ABSTRACT

Title of Thesis: TEXT RECOGNITION SKILLS AND TEXT RECALL ABILITY OF ADULTS WHO DO AND DO NOT STUTTER IN ORAL AND TYPED MODES

Erica Lescht, Master of Arts, 2017

Thesis Directed By: Professor Nan Bernstein Ratner
Department of Hearing and Speech Sciences

BACKGROUND: Growing evidence suggests subtle weak language and memory skills in adults who stutter (AWS). It is unknown if differences are due to oral task demands; prior memory studies have focused on phonological short-term memory. This study further investigates story retelling and word fluency skills in AWS and adults who do not stutter (AWNS).

METHOD: Fifteen pairs of AWS and AWNS completed story retell and recognition tasks conducted in oral and typed modes. Word fluency (lexical retrieval) skills, digit span in both modes, and vocabulary were also assessed.

RESULTS: No significant differences were found between groups on most measures, across modes, with the exception of word fluency, where AWS produced fewer items in both modes. Vocabulary and word fluency skills correlated with story retell ability in AWNS, but not in AWS. Findings suggest potential lexical access and retrieval differences in AWS, but no specific disadvantages posed by oral responses.

TEXT RECOGNITION SKILLS AND TEXT RECALL ABILITY OF ADULTS WHO DO AND DO NOT STUTTER IN ORAL AND TYPED MODES

by

Erica Lescht

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Arts
2017

Advisory Committee:
Professor Nan Bernstein Ratner, Chair
Professor Rochelle Newman
Professor Jan Edwards

Acknowledgements

I am very appreciative of Dr. Nan Bernstein Ratner for all of her time spent, guidance, and support for this master's thesis and over the past several years. I would like to thank Dr. Jan Edwards and Dr. Rochelle Newman for their direction, feedback, and for serving on my master's thesis committee. Further, I would like to thank Melissa Stockbridge for her assistance and guidance in making a script and allowing us to use her computer software. I would also like to thank Courtney Luckman, Vivian Sisskin, Sara MacIntyre and Heather Grossman at the American Institute for Stuttering, the Rockville National Stuttering Association chapter, and the Northern Virginia National Stuttering Association chapter for their help with participant recruitment. I am thankful for being awarded funding from the MCM Fund for Student Research Excellence, which helped to support this research project. Lastly, I would like to thank my parents, sisters, family, friends, and classmates for their constant support and encouragement.

Table of Contents

List of Tables

List of Figures

Introduction

Linguistic Factors in Stuttering

In recent years, investigations into the potential causes of stuttering have gone beyond earlier explanations, which tended to emphasize potential motor coordination differences between adults who do and do not stutter (Bloodstein & Bernstein Ratner, 2008). An area of large interest has been language skills in people who stutter (PWS). Numerous investigations have found differences in semantic, syntactic, and phonological processing skills in adults who stutter (AWS) compared to adults who do not stutter (AWNS; e.g., Bosshardt, 1993; Bosshardt & Fransen, 1996; Bosshardt, Ballmer, & De Nil, 2002; Byrd, McGill, & Usler, 2015; Byrd, Vallely, Anderson, & Sussman, 2012; Ludlow, Siren, & Zikria, 1997; McGill, Sussman, & Byrd, 2016; Newman & Bernstein Ratner, 2007; Weber-Fox, Spencer, Spruill, & Smith, 2004; Wingate, 1988). Lexical access has been a major focus of much of this work.

Influence of Lexical Factors in Adult Stuttering

One way in which linguistic attributes appear to interact with stuttering profiles is in the distribution of stuttered events in spoken language. For instance, lexical features influence where disfluencies occur. AWS are more likely to stutter on uncommon or low-frequency words, on content words rather than function words, and on longer words (e.g., Brown, 1937; Brown & Moren, 1942; Danzger & Halpern, 1973; Eisenson & Horowitz, 1945; Griggs & Still, 1979; Hahn, 1942; Hubbard & Prins, 1994; Lanyon & Deprez, 1970; Newman & Bernstein Ratner, 2007; Palen & Peterson, 1982; Prins, Main, & Wampler, 1997; Quarrington, Conway, & Siegel, 1962; Ronson, 1976; Schlesinger, Melkman, & Levy, 1966; Silverman, 1972; Soderberg, 1966; Taylor, 1966; Wingate, 1967).

Do AWS demonstrate atypical lexical access compared to AWNS?

Such profiles have long given rise to the question of whether AWS have difficulty with lexical storage or access that might lead to fluency breakdown. In studies that have focused on production of individual words using a variety of behavioral measures, AWS had lower accuracy on word naming tasks (e.g., Newman & Bernstein Ratner, 2007), showed more variability in responses on word association tasks (e.g., Crowe & Kroll, 1991; Wingate, 1988), used fewer synonyms and lengthier definitions when defining vocabulary words (e.g., Wingate, 1988), and generated fewer words in a written word fluency task (e.g., Wingate, 1988). While these findings indicate differences in lexical access between AWS and AWNS, it is unknown whether storage or retrieval differences are factors in these profiles.

Research evaluating potential word retrieval difficulties in AWS using response time latencies has yielded conflicting results. On word association and word naming tasks, AWS appear to respond just as quickly as AWNS (Crowe & Kroll, 1991; Newman & Bernstein Ratner, 2007; Taylor, Lore, & Waldman, 1970; Wingate, 1988), once responses impacted by stuttering are eliminated. In contrast, another study using a word naming task found that AWS had slower latencies, especially for naming verbs (Prins, Main, & Wampler, 1997). Similarly, AWS have been found to have slower reaction times on lexical decision tasks, which ask participants to identify if a stimulus is a real word or a nonsense word (Hand & Haynes, 1983; Rastatter & Dell, 1987).

Other research has suggested that, depending upon what is measured, lexical access abilities in AWS do not differ from AWNS (Newman & Bernstein Ratner, 2007; Packman, Onslow, Coombes, & Goodwin, 2001). As noted, Newman and Bernstein Ratner (2007) evaluated lexical access in AWS and AWNS in a confrontation naming task**.** Stimuli were selected based on word frequency, phonological neighborhood density, and phonological neighborhood frequency features since these lexical factors have been shown to impact lexical retrieval. As previously mentioned, AWS made more naming errors. However, lexical features influenced performance similarly between groups. This finding suggests that AWS and AWNS have a comparable lexical organization, but less efficiency in retrieving words. Subtle differences in lexical access processing may thus impact fluency breakdown and word retrieval in continuous speech.

Potential learned profiles of speaking and stuttering might emerge over the lifespan. For instance, oral tasks may create anxiety or disadvantage in AWS by asking them to provide oral responses. Thus, it is valuable to examine how language is processed when AWS do not need to respond to a task orally and potentially stutter. Accordingly, work has been done to evaluate lexical access abilities in AWS using passive listening paradigms. Such research suggests that differences found during linguistic processing are not merely the result of speech production limitations that logically would accompany stuttering. For example, research has suggested that AWS show atypical cortical activation patterns when listening to linguistic stimuli, as measured by Event-Related Potentials (ERP; e.g., Bosshardt & Fransen, 1996; Cuadrado & Weber-Fox, 2003; Sasisekaran & De Nil, 2006; Weber-Fox & Hampton, 2008).

<u>Do lexical access differences in AWS occur in processing of discourse?</u>

There are no reports in the literature of how well AWS process or remember larger linguistic units, such as connected text or discourse. Recently, Lescht, Bernstein Ratner, Chow, and Braun (2015) and Lescht, Chow, Liu, and Bernstein Ratner (2016) sought to investigate brain activity of AWS and AWNS during lexical and syntactic processing tasks. They had AWS and AWNS complete a story retell task, a word fluency task, and the *Working Memory* subtest of the *Behavioral Rating Inventory of Executive Function- Adult Version (BRIEF-A;* Roth, Isquith, & Gioia, 2005*).* While the intent was to see if groups differed in how language tasks were processed, unexpected differences in accuracy on the tasks emerged.

The story retell task involved asking participants to read and memorize the plots of a series of 12 written short stories the day before their visits to the lab. The 12 written stories were adapted from narrative picture cards (Helm-Estabrooks & Nicholas, 2003). Participants were asked to retell the stories using their own words when the story's title appeared on the screen. It was found that the AWS recalled significantly fewer propositions, recalled significantly fewer key words, used significantly fewer utterances, and showed lower vocabulary diversity averages compared to AWNS during story retell.

Participants in the Lescht et al. study also completed a conventional word fluency task to examine brain circuits used in lexical retrieval. In a conventional "word fluency" task, people are asked to name as many words as they can that start with a given letter of the alphabet, or in a given semantic category, such as animals. Word fluency tasks appear to be related to vocabulary knowledge (Ruff et al., 1997). AWS produced significantly fewer words on the word fluency task, which can be viewed as a time-sensitive word retrieval challenge.

In order to evaluate self-reported working memory, participants completed the *Working Memory* subtest of the *BRIEF-A.* The *Working Memory* subtest of the *BRIEF-A* has been shown to be an effective measure of working memory in both typical and impaired populations (e.g., Garlinghouse, Roth, Isquith, Flashman, & Saykin, 2010; Hocking, Reeve, & Porter, 2015). It was found that AWS and AWNS did not differ in self-reported memory on the *Working Memory* subtest of the *BRIEF-A*.

In a post-hoc analysis, word fluency task scores of AWS were found to be significantly and highly correlated with the number of propositions and key words recalled in the story retell task. This can be taken to suggest that AWS might have been less proficient at retrieving vocabulary during the story retell task. Performance on the *Working Memory* subtest of the *BRIEF-A* was also significantly and highly correlated with the number of key words recalled in the stories, suggesting that AWS may have had difficulty recalling the original vocabulary used in the stories. In contrast, for AWNS, neither word fluency skill nor self-reported working memory skills were significantly correlated with any of the behavioral and story measures.

One limitation of the Lescht et al. study is that there was relatively loose control over story presentation and retell. The participants were provided with the stories the night before their brain imaging visits and asked to read and memorize them. The amount of time that each participant spent reading and memorizing the stories was not controlled, and this may have affected participants' performance.

A second limitation is that findings are inconclusive because they were obtained only in the oral mode. It is possible that AWS experience task demand that limits their ability to perform well on a spoken task. Oral formulation may be difficult in AWS for many reasons, such as that they avoid feared words by using alternate vocabulary or that they circumlocute by using more words than necessary to avoid a particular word (e.g., Martens & Engel, 1986; Van Riper, 1963). Also, time spent during disfluent productions may tax working or long-term memory by stretching the response time window, thus making recall more difficult. Additionally, it is also possible that in spoken language tasks, AWS used fewer propositions and simpler language as a way to limit their stuttering during the retell task, a possibility raised by Bosshardt et al. (2002).

An alternative explanation is that, since the word fluency scores of the AWS were significantly lower than AWNS, that the AWS had difficulty with *lexical storage or access,* and were generally unable to "pull" words when given a prompt, such as to generate words that start with a given letter or semantic category, or retell a story that had specific targeted vocabulary and propositions. It may be that the semantic organization or lexical retrieval systems in AWS make it less likely that they will retrieve content words efficiently. Inefficient retrieval of content words would also impact the ability of AWS to recall key words and propositions in a story retell task.

A third possible explanation is that AWS may have differences in memory for discourse. To establish whether poor memory for discourse is the result of memory differences between AWS and AWNS, further investigation is needed using more objective measures. A limitation of the prior study was that a self-report measure of working memory was used instead of a more objective measure.

What is known about memory in AWS?

While early research evaluating memory abilities in PWS was focused on memory for motor sequences (e.g., Ludlow, Siren, & Zikria, 1997; Smits-Bandstra, De Nil, & Saint-Cyr, 2006), work has expanded to evaluate the relationship between memory and language targets in PWS. Due to its strong associations with language proficiency, the vast majority of studies examining memory function in AWS have focused on phonological working memory (Bosshardt, 1993; Byrd, McGill, & Usler, 2015; Byrd, Vallely, Anderson, & Sussman, 2012; Byrd, Sheng, Bernstein Ratner, & Gkalitsiou, 2015; Ludlow, Siren, & Zikria, 1997; McGill, Sussman, and Byrd, 2016; Sasisekaran, 2013; see Bajaj, 2007, for a review). Findings are suggestive of the premise that phonological working memory in AWS is less efficient than in AWNS and that AWS have less proficient subvocal rehearsal systems. More work needs to be done to investigate memory skills in AWS and to determine if memory differences in AWS exist beyond the domain of phonological working memory.

Current Study

In summary, a recent study found large differences in the ability of AWS to recall story content. This study also found differences in word fluency (time-sensitive lexical retrieval) between AWS and AWNS. This latter finding has been reported in a few prior studies. However, its interpretation is unclear, as word fluency results may be impacted by strategies used by AWS to avoid stuttering. Findings of depressed story recall in the same study could also be attributed to memory differences in AWS compared to AWNS. Although many studies have examined phonological working memory in AWS, no prior studies have attempted to identify potential long-term memory differences for larger units, such as text or discourse.

However, the design to test story recall and memory used by the prior study was not robust, suggesting a need to replicate the study using more controlled tasks and procedures. For instance, conditions of story presentation and retell were not well controlled. Additionally, the memory assessment relied on self-report. Thus, it is unclear whether the differences in story memory that were seen can be replicated under more controlled conditions.

Finally, it is impossible to separate spoken language task demands from memory or vocabulary limitations using an oral task with AWS. In order to do this, we would need to offer a way for the AWS to respond without speaking. Alternatives would be to appraise memory using typed responses (which could separate recall of content from speaking demand) or recognition tasks (which more clearly separate memory from lexical retrieval difficulty).

