, Percent Words with Grammatical Morphemes ............... 38
6. Participant B, Percent Words with Grammatical Morphemes ............... 38
7. Participant A, Disfluencies per Word ............................................ 39
8. Participant B, Disfluencies per Word ............................................ 40
I. INTRODUCTION

Aphasia is an acquired neurologic communication disorder in which the primary deficits are in language. It can impair a person’s comprehension, fluency, grammaticality, repetition skills, and, almost always, word finding abilities. Aphasia affects about one million people in the United States (National Institute for Deafness and Other Communication Disorders, 2006). Approximately 40% of stroke survivors present with aphasia in the acute phase of recovery (Pedersen et al., 1995). Many intervention techniques to treat aphasia have been developed and studied. They are wide ranging and address various types of impairment by targeting semantic networks, grammatical processing, or increasing communicative competence via gesturing. There are treatment programs that aim to improve auditory comprehension, reading, writing, fluency, and prosody. Evidence for most aphasia therapy is at the phase I and phase II levels of outcomes research (Robey, 2004). These are the preliminary phases of a five-phase model for defining forms of clinical research. Phase I research aims to begin describing the protocol, candidacy, and treatment effects. Phase II research begins to refine these parameters to determine experimental design for optimal outcomes.

Aphasia treatments are often described as impairment-based or consequence-based. Impairment-based approaches specifically target the aspects of language that are impaired. They are justified by the assertion that aphasia is a language disorder at its core and that it is out of the scope of practice for speech-language pathologists to target psychosocial issues related to aphasia. In contrast, consequence-based treatments reflects the model asserted by the World Health Organization’s International Classification of Functioning, Disability and Health (World Health Organization, 2007), which emphasizes the impact of a disability on a person’s full participation in his or her environment. Consequence-based approaches to aphasia emphasize the psychosocial aspects of impaired language and encourage a greater voice for people with the disorder in functional situations.
II. SCRIPT TRAINING

Many aphasia treatments incorporate principles of impairment-based and consequence-based approaches. This blend of justifications for treatment allows speech-language pathologists to target both the underlying language deficits in aphasia and the psychosocial implications. Script training is one such treatment. In script training intervention, people with aphasia develop and learn short scripts about topics with real-world and personal significance. In this way, script training targets aphasia deficits on the level of discourse, but also encourages linguistic accuracy. By addressing discourse as a whole, script-training focuses on language deficits. Articulation, morphology, and word finding are explicitly addressed in script training intervention. On the other hand, by giving the patient scripts to use in real-life situations, script training helps people with aphasia directly improve their communication in functional situations.

In the script training method, dialogues or monologues are learned by repetitive drilling of words, followed by sentences, and then conversational turns. A speech-language pathologist aids in script acquisition by following a cueing hierarchy on each line of the script. First the participant repeats the line after the therapist, then the participant reads the line while the therapist simultaneously reads it (choral reading), and finally the participant practices production of the line independently. Participants progress through this hierarchy, based on a determination of mastery made by either the clinician or the participant.

An accurate, albeit short, expression of fluent and accurate speech can have profound implications for a person with aphasia and his or her family. In previous uses of script training, scripts have allowed persons with aphasia, who otherwise communicate with great effort and frustration, to deliver a speech at a graduation party, to interact with co-workers confidently, to explain their disability, to ask questions at the grocery store, or to express affection and gratitude for loved ones.
These scripts give people with aphasia greater ability to gain social closeness, improve work performance, and exchange information that, while limited, can improve their daily lives. Script training relies on an instance theory of automatization, which suggests that automaticity can be achieved by repeated practice of a whole unit, as opposed to its components (Logan, 1988). Script training encompasses this principle by working on phrases and sentences primarily, with less of an emphasis on the word level. Also, instead of addressing rules for combining words, patients learn word combinations in a rote manner. The goal is for utterances to become over-learned and automatic.
III. PREVIOUS OUTCOMES

Holland and colleagues introduced script learning for people with aphasia and demonstrated that, with this intervention, people with aphasia could accurately deliver scripts (Holland, Milman, Muñoz, & Bays, 2002). All six participants in this initial study mastered their individualized scripts in 4-34 sessions. This range is notable for the high degree of variability among participants’ rates of script learning.

These promising findings have since been extended through a series of systematic replication studies. Cassette tapes were the original modality for script recordings (Holland et al., 2002). Youmans and colleagues trained two participants on three 3-4 sentence scripts over the course of about 10 weeks using cassette tape recordings (Youmans, Holland, Muñoz, & Bourgeois, 2005). Participants engaged in 30-45 minute treatment sessions with a speech-language pathologist three times each weekly and practiced independently at least 15 minutes each day. Script recordings were made by each participant at the start of every treatment session. Both participants demonstrated script mastery based on accuracy and rate of speech. Accuracy was defined by two separate variables: percent script correct and frequency of errors. Percent script correct was calculated as the number of produced script words, divided by the number of words in the target script. Errors were defined as “non-communicative words or phrase repetitions, fillers, pauses of three seconds or more, and unrecognizable utterances” (p. 438). Real-word substitutions and circumlocutions were not counted as errors. Experimental control was demonstrated in the context of a multiple baseline design across behaviors.

To facilitate independent practice, computer software called AphasiaScripts was developed (Cherney, Halper, Holland, & Cole, 2008). This software is used to implement script training intervention in a personal setting. AphasiaScripts provides users with the text and audio of the script,
and features an animated speech therapist, the Personal Automated Teacher (PAT). PAT provides a visual cue by showing the oral movements associated with each word and also provides an element of personal interaction. Users may listen to their script, or practice one sentence or conversational turn at a time. Participants also can highlight sections of the text and hear individual words. Furthermore, AphasiaScripts utilizes a cueing hierarchy consistent with the script training protocol previously discussed; PAT provides a model, encourages choral reading, and eventually allows the user to produce each utterance independently. The software records the user’s utterances so that they can be replayed for the user. The length of time given for a user to take his or her conversational turn and the time between utterances are adjustable.

Cherney and colleagues introduced AphasiaScripts in a multiple baseline treatment study (Cherney et al., 2008). Each participant used a laptop computer with three scripts recorded and typed into AphasiaScripts. The scripts were developed during four sessions with a speech-language pathologist over the course of four weeks. After a stable baseline was established, participants practiced each script for three weeks before moving on to the next one. Participants agreed to practice at home for at least 30 minutes each day and were responsible for decreasing cueing as they saw fit. Speech language pathologists were able to monitor progress through hand-written and software automated logs, and progress was also monitored in weekly meetings with the participants. Still, this complete independence in determining level of cueing and intensity of practice were new considerations in using this computerized intervention. The authors reported improvement in content, grammatical productivity, and rate of production of script related words for all scripts after the 9 week treatment.

Candidacy

Although script training has been effective for many people with aphasia, there is reason to believe that aphasia severity could affect the level of benefit from this intervention, or the way in which the intervention is implemented. Identifying the ideal candidates for this intervention may help inform how it is used clinically, to best allocate time and resources. Script training requires
participants to build upon existing language. Lee and colleagues identified various factors that may affect the utility of script training for a person with aphasia (Lee, Rosalind & Cherney, 2009). They determined aphasia severity for each participant on the basis of the Western Aphasia Battery Aphasia Quotient (Kertesz, 1988).

The Aphasia Quotient (AQ) is a composite score determined from ten subtests that cover four domains of language: spontaneous speech, auditory comprehension, repetition, and naming. Shewan and Kertesz (1980) explain that the Aphasia Quotient is a “functional measure of severity of the spoken language deficit in aphasia” (p. 309). Indeed, the AQ is heavily weighted towards expressive language tasks. In the Lee et al. study, script lengths were adjusted based on the AQ score. For instance, the participant with the most severe aphasia (WAB AQ 30.5) had scripts with 40-50 words, and participants with less severe aphasia (defined here as having a WAB AQ of 76.8 to 85.3) had scripts with 96 to 141 words in previous script training studies in the studies described by Lee et al.

Participants were given a suggested or required minimum weekly practice time (30-60 minutes per day, 5 days per week, depending on the study); however, the actual amount of practice on scripts varied widely. In Lee et al., actual script practice time ranged from 1.93 hours per week to 16.87 hours per week. Not surprisingly, improvement in accuracy and rate correlated significantly with amount of practice. The researchers also determined that individuals with more severe aphasia practiced more and showed greater improvement in accuracy than those with milder aphasia. Participants with less severe aphasia who practiced more hours showed greater improvement in rate of speech than those with more severe aphasia. However, this relationship between amount of practice and amount of improvement was not observed in each individual participant; some participants were able to make great gains with relatively little practice, while others practiced more diligently and saw less improvement. Individual variations may have been due to a number of factors, including differences in language profile, motivation, and overall learning style.