Research Questions and Hypothesis

1. Do AWS differ from AWNS on measures of memory for discourse (such as propositional content or key words)?
    a. Can the phenomenon of depressed propositional and key word recall in AWS be replicated on a more carefully controlled oral recall task?
    b. Is this same difference apparent in the written modality using typed responses?
    c. Does memory for lexical content in AWS differ in recognition memory tasks as well, where no lexical retrieval or linguistic formulation is required? Specifically, might AWS "false alarm" to changes in vocabulary (e.g., paraphrases or inferences) on recognition tasks for stories?
    d. If any differences in recall or recognition are detected, can their possible basis be identified? For instance, do AWS differ from AWNS on digit span measures, often used as a non-linguistic measure of memory?
2. Do AWS differ from AWNS on measures of lexical access?
    a. Can the prior observed phenomenon of depressed oral word fluency performance be replicated?
    b. Is this difference in word fluency performance apparent in the written modality using typed responses?
    c. Is this difficulty also apparent on standardized measures of expressive and receptive vocabulary in the predicted direction?

d. Does lexical access as measured by word fluency or standardized vocabulary test performance relate to measures of story recall skill in adults who do and do not stutter?

Our hypotheses are as follows:

If we believe that depressed story recall scores and word fluency skills result only from spoken task demands that unfairly penalize AWS, then we will expect:

- AWS may produce fewer propositions and key words than AWNS in their oral retell of stories. However, the number of propositions and key words in the typed story retells of AWS and AWNS will not differ.
- The accuracy of AWS and the AWNS on the story recognition task will not differ.
- AWS may produce fewer words on an oral word fluency task compared to AWNS, but AWS will not differ from AWNS in performance on a typed word fluency task.
- AWS and AWNS may differ in expressive vocabulary (*EVT-2)* scores, but will not differ in receptive vocabulary (*PPVT-4)* scores.
- AWS and AWNS may differ in digit span in the oral mode, but will not differ in the typed mode.

If we believe that AWS do have limitations on memory for verbal materials, then we will expect:

- AWS will produce fewer propositions and key words than AWNS in their retell of stories both in the oral and typed mode.
- AWS will have a lower accuracy score on the recognition task.
- AWS may show depressed digit span scores compared to AWNS in both the oral mode and typed mode.

If we believe that both propositional and key word recall differences in AWS and AWNS will emerge due to lexical access difficulties, then we will expect:

- AWS will use fewer original key words than AWNS in their oral retell of stories, but the number of propositions used may not differ.
- AWS will use fewer original key words than AWNS in their typed retell of stories, but the number of propositions used may not differ.
- AWS should not have a lower accuracy score on the recognition task.
- AWS will perform more poorly on both the oral and typed word fluency tasks.
- AWS will perform more poorly on the *EVT-2,* but will not perform poorly on the *PPVT*, which does not require word retrieval.

Method

Participants

Participants were recruited from a variety of sources, including the University of Maryland, College Park campus, local clinics, and the National Stuttering Association (NSA). Fifteen AWS were recruited first, with fifteen AWNS recruited to form a gender-, age- and educationally-matched control group. Matching criteria were age (within one year), gender, handedness, and education (within general categories, such as associate degree, bachelor's degree, and graduate degree (+/- 2 years duration)). All participants were monolingual and native English speakers. By self-report, they had no significant cognitive disorders, communication disorders (other than stuttering for the AWS), psycho-educational concerns, history of learning disability, or history of hearing loss. Participants completed a phone case history form prior to participating in the study where this information was recorded (See Appendix A). All participants received a $25 gift card on completion of their participation.

Stuttering Participants

The AWS group included six females and nine males (age range 19 - 57 years). One AWS was left-handed and the rest were right-handed. Education levels ranged from high school degrees to a doctoral degree. All of the AWS completed the *Stuttering Severity Instrument, 4th edition* (SSI-4; Riley, 2009) to verify group status as AWS, which was confirmed. The speaking sample used an interview discussing personal experiences with stuttering; the standard reading sample was the "Friuli" passage. The *SSI-4* was both audio- and videotaped. Scoring followed the *SSI-4* protocol.

Typically Fluent Participants

The AWNS group also included six females and nine males, ranging in age from 19 to 56 years old. One of the AWNS was also left-handed and the rest were right-handed. Education levels ranged from high school degrees to a doctoral degree. A t-test was computed for age and found that there was no significant difference between groups, $t(28) = -0.0437$, $p = .9655$. A t-test was also found that there were no significant differences between groups for education level in years, $t(28) = 0.0000$, $p = 1.00000$.

Consent

All participants were provided with a consent form, which explained the study aims, requirements, and the storage of personal information. Time was provided for the participants to read over the consent form and ask any questions. After, participants signed the consent form and were given a signed copy to keep.

Background Measures

Vocabulary: The *Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4*; Dunn & Dunn, 2007) and the *Expressive Vocabulary Test, Second Edition (EVT-2*; Williams, 2007) were administered to all participants to assess participants' expressive vocabulary and word recognition skills.

Digit span: The *Memory for Digits* subtest of the *Comprehensive Test of Phonological Processing, Second Edition (CTOPP-2)* was used to assess digit span (Wagner, Torgesen, Rashotte, & Pearson, 2013). The subtest was adapted to include 10 digit strings from its original form and to create a second one that used the same digits, but in a different order. This was done so that the task could be administered in both oral and typed formats. The digit strings ranged from 2 digits to 8 digits.

Experimental Tasks

Stimuli

***Story retell task**:* Six of the twelve stories used in Lescht, Bernstein Ratner, Chow, and Braun (2015) and Lescht, Chow, Liu, and Bernstein Ratner (2016) that were previously adapted from Helm-Estabrooks and Nicholas (2003) were used to assess oral and typed memory. Pairs of stories were balanced in propositional and key word content, and thus assignment to typed or oral retell condition (below) was random within story pairs. Two stories had 39 propositions and 15 key words, two stories had 31 propositions and 15 key words, and two stories had 43 propositions and 17 key words. Therefore, in each mode, the total number of propositions in the stories was 113 and the total number of key words was 47. Distractor trail-making tests were used in between story reading and story retell (Motus Design Group, 2016). Patterns of completion of distractor tasks were not analyzed for the purposes of this thesis.

Story recognition task: For the recognition task, four stories and a recognition task were adapted from Radvansky, Curial, Zwaan, and Copeland's (2001) study of propositional memory. The four narrative texts were 37-50 sentences long each. Following each story, a word jumble task adapted from Brubaker (1987) and Tomlin (2002) was administered. Responses to the word jumble distractor task were not analyzed for this thesis.

The recognition task had 15 questions per story with a total of 60 multiple-choice recognition questions. Each recognition question contained both an original sentence and three foils: a lexical foil, an inference foil, and a mixed foil. Table 1 provides examples of each response type. The **original sentence or verbatim response** was the exact sentence from the passage. The **lexical foil** had the same number of propositions as the original sentence. However, it had two vocabulary changes. The **inference foil** included information not in the original wording that made sense as an interpretation of the passage. The **mixed foil** included a grammatical re-wording of the inference foil and a blend of vocabulary words from the verbatim response, lexical foil, and inference foil. Scores that were derived included total accuracy score (correct verbatim choices), as well as propensity to select lexical foils, inference foils, mixed foils in error, or failure to respond.

Table 1. Recognition Task Response Examples

Response Type	Example
Original Sentence/Verbatim Response	People in Lakewood, like Mary, felt that the great demand for beanie babies would hold forever.
Lexical Foil	People in Lakewood, like Mary, felt that the high demand for beanie babies would last forever.
Inference Foil	To many collectors, beanie babies seemed to have an enduring appeal.
Mixed Foil	Beanie Babies have enduring appeal to most collectors in Lakewood

Procedures

Other than the background measures, all tasks were computer-administered. Responses were audio-recorded. Tasks were presented in a set order, as listed below, in alternating mode (oral/typed).

Word fluency task: This task was adapted from Tombaugh, Kozak, & Rees (1999). Half of the participants were asked to orally generate words starting with the letters “F” and “S” in one minute and then type as many words that start with the letters “A” and “P” in one minute. Letter prompts and mode were reversed for the other half of participants, within each group. A count-down timer was shown as a prompt.

Digit span task: As described under background measures, each participant completed one version of the digit span task orally, and a second condition in typed format, counter-balanced within groups.

Story retell task*:* Each participant was shown six stories, which had paired random assignment to typed/oral retell ordering; half of the participants in each group retold three stories orally, while half of the participants in each group typed these same stories. The order of presentation of the six stories was randomized for each participant. The story appeared on a screen for 2 minutes and participants were instructed to read and memorize the story before automatically moving on. A count-down timer was shown on the screen. When time expired, the story disappeared from the screen and participants completed a trail-making distractor task for 30 seconds, also guided by a timer. Next, participants were asked to retell the story either orally or by typing, and prompted by presentation of the story title. Typed responses were saved into the system by the *Key Time-Stamping for Qualtrics* computer software (Stockbridge, 2017). Participants were given 2 minutes to retell each story, but could move on to the next story if they finished retelling the story in less than 2 minutes.

Story recognition task*:* In this condition, participants were asked to read the story shown on the screen and to try to remember the story because they would be answering questions about the story afterwards. After the story disappeared from the screen, participants were given one minute to complete a word completion distractor task using paper and pencil. For this task, participants were shown the alphabet and a list of words with 1-3 letters missing. They were instructed to fill in the missing letter(s) to complete each word.

Next, participants completed a 15-question multiple-choice recognition task on the screen. Participants were presented with the recognition tasks where each type of response was pre-assigned to a different order (e.g., choice A, B, C or D) to prevent an order of response bias. Participants were given 15 seconds to click on the phrase they thought was originally in the story, accompanied by a count-down timer.

After participants read the first story, completed the word jumble distractor task, and completed the recognition task for the first story, they moved on to the second story, until all four stories had been read and questioned. Participants were given 2 minutes and 30 seconds to read the first three stories and 2 minutes to read the fourth, shorter story.

Scoring and Analysis

***Story retell task**:* Oral stories were transcribed in CLAN (MacWhinney, 2000), and typed stories were converted to transcript using CLAN TextIn. The variables of interest were: propositional count and key word count in both spoken and typed modes. CLAN EVAL computes propositions automatically based on formulas developed by Brown, Snodgrass, Kemper, Herman, and Covington (2008) and Turner and Greene (1977). CLAN proposition count, computed on the %mor tier that performs grammatical tagging, includes verbs, participles, adjectives, adverbs, prepositions, conjunctions, possessive pronouns, interrogatives/relatives, quantifiers, negatives, a copula [only if followed by a noun phrase], determiners, and numbers [only if followed by a noun]. Participants only received credit for using an original story proposition if the word on the %mor tier was identically matched to the one from the original story. Participants' use of key words was calculated by using the FREQ utility to search for exact words appearing in the original story. Inflected forms were not considered matches in this task. For example, if a participant said, "roast," for the original form "roasting", this was not counted as a correct keyword match.

A participant's proposition score was calculated by dividing the proposition count of the retold stories by 113, the stories' total original proposition value. This was done separately for story retells in the oral and typed mode. Key word counts were calculated in the same way by dividing the key word count of the retold stories by 47, the stories' total original key word count.

Results

Background Measures

Vocabulary: The mean *PPVT-4* score for AWS was 107.53 (range = 40; SD = 11.03) and the mean *PPVT-4* score for AWNS was 110.53 (range = 26; SD = 10.16). The mean *EVT-2* score for AWS was 116.73 (range = 34; SD = 8.54) and the mean *EVT-2* score for AWNS was 119.13 (range = 22; SD = 7.32). A two-sample t-test showed no significant difference between groups on either test: *PPVT-4*, $t(28) = .77$, $p = .44$; *EVT-2*, $t(28) = .83$, $p = .42$. Thus, the mean *PPVT-4* and *EVT-2* scores for the AWNS group were non-significantly higher than the AWS. Mean scores are plotted in Figure 1.

Digit span (memory): Digit span scores were computed based on the conventions used by the *CTOPP-2* for the oral and typed modes separately. Participants could get a total score of 10 for each of the oral and typed digit span tasks. A 2x2 repeated measures ANOVA was used to assess accuracy by group (AWS/AWNS) and condition (oral/typed) and any possible interaction and found no significant effect of group, $F(1,28) = 1.63$, $p = .21$, no significant effect of mode, $F(1,28) = 2.44$, $p = .13$, and no significant interaction, $F(1,28) = .07$, $p = .8$. Thus, AWS and AWNS did not differ on digit span, our short-term memory assessment, by group or by mode. Mean scores are plotted in Figure 2.

Word fluency (word retrieval): The total number of words that participants generated in the oral mode and typed mode was calculated separately. A 2x2 repeated measures ANOVA was used to assess accuracy by group (AWS/AWNS) and condition (oral/typed) and any possible interaction ($p = .05$). A significant main effect for group was found, $F(1,28) = 4.50$, $p = .04$; AWNS generated significantly more words than did AWS. No significant main effect was found for mode, $F(1,28) = .08$, $p = .77$ and no significant interaction was found, $F(1,28) = .46$, $p = .5$. Thus, groups did not differ on the word fluency task by mode. Mean scores are plotted in Figure 3.

Figure 1.

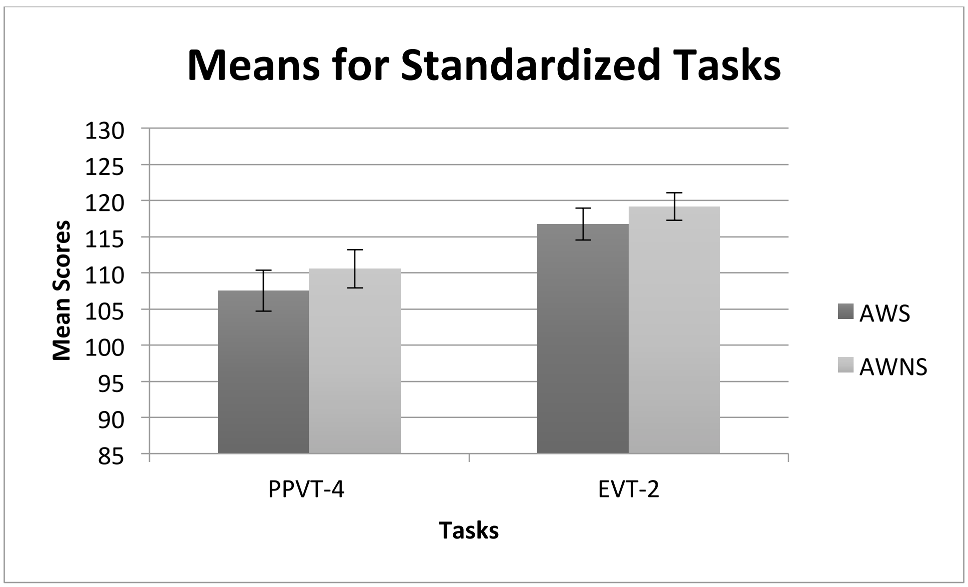


Figure 2.

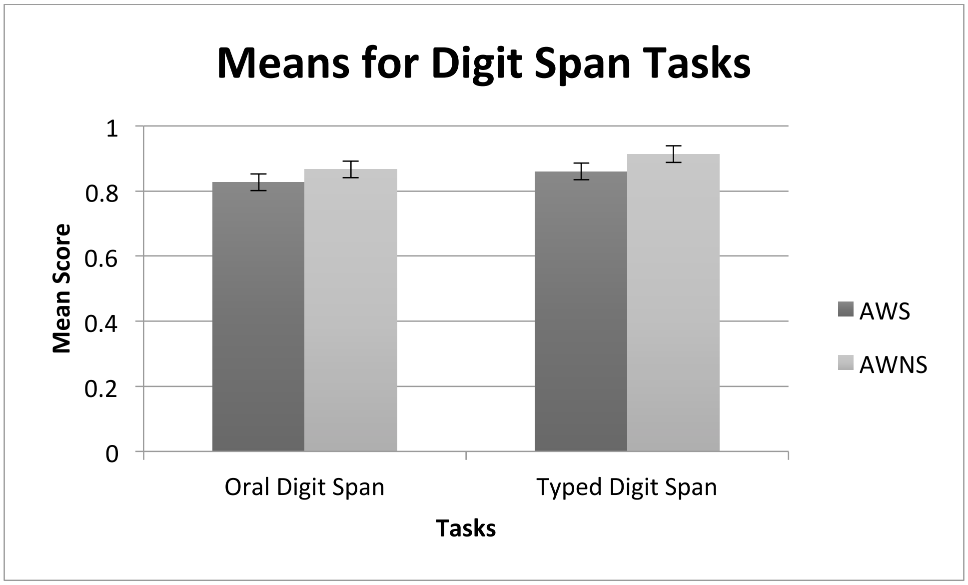
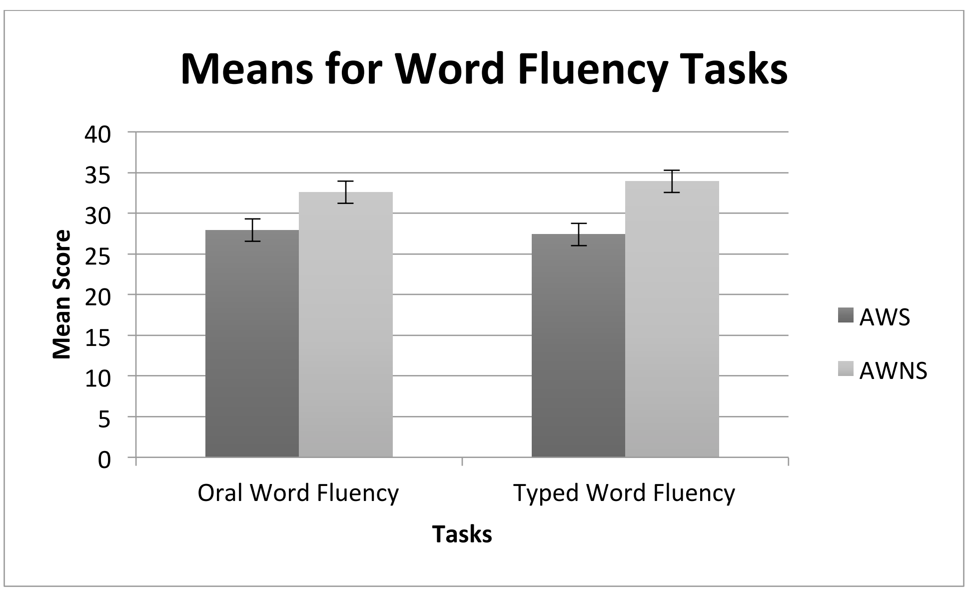


Figure 3.



Experimental Tasks

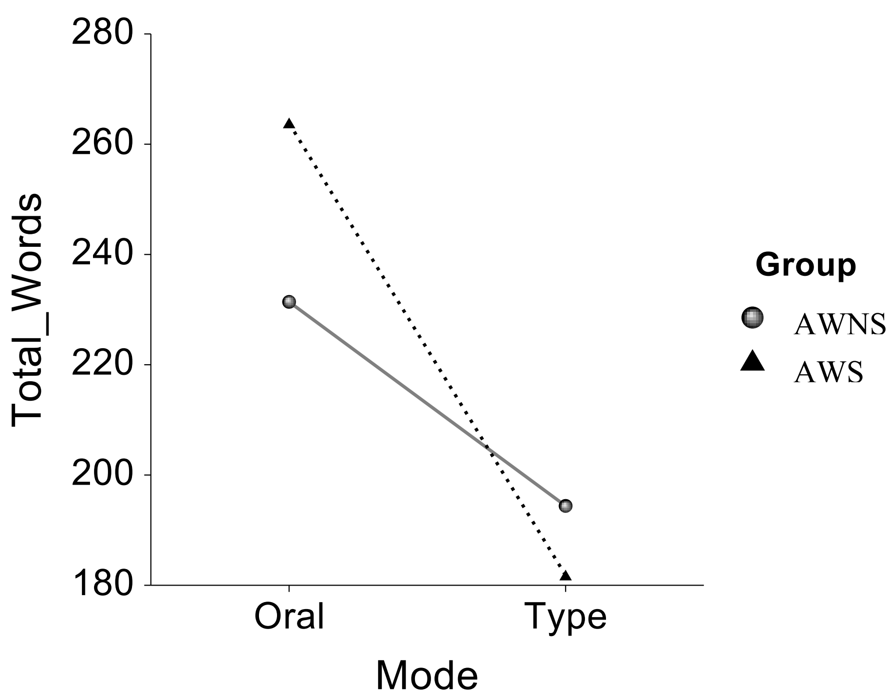
Story retell task:

Preliminary analyses: Before doing any other story analyses, it was important to ensure that the AWS and AWNS produced oral/typed stories of comparable length. A 2x2 repeated measures ANOVA was done to assess story length in words by group (AWS/AWNS) and condition (oral/typed).

No significant main effect was found for group, $F(1,56) = .79$, $p = 0.38$. A significant effect was found for mode, $F(1,28) = 29.25$, $p = .000009$; both AWS and AWNS used more words in their oral stories than in their typed stories. A significant interaction was found between group and mode, $F(1,28) = 4.16$, $p = .05$, as shown in Figure 4. The oral stories of the AWS contained significantly more words than the oral stories of the AWNS.

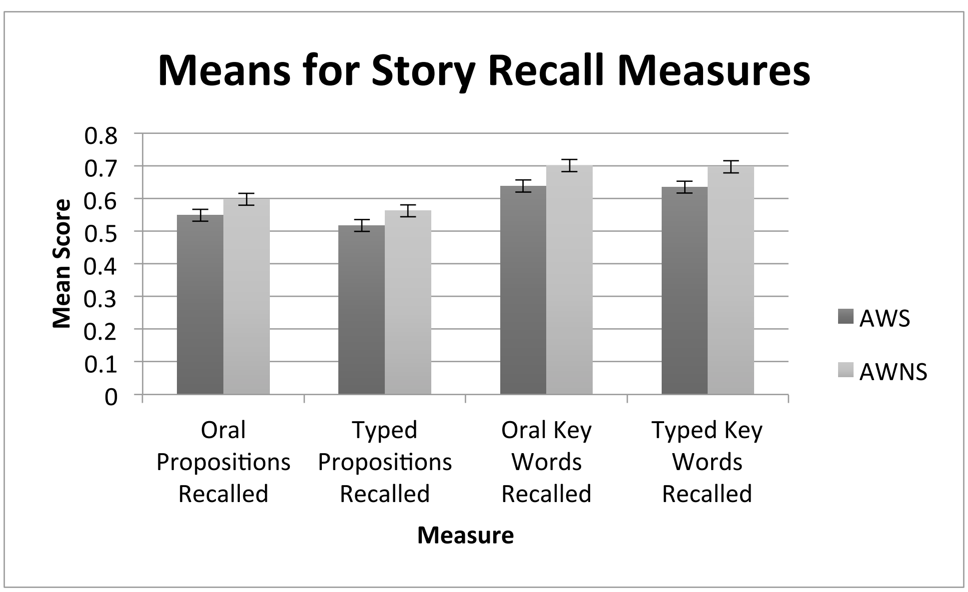
Figure 4.

Means Plot of Total Words by Group and Mode



Story retell analyses: Two 2x2 repeated measures ANOVAs were done to assess accuracy by group (AWS/AWNS) and condition (oral/typed) and any possible interaction for propositions recalled and key words recalled from the stories (**adjusted alpha = .025).** Results are plotted in Figure 5. For propositions recalled, no significant main effect was found for group, $F(1,28) = 1.24$, $p = .27$, no significant main effect was found for mode, $F(1,28) = 3.22$, $p = .08$, and no significant interaction was found, $F(1,28) = .01$, $p = .91$. Thus, groups did not differ in the number of propositions recalled and no differences were found within groups on typing versus speaking. For key words recalled, no significant main effect was found for group, $F(1, 28) = 1.97$, $p = .17$, no significant main effect was found for mode, $F(1,28) = .05$, $p = .83$, and no significant interaction was found between group and mode, $F(1,28) = .00$, $p = .97$. Thus, AWS and AWNS did not differ in accuracy of proposition recall by group or within groups by mode. However, the mean for key word recall by AWNS was non-significantly higher than that seen in AWS.

Figure 5.



Recognition task: Accuracy, number of lexical, inference, and mixed false alarm errors, as well as number of null responses, were calculated. Participants could earn a total accuracy score of 60. A participant's accuracy score was calculated by dividing the number of correct answers by 60 to get a percentage. Post-hoc analysis of false alarm to lexical, inference, and mixed foils and no response to questions was calculated separately to determine profiles of group response type (**adjusted p = .01**). Means are shown in Table 2.

Table 2. Means of Recognition Memory Task Responses

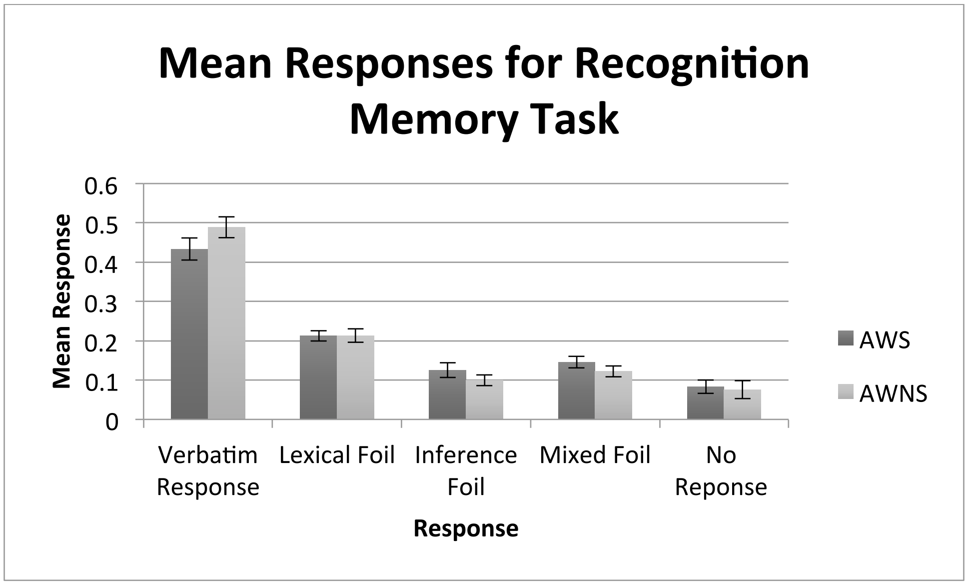
	Verbatim Response	Lexical Foil	Inference Foil	Mixed Foil	No Response
AWS	.43	.21	.13	.15	.08
AWNS	.49	.21	.1	.12	.08

Results are displayed in Figure 6. A two sample t-test showed that there were no differences between groups on overall accuracy, $t(28) = 1.44$, $p = .16$. No significant differences were found between groups for selecting lexical foils, $t(28) = .05$, $p = .96$, inference foils, $t(28) = -1.07$, $p = .29$ or mixed foils, $t(28) = -1.14$, $p = 0.26$. Finally, no significant difference was found for the number of no responses on the recognition task, $t(28) = -.28$, $p = 0.79$.

<u>Relationships Among Measures</u>

Oral story retell: To appraise potential relationships among retell accuracy and vocabulary, memory and word fluency measures, Pearson correlations were computed for the oral mode and included *EVT-2* performance, propositions and

Figure 6.



key words recalled in the oral stories, total digit span performance, and oral word fluency performance **(initial adjusted p = .008)**. Total digit span score was used as opposed to just oral digit span, since as reported earlier, no major differences were found between groups and modes. Correlations were computed for groups together and separately. Results are plotted in Figure 7, Figure 8, Figure 9, Figure 10, Figure 11, and Figure 12.