*Script development*
Script topic may also affect treatment outcome. It seems that when participants are more motivated by their script, they are more likely to perform better, perhaps due to increased practice time. Lee et al. (2009) and Holland et al. (2010) suggested that participants demonstrate increased learning on highly motivating scripts, such as a graduation speech for a son or a personally important work-related conversation. Holland, Halper, and Cherney (2010) described the collaborative scripting process that they used in order to ensure that participants created scripts that were personally motivating. After explaining the goals, limitations, and process of script training, the importance of selecting a personally significant and interesting topic was emphasized. In drafting the script, the speech-language pathologist’s goal was to approximate the participant’s pre-stroke speaking style, although it is unclear with what method they accomplished this. Monologue and dialogue scripts were created to be functionally useful in everyday situations. In order of frequency, dialogue topics included conversations with family members, seeking or providing an explanation, outside interests, ordering in a restaurant, phoning, personal experiences, work talk, and making plans. The majority of monologues told the story of the participant’s stroke. Other monologue topics were introductions, life stories, prayers, speeches, outside interests, and making social plans.

**Generalization effects**

Script training was developed so that a person with aphasia could follow the learned script in the exact context for which it was learned. The learning of content-specific, automatic speech was not anticipated to affect overall language. Instead, the approach was developed as a “whole task” procedure that might only generalize to similar tasks (Cherney et al. 2008). Accordingly, most script training studies have not explored the possibility of generalization. However, functional use of scripts with novel conversation partners who are unaware of the script is a practical extension from script training intervention. In order to accomplish this, people with aphasia must be able to generalize the script reading task to other situations.

There is preliminary evidence that scripts could be used as an impetus for improved discourse in more general conversational settings. For instance, in the aforementioned study by Cherney et al.
(2008), participants reported increased communicative confidence and overall increased ease of communication in exit interviews. These results seem to indicate that script training had an effect on the participant’s perception of their language and communication outside of a specific script. Although this improvement does not necessarily indicate improved linguistic competence, it is a promising indication that script training could influence language usage outside of the target script. Also, the same study concluded that the learned scripts were useful in situations such as talking on the phone, reading, and writing. The authors attribute these generalized gains to “the multimodal nature of training that engaged auditory and reading comprehension as well as verbal production” (p. 27).

Cherney et al. (2008) report minimal to no change on the Quality of Communication Life Scale (QCL) or on the Communication Activities of Daily Living- Second Edition (CADL-2) (Holland, 1980), given scores from before and after script training intervention. However, other studies have demonstrated that script training can have a positive effect on communicative confidence. In particular, scores on the communication difficulty subscale of the Burden of Stroke Scale (Doyle et al., 2003) decreased significantly after a typical 9-week script training intervention (Manheim, Halpher, & Cherney, 2009). The communication difficulty subscale investigates 7 domains of communication, each on a 5 point rating scale. The seven items address the following questions: “Because of your stroke, how difficult is it for you to (1) talk; (2) understand what people say to you; (3) understand what you read; (4) write a letter; (5) talk with a group of people; (6) be understood by others; and (7) find the words you want to say?”. An average decrease of 6.79 points on these ratings from before intervention to after intervention seems to indicate improved communicative confidence after script training intervention. However, one must consider that the results may be skewed by participants feeling obliged to demonstrate to the researchers that script training was a positive experience for them. This potential rater artifact may have affected the results of post-treatment communication difficulty subscale ratings.

There are conflicting results regarding the effect of script training on underlying linguistic competence. In the initial study by Holland et al. (2002), participants showed little improvement on a
measure of linguistic competence, the WAB. However, in the first computerized script training study (Cherney et al., 2008), WAB Aphasia Quotient scores increased for two of the three participants by more than five points, which is more than the standard error of measurement for this test.

Youmans et al. (2005) included generalization training as a component of the intervention process. After each 3-4 sentence script was mastered, it was practiced with a variety of deviations from the learned script. The clinician working with the participant varied her lines in the dialogue. The script was also practiced with seven new conversation partners who had been trained in supported conversation (Kagan et al., 2001), but who had no precise knowledge of the content of the script. Finally, novel partners were instructed to say something appropriate but surprising in an attempt to throw off the person with aphasia. Novel partners were students and faculty in speech-language pathology who had been trained in supported conversation and had experience communicating with people with aphasia.

After mastery and generalization training, generalization probes were administered and measures were taken and compared to measures taken from the original scripts. These generalization probes are not explicitly described, but presumably they paralleled the generalization training conditions. Both participants were able to produce all three scripts with 80% to 100% accuracy during the generalization trials. Furthermore, both participants demonstrated the most rapid rate of speech in response to generalization probes on two of their three scripts. The authors concluded that “automatization of scripts combined with specific generalization practice promoted an ability to use scripts flexibly and functionally in conversation” (p. 447).

Summary of previous research

Research to date has shown consistently that script training is a valuable and effective tool in teaching people with aphasia very specific monologues or dialogues. Although length, content, and complexity of scripts are variables that are still being optimized, there is strong evidence purporting that people with aphasia are good candidates for learning scripts. There is also evidence that succeeding with this intervention provides people with aphasia more confidence in their
communication. However, data and methodology regarding generalization of learned scripts to more novel and functional situations have been mixed and vague. There is preliminary evidence that script training could improve communicative competence in circumstances somewhat extended from the limited context of the script.

**Rationale for further research**

If it is found that learning a script improves the ability of people with aphasia to discuss topics in functional settings, then script training may have wider application than already documented. For instance, a person may be able to learn a script about grocery shopping, and become more fluent, confident, and communicative in any grocery shopping situation. While encouraging, the established data fall short of being truly compelling. In the single study in which generalization probes were used, they were not clearly explained, and the study included specific generalization training in the treatment. They also used very short, limited scripts. Furthermore, the various degrees of generalization are not explored. It is unclear how much a generalization trial can deviate from the established script and still result in improved accuracy compared to baseline measurements.

While many dependent variables have been examined in script training research treatment effects on speech fluency have not been considered.. As scripts become more automatic, one might expect disfluencies to decrease and speech intelligibility to improve. This may be the result of more automatic motor sequencing programs or improved automaticity of particular word and grammatical structures. A recent study (Youmans, Youmans, & Hancock, 2010) found that adults with apraxia of speech benefited from script training intervention. The participants all produced 90% of the words in their target scripts. Only words with no more than one sound production error were included. Also, the participants with apraxia decreased the number of errors in their script productions during the maintenance phase. Errors included repetitions, pauses, unintelligible utterances, and interjections. Applying techniques borrowed from motor learning theory, Youmans et al. modified the feedback schedule and pace of script practice in hopes of optimizing oral motor planning ability. The
achievement of script learning for people with apraxia of speech suggests that script training may have an effect on ease of articulation and this possibility should be examined empirically.
IV. PURPOSE

The research on script training to date describes positive outcomes of the production of a target script. The research base is still emerging, thus the first aim of this study was to replicate a positive effect of script training on script learning, based on commonly used variables such as content, rate of speech, and grammatical productivity. Fluency of speech production which is often impaired for people with aphasia, was also included as a dependent variable in this study. We wanted to see whether achieving automaticity could reduce the use of disfluencies. Previous research has incorporated measures of disfluencies in determining number of errors. In contrast, number of disfluencies was a variable in and of itself in this study. A second purpose of this study was to determine whether people with aphasia can use learned scripts in situations more functional and less structured than simply being asked to produce the script. Generalization was examined in three contexts in which participants had conversations about their script topics with partners who did not “stick to the script”. A third aim was to explore the usefulness of remote script training sessions via Skype. Remote script training was not isolated as an independent variable, but most treatment sessions were conducted via Skype.

Specifically, the following research questions were targeted:

1. Does non-computerized script training intervention improves accuracy, grammatical competence, and rate of speech in a target script?

2. Does script training decrease disfluencies in a given script?

3. Does script training intervention affects the accuracy, grammatical competence, fluency, and rate of speech in more general conversational contexts and when speaking with a novel partner?
4. Is remote script training via Skype combined with in-person sessions a viable option to produce script learning effects?

It was hypothesized that script training intervention would have a positive effect on accuracy, grammatical complexity, and fluency in a script. It was hypothesized that these variables would show a positive, albeit smaller effect during generalization trials. Furthermore, it was hypothesized that number of disfluencies would decrease during script training intervention. All of these effects were expected given most intervention sessions occurring via Skype.
V. METHODS

Participants

Participant A was a 41-year-old right-handed female. She was monolingual and her native language was English. Her exact medical history was unclear, but her aphasia was caused by a gunshot wound to the left cerebral hemisphere. She was 7 years post-onset of aphasia. Participant A profiled with Broca’s aphasia on the Western Aphasia Battery. Her spontaneous speech score was 13 out of 20, the auditory verbal comprehension score was 5.4 out of 10, the repetition score was 4.5 out of 10, and the naming and word finding score was 5.7 out of 10. Overall, her aphasia quotient was 57.2 out of 100. An oral motor exam revealed mild-moderate right lower facial weakness and evidence of nonverbal oral apraxia. Participant A had hearing acuity within normal limits, with thresholds less than 25dB HL at all tested frequencies.