<u>Analysis combined across groups</u>

For all multiple correlations reported in this and the following sections, we adjusted alpha using the Holm's Sequential Bonferroni Procedure (Abdi, 2010). As a result of this adjustment, the first level of significance for six comparisons is p = .008; the next critical value for p is set at p = .01, the test with the third lowest p-value is p

= .0125, the test with the fourth lowest p-value is p = .0167, the test with the fifth lowest p-value and the test with the sixth lowest p-value is p = .025.

For both groups combined, no significant correlations were found between word fluency and propositions recalled, r*(28*)= .32, p = .08 or between word fluency and key words recalled, r(*28*)= .33, p = .07. Thus, oral word fluency, or oral word retrieval abilities, was not associated with story retell measures for both groups considered as a whole. For the total sample, no significant correlations were found between total digit span score and propositions recalled, r(*28*)= .17, p = .38 or between total digit span score and key words recalled, r(*28*)= .18, p = .35. Thus, memory, as measured by the digit span task, was not found to be associated with oral story retell measures. No significant correlations were found between *EVT-2* performance and propositions recalled, r(*28*)= .30, p = .11 or between *EVT-2* performance and key words recalled, r(*28*)= .36, p = .05. Thus, expressive vocabulary knowledge was not found to be associated with oral story retell measures.

AWS

For AWS, no significant correlations were found between word fluency performance and propositions recalled, r(*14*)= .19, p = .49 or between word fluency and key words recalled, r(*14*)= .25, p =.37. Thus, oral word fluency performance was not associated with any story retell measures. For AWS, no significant correlations were found between total digit span score and propositions recalled, r(*14*)= .09, p = .76 or between total digit span score and key words recalled, r*(14*)= .02, p = .94. Thus, memory, as measured by the digit span task, was not found to be associated with oral story retell measures. For AWS, no significant correlations were found between *EVT-2* performance and propositions recalled, r(*14*)= .0002, p = 1 or between *EVT-2* performance and key words recalled, r(*14*)= .26, p = .35. Thus, expressive vocabulary knowledge was not associated with any oral story measures.

AWNS

For AWNS, no significant correlations were found between word fluency performance and propositions recalled, r(*14*)= .39, p = .15 or between word fluency performance and key words recalled, r(*14*)= .35, p = .19. Thus, expressive word retrieval abilities were not found to be related to story retell measures. For AWNS, no significant correlations were found between total digit span score and propositions recalled, r(*14*)= .15, p = .60 or between total digit span score and key words recalled, r(*14*)= .22, p = .44. Thus, memory as measured by the digit span task was not associated with story retell measures. For AWNS, no significant correlations were found between *EVT-2* performance and propositions recalled, r(*14*)= .53, p = .04 or between *EVT-2* performance and key words recalled, r(*14*)= .42, p = .12. Thus, expressive vocabulary knowledge was not associated with story retell measures.

Typed story retell: Pearson correlations were computed for the typed mode and included *EVT-2* performance, propositions and key words recalled in the typed stories, total digit span performance (oral and typed digit span), and typed word fluency performance **(initial adjusted p = .008).** Total digit span score was used as opposed to typed digit span, since as reported earlier, no major differences were found between groups and modes. Correlations were computed for groups together and separately. Results are plotted in Figure 13, Figure 14, Figure 15, Figure 16, Figure 17, and Figure 18.

Results across groups

Performance on typed stories showed slightly different relationships than those seen for oral stories. For both AWS and AWNS, a significant correlation was found between word fluency and propositions recalled in the typed stories, $r(28) = .52$, $p = .004$, which suggests that typed word retrieval abilities were associated with the number of propositions recalled in the typed stories. A significant correlation was also found between word fluency and key words recalled, $r(28) = .57$, $p = .0009$, suggesting that typed word retrieval abilities were associated with the number of key words recalled in the typed stories. For both AWS and AWNS, no significant correlations were found between total digit span score and propositions recalled, $r(28) = .34$, $p = .07$ or between total digit span score and key words recalled, $r(28) = .38$, $p = .04$, suggesting that memory ability, as measured by the digit span task, was not associated with story retell measures. A significant correlation was found between *EVT-2* performance and propositions recalled, $r(28) = .46$, $p = .01$, suggesting that expressive vocabulary knowledge was associated with the number of propositions recalled in the typed stories. Finally, a significant correlation was found between *EVT-2* performance and key words recalled, $r(28) = 0.43$, $p = .02$, suggesting that *EVT-2* performance was associated with the number of key words recalled in the typed stories.

AWS

For this group, no relationships achieved significance after Holm Bonferroni adjustment. For AWS, no significant correlations were found between word fluency performance and propositions recalled, r(*14*)= .38, p = .17 or between word fluency performance and key words recalled, r(*14*)= .48, p = .07. Thus, typed word retrieval abilities were not found to be associated with story retell measures. For AWS, no significant correlations were found between total digit span score and propositions recalled, r(*14*)= .23, p = .40 or between total digit span score and key words recalled, r(*14*)= .25, p = .36, suggesting that memory, as measured by the digit span task, was not associated with story retell measures. For AWS, no significant correlations were found between *EVT-2* performance and propositions recalled, r(*14*)= .47, p = 0.08 or between *EVT-2* performance and key words recalled, r(*14*)= .33, p = .23. Thus, expressive word knowledge was not found to be associated with story recall measures for AWS participants.

AWNS

Many correlations remained similar when AWNS were considered separately from the total group of participants, but did not reach significance with smaller sample size, even after using the Holm's Bonferroni adjustment. All suggested some positive relationship between basic vocabulary, word retrieval and digit span and story retell quality; these relationships were stronger than those seen in the correlation coefficients obtained for the same relationships in the AWS group.

For AWNS, no significant correlations were found between word fluency performance and propositions recalled, r(*14*)= .64, p = .01 or between word fluency performance and key words recalled, r(*14*)= .61, p = .01, suggesting that typed word retrieval abilities were not associated with story retell measures. No significant correlations were found between total digit span score and propositions recalled, r(*14*)= .40, p = .14 or between total digit span score and key words recalled, r(*14*)= .45, p = .09, suggesting that memory abilities, as measured by the digit span task, were not associated with story retell measures. No significant correlations were found between *EVT-2* performance and propositions recalled, r(*14*)= 0.42, p = 0.12 or between *EVT-2* performance and key words recalled, r(*14*)= .51, p = 0.05. Thus, expressive vocabulary knowledge was not found to be associated with story recall measures. In summary, for AWNS, correlations between basic vocabulary, memory and word fluency measures and story retell were reasonably high, unlike those seen in the AWS group, but did not achieve significance after adjustment for multiple comparisons.

Figure 7.

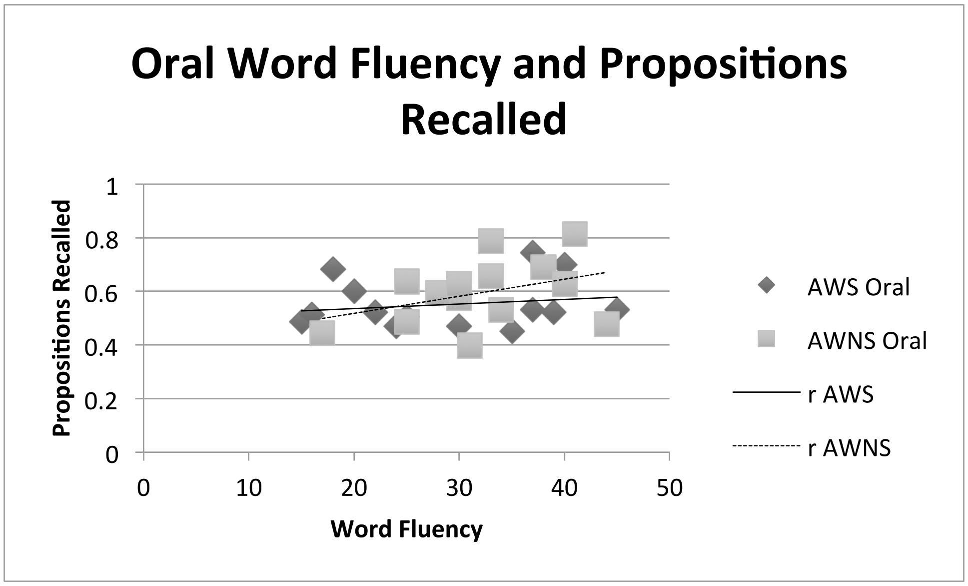


Figure 8.

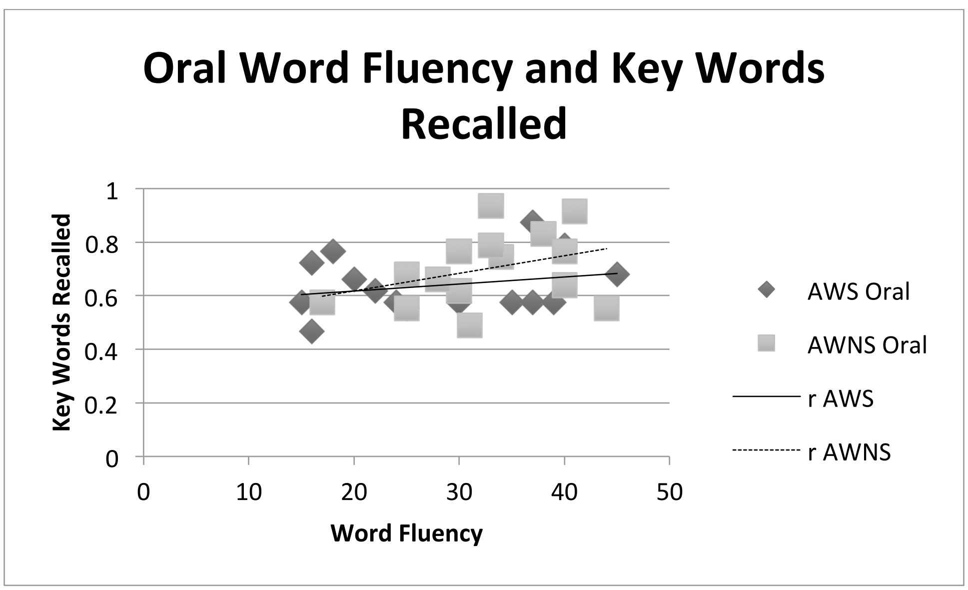


Figure 9.

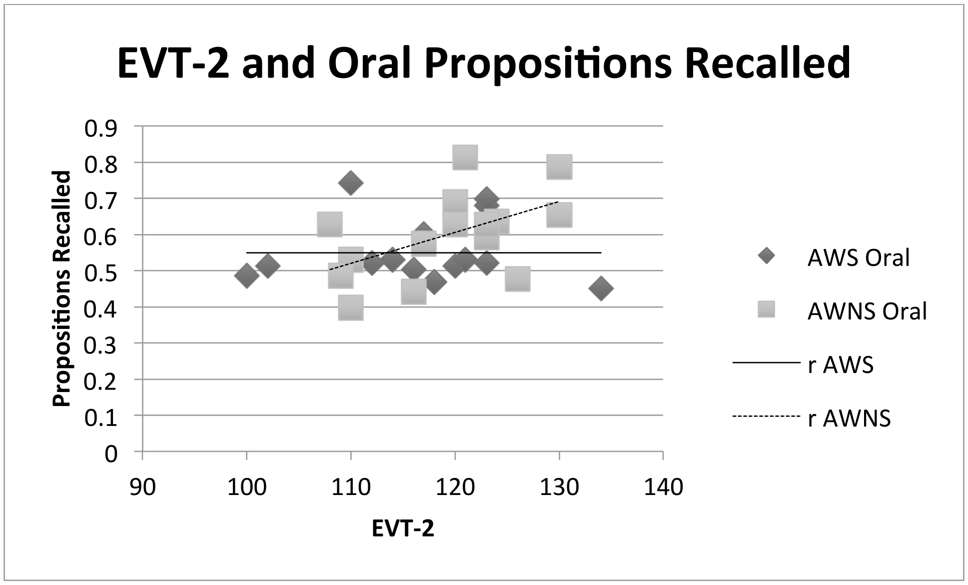


Figure 10.

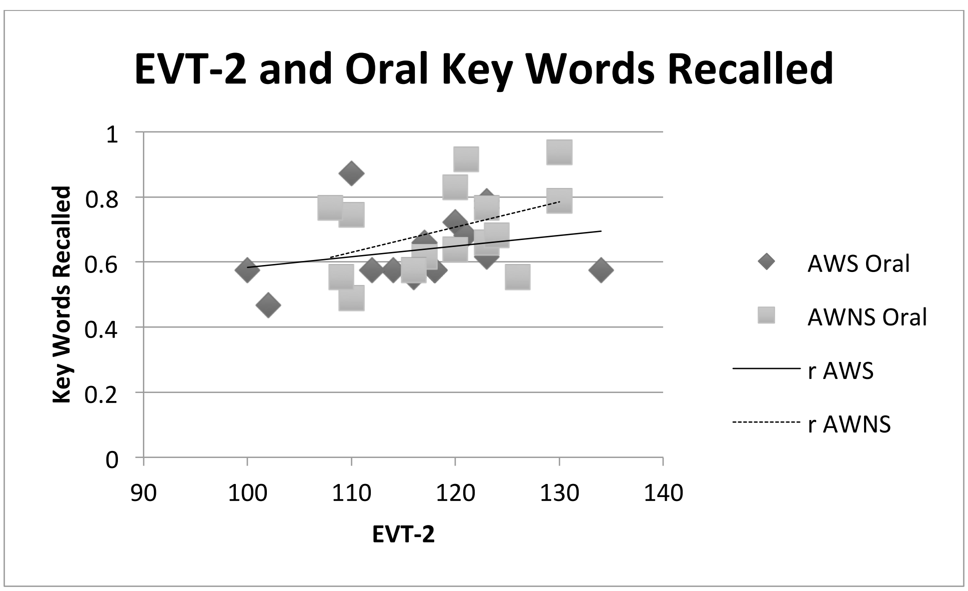


Figure 11.

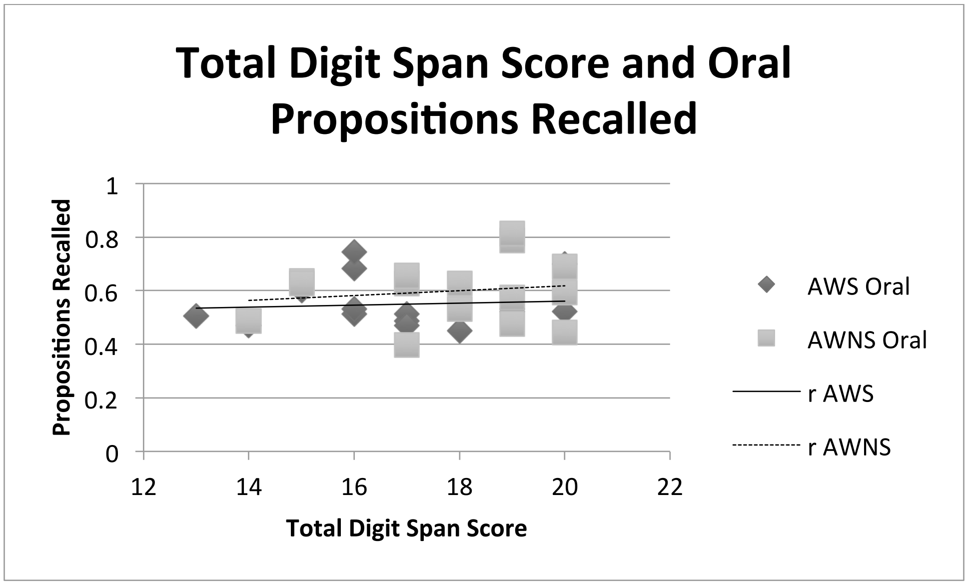


Figure 12.

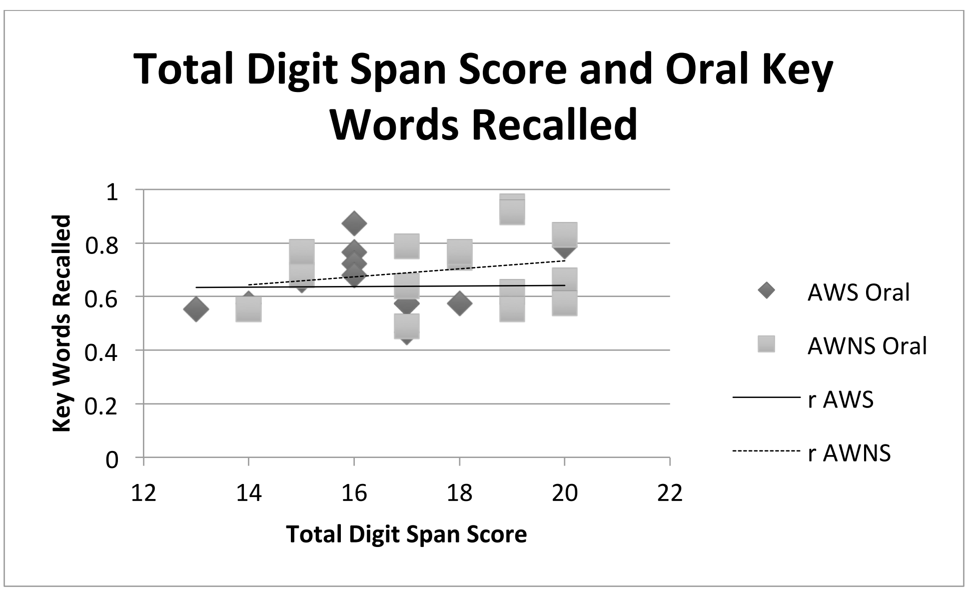


Figure 13.

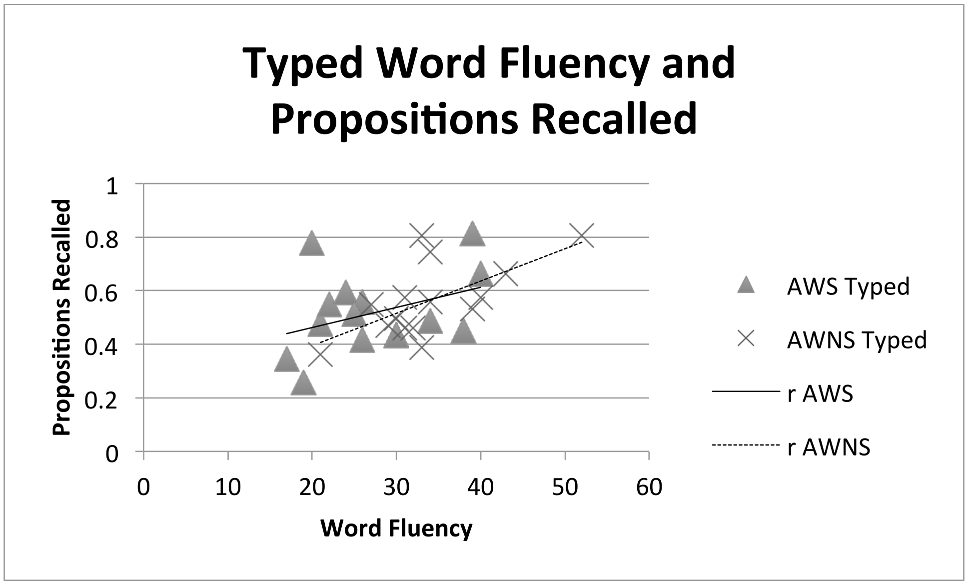


Figure 14

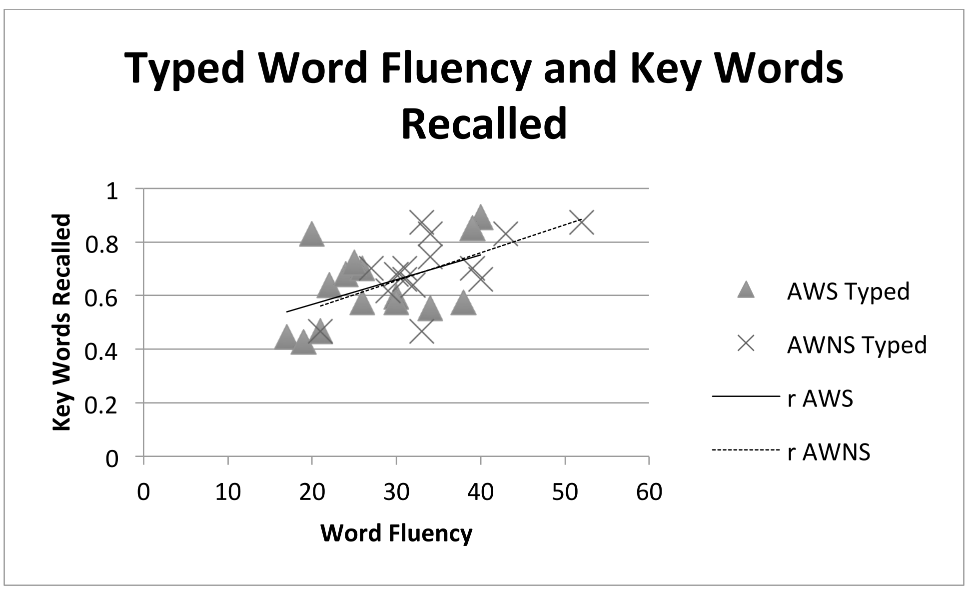


Figure 15

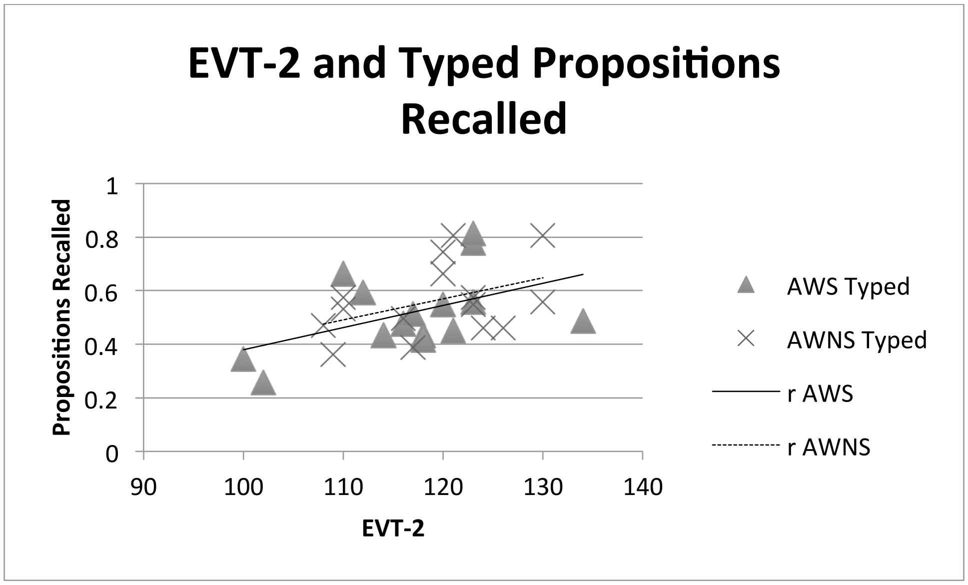


Figure 16

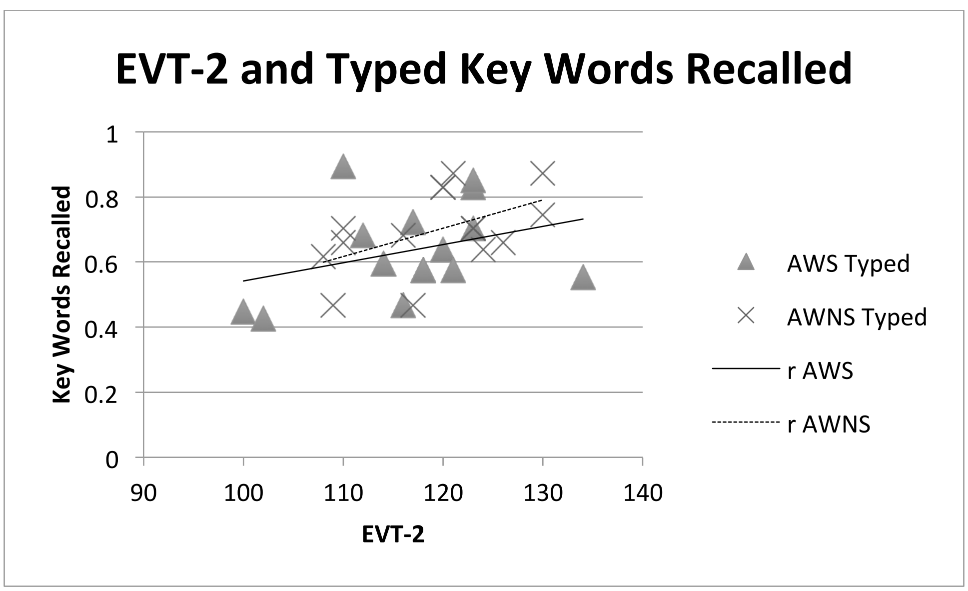


Figure 17

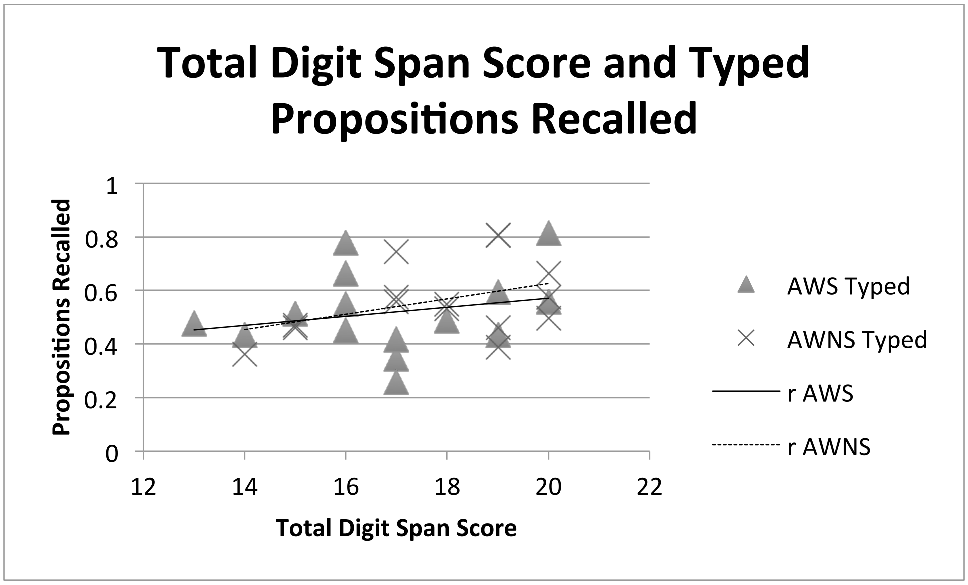
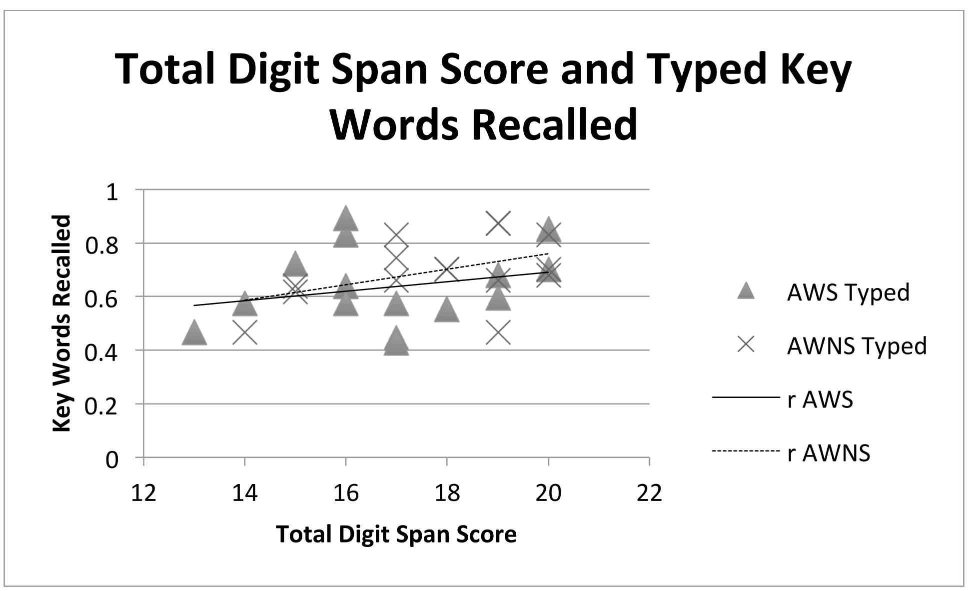


Figure 18



Recognition task: Pearson correlations were computed for *PPVT-4* performance (vocabulary), total digit span score (memory), and story recognition total accuracy score **(adjusted p = .0167)**. Correlations were computed for groups together and separately. Results are plotted in Figure 19 and Figure 20.