Participant B was a 49-year-old right-handed male, 5 years post a large ischemic stroke to the left cerebral hemisphere. He was a monolingual native speaker of English. On the WAB, he profiled with conduction aphasia. His spontaneous speech score was 14 out of 20, the auditory verbal comprehension score was 7.75 out of 10, the repetition score was 6.5 out of 10, and the naming and word finding score was 7 out of 10. Overall, his aphasia quotient was 70.5 out of 100. Participant B presented with very mild lower facial weakness on the right side. Participant B had auditory acuity at 30 dB HL for 500, 1000, and 2000 Hz, but a steeply sloping moderate to severe sensorineural hearing loss at higher frequencies. His hearing thresholds were 70-75 dB at 4000Hz and 50-60dB at 8000Hz. While this hearing loss may have implications on communication, Participant B’s hearing thresholds were in the range for qualification for this study, and it appears that it did not negatively impact his success with script training.

Script development
One priority of this study was to follow principles of the life participation approach to aphasia therapy. Script training was guided by this ideal. In this case, the life participation approach was upheld by empowering participants to select their own script topics and to play a role in constructing the scripts. Participants were encouraged to think of situations in their own life in which they would like greater communicative competence and confidence.

Script length was intended to be determined by the participant’s aphasia quotient on the Western Aphasia Battery. For an AQ less than 50, there was to be a total of 20-30 words in each script, for an AQ 50 to 75, there was to be 30-40 words in each script, and for an AQ greater than 75, the goal was to include 40-50 words in each script. In practice, however, script length was determined with these guidelines in mind, but the total number of words produced during conversations recorded prior to finalizing the scripts was also a factor. These conversations were informal and unscripted. They were based on the topics participants had chosen and they were conducted prior to the completion of script writing. Participant A used 72 words in her conversation sample on the topic of her aphasia, including 38 unique words. This information, taken in consideration with her aphasia quotient, modified the target number of words per scripts to 19 and 21 words, considerably fewer than the 30 to 40 words that would be used based on her aphasia quotient. Participant B used 97 words in conversation on the script 1 topic and 126 words in conversation on the script 2 topic. His Aphasia Quotient alone would have predicted only 40 to 50 words in his script; however, that seemed limiting for this particular participant, so longer scripts were deemed more appropriate. His scripts were 89 word and 87 words long.

**Participant A**

Script development sessions were held in the participant’s home. The first session began with a general conversation about her interests. This conversation was facilitated by objects in the participant’s home. Next, the task of identifying two topics for scripts was more specifically addressed. The concept of a dialogue was explained and the participant was given examples. The participant was asked to think about everyday situations and special situations during which she
wanted to become more communicative. She was shown 27 clip art pictures and simple photographs depicting topics chosen for scripts in previous research. These included meetings, talking to children, talking to loved ones, participating in work-related activities, talking on the phone, talking about aphasia, talking about travel, and communicating in various community settings. Participant A immediately identified two topics that were of high priority to her: talking about aphasia and talking to children.

Following this initial presentation of potential topics, Participant A participated in a card sort using the Life Interest and Values (LIV) cards (Haley et al., 2010). Thirty-three cards were presented from the total ninety-five LIV cards. The selected cards either depicted situations in which one might have to communicate or a topic on which one might choose to converse. Cards that did not lend themselves to either of these categories were excluded. For example, cards for taking out the trash, doing puzzles, and sitting and thinking were not used. Given a LIV card, the participant was asked whether it depicted something she wanted to talk about. Following LIV card protocol, a checkmark was used to mark the pile for affirmative answers, while an ‘X’ was used to mark the pile for negative responses. After participant A completed the sort, topics selected during the card sort as well as the topics she previously identified were discussed individually. Ultimately, the participant decided her script topics should be talking about aphasia to a new acquaintance and talking to her 15-year-old son.

At the conclusion of the first script development session, the participant was asked to think about what she would like to include in her scripts, write some notes, and bring any materials that were relevant to the topic subjects, such as pictures, letters, or items to the next script development session.

The second script development session took place at Participant A’s home as well. This time her son was present. Each topic was discussed in more depth. The participant did not bring any materials to the session, but the clinician offered a list of published suggestions for talking to people with aphasia from the National Aphasia Association website as a starting point. Open-ended questions such as “what is aphasia?” and “what do you want people to know when they meet you?”
helped elicit information about what the participant wanted to convey. Participant A’s son provided information about himself to better inform the discussion on the script about him.

Throughout conversation, the clinician took notes. The participant verified which pieces of information she wanted to include in her script. Before anything was written for the script, she approved the content. By the end of the conversations, two scripts were drafted. These scripts were read to the participant and she had another opportunity to make any changes. Participant A stated she was satisfied with her script at this point and did not suggest any changes.

Ultimately, two scripts were finalized. The participant’s lines in script 1 had 29 total words, 5 grammatical morphemes, and 19 content words. Her lines in script 2 had 30 total participant words, 7 grammatical morphemes, and 21 content words.

Participant B

Script development for Participant B began with a meeting with him and his wife in their home. After briefly discussing the purpose of creating dialogues, the participant’s most prevalent communication difficulties were discussed. The participant identified talking on the phone to make appointments, talking on the phone to wife, and talking to his children as situations for which he wished to improve his communicative competence. Next, the participant was presented the same clip art drawings depicting script topics selected from previous research. Participant B reiterated his desire to address the situations he had already discussed. To introduce other ideas, the participant completed a card sort using the LIV card set described above.

He completed the sort very independently and became very excited about certain situations, holding up the cards, nodding, and talking about the ones that interested him most. In particular, he identified talking about politics, conversing with friends over coffee, talking about travel, and, again, talking on the phone as topics of interest for his script. The participant was asked to choose only two situations from everything we had discussed, and decided on talking about politics and making an appointment over the phone. Participant B explained that he had majored in political science in college and loved politics, but found his aphasia to be a large barrier to participating in political
conversations. He also explained that making appointments over the phone was a stressful task for him, and one he had to engage in frequently.

Once the topics were selected, Participant B had a conversation about each topic with the clinician. Additionally, the participant was asked directly what information he would like to include in the script. The clinician took notes from the conversations about what he identified as important information to convey.

At the next script training session, Participant B was presented with rough drafts of his two scripts. After discussing both topics further, the scripts were read to Participant B and he was given a written copy. He was asked to make changes to the overall script as he saw fit. Then, for each line the participant was asked whether he would like to omit the line or make any changes to it. He added lines, altered wording, and made some of the details more accurate and personal. At the end of the second script development session, Participant B expressed satisfaction with both of his scripts.

Script 1 consisted of 87 words, 15 grammatical morphemes, and 64 content words. Script 2 consisted of 87 words, 30 grammatical morphemes, and 55 content words.

*Baseline conversation*

At the start of the second script development session, a general conversation probe was taken. The participants were asked to engage in conversations on each of the topics of their scripts. At this point, participants had chosen their script topics, but had not yet seen even drafts of their scripts. These short conversation samples corresponded loosely with the expected length of the script. The samples were analyzed based on the same dependent variables used to describe all probes, except for percent script related words, because the target script had not yet been finalized. These samples were obtained to provide a comparison for conversation samples obtained during generalization probes after the intervention and maintenance phases of script training were completed.

*Research design*

This study used a multiple baseline design across behaviors. In this case, each “behavior” was a different script that was learned. This design was most appropriate for this study because of the
small sample and because the intervention had long-standing effects; a learned script cannot be unlearned immediately. In a multiple baseline design, an intervention is applied to more than one behavior or participant. In this case, the intervention (script training) was applied to two different scripts for each participant. After baseline measurements were recorded, the intervention began for the first behavior while baseline measures were continued for the other behaviors. The purpose of this delay was to demonstrate that the dependent variables remained constant for the untreated behavior while changing for the treated behavior. In this way, it can be inferred more conclusively that the intervention is the factor causing the change in behavior. Intervention is supposed to begin for the subsequent behavior when a change has been demonstrated in the previous behavior. Experimental control is demonstrated if the treatment effect is replicated in the second behavior.

In this case, participants made baseline recordings of their scripts prior to intervention. After three baseline probes, script 1 entered the intervention phase, while script 2 remained in baseline phase with no intervention. Participants were instructed not to practice script 2 and were not given audio recordings of script 2 until the start of the intervention phase for that script. Script 2 intervention began two weeks after script 1 intervention had been initiated. The participants had one week when both scripts were in the intervention phase.

In-person treatment sessions took place either in the Center for Aphasia and Related Disorders at UNC Chapel Hill, or in the participants’ homes. For Skype sessions, the participants were in their home and the clinician was in a private room. The participant and clinician logged into Skype at a previously agreed upon time. This system was successful, as all parties were consistently reliable and punctual.

*Script training intervention*

Script training procedures in this study were based on those reported in previous script training studies ([Lee et al., 2009; Youmans et al., 2005]. Scripts were trained one line at a time using a cueing hierarchy that began with phrase repetition, continued on to choral reading, and ended
with independent production. A new line was learned after the previous line has been produced at least ten consecutive times at the given cueing level.