<u>Both Groups</u>

For both groups combined, no significant correlations were found between total digit span score and overall accuracy, *r*(28) = .24, p = .19. Thus, memory as measured by the digit span task was not found to be associated with overall accuracy on the recognition memory task. For both groups, no significant correlation was found between *PPVT-4* performance and overall accuracy, r*(28*) = .37, p = .04. Thus, *PPVT-4*, or receptive vocabulary knowledge, was not found to be associated with overall accuracy on the recognition memory task.

<u>AWS</u>

For AWS only, no significant correlation was found between total digit span score and overall accuracy, r(*14*)= .01, p = .97. Thus, memory, as measured by total digit span performance, was not found to be associated with overall accuracy on the recognition task for AWS. For AWS only, no significant correlation was found between *PPVT-4* performance and overall accuracy, r(*14*)= .05, p = .86. Thus, in AWS, vocabulary knowledge was not found to be associated with accuracy on the recognition memory task.

AWNS

For AWNS only, no significant correlation was found between total digit span score and overall accuracy, r(*14*)= .40, p = .14. Thus, memory, as measured by total digit span score was not found to be associated with overall accuracy on the recognition task for AWNS.

However, for AWNS only, a significant correlation was found between *PPVT-4* performance and overall accuracy, r(*14*) = .68, p = .0052, suggesting that richer vocabulary knowledge was associated with likelihood to select an accurate response on the recognition memory task. As in the story retell analyses, AWNS showed stronger relationships between measures of vocabulary skill and memory and their story recognition scores than did AWS.

Figure 19.

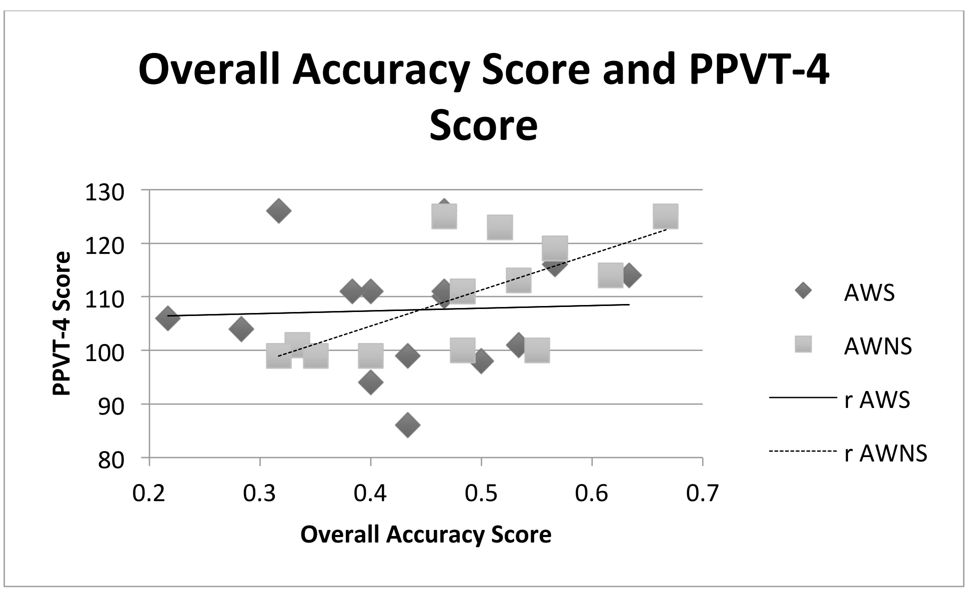
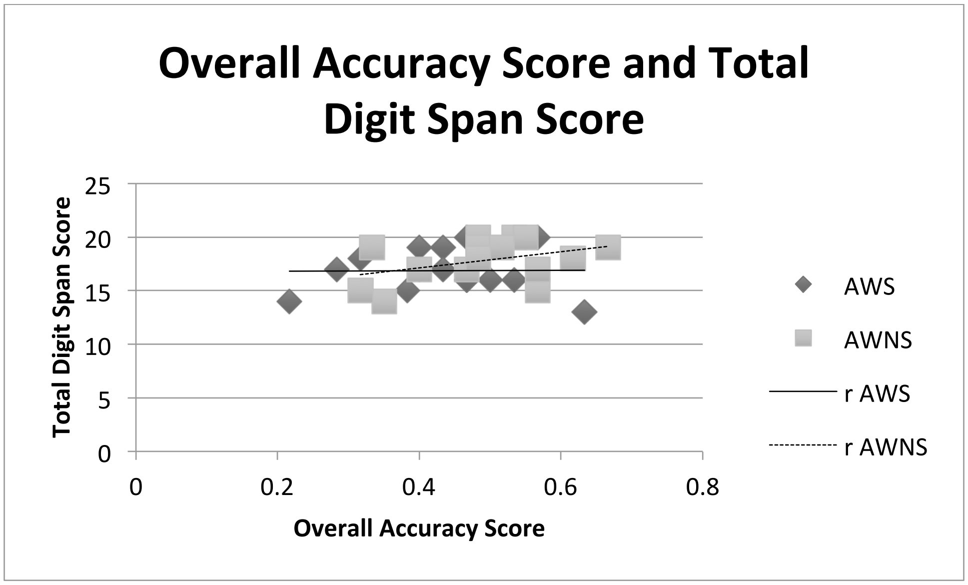


Figure 20.



Discussion

The primary goal of this study was to explore profiles of performance by AWS on story retelling and word fluency tasks compared to AWNS as found in an earlier study by Lescht, Bernstein Ratner, Chow, & Braun (2015) and Lescht, Chow, Liu, and Bernstein Ratner (2016). Our first task was to see if we could replicate the original finding, and if so, determine possible mechanisms. Thus, the current study aimed to find if AWS differed from AWNS on oral story retelling measures (propositional content and key words) for discourse because of spoken demands, memory constraints, or lexical retrieval abilities. A second aim was to appraise whether oral task demands would produce different findings than typed responses. A third aim of this study was to determine if any differences between groups were constrained to the productive/expressive mode. To this end, we asked if AWS differ on measures of recognition memory compared to measures of recall memory.

Did the current study replicate the previous study?

The current study did not replicate findings of the previous study. Lescht, Bernstein Ratner, Chow, & Braun (2015) and Lescht, Chow, Liu, and Bernstein Ratner (2016) found that AWS recalled significantly fewer propositions and key words in their stories compared to AWNS. The previous study also found that the oral word fluency task scores of AWS were significantly and highly correlated with the number of propositions and key words recalled in the story retell task. The current study found that AWS and AWNS did not significantly differ in the number of propositions they recalled in their oral and typed stories. AWS used fewer key words in their oral and typed stories. However, this difference was not significant. No significant correlations were found between word fluency and propositions and key words recalled in the oral or typed stories of AWS in the current study. Additionally, this study added a recognition memory task and found no overall accuracy differences between groups. The only replicated finding from the previous study was that the AWS generated significantly fewer words in the word fluency task compared to the AWNS. Possible reasons why this study did not replicate findings from the previous study will be explored in later sections.

Did AWS have diminished lexical access skills, as measured in this study, that might impact retell and recognition?

Four background measures were used in the study: the *PPVT-4*, *EVT-2*, oral digit span, and oral word fluency. We also added an option of a typed response for digit span and word fluency. These background measures were carefully selected to determine if AWS have weaknesses in skills that might lead to poor story retell or recognition. Logically, if AWS showed weaknesses on these measures, it could explain why they might show difficulties with story recall and recognition. The only significant difference between groups was on the word fluency task. Interestingly, the AWS generated significantly fewer words on both the spoken and typed word fluency tasks. These findings are consistent with older previous results (Okasha, Bisry, Kamel, & Hassan, 1974; Wingate, 1988). These findings are suggestive that AWS have difficulty with some aspect of *lexical storage or access* and the ability to retrieve content words efficiently and quickly.

Were the AWS disadvantaged by oral tasks?

We felt that it was important to add a typed component to the current study. As previously mentioned, oral formulation may be difficult for AWS and they may use different vocabulary or circumlocute by adding in additional words than required to avoid saying a particular word (e.g., Martens & Engel, 1986; Van Riper, 1963). Thus, it was important to understand in this study if the prior findings of depressed story retell and word fluency skills were consequences of the tasks requiring an oral response.

However, results from the current study suggest that oral story telling and word fluency tasks do not penalize AWS compared to AWNS. Not only did AWS and AWNS not differ in the total number of words they used in their stories, but the AWS actually used slightly more words in their oral stories compared to AWNS. These findings are consistent with Wingate (1988) who used a spoken and written story telling task and found that AWS used non-significantly more words in their oral stories than in their written stories. Time was also not a factor with the oral stories in the current study. No participants ran out of time because of delays in production caused by moments of stuttering. Additionally, no differences were found between the propositions and key words recalled in the oral and typed stories of AWS. Thus, oral formulation, alone, does not appear to impact story retelling and word retrieval abilities in AWS. Slight but similar disadvantages emerge when the task is typed as well.

<u>Do some vocabulary, word fluency, and memory skills relate to measures of story retelling and recognition?</u>

No significant correlations were found among oral word fluency, oral story retelling measures, and vocabulary for both groups considered together and separately. However, several significant correlations were found between expressive vocabulary knowledge, typed word fluency, and propositions and key words recalled from the typed stories. Across groups, typed word fluency performance was significantly correlated with number of propositions and number of key words recalled in the typed stories. These correlations remained similar for the AWNS when they were looked at separately; however, some were not significant because of loss of power. Similarly, for both groups, *EVT-2* performance was found to be significantly associated with the number of propositions and key words recalled in the stories. These correlations remained similar for AWNS when they were looked at separately, again moderated by loss of power.

However, unlike the patterns seen in AWNS, correlations were not significant and did not look similar for AWS considered separately. In AWNS, the ability to retrieve words given a prompt or targeted key words and propositions were related. Results also suggest that for AWNS, richer expressive vocabulary knowledge is associated with the ability to recall propositions and key words. These relationships seemed diminished in AWS; the AWS also recalled significantly fewer words on the word fluency task, and the AWS recalled fewer key words. Thus, this may be suggestive of subtle differences in *lexical storage or access* abilities in AWS compared to AWNS. The nature of these differences is unknown, and should be explored in further investigations. These findings are supportive of previous studies that have reported lexical access and retrieval difficulties in AWS (e.g., Newman & Bernstein Ratner, 2007; Crowe & Kroll, 1991; Wingate, 1988), as well as atypical ERP responses during the processing of lexical anomalies (Weber-Fox & Hampton, 2008).

On the recognition memory task, AWS did not differ from AWNS on overall accuracy. However, interestingly, a significant correlation was found between *PPVT-4* performance and overall accuracy on the recognition task in AWNS only. This is suggestive that richer receptive vocabulary knowledge may be associated with the likelihood of selecting an accurate response on the recognition task. The *PPVT-4* is a measure of vocabulary knowledge and not word retrieval. However, greater word knowledge may make it easier to store and retrieve words. Comparable to profiles seen in story retell ability and vocabulary skill, AWS again appear to show different relationships between vocabulary measures and language recognition tasks than are seen in typically fluent peers.

The majority of between-group comparisons between fluent and stuttering participants did not show significant differences. No significant differences were found between groups on the number of propositions and key words recalled and on overall accuracy on the recognition task. Additionally, no significant correlations were found between digit span performance, a measure of memory, and number of propositions and key words recalled in oral and typed stories and between digit span performance and overall accuracy on the recognition task for both AWS and AWNS.

It is important to note, however, that we measured memory for semantic content in our experimental tasks. As noted by Byrd, Sheng, Bernstein Ratner, and Gkalitsiou (2015), the majority of studies analyzing memory in PWS have focused on phonological working memory. Results of most studies suggest phonological working memory differences in AWS compared to AWNS (Bosshardt, 1993; Byrd, McGill, & Usler, 2015; Byrd, Vallely, Anderson, & Sussman, 2012; Byrd, Sheng, Bernstein Ratner, & Gkalitsiou, 2015; Ludlow, Siren, & Zikria, 1997; McGill, Sussman, and Byrd, 2016; Sasisekaran, 2013; see Bajaj, 2007, for a review). Thus, any marginal differences observed in this study may be due to factors other than lexical or semantic representation of the stimulus stories.