The initial goal was for participants to produce ten accurate productions of a line at the current cueing level before decreasing cueing to the next step in the hierarchy. This was the procedure for learning most lines in the initial intervention sessions. It was followed for approximately the first three sessions for Participant B’s scripts and the first six sessions for Participant A’s scripts. However, as participants became more familiar with the scripts, this ideal was often not met. Participants were told that if they became frustrated or bored, they could opt to move to a different line before the criteria for the previous line had been passed. Similarly, participants could choose to move past a cueing level before demonstrating 10 accurate responses at that level. In practice, both participants were aware of what level of cueing was necessary for achieving accuracy. If the participant could repeat the line perfectly, he or she often opted to work at the level of independent production, rather than repetition or choral reading. Between 5 and 10 utterances of the phrase were required before moving to the next line, depending on the complexity of the line, the confidence of the participant, how long the participant had been working on the script, and whether that line had been consistently problematic. The clinician suggested practicing task (i.e. “Repeat this line 10 times.”) based on these factors. However, the participants had the final say as to how they practiced a given part of their scripts. At any point they were able to work on a specific word or phrase in isolation.

Skype posed a challenge to choral reading. Imperfect internet connections created imperfect timing conditions. Even if the delay was only a few tens of milliseconds (and often it was much more), it was impossible to perform choral reading. Therefore, repetition and independent production were the only two levels of cueing available in that modality. In order to add intermediate steps of difficulty and more support in learning new lines, the clinician requested that participants practice lines in segments, even to the level of individual words. In this way, phrase length became the more common mode of hierarchical support for Skype sessions. When a participant could produce a single
word ten times, he or she would be asked to produce a two or three word phrase including that word, and so on until the entire line was produced accurately.

Particular methods were introduced to address problems that arose as participants were learning their scripts. For example, at one point in the intervention process, Participant A could produce the lines of her script 2 with no difficulty when cued, but when simply reading the script, she had difficulty saying her lines. To address this problem, part of Participant A’s intervention for script 2 focused on practicing her lines in pairs, with the partner’s line in between. An example of an exchange that was practice like this would be:

PWA: How are you doing?
Clinician: Okay. I have a Russian test tomorrow.
PWA: Good. Are classes hard?

This contextual practice was useful in emphasizing the initiation of lines in context. The initiation problem was also addressed by explicitly encouraging self-cuing strategies. For instance, thinking of words or gestures to associate with the first word of a line proved to be useful. The observed benefit of self-cueing in script learning will be discussed in more detail below. The last ten minutes of intervention sessions were dedicated to practicing the learned lines in the context of the script. Participants were encouraged to produce lines without cues, but lines were modeled as needed. Written scripts were provided at all levels of cueing.

In addition to training with the clinician, participants were given a digital recording of their script made by the clinician. Participants agreed to 15-30 minutes of script practice, five days per week. Three out of four of the scripts were recorded directly onto the participants’ personal computers. Participant B had his second script recorded line by line onto his cell phone. Participants were asked to practice their script in accordance with their current level of cueing. They were asked to keep track of practice time on log sheets given to them at the start of intervention, but neither participant kept consistent data. In an interview following the study, Participant A and her family and caretaker reported that she practiced formally with recordings one or two times each week, and
Informally daily. Informal practice for her consisted of practicing the lines in her head or out loud, but without the model. Participant A reported practicing the script with her caregiver, mother and son, in addition to independently. Participant B reported practicing one script every day of the week. He used the computer or cell phone recordings as necessary. He also reported that for about half of his practice sessions, he was assisted by his wife or his mother.

After 3 weeks of intervention, scripts entered the maintenance phase. During this phase, participants were asked to continue practicing the scripts on their own. However, scripts in maintenance phase were no longer addressed during intervention sessions. After three probes were taken in the maintenance phase (about one week), participants were told they no longer needed to practice that script. However, since script practice at home was not formally documented, it is possible participants continued to practice both scripts. When questioned during maintenance probe sessions, Participant A reported continued practice of both scripts. Participant B reported that he did not practice Script 1 after the first week of the maintenance phase.

*Frequency and intensity of intervention*

Each participant participated in about three script training sessions each week. Each script remained in the intervention phase for three weeks. About 30 minutes of intervention time was allotted for each script, each session. Overall, participants received 4 hours of intervention on their first script and 4.5 hours of intervention on their second script. Script training sessions took place either in person or via Skype. In all, Participant A took part in two intervention sessions in person (sessions 5 and 7) and six interventions via Skype for script 1. For script 2, she took part in two intervention sessions in person (sessions 3 and 6) and seven interventions via Skype. Participant B took part in three (sessions 1, 3, and 8) intervention sessions in person and five interventions via Skype for script 1. For script 2, he took part in three intervention sessions in person (sessions 3, 4, and 6) and six interventions via Skype.

*Generalization probes*
Generalization probes took place over two sessions for each participant. The probes were the same for both sessions, except for a change in the novel partner. There were three generalization probes. Each probe was repeated for both script topics. First, a novel partner engaged in a conversation with the participant on the topic of his or her script. The novel partner had no or limited prior knowledge of the content of the participant’s script and was only told the topic and situation. For instance, the novel partner was told, “You are the receptionist at a doctor’s office. The participant is calling you.” For the second probe, the clinician who worked with the participant throughout the script training intervention engaged the participant in a conversation using lines that had been written prior to the session. The lines were constructed so that the participant could use his or her script lines to respond. However, they did not elicit the script in order and did not mandate the use of script lines.

For the third generalization probe, the novel partner from the first probe used a new set of lines written before the session to potentially elicit the participant’s script lines, similar to the second probe. The second and third probes were the same, except that the conversation partner was different and the lines the conversation partner used were different. In both probes, the conversation partner was able to modify her lines to make the conversation more natural, so not every script line was necessarily targeted on every trial, although this was the intent. For all probes, the participant was informed of the extent of the conversation partner’s knowledge of the script. The participant was told he or she may choose to use his or her script lines, but that it was not mandatory. The participant did not have the option of using the written script during generalization trials.

Probe recordings

Recordings for probes were taken at the start of the subsequent session. For instance, the first “Intervention” probe was elicited at the start of the second intervention session. Regardless of their current level of cueing, participants were asked to produce the script independently, with only the written script in front of them. They were able to use their own copy of the scripts, even if they had written notes on it to help themselves. No cues were given to participants during recordings, except occasionally participants were prompted to move on from a line they could not produce. Recordings
were made on an Olympus Digital Voice Recorder WS-300M using an external microphone, or
directly on a laptop computer using the program Audacity. All Skype sessions were recorded using
Audacity.
VI. DEPENDENT VARIABLES

Percent script related words

Script related words were the content words found in the target script. They included nouns, verbs, adjectives, negation (not), affirmatives (yes), and greetings. This variable was a measure of the percentage of these words the participants used compared to the total number of possible content words in the target script. Content words for each script were identified and totaled prior to script analysis. For each probe recording, the number of target content words that the participant used was divided by the number of target content words and multiplied by 100, yielding a percentage. If a content word was used more than once in a probe recording but only once in the target script, it was only counted as one script related word.

Percent words with grammatical morphemes

The definition of grammatical morphemes was based on the work of Brown (1973) in language acquisition. Each of the morphemes he identified (bound and unbound) was counted. Morphemes were counted each time they occurred in the script. For example, in the phrase, “adults and kids,” the plural ‘s’ morpheme would be counted twice. Once the total number of morphemes used by the participant in a script was determined, this number was divided by the total number of complete and intelligible words and multiplied by 100. Brown’s grammatical morphemes are present progressive (ing), prepositions, plural marker or irregular plural, past tense marker or irregular past tense, possessive (‘s), copula (contracted or uncontracted), articles, 3rd person present tense marker (regular and irregular), and auxiliary (contracted or uncontracted). Of note, the “to” that marks an infinitive was not included as a morpheme. Grammatical morphemes were only counted if they were used correctly in terms of its position in the sentence. The morpheme had to be in an appropriate place in the syntax, but it did not have to agree in number, person, and tense. For example, in the
sentence, “we is starting to make progress”, “is” would be counted. But in the sentence, “I can come Tuesdays mornings”, only one plural “s” would be counted. Only grammatical morphemes from words included in the total word count were included. See more on total word count below.

**Disfluencies per word**

Disfluencies were defined as repetitions, revisions, and self-interruptions. A repeated part-word, whole-word, or phrase counted as one repetition. If any of these units were repeated multiple times, each repetition was scored as a separate disfluency. Revisions were defined as more than one attempt at a word. Again, multiple attempts at the same word were counted individually. Self-interruptions occurred when the participant stopped mid-word or mid-phrase to insert an interjection or filler word. If the filler word occurred at the beginning or end of a phrase, it was not scored as a disfluency. Prolongations were not scored as a disfluency. The total number of disfluencies in a script was divided by the total number of participant words, defined below as total words, to yield a ratio.

**Rate of speech**

Rate of speech was defined as the total number of words used by the participant divided by the total time in minutes of the participant’s turns in a given script. This measure included all participant words, not only script-related words. In this case, rate of speech was a measure of the productivity of speech. In other words, it measured how much time it took for a participant to convey information. Total time was measured by recording the starting and ending time of each of the participant’s turns, taking the difference, and adding them together. This measurement was taken in seconds by listening to the audio in real time and recording the time from the program’s counter at the start and finish of each participant line. The differences were summed and then divided by 60 to convert it to minutes for the rate of speech measurement. Total time was defined as a measure of the participant’s rate of speech, so the conversation partner’s lines were not included. A turn was defined as beginning approximately half a second after the conversation partner finished speaking and ending after the last word the participant said, including pauses, filler words, and tangents. In some cases, the participant’s turn ended with silence that was clearly indicative of the participant thinking because he
or she was unable to complete the line but unwilling to move on. In these cases, the length of the turn
was calculated to include those silences, even though they occurred at the end of the turn. On rare
occasions, the participant did not concede his or her turn until prompted by the clinician with the
question, “should we move on?” Admittedly, this subjective determination that the participant had
spent enough time thinking about the line without retrieving the target influenced total time.
However, all probes were given by the same clinician with approximately the same threshold for
retrieval time, which was about 15 seconds.

Filler words

Filler words could be a single word (“um”, “no”, “wait”, “sorry”), or an entire phrase (“I
know it!”, “hold on”, “just a minute”). These are words or phrases used at the beginning, middle or
end of a line. They are not attempts at the target script, but instead are interjections and disruptions in
the speech. However, exchanges clearly not related to attempting the script were not counted as filler
words, such as a participant telling the clinician about an unrelated event. On occasion, a participant
used a filler word intentionally for inflectional purposes. For instance, Participant B often said “ah,
no” instead of simple “no”. The “ah” was not considered a filler word, as it was used functionally to
create an apologetic intonation.

Total number of words

Total number of words was not a dependent variable by itself, but the definition is important
for determining disfluencies per word, percent words with grammatical morphemes, and rate of
speech. Total words referred to the number of words the participant used. In order to be included,
words had to be intelligible in context. Contractions were counted as two words. Filler words, as
described above, were not included. Similarly, interjections and false starts were not included, nor
were repetitions or revisions. The only exception was when a second attempt at a phrase included
different words. In that case, these words were counted. Similarly, if a phrase was repeated in a
different line, both repetitions were counted towards the total words. Words quietly uttered by the
participant for the purposes of self-cueing were not included in the total word count, nor were exchanges that were clearly not attempts at the script.
VII. SCRIPT ANALYSIS

All probe recordings were analyzed by the principal investigator. Analyses were completed on a personal computer using the digital audio recordings from the probes. First, all recordings were transcribed verbatim. Variables were derived from the written transcription and number of script related words was compared to the target script. Next, scripts were analyzed from the original recording a second time to calculate total time as described above (see “Rate of speech”).
VIII. RELIABILITY

Inter-rater reliability scores were calculated for each of the four main dependent variables. Twenty percent of the probes were re-scored by two research assistants. Probes from all phases of the study were included, except for generalization. One research assistant scored script recordings for Participant A, while the other scored script recordings for participant B. Research assistants were given recordings of the scripts and written instructions on how to analyze scripts for each variable. Using the Pearson product moment correlation, inter-rater reliability was .98 for percent script related words, .97 for rate of speech, .61 for words with grammatical morphemes, and .07 for disfluencies per word.

Inter-rater reliability was excellent for percent script related words and rate of speech, low for words with grammatical morphemes, and very poor for disfluencies per word. Still, we hypothesize that all data measurements are reliable. The process of obtaining inter-rater reliability was rushed, and the second raters did not receive adequate training on how to analyze the recordings. The variables with poor inter-rater reliability were harder to quantify. They required adherence to definitions that were less intuitive. Had raters received more supervised training, it is likely that inter-rater reliability would have been better. Nevertheless, conclusions from the data on grammatical morphemes and especially disfluencies cannot be made definitively.
IX. RESULTS

The results of a multiple baseline study can be described by two criteria: experimental criterion and therapeutic criterion. Therapeutic criterion refers to the clinical significance of the results. Experimental criterion refers to the ways data are determined to show an effect. Typically, for single-case research, this is accomplished by visually inspecting a graphic display of the data. The explanation of results is justified by four characteristics of the figure. The magnitude of the effect is determined by mean and level, while the rate of change is determined by trend and latency. Change in mean is the change in average rate of performance between phases. Change in level is the shift from the end of one phase to the start of another. Change in trend is the systematic increase or decrease of performance over time between phases, and latency of the change refers to how long after the onset of a condition a change occurred. Visual inspection as a means of data analysis is justified because it is insensitive to smaller effects; only large effects will be identified as noteworthy and indicative of intervention producing a change (Kazdin, 1982).

Due to the nature of script training, one would expect to see a learning curve in the outcomes once intervention has begun. Therefore, one would not expect large changes in level, as it is unlikely that a participant would drastically improve or deteriorate on a variable after only one intervention session. Instead, the trend of the data points between phases is the most telling characteristic related to determining whether the intervention had an effect. Changes in mean between phases and between scripts are informative for quantifying magnitude of outcomes. However, it is important to bear in mind the aforementioned learning curve, as averages from the intervention phases encompass probes from the very beginning of script training where a large change would be unlikely. Observations of latency of change were useful in many cases and will be discussed below.
The results are presented in Appendices C and D. It is important to point out some missing data points. Due to errors made while saving recorded probes on the computer, some raw data became irretrievable prior to being scored. Thus, unfortunately, some script samples were not included in the data. These sessions are marked with an ‘x’ on all graphs. Furthermore, Participant A was unable to converse as if she were talking to her son during the conversation probe prior to script training. Therefore, the pre-script training conversation probe is null for all dependent variables. This is also represented by an ‘x’ for the first trial of her script 2 variables. Lastly, percent script-related words was not calculated for pre-script training conversations, and this data point is likewise indicated with an ‘x’.

In all of the associated graphs, each variable for each participant is presented separately. Both scripts are presented together for each of these variables in order to show stability in the untreated script. The first phase consists of only the single probe that elicited a general conversation on the topic of the scripts before scripts were finalized. The second phase is a baseline phase in which the script was not treated. The third phase is the intervention phase in which the script was actively being addressed in treatment sessions. The fourth phase represents maintenance, during which the participants practiced the script on their own time, and the fifth phase represents the generalization probes. The three different generalization probes are represented by different shapes on the graphs. The general conversation probe is represented by a triangle, the structured conversation with the same clinician is represented by a square, and the structured conversation with the novel partner is represented by a circle.

Overall, both participants showed improvement on all variables for both scripts. For Participant A, the variables that showed the largest effects from script training were percent script-related words and rate of speech. Effect size for percent script related words was 4.04% and 5.07% for script 1 and script 2, respectively. Effect size for rate of speech was 18.88 and 8.62 words per minute for script 1 and script 2, respectively. For Participant B, the variables that showed the largest effects from script training were rate of speech and disfluencies per word. Effect size for rate of
speech 7.97 and 4.42 words per minute for script 1 and script 2, respectively. Effect size for disfluencies per word was -1.595 and -1.797 disfluencies per word for script 1 and script 2, respectively. As a whole, participants improved communicative competence in the areas of grammatical morpheme use and rate of speech during generalization trials, as compared to conversation before script training. Number of disfluencies during generalization trials was not significantly lower than during conversation prior to script training. Use of script related words during generalization trials was variable for both participants. Below, the effect of script training on each dependent variable is discussed. Then, the results of generalization probes are explained in greater detail.

Effect size for each of the four main dependent variables was determined for each script in order to quantitatively describe the overall magnitude of change in a variable from before script training to after script training. Effect size was calculated by subtracting the mean score of baseline probes from the mean score of maintenance probes, and dividing by the standard deviation of the baseline scores. Standard interpretation of magnitude of effect size is still emerging. Beeson and Robey (2006) suggested the benchmarks 4, 7, and 10.1 to constitute small, medium, and large effect sizes, respectively, for a lexical retrieval task. Most of the effect sizes in this study are considerably smaller, partially due to the amount of variance during baseline probes. The significant decrease in variance during the maintenance probes was not taken into consideration by this calculation of effect size.

Overall, effect size for rate of speech was medium to large. Effect size for script related words was small, but robust, as evidenced by the stabilization of variability during maintenance, the absence of overlap between phase measurements, and the data approaching to maximum performance on this variable. Effect size for grammatical morphemes was small but robust, as variance decreased during maintenance and there was no overlap between phase measurements. Effect size for disfluencies was small and not as robust, although there was a noticeable reduction in variation during
the maintenance phase. Effect sizes are listed in Table 1 for Participant A and Table 2 for Participant B.