Why didn't the current study replicate the previous study?

There are a few major differences between the design of the previous study and the current study. One major difference is that the previous study did not control for time spent reading and memorizing the stories, and created a longer latency between story presentation and demand for recall. The participants were given the stories the night before they came into the lab and asked to read and memorize them. However, in the current study, the participants read each story, completed a distractor task for 30 seconds and then retold each story. It may be that a longer time delay between the reading of each story and retell would result in proposition and key word recall profiles that would more obviously differentiate AWS and AWNS.

An additional difference between the current study and the previous study was the way that the story measures were scored. Even though Lescht, Bernstein Ratner, Chow, and Braun (2015) and Lescht, Chow, Liu, and Bernstein Ratner (2016) found differences in proposition recall, it is important to note that they analyzed how *many* propositions the participants used over the total number of propositions in the original stories. This means that, unlike the current study, the prior study did not look at *exact* proposition recall from the original stories. A problem with this approach is that participants may have used more and/or different propositions in their story retells than in the original stories. This behavior may have impacted results. For example, the word “running” is an original proposition in one of the original stories. If a participant used the proposition “ran” instead of the proposition “running,” a participant did not get credit in the current study. However, in the previous study, a participant would have received credit for using this item. However, since the current scoring was more stringent, one might have expected greater, not fewer differences between groups.

A third difference between the previous study and the current study is that the previous study included twelve short stories. In contrast, the current study only used six of the original twelve stories. In the previous study, participants had more opportunities to retell stories and therefore more opportunities to recall propositions and key words that might have become confused with one another. Thus, it may be that more stories to be held in memory impair talk performance, and the current study did not provide as difficult a memory burden for our participants.

Weaknesses of the current study

Several weaknesses can be detected in this study. Although all participants had the same amount of time to read each story, it is unknown if AWS and AWNS differed in reading speed. If this was the case, one group may have been able to read items from the passages more times than the other group. Another limitation was the scoring of proposition and key word recall. It is possible that either AWS or AWNS participants may have included more propositions and key words in their stories that were not credited because participants only got credit for using exact wording of original propositions and key words. Perhaps more differences may have been found if participants got credit for recalling propositions and key words that share the same lemma, or underlying root form, as the original propositions and key words.

Our lexical retrieval task, word fluency, may have maximized differences between groups that do not inform story recall or recognition. Our word fluency task could potentially be viewed as a phonological task, because it required people to name as many words as they could that began with specific letters/sounds. It is possible that differences between groups on this measure could be attributed to differences in the phonological organization of the mental lexicon in AWS. Interestingly, Newman and Bernstein Ratner (2007) found that lexical features of word frequency, phonological neighborhood density, and phonological neighborhood influenced word naming similarly in both AWS and AWNS. However, AWS had lower word naming accuracy scores. Thus, it is unknown if differences in the current word fluency task are due to differences in phonological organization or retrieval that do not bear on memory for larger units of language, such as text.

Even though no significant differences were found between groups on story recall tasks in this study, it may be that such differences exist in AWS but are subtle and only appear under significant stress or cognitive loading. The current study intended to increase processing stress on participants sufficient enough to magnify any underlying differences. However, it is possible that the story re-tell tasks were not difficult enough to evoke potential differences. Future studies may want to use more difficult stimuli, such as more complex stories in the story retelling task, a longer latency between presentation and response, a more cognitively challenging interim distractor task, or alternative paradigms based on past psycholinguistic work (to be discussed further, below).

Finally, given the small sample size of the current study, we can only report mean performance by group; it may be worthwhile to determine if there are clusters of AWS who are better able at recalling propositions and key words in their stories than other subgroups of AWS. Variability among AWS may potentially impact results when they are measured as mean profiles of performance. The notion that there are potentially subtypes of people who stutter with differing levels of language skill was discussed by Yairi and Ambrose (2013) based on profiles seen in children, and has been observed for adults who stutter by Watson et al. (1994).

Future Research Directions

A major finding from the study was that mode does not seem to determine lexical retrieval and story recall ability in AWS compared to AWNS. AWS generated significantly fewer words on the word fluency task regardless of mode. Also, none of the oral tasks appeared to disadvantage AWS. No significant differences between groups were found on the other measures. However, AWS non-significantly performed more poorly on these tasks compared to AWNS. Thus, there are many future directions that can be taken based on the findings of this study.

The group differences found on the word fluency task suggest that more work needs to be done to investigate lexical access and retrieval abilities in AWS. As mentioned previously, the word fluency task used could be viewed as a phonological task. One option for future work is to use alternative word fluency tasks, such as an action (verb) fluency prompt (Piatt, Fields, Paolo, & Tröster, 2004; Woods et al., 2005) or a semantic prompt. Future work could also further look at the internal organization of the words that participants generated, such as word frequency, phonological neighborhood density, or other features that inform the organization of the mental lexicon. Newman and Bernstein Ratner (2007) found that these factors influenced word naming similarly in both AWS and AWNS. However, it may be worthwhile to determine if these features influence word fluency tasks similarly in AWS and AWNS.

As previously mentioned, studies investigating response time latencies in people who do and do not stutter for a variety of tasks have yielded conflicting results. Future studies could also evaluate response time latencies during word recognition (e.g., *PPVT*) and sentence recognition tasks. Moreover, future work could also use alternative paradigms, such as more challenging lexical batteries, classic word association tasks (e.g., Fitzpatrick, Playfoot, Wray, & Wright, 2015), or evaluate brain activity during lexical recognition, judgment and retrieval tasks. A greater variety of lexical priming tasks could also be used, since a few studies have shown differences in AWS compared to AWNS (e.g., Maxfield, Morris, Frisch, Morphew, & Constantine, 2015; Maxfield Pizon-Moore, Frisch, & Constantine, 2012). Another future direction for this research includes expanding the study to younger participants, which would potentially be more challenging and reveal more differences between groups. Previous research has indicated that lexical access and retrieval differences exist in children who do and do not stutter (e.g, Bauman, Hall, Wagovich, Weber-Fox, & Bernstein Ratner, 2012; Okasha, Bisry, Kamel, & Hassan, 1974; Silverman & Bernstein Ratner, 2002).

Verbal word fluency tasks also appear to tap executive functioning skills. For instance, children with attention-deficit/hyperactivity disorder (ADHD; e.g., Takács, Kóbor, Tárnok, & Csépe, 2014), adults with ADHD (Tucha et al., 2005), and individuals with frontal lobe lesions (e.g., Baldo & Shimamura, 1988) have been found to generate fewer words on word fluency tasks than typical peers. Past research investigating executive functioning abilities in PWS has suggested differences in attention processes (e.g., Anderson & Wagovich, 2010; Heitmann, Asbjørnsen, &

Helland, 2004; Kaganovich, Wray, Weber-Fox, 2010), cognitive flexibility (e.g., Smits-Bandstra, De Nil, & Rochon, 2006; Smits-Bandstra, De Nil, & Saint-Cyr, 2006), and response inhibition (e.g., Anderson & Wagovich, 2017; Eggers, DeNil, & Van den Bergh, 2013). Thus, further investigations may explore whether group differences on verbal word fluency tasks are attributable to differences in executive functioning skills and use tasks that directly target executive functioning.

On the *PPVT-4* and *EVT-2*, the AWNS scored non-significantly higher than the AWS. However, the mean scores were not that different. It may be that these standardized measures are not sensitive enough or the most appropriate measures to determine differences between AWS and AWNS. These standardized measures are often used to determine a diagnosis of a language disorder (Bernstein Ratner, 1997; Bloodstein & Bernstein Ratner, 2008). Past work has indicated subtle linguistic differences and weaknesses in AWS. Future work may want to use more complex standardized tests, such as college or graduate achievement, aptitude tests, or the *Test of Adolescent/Adult Word Finding- Second Edition (TAWF-2;* German, 2016). Future studies could also evaluate performance on these standardized vocabulary assessments by looking at measures beyond just overall accuracy. For instance, future studies could evaluate latency response time, eye gaze to stimuli through eye-tracking software, the use of semantic priming for responses, or neurological brain activity.

The story recall measures found that AWNS non-significantly recalled more propositions and key words than the AWS. The stories used in this task were very short, simple, and lacked true story grammar (e.g., Stein & Glenn, 1979). In the future, it may be worthwhile to replicate more classic psycholinguistic studies of memory for discourse that have been conducted using typical adults. The use of these studies could better help differentiate memory for discourse in AWS compared to AWNS. Is memory for discourse impacted in the same ways in people who do and do not stutter? For instance, Thorndyke (1977) used stories with clear story grammar structures and unclear story grammar structures and found that recall was much better for the stories with clear story grammar structures. Bransford and Johnson (1972) found that participants better recalled passages when they were given context before presentation of a story as opposed to being given no context or the context of the story after reading the story. We do not currently know if AWS are helped equally by such factors that improve memory for story detail.

Additionally, many studies have evaluated what people remember from text. Our recognition task attempted to replicate aspects of Jacqueline Sach's (1967; 1974) studies. Sachs had participants read short stories and showed them semantic and non-semantic foils individually. Sachs found that people remember the main idea and the gist of sentences from discourse. However, people discard the exact word order and syntactic structure of sentences immediately. The recognition task used in the current study showed that AWS performed slightly worse than AWNS. However, differences between groups were not significant and the means for overall accuracy for both groups were close to chance. Thus, it seems that the recognition memory task used in this study may have been too difficult. Future studies could make this task easier by reducing the amount of time for the distractor task, using a different distractor task, using easier or shorter stories, or changing the foils so that they are less alike. Future studies could also change the recognition task paradigm to be more similar to Sachs by presenting the verbatim responses and foils individually.

Further, many classic psycholinguistic studies have investigated the construction-integration model as a model of memory for discourse (e.g., Fletcher & Chrysler, 1990; Kintsch, 1979, 1988, 1998; Kintsch & van Dijk, 1978; Kintsch, Welsch, Schmalhofer, & Zimny, 1990). These studies have used story recognition tasks and have found that people remember the situational model best, which is the constant updating of the overall meaning of the text by integrating new propositions and old propositions, and making inferences. People remember the propositional information, which is the overall meaning, less well and the original syntactic form of a text passage is barely remembered. Thus, future studies could use paradigms that replicate past construction-integration model work to better understand how AWS remember the content as well as the form of stories.

It could be argued that it would be important to control for typing ability by using a baseline typing task. It did not seem to be an issue in the current study, since typed responses did not appear to be disadvantaged for any group of participants. Previous research has suggested that AWS have generalized motor slowing (see Namasivayam & van Lieshout, 2011; Bloodstein & Bernstein Ratner, 2008, for reviews). For instance, on syllable-sequence reading and finger-tap sequence tasks, the reactions times of AWS did not improve over practice as much as seen in AWNS (Smits-Bandstra, De Nil, & Rochon, 2006; Smits-Bandstra, De Nil, & Saint-Cyr, 2006). AWNS also had faster reaction times when doing these tasks individually, as opposed to in dual task conditions. However, the AWS did not show this difference.

Thus, future studies may want to include a baseline typing task, or include other measures of motor skill to control for the possibility that AWS have generally slower performance across both the oral and manual modes.

Conclusions

In this study, spoken task demands and memory differences do not appear to differentiate AWS from AWNS on story retelling tasks and on a recognition memory task, where both groups performed similarly. However, as in many other studies, slight decrements in performance were seen in the AWS, across modes (typed/oral) and across demands (recall vs. recognition). The need to provide an oral response did not appear to disadvantage the AWS, and similar decrements in performance were seen in the typed and recognition tasks. The major difference observed between the matched samples of adults who do and do not stutter was in word fluency (lexical retrieval) to a phonological prompt, whether the response was oral or typed. AWS performed more poorly, regardless of mode.

Since no differences were found across mode, findings suggest that future studies should compare AWS and AWNS on tasks that do not require oral production. There is a large amount of well-researched psycholinguistic research that has investigated lexical access and memory for discourse in typical adults, such as word association tasks and recognition memory reading tasks. It may be valuable to further look at language processing in PWS using these well-researched paradigms.

Appendix A. Case History Screening Form Over the Phone

1. Name _______________________________

2. Age _______________________________

3. Gender _______________________________

4. Native Language _______________________________

5. Any other languages spoken _______________________________

6. Highest level of education _______________________________

7. Current employment _______________________________

8. Person who stutters? _______________________________

9. Do you have any

   a. Speech problems other than stuttering _______________________________

   b. Language delay or disorder _______________________________

   c. Learning disability _______________________________

   d. Other major medical or psycho-educational concerns

      ____

10. What is your preferred handedness?

11. Do you have any hearing loss?

References

Abdi, H. (2010). Holm's sequential Bonferroni procedure. In N.J. Salkind, D.M., Dougherty, & B. Frey (Eds.): *Encyclopedia of research design* (pp. 573-577). Thousand Oaks, CA: Sage.

Anderson, J. D., & Wagovich, S. A. (2010). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of Fluency Disorders*, *35*, 216–234.

Anderson, J. D., & Wagovich, S. A. (2017). Explicit and implicit verbal response inhibition in preschool-age children who stutter. *Journal of Speech, Language & Hearing Research*, *60,* 836-852.

Bajaj, A. (2007). Working memory involvement in stuttering: Exploring the evidence and research implications. *Journal of Fluency Disorders*, *32*(3), 218-238.

Baldo, J. V., & Shimamura, A. P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology*, *12*(2), 259-267.

Bauman, J., Hall, N., Wagovich, S., Weber-Fox, C., & Ratner, N. (2012). Past tense marking in the spontaneous speech of preschool children who do and do not stutter. *Journal of Fluency Disorders, 37*(4), 314-324

Bernstein Ratner, N. (1997). Stuttering:A psycholin- guistic perspective. In R. Curlee & G. Siegel (Eds.), *Nature and treatment of stuttering: New directions* (2nd ed.) (pp. 99 – 127) Needham, MA: Allyn & Bacon.