<table>
<thead>
<tr>
<th>Participant A</th>
<th>Script 1</th>
<th>Script 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent script related words</td>
<td>4.04</td>
<td>5.97</td>
</tr>
<tr>
<td>Rate of Speech (words/minute)</td>
<td>18.88</td>
<td>8.62</td>
</tr>
<tr>
<td>Percent words with grammatical morphemes</td>
<td>1.96</td>
<td>3.78</td>
</tr>
<tr>
<td>Disfluencies per word</td>
<td>-0.42</td>
<td>-0.225</td>
</tr>
</tbody>
</table>

Table 1: Effect sizes for Participant A

<table>
<thead>
<tr>
<th>Participant B</th>
<th>Script 1</th>
<th>Script 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent script related words</td>
<td>1.94</td>
<td>2.68</td>
</tr>
<tr>
<td>Rate of Speech (words/minute)</td>
<td>7.97</td>
<td>4.42</td>
</tr>
<tr>
<td>Percent words with grammatical morphemes</td>
<td>1.84</td>
<td>2.08</td>
</tr>
<tr>
<td>Disfluencies per word</td>
<td>-1.595</td>
<td>-1.797</td>
</tr>
</tbody>
</table>

Table 2: Effect sizes for Participant B

**Percent script-related words**

Visual inspection of the data clearly shows that percent script-related words used was sensitive to script training intervention for Participant A (see figure 1). Percent script-related word scores remained relatively stable during baseline phases for both Participant A scripts, followed by an unmistakable upward trend beginning immediately after the start of intervention. These gains were sustained during the maintenance phase at an average of 89.47% for script 1 and 93.65% for script 2.

For Participant B, script 1 scores suggest intervention produced an effect on percent script related words (see figure 2). However, this assertion is not strongly supported because the baseline for script 2 was not stable. Instead, the last few trials indicated an increase in the percentage of script related words used by Participant B. It is likely that there was a small learning effect produced by repeated baseline probes. Participant B produced the script 8 times over the course of the baseline recordings. It appears that by the sixth elicitation, participant B had begun to learn more words in the script. Nevertheless, after intervention was initiated, percent script related words quickly neared 100%. Therefore, even if learning from repeated probes occurred, it appears that the participant learned script related words were efficiently during script training intervention than before it.
Figure 1: Participant A, Percent Script Related Words

Figure 2: Participant B, Percent Script Related Words
Rate of speech

The effect of script training intervention on rate of speech is obvious upon visual inspection of the figures for both scripts of both participants. Scores taken during baseline phases remained very stable, and there was a marked change in trend between phases. Given the increase during the maintenance phase, longer intervention would likely have yielded a greater effect. Participant A increased her rate of speech by almost 46 words per minute between baseline and maintenance for script 1, and by over 20 words per minute between baseline and maintenance for script 2 (see figure 3).

For Participant B, a latency effect occurred, as rate of speech did not increase until after two to four intervention sessions (see figure 4). This phenomena is likely similar to the disfluency outcomes in origin; in the stage where participants have some knowledge of their scripts but are not yet competent in them, it appeared they slowed down to achieve more accuracy.
Figure 4: Participant B, Rate of Speech

Percent words with grammatical morphemes

Given the marginal inter-rater reliability, the results for this variable remain unconfirmed. However, by observation of the data as it is, Participant A demonstrated improvement on the percent of words with grammatical morphemes used, as demonstrated by the effect size. Nevertheless, baseline measurements for script 1 were not stable for this variable, which weakens the argument that this effect is the result of script training (see figure 5). For script 2, the intervention effect was demonstrated by a change in trend between baseline and intervention.

For Participant B, positive trends were observed during intervention phases for both scripts and the increase in grammatical morphemes was sustained during generalization trials (see figure 6). However, both baseline phases exhibited positive trends as well, particularly in learning script 1. It is possible that intervention facilitated the increase in grammatical morphemes because the increase during baseline may have exhausted the potential improvement possible from repeated baseline trials. Still, this is speculation and one cannot claim an effect from these figures. In any case, effect size
calculations indicated Participant B improved on this variable for both scripts.

Figure 5: Participant A, Percent Words with Grammatical Morphemes

Figure 6: Participant B, Percent Words with Grammatical Morphemes
**Disfluencies per word**

The effect of script training on disfluencies was of particular interest in this study. This variable had not been studied in isolation in previous research. The results for disfluencies per word in this study are suspect due to extremely poor inter-rater reliability. Given the data as it was analyzed, disfluencies per word varied greatly throughout all phases for Participant A (see figure 7). Based on the means of each phase, disfluencies per word increased from baseline to intervention, but decreased during maintenance and generalization trials. If the recorded data are indicative of true behavior, a likely explanation is that the number of disfluencies per word increased for both participants as they became more familiar with the script and worked on producing it accurately, before eventually decreasing. In practice, the participants were making more attempts at target words and phrases, whereas with less knowledge of the script, they would simply skip those lines or words. However, once participants became more competent in their scripts, the number of decreased to lower than the baseline levels. During maintenance phases, a marked decrease in disfluencies was noted for both participants. Participant B produced few disfluencies during maintenance, although number of disfluencies per word was beginning to increase by the end of the maintenance phases (see figure 8).

![Figure 7: Participant A, Disfluencies per Word](image-url)
Generalization effects were varied. Between baseline probes, Participant B’s scripts eventually followed a negative trend. Filler words used by participant B decreased immediately after script training was initiated and then increased for a few sessions immediately after script training was initiated, and finally began to decrease. Use of filler words during script 2 probes was variable throughout the phases.

**Generalization effects**

For data from generalization probes, refer to the generalization phase of figures 1-8. Generalization effects were varied. Of note, prior to script training, it was very difficult to elicit a
conversation with Participant A in which the conversation partner played the role of her son. Therefore, there is no direct comparison for generalization measures for her second script. On the other hand, the fact that she participated in 6 probes of this nature is indicative of greater communicative competence or perhaps greater communicative confidence.

Surprisingly, no trends were consistently observed to distinguish the three different generalization contexts. The variable that showed the greatest difference from pre-script training conversation to generalization probes was the percentage of words with grammatical morphemes. Participant A had an average of 13.43% words with grammatical morphemes during generalization trials, as compared to 8.33% during conversation before script training intervention. Similarly, between these two phases, Participant B increased his percentage of words with grammatical morpheme usage from 10.31% of words to an average of 22.43% of words for script 1 and from 8.25% of words to an average of 14.39% of words for script 2. Interestingly, Participant B increased his percent of words with grammatical morphemes steadily during generalization trials for script 1 after the first probe. Script 2 scores was more variable, as were scores from Participant A’s generalization trials. Recall, however, that inter-rater reliability for this variable was .61 and generalization probes were not included in the reliability estimation.

Participant A demonstrated increased rate of speech during generalization trials. She spoke over 30 words per minute faster during generalization conversations than during conversation before script training intervention. A moderate improvement of 11.48 words per minute was noted for Participant B’s first script, on average. On the contrary, rate of speech during script 2 generalization trials for Participant B was on average similar to rate of speech in conversation before script training.

Since the script was not created at the time of the initial conversation recordings, and because much of the script content was taken from those original conversations, percent script related words was not deemed a useful calculation for the original conversations. However, the percentage of script-related words used during generalization trials is a useful measure of the participants’ ability to use their scripts functionally in more generalized contexts. There was considerable variety in how many
target script words participants used on each trial. Participant A used between 14% and 89% of the content words from her scripts during generalization trials. Participant B used more script-related words during generalization trials, but still showed variation, using between 42.27% and 73.44% of the content words from his script during generalization trials. On average, Participant A used 61.05% of her script 1 words and an average of 32.96% of her script 2 words during generalization trials. Participant B used, on average, 64.95% of the script related words from script 1 and 54.56% of the script related words from script 2.

Disfluencies per word were variable during generalization trials. For Participant A, script 1 generalization trials yielded an average of .039 disfluencies per word, which was not significantly changed from the .042 disfluencies per word produced during pre-script training conversation on the same topic. During generalization trials for script 2, she produced an average of .065 disfluencies per word. For Participant B, number of disfluencies in generalization probes was variable, but similar to the number of disfluencies produced in conversation before script training. Taken as a whole, disfluencies did not decrease in generalized conversation after script training. In any case, these results are highly suspect due to the poor inter-rater reliability for this variable. There was no significant change between number of filler words in pre-script training conversations and number of filler words used during generalization trials by Participant B. However, Participant A decreased her usage of filler words by 4.34 times from pre-script training conversation to generalization conversations.
X. DISCUSSION

The first aims of this study were to replicate positive effects of script training intervention on use of target script words, grammatical complexity, and rate of speech, as well as determine effect of script training intervention on disfluency. An additional aim was to explore how effectively participants used learned scripts in more functional communicative contexts. Finally, the outcome of combining Skype sessions with traditional face-to-face sessions for intervention was investigated. This study satisfied these aims. The effects of script training intervention were replicated most robustly for rate of speech, but strong effects were also noted for percent of target script words used. An effect on disfluencies per word was also evident; however the poor reliability makes that conclusion suspect. By introducing a new protocol for probing generalization, we contributed useful information towards understanding how people with aphasia are able to use learned scripts functionally. Lastly, we demonstrated that script training sessions can be performed via Skype were feasible and yielded positive outcomes.

Script training outcomes

Outcomes related to percent words with grammatical morphemes and disfluencies per word are tainted by poor inter-rater reliability scores. While it is likely that the scores were in fact reliable and consistent, we cannot draw conclusions based on them. Overall, both participants achieved a level of competence on their scripts that allowed them to communicate the entire message conveyed in their scripts. There is strong evidence that both participants were not only able to attain this level, but also maintained it after the treatment had ended. By the end of script training, Participant B’s errors involved occasional pauses and revisions, but he produced all script related words and grammatical morphemes. This is evidenced by the six separate probes for script 1 and the seven separate probes for script 2 in which Participant B stated 100% of the script related words. He first attained this level
after the final (eighth) intervention session for script 1, and after the sixth intervention session for script 2. He attained the percentage of words with grammatical morphemes in the target script by the fourth maintenance probe for scripts 1 and maintained that level of accuracy for the remainder of the probes. He achieved the target percentage of words with grammatical morphemes for script 2 after the final intervention session.

Beginning at the start of the maintenance phase, Participant A produced script 1 almost identically in ten consecutive probes, producing 89.47% of the script-related words and only 1 to 2 disfluencies. This sustained level of performance is strong evidence that scripts had become over-learned and automatic. Overall, the variables that were affected the most by script training became stable by nine script training sessions for both participants. Participant B achieved a stable level of competence on all variables for both scripts in seven to nine sessions. Participant A achieved a stable level of performance for her first script three to five weeks (nine to fifteen sessions) after script training was initiated.

There was one exception to this trend of achieving automaticity. While Participant A demonstrated significant improvement for all dependent variables during script 2 probes, her variable performance even towards the end of the intervention phase and during maintenance suggests that this script had not become automatic. However, scores for every probe taken of every variable during the maintenance period for script 2 shows improvement over the probes taken during intervention. Therefore, it is evident that Participant A was continuing to learn the second script, even during maintenance. It is likely that Participant A would have benefited from a longer intervention phase for this script. Implications of differences between the scripts will be discussed in further detail below.

**Observations from generalization probes**

The generalization probes administered in this study offered a crude but informative assessment of participants’ abilities to utilize learned scripts in functional situations. The dependent variable most sensitive to the comparison between these probes and pre-script training conversation was percent words with grammatical morphemes. The best explanation for this effect is that both
participants utilized many of their script lines during generalization probes. These lines had been memorized in full, with grammatical morphemes, so using them increased the number of words with grammatical morphemes. Furthermore, these probes were the first time the participants were asked to retrieve lines out of context and without a written copy of the script in front of them. This retrieval time and attempts can explain why disfluencies per word did not decrease and rate of speech did not increase for the generalization probes in relation to pre-script training conversations.

Surprisingly, none of the generalization contexts seemed consistently more difficult or easier for participants, yet two of the conditions were constructed in such a way that participants could interact in the dialogue simply by using lines from his or her script. It was observed that participants often opted to use more spontaneous utterances instead of scripted lines. It is impossible to determine whether this was due to participant’s not recalling their lines in the moment, or if it was an intentional attempt to generate novel utterances. Although participants helped write their scripts and had the final say on the wording, in real conversation, they sometimes chose to create a novel string of language.

That is not to say that participants did not use their scripts functionally when responding to generalization probes. In some cases, participants seemed to use script lines as a springboard for beginning novel ideas. For example, Participant B used lines from his script about politics to introduce new topics during conversation with a partner less versed in current events. Other times, participants used script lines simply because they were appropriate and expressed what they wanted to say. All usage of script lines was appropriate within the conversation; neither participant used script lines in a way that did not make sense. Over the course of the six probes, Participant A used elements of all four lines from her first script and elements of four out of six lines from her second script. Participant B used elements from all of his lines from script 1 and all but one of his lines in script 2 during generalization trials.

\textit{Self-cueing}

Through script training sessions, it became evident with both participants that self-cueing was an invaluable skill for script learning. Participant B demonstrated the utility of self-cueing by
successfully using it as his main strategy for word finding. He was very skilled at connecting less familiar words with more familiar words that were phonemically or semantically related, to overcome word finding and motor planning deficits. Often the less familiar words were functor words. For example, to elicit the word “they”, Participant B would whisper “hey” to himself, and then retrieve the word immediately. He reported thinking of the word “half” in order to say “have”. When he had difficulty with the word “has”, we came up with the word “hazard” to help him cue himself. As a more semantically related example, Participant B would quietly say to himself “she”, before saying the target word “he” aloud. Some self-cues were paired with limb movements as well. For example, to elicit the word “that”, Participant B pantomimed putting a hat on his head while quietly saying “hat”. This strategy was always successful for the participant and never required clinician input.

Participant A elucidated the importance of self-cueing because initiating script lines and retrieving words was challenging for her. A few sessions into intervention, she was unable to progress from choral speaking to independent production, but when cued with the initial word she could produce the entire line. In fact, on the third intervention probe for script 2, Participant A produced only 10 content words. But when she was given the first word of each line, she produced 15 out of 19 content words. She had more difficulty with initiating lines in her second script for reasons discussed below.

As described above, the inability to self-cue was remediated by practicing lines in chunks and by explicitly teaching self-cueing strategies. Unfortunately, many of the participant’s lines in script 2 started with other functor words that were harder to link with familiar words in a concrete way. But Participant A successfully used self-cues with some assistance to initiate two lines in her second script. First, her caregiver suggested gesturing like she was making a Native American greeting to elicit the word “how”. Participant A used this strategy with about 50% accuracy. Accuracy increased as she learned to initiate this cue, and eventually she was able to remember the line “how are you doing” with no self-cueing at all. She was also able to use the name “William” as a cue for a line beginning with the auxiliary “will”. With the exception of one conversation in which she asked
“William you see you friends this weekend?”, Participant A used this strategy successfully and again was able to stop using the cue once the line became more automatic. While strong self-cueing skills are not imperative to script-learning, they are a very valuable tool that helps facilitate the process. Teaching self-cueing skills in conjunction with script training may result in faster, easier script learning.

Candidacy

Both of the people with aphasia who participated in this study were motivated to improve their expressive language skills and resilient in the face of frustration. However, they differed in their type and severity of aphasia. Participant A had Broca’s aphasia and Participant B profiled with conduction aphasia, although his repetition skills were functional in practice. As described above, Participant B’s success was largely attributed to his strong self-cueing abilities. He also had script-learning experience both prior to and after onset of aphasia. Still, it was the first time either participant had attempted to learn a grammatically correct script since the onset of aphasia.

Participant B also had moderately preserved reading skills, so he was able to use the written script as a cue. The written script was helpful to Participant A once she learned the lines, but she was unable to decipher the words before she had a more developed knowledge of them.

These differences had implications for the way participants learned their scripts. Word retrieval was a challenge to Participant A, while Participant B spent more time working at the level of fluency. After three intervention sessions, Participant B reported that he knew the words but wanted to practice making it easier to say. Fluency was a target for Participant A as well, but it was achieved by improving word retrieval. This was possible because participant A’s lines were shorter and fewer, and she did not use as many disfluencies as Participant B at baseline. Intervention sessions emphasized the type of errors the participant was making. When fluency was the target, lines were practiced in context for more than just the prescribed ten minutes at the end of the session. Given the participants’ focus on either the word level or fluency, it is a logical prediction that percent script-related words would be sensitive to script training for Participant A, that number of disfluencies per
word would be very sensitive to script training for Participant B, and that rate of speech would be sensitive to script training for both participants. Regardless of their differences, both participants reported benefit from script training (see below) and improved in all dependent variables on both scripts. No contraindications for script training have been evidenced from this study. Instead, certain skills such as ability to self-cue, read target phrases, and quickly learn to retrieve words were demonstrated to make script training easier.

Differences between scripts

Although the participant scripts were written to be loosely comparable to one another in complexity and length, more subtle differences between a participant’s scripts may have influenced how the scripts were learned. Both participants reported that their first script was easier. There are many possibilities, in addition to number of words, grammatical morphemes, and content words, for this perception. To start, participants’ first scripts were about a communicative exchange with a relatively closed set of content. The scenarios were telling a new acquaintance about your aphasia and calling a doctor’s office to make an appointment. In both cases, it is easier to predict what sort of content will be included. This may have affected generalization trials as well, as conversation about script 1 topics were more likely than conversation about script 2 topics to elicit script-related words and phrases.