Bloodstein O., & Bernstein Ratner N. (2008). *A handbook on stuttering* (6th ed.). Clifton Park, NY: Cengage Learning.

Bosshardt, H. (1993). Differences between stutterers' and nonstutterers' short term recall and recognition performance. *Journal of Speech & Hearing Research*, *36*(2), 286.

Bosshardt, H., Ballmer, W., & Nil, L. d. (2002). Effects of category and rhyme decisions on sentence production. *Journal of Speech, Language & Hearing Research*, *45*(5), 844-857.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior*, *11*(6), 717-726.

Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., & Covington, M. A. (2008). Automatic measurement of propositional idea density from part-of-speech tagging. *Behavior Research Methods, 40,* 540–545.

Brown, S. F. (1937).The influence of grammatical function on the incidence of stuttering. *Journal of Speech Disorders, 2,* 207–15.

Brown, S. F., & Moren, A. (1942).The frequency of stuttering in relation to word length during oral reading. *Journal of Speech Disorders, 7,* 153 – 59.

Brubaker, S. (1987). *Workbook for cognitive skills: Exercises for thought processing and word retrieval*. Detroit: Wayne State University.

Byrd, C. T., McGill, M., & Usler, E. (2015). Nonword repetition and phoneme elision in adults who do and do not stutter: Vocal versus nonvocal performance differences. *Journal of Fluency Disorders*, *44,* 17-31.

Byrd, C. T., Sheng, L., Bernstein Ratner, N., & Gkalitsiou, Z. (2015) Veridical and false recall in adults who stutter. *Journal of Speech, Language and Hearing Research. 58*(1), 28-42.

Byrd, C. T., Vallely, M., Anderson, J. D., & Sussman, H. (2012). Nonword repetition and phoneme elision in adults who do and do not stutter. *Journal of Fluency Disorders, 37(*3), 188-201.

Crowe, K. M., & Kroll, R. M. (1991). Response latency and response class for stutterers and nonstutterers as measured by a word-association task. *Journal of Fluency Disorders, 16,* 35 – 54.

Cuadrado, E., & Weber-Fox, C. (2003). Atypical syntactic processing in individuals who stutter: Evidence from event-related brain potentials and behavioral measures. *Journal of Speech, Language & Hearing Research*, *46*(4), 960-976.

Danzger, M., & Halpern, H. (1973). Relation of stuttering to word abstraction, part of speech, word length, and word frequency. *Perceptual and Motor Skills, 37*, 959–963.

Dunn, L., & Dunn, D. (2007) Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4). Minneapolis, MN: NCS Pearson Psychcorp.

Eggers, K., De Nil, L. F., & Van den Bergh, B. R. (2013). Inhibitory control in childhood stuttering. *Journal of Fluency Disorders, 38*, 1–13.

Eisenson, J., & Horowitz, E. (1945). The influence of propositionality on stuttering. *Journal of Speech Disorders, 10,* 193–97.

Fitzpatrick, T., Playfoot, D., Wray, A., & Wright, M. (2015). Establishing the reliability of word association data for investigating individual and group differences. *Applied Linguistics, 36*(1), 23-50.

Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situation models: recognition memory for three types of textual information. *Discourse Processes*, *13*(2), 175-90.

German, D. J. (2016). *Test of Adolescent/Adult Word Finding, Second Edition (TAWF-2).* Austin, Texas: Pro-ed.

Griggs, S., & Still, A.W. (1979). An analysis of individual differences in words stuttered. *Journal of Speech and Hearing Research, 22,* 572–80.

Hahn, E. F. (1942). A study of the relationship between stuttering occurrence and grammatical factors in oral reading. *Journal of Speech Disorders, 7,* 329 – 35.

Hand, C. R., & Haynes, W. (1983). Linguistic processing and reaction time differences in stutterers and nonstutterers. *Journal of Speech and Hearing Research, 26*, 181–185.

Heitmann, R. R., Asbjørnsen, A., & Helland, T. (2004). Attentional functions in speech fluency disorders. *Logopedics Phoniatrics Vocology, 29,* 119 – 27.

Helm-Estabrooks, N., & Nicholas, M. (2003). *Narrative story cards.* Austin, Tex.: PRO-ED, Inc..

Hocking, D. R., Reeve, J., & Porter, M. A. (2015). Characterising the Profile of Everyday Executive Functioning and Relation to IQ in Adults with Williams Syndrome: Is the BRIEF Adult Version a Valid Rating Scale?. *Plos ONE*, *10*(9), 1-18.

Hubbard, C. P., & Prins, D. (1994). Word familiarity, syllabic stress pattern, and stuttering. *Journal of Speech and Hearing Research, 37*, 564–571.

Kaganovich, N., Wray, A. H., & Weber-Fox, C. (2010). Non- linguistic auditory processing and working memory update in pre-school children who stutter: An electrophysiological study. *Developmental Neuropsychology, 35*, 712–736.

Kintsch, W. (1979). On modeling comprehension. *Educational Psychologist*, *14*3-14.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, England: Cambridge University Press.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363-394.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory And Language*, *29*(2), 133-159.

Lanyon, R. I., & Duprez, D. A. (1970). Nonfluency, information, and word length. *Journal of Abnormal Psychology, 76,* 93–97.

Lescht, E.,Bernstein Ratner, N., Chow, H. M., & Braun, A. (2015, November). Long-term memory for language in adults who stutter. American Speech-Language-Hearing Association Annual Convention, Denver.

Lescht, E., Chow, H. M., Liu, S., & Bernstein Ratner, N. (2016, November). Lexical Retrieval & Long-term memory for language in adults who stutter. American Speech-Language-Hearing Association Annual Convention, Philadelphia.

Ludlow, C., Siren, K., & Zikria, M. (1997). Speech production learning in adults with chronic developmental stuttering. In W. Hulstijn, H. F. M. Peters, & P. van Lieshout (Eds.), *Speech production: Motor control, brain research and fluency disorders* (pp. 221–230). NY: Elsevier.

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Martens, C. F., & Engel, D. C. (1986). Measurement of the sound-based word avoidance of persons who stutter. *Journal of Fluency Disorders, 11,* 241 – 50.

Maxfield, N. D., Morris, K., Frisch, S., Morphew, K., & Constantine, J. L. (2015). Real-time processing in picture naming in adults who stutter: ERP evidence. *Clinical Neurophysiology, 126*(2), 284-296.

Maxfield, N. D., Pizon-Moore, A. A. Frisch, S. A., & Constantine, J. L. (2012) Exploring semantic and phonological picture-word priming in adults who stutter using event-related potentials. *Clinical Neurophysiology, 123*(6), 1131-1146.

McGill, M., Sussman, H., & Byrd, C. T. (2016). From grapheme to phonological output: performance of adults who stutter on a word jumble task. *Plos ONE*, *11*(3), 1-19.

Motus Design Group. (2016). INPL Trail Making Test (Version 1.2) [Mobile application software]. Retrieved from http://itunes.apple.com

Namasivayam, A. K., & van Lieshout, P. (2011). Speech motor and skill learning. *Journal of Motor Behavior, 43*(6), 477-489.

Newman, R. & Bernstein Ratner, N. (2007). The role of selected lexical factors on confrontation naming accuracy, speed and fluency in adults who do and do not stutter. *Journal of Speech, Language and Hearing Research, 50*, 196-213.

Okasha, A., Bishry, Z., Kamel, M., & Hassan, A. H. (1974). Psychosocial study of stammering in Egyptian children. *British Journal of Psychiatry, 124,* 531 – 33.

Palen, C., & Peterson, J. M. (1982). Word frequency and children's stuttering: The relationship to sentence structure. *Journal of Fluency Disorders, 7*, 55–62.

Piatt, A. L., Fields, J. A., Paolo, A. M., & Tröster, A. I. (2004). Action verbal fluency normative data for the elderly. *Brain and Language, 89*(3), 580-583.

Prins, D., Main, V., & Wampler, S. (1997). Lexicalization in adults who stutter. Journal of Speech, Language, and Hearing Research, 40, 373–384.

Quarrington, B., Conway, J., & Siegel, N. (1962). An experimental study of some properties of stuttered words. *Journal of Speech and Hearing Research, 5,* 387 – 94.

Radvansky G.A., Curiel J.M., Zwaan R.A., & Copeland D.E. (2001). Situation models and aging. *Psychology and Aging*, *16*, 145-60.

Rastatter, M., & Dell, C. (1987). Vocal reaction times of stuttering subjects to tachistoscopically presented concrete and abstract words: A closer look at cerebral dominance and language processing. *Journal of Speech and Hearing Research, 30*, 306–310.

Riley, G. (2009). *Stuttering Severity Instrument for Children and Adults* (4th ed.). Austin, TX: Pro-Ed.

Ronson, I. (1976). Word frequency and stuttering: The relationship to sentence structure. *Speech and Hearing Research, 19*, 813–819.

Roth, R. M., Isquith, P. K., & Gioia, G. A. (2005). Behavior Rating Inventory of Executive Function—Adult Version: Professional manual. Lutz, FL: Psychological Assessment Resources.

Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1997). The psychological construct of word fluency. *Brain and language*, *57*(3), 394-405.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, *2*(9), 437-442.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition*, *2*(1-A), 95-100.

Sasisekaran, J. (2013). Nonword repetition and nonword reading abilities in adults who do and do not stutter. *Journal of Fluency Disorders*, *38*(3), 275-289.

Sasisekaran, J., & De Nil, L. F. (2006). Phoneme monitoring in silent naming and perception in adults who stutter. *Journal of Fluency Disorders, 31,* 284-302.

Schlesinger, I. M., Melkman, R., & Levy, R. (1966). Word length and frequency as determinants of stuttering. *Psychonomic Science, 6,* 255 – 56.

Silverman, F. H. (1972). Disfluency and word length. *Journal of Speech and Hearing Research, 15,* 788–91.

Silverman, S., & Bernstein Ratner, N. (2002). Measuring lexical diversity in children who stutter: Application of VOCD. *Journal of Fluency Disorders*, *27*, 289–305.

Smits-Bandstra, S., De Nil, L., & Rochon, E. (2006). The transition to increased automaticity during finger sequence learning in adult males who stutter. *Journal of Fluency Disorders, 31,* 22 – 42.

Smits-Bandstra, S., De Nil, L., & Saint-Cyr, J. A. (2006). Speech and nonspeech sequence skill learning in adults who stutter. *Journal of Fluency Disorders*, *31*(2), 116-136.

Soderberg, G. (1966). The relations of stuttering to word length and word frequency. *Journal of Speech and Hearing Research, 9*, 584–589.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 53-120). Norwood, NJ: Ablex.

Stockbridge, M.D. (2017). Key time-stamping for Qualtrics [Computer software]. Available from the author.

Takács, Á., Kóbor, A., Tárnok, Z., and Csépe, V. (2014). Verbal fluency in children with ADHD: strategy using and temporal properties. *Child Neuropsychology*. *20*, 415–429.

Taylor, I. K. (1966).The properties of stuttered words. *Journal of Verbal Learning and Verbal Behavior, 5,* 112 – 18.

Taylor, W. L., Lore, J. I., & Waldman, I. N. (1970). Latencies of semantic aphasics, stutterers and normal controls to cloze items requiring unique and non-unique oral responses. *Proceedings of the Annual Convention of the American Psychological Association, 78,* 75–76.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9*, 77-110.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives Of Clinical Neuropsychology*, *14*(2), 167-177.

Tomlin, K. (2002). *WALC 2: Workbook of activities for language and cognition: Cognitive rehab*. Austin, Texas: Pro-Ed.

Tucha, O., Mecklinger, L., Laufkötter, R., Kauzinger, I., Paul, G. M., Klein, H. E. & Lange, K. W. (2005). Clustering and switching on verbal and figural fluency functions in adults with attention deficit hyperactivity disorder. *Cognitive Neuropsychiatry, 10*(3), 231-248.

Turner, A., & Greene, E. (1977, April). *The construction and use of a propositional text base* (Technical Report No. 63). Boulder, CO: University of Colorado, Institute for the Study of Intellectual Behavior.

Van Riper, C. (1963). *Speech correction: Principles and meth- ods, 4th ed.* Englewood Cliffs, NJ: Prentice-Hall.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing–Second edition*. Austin, TX: PRO-ED.

Watson, B., Freeman, F., Devous, M., Chapman, S., Finitzo, T., & Pool, K. (1994). Linguistic performance and regional cerebral blood flow in persons who stutter. *Journal of Speech, Language and Hearing Research, 37*(6), 1221-1221.

Weber-Fox, C., & Hampton, A. (2008). Stuttering and natural speech processing of semantic and syntactic constraints on verbs. *Journal of Speech, Language & Hearing Research*, *51*(5), 1058-1071.

Weber-Fox, C., Spencer, R., Spruill, J. I., & Smith, A. (2004). Phonologic processing in adults who stutter: electrophysiological and behavioral evidence. *Journal of Speech, Language & Hearing Research*, *47*(6), 1244-1258.

Williams, K. T. (2007). *Expressive vocabulary test, Second Edition (EVT-2)*. Minneapolis, MN: NCS Pearson.

Wingate, M. E. (1967). Stuttering and word length. *Journal of Speech and Hearing Research, 10,* 146–52.

Wingate, M. E. (1988). *The structure of stuttering: A psycholinguistic analysis.* New

York: Springer-Verlag.

Woods, S. P., Scott, J. C., Sires, D. A., Grant, I., Heaton, R. K., & Tröster, A. I. (2005). Action (verb) fluency: test-retest reliability, normative standards, and construct validity. *Journal of the International Neuropsychological Society, 11*(4), 408-415.

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders, 38*(2), 66-87.