Furthermore, Participant A’s second script was difficult because it was written so that she drove the conversation. That script began with her line, and continued with a series of questions for her son. Therefore, the partner’s lines did not set her up for her line; there was not intrinsic cueing in the conversation. Also, although it was only one word longer than script 1, script 2 had two additional lines for the participant. As mentioned previously, initiating lines was a challenge for Participant A, so in that sense, this script was one and half times as difficult as script 1. Participant B’s script 2 was challenging for him because it included 26 multisyllabic words, as compared to 16 in the first script. Many of the words in script 2 had three or more syllables. Multisyllabic words were difficult for Participant B, so more script training intervention time had to be allotted for practicing motor speech.
patterns for these words. These aspects of target scripts should be considered in future research.

Perhaps eventually, script development could be refined to the point where clinicians could design scripts that specifically target desired skills.

**Skype**

Skype appears to be an efficient, viable method of delivering script training intervention, when supported by face-to-face intervention sessions. Even with the majority of intervention conducted via Skype, script learning effects were observed. No overt differences were observed between recordings elicited in person and recordings elicited via Skype. Skype sessions allowed patients to meet with the clinician regularly without either party spending time, energy, and money on transportation. Nevertheless, Skype sessions had unique challenges. As described in the methods section, some aspects of script training intervention had to be altered due to timing differences during Skype sessions. In particular, choral reading was difficult via Skype. The quality of visual cues was also poorer; it was more difficult to see small facial movements. Often when a participant was having a particularly difficult time with a word or phrase, we would decide to hold off on working on it until the next in-person session. In-person sessions once each week were vital for script learning success.

**Social validity**

After all phases and generalization trials were complete, participants were asked about their feelings regarding script training. Both participants had similar positive responses. They exhibited a sense of pride and accomplishment in learning their scripts. Interestingly, when asked if they have used their scripts in real-life situations, both participants reported using their second script, the one that was more difficult for them. This may be simply because these scripts addressed more commonly occurring situations. Participant A reported using her second script to talk with her son. Her son reported that the scripts have helped her with “pronunciation”. Participant B reported using lines from his script during conversational interactions at a party.

Both participants seemed to latch onto phrases from their scripts that described their language impairment. This was evidenced by anecdotal observations of the participants using these phrases in
general conversation during their participation in unrelated projects with other researchers. For
Participant A, the phrase was “I know what to say, but I can’t get it out.” For Participant B, the phrase
was “Numbers are hard for me. Please slow down and repeat them.” These lines both come from the
participants’ first scripts.

Both participants reported they would be interested in learning more scripts in the future. We
hypothesize that they would be successful conducting much of this practice independently, especially
if they could consult with a speech-language pathologist at the start of script training and then
intermittently in order to identify useful script training strategies for the participant. Because they
have difficulty with functors in spontaneous speech, participants were asked whether they would be
interested in learning scripts without worrying about morphosyntactic precision. Both participants
answered they would prefer to practice scripts that were grammatically complete.

Study limitations

The area of this study that would benefit most from modification was the generalization
probes. The probes were useful as a basis for exploring participant’s abilities to use learned scripts
functionally. However, they may not have been thorough enough. For example not all script lines
were given a chance for elicitation in the generalization probes that aimed to allow participants to use
their script lines. This occurred when the conversation partner altered her lines in order to be more
appropriate in the conversation. Furthermore, the generalization probes were artificial and limited in
time. Also, both novel partners were people who knew and had worked with the participants on other
projects. It would have been more informative to measure real-life interactions in which the
participant used his or her script. The generalization probes, as constructed and measured here, may
not have been a true estimation of use. Lastly, improvement in generalization conversation was
calculated in comparison to a conversation probe elicited before script training began. This would
have been a more robust comparison if there had been multiple conversation probes on each script
topic because the initial conversation sample may not have been representative of the participants’
spontaneous speech.
Another limitation of this study was that probing procedures were not as clearly defined as they could have been. For instance, operationalizing the amount of time allotted for the participant to search for a word or a line would have made script recording more validly comparable to one other. In many cases, the clinician intervened when she felt it was appropriate and asked if the participant would like to move on. Although the clinician attempted to be consistent, the amount of time in each situation was not measured. Inconsistencies in this regard likely affected rate of speech measurements.

The missing data points limit this study as well. It is particularly unfortunate that data from the first session of the intervention phase was lost for Participant B’s second script, as this phase would have provided useful information about changes in level, one of the criteria used for visual inspection of time series graphs generated in single-subject research designs. However, due to the nature of script training and data from other scripts, it is unlikely that there would be large changes in level between treatment phases. Tracking participants’ independent practice more precisely would have provided more information about how much time participants truly put into learning their scripts. It would have also been helpful to know in more detail how the participants practiced and how often they enlisted friends or family in their practice sessions. This information might provide insight into the outcomes of this study, and would better inform standards for future script training.

*Future research*

Future studies could more thoroughly describe and more accurately quantify participant’s use of learned scripts in more general and functional contexts. One way to accomplish this would be to ask participants to wear a voice activated recorder throughout their day for some time during and following script training intervention. This would allow researchers to collect data in a very natural environment. This data would be a more accurate measure of how script training can affect the functional communication of people with aphasia.

Further examining the impact of self-cueing on script training would offer useful information about candidacy and parallel treatments for optimal success. Future research could compare script
training before and after specific self-cueing strategies were taught. An alternative would be to examine whether people who learned self-cueing strategies learned scripts more independently and efficiently than people who were not taught self-cueing strategies. Perhaps teaching self-cueing strategies, in addition to the cueing hierarchy, would provide enough information for otherwise untrained caregivers to assist people with aphasia in script learning. These findings may have implications for increasing the efficiency of script-training. Another study that could improve script training efficiency might probe whether script training was effective when all sessions were conducted via Skype. Skype-only script training could allow people who are unable to see a speech-language pathologist face-to-face benefit from script training intervention.
XI. CONCLUSIONS

The findings in this study support the assertion that script training intervention is effective in improving discourse for people with aphasia on a chosen topic. Factors that appear to facilitate script learning include strong self-cueing skills, high level of motivation, and script that provide inherent cues from the partner’s lines. The result of generalization probes indicates that people with aphasia are able to use learned scripts in functional situations. There was evidence that script learning generalized beyond producing the target script. Furthermore, this study highlighted how people with aphasia focus on different aspects of script training, relative to their impairments and goals. This study also suggests that Skype may be a useful avenue for script training, especially if in-person sessions are impractical. Skype appears to be an effective way to supplement face-to-face interaction with a clinician, which may make script training more practical for more patients. It is this author’s hope that, as treatment protocols become more refined, script training will become increasingly effective at improving the functional speech of people with aphasia in situations that are important to their interests, goals, and interaction with others.
APPENDIX A: PARTICIPANT A SCRIPTS

Script 1

Clinician: Hi, nice to meet you.
Person With Aphasia: Hi. Please be patient. I was shot in the head.
C: Oh my. I’m sorry to hear that.
PWA: I have aphasia.
C: What is aphasia?
PWA: I know what to say, but I can’t get it out.
C: Okay. Can you understand me alright?
PWA: Yes, I understand well.

Script 2

Person With Aphasia: Come here. I need your help.
Clinician: Okay.
PWA: How are you doing?
C: Okay. I have a Russian test tomorrow.
PWA: Good. Are classes hard?
C: Yes.
PWA: Do you have homework?
C: No.
PWA: Will you see your friends this weekend?
C: Yeah, we’re going to the movies.
PWA: How is art and guitar?
C: Good, I just started a new piece.
APPENDIX B: PARTICIPANT B SCRIPTS

Script 1

Clinician: Hello, how may I help you?
Person With Aphasia: I want to make an appointment.
C: Okay.
PWA: I had a stroke and do not speak well, but I understand everything. Numbers are hard for me. Please slow down and repeat them.
C: No problem. What is your name?
PWA: **. I need my 1 year check-up.
C: Which doctor do you see?
PWA: Dr. Watson.
C: Okay. When are you available?
PWA: I can come Tuesday, Wednesday, or Thursday mornings.
C: How about this Tuesday at 8 am?
PWA: No, I live two hours away. I need time to drive there and after I have to pick-up my children at two thirty.
C: How about Wednesday morning?
PWA: Can I come at ten or eleven o’clock?
C: I have an appointment at 11 o’clock.
PWA: Can you repeat that?
C: 11 o’clock.
PWA: Yes, that’s fine. Thank you.
C: You’re welcome. Goodbye.

Script 2

Clinician: Do you follow politics?
Person With Aphasia: Yes, I am very interested in politics.
C: Are you a democrat or a republican?
PWA: I am a democrat.
C: What do you think of Obama?
PWA: I like Obama because he’s smart and thinks about everything. He listens to both sides.
C: I agree.
PWA: He has not fixed our problems, but we are starting to make progress.
C: So are you happy with the political scene right now?
PWA: No, I am frustrated with our government.
C: Why are you frustrated?
PWA: There is friction between the democrats in the senate and the republicans in the house.
C: What do you think will happen?
PWA: Absolutely nothing! They need to compromise, but they are too stubborn.
C: Are you for or against the health care bill?
PWA: I like the health care bill. It still has problems, but it is better than nothing.


