MEMORY AND LANGUAGE: EVIDENCE OF RELATIONSHIPS FROM THREE STUDIES OF SCHOOL-AGE CHILDREN

by

Michele A. Anderson

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Interdisciplinary Health Sciences
Advisor: Nickola Nelson, Ph.D.

Western Michigan University
Kalamazoo, Michigan
June 2010
NOTE TO USERS

This reproduction is the best copy available.

UMI
The purpose of this three-paper dissertation was to explore the relationship between performance on verbal memory and language assessment tasks in school-age children. Study one used data from whole-class administration of three language assessment tasks to 2nd, 4th, 6th, and 9th graders (N = 187). Studies two and three used data from individualized administration of language assessment and verbal memory tasks to comparable samples of children between the ages of 6 and 18 years with typical language (TL) and language impairment (LI) (n = 37 per group).

The first study addressed practical and theoretical questions regarding memory and language demands of language tasks assessing nonword spelling, listening comprehension, and procedural direction following. Correlational analysis of the three tasks showed that they were only moderately correlated, suggesting that they were measuring different constructs. It was hypothesized that verbal memory might play a significant role in mediating each of the language tasks, but in different ways.

The second study was designed to explore the hypothesis developed in study one regarding the role of verbal memory in language assessment by incorporating measures of short-term and working memory with low language loads (digit span forward and reversed). Vocabulary awareness scores were combined to create a composite language
variable with low demands on memory. Regression models showed different patterns of variables (age, vocabulary awareness, and memory) predicting nonword spelling, listening comprehension, and direction following performances of children with and without LI. Memory contributed unique variance to nonword spelling for both groups and to direction following for the students with TL but not LI.

The third study investigated patterns of association between memory skills and language skills. Four theory-driven hypotheses were presented. Results of classifications based on cut-off scores for a memory composite and language composite created profiles (high memory/low language, high memory/high language, low memory/low language, and low memory/high language). Analysis of these profiles using chi-square indicated statistically significant results disproving the null hypothesis of no association between memory and language. Further analysis concluded that the hypothesis that best fit the data was that memory is necessary, but not sufficient for language.
ACKNOWLEDGEMENTS

I am so grateful for the manner in which this process unfolded. As with many of my journeys, I knew where I was beginning this Ph.D. process, and I had an idea of what I hoped to find at the destination, but I'm always aware that the path may offer up unforeseen challenges, as well as, unplanned opportunities along the way. Some of the best trips have been those with unintended detours.

Beginning with a cohort of eleven other people embarking at the same point and walking in lockstep those first two years was a wonderful way to gain a feel for the road. Learning in an interdisciplinary environment with such diversity enriched the process. I thank each of you for your insights, encouragement, friendships and much needed humor that made this journey "doable".

I am indebted to and wish to acknowledge the generous support provided by my dissertation chair, Dr. Nickola Nelson. Although this may have started as part of her "life's work", it has morphed unexpectedly into my own. I would have been the first to state quite emphatically that I am a clinician who works with adults, not kids. In our working together during these past few years, I have found an unbeknownst interest in applying my years of interest in cognitive processing to questions of assessment of language in children. I appreciate Dr. Nelson's patience as my attitudes became less dogmatic, benefitting from her love of learning and open-mindedness and our--at times--passionate debates. And I would be remiss if I did not mention my gratitude for the persistent yet kindly proffered feedback from my other committee members, Dr. Amy Curtis and Dr. Diane Dirette. With their assistance, I have gained more confidence in my abilities to design, conduct, and disseminate the results of research.
Acknowledgements—Continued

Family and friends provided the fortitude to keep me going. I wish to give a special "shout out" to neighbors, Susie Korstange and Lucia Leonardelli, both of whom provided not only moral support, but offered their time and energy to assist with data verification. To my grandmother, Betty Fetrow, I say thank you for setting an example for our family of the importance of education and being a life-long learner. To my mother, thanks for answering the phone during some of those dark moments each of us encounter. And to Carla, without your support and encouragement, I might have opted for a detour or turned the car around altogether. Now that this journey is over, I can say "it's been a great ride"!

Michele A. Anderson
TABLE OF CONTENTS

ACKNOWLEDGEMENTS.............................................................................................................. ii

LIST OF TABLES...................................................................................................................... ix

LIST OF FIGURES...................................................................................................................... x

CHAPTER

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

Memory Definitions .................................................................................................................. 1

Capacity and Processing in Assessment............................................................................... 3

Introduction to the Three Studies ......................................................................................... 4

References ............................................................................................................................... 5

II. A PILOT STUDY OF THREE INTEGRATED TASKS OF LANGUAGE AND VERBAL MEMORY: PSYCHOMETRIC PROPERTIES AND CORRELATIONS......................................................................................................................... 7

Short-term Memory and Working Memory .......................................................................... 8

A Model of Working Memory................................................................................................. 10

Working Memory and Language Skills Integrated in the Classroom ............................. 11

Working Memory and Language Impairment .................................................................... 13

Language Assessment and Verbal Memory ........................................................................ 14

Purpose of the Current Research Study ............................................................................. 15

Methods ................................................................................................................................. 16
Table of Contents—Continued

CHAPTER

Participants ............................................................................................................. 16

Measures ............................................................................................................. 17

General Testing Procedures .............................................................................. 21

Data Preparation and Analysis .......................................................................... 22

Results ................................................................................................................ 23

Psychometric Properties of the Measures ....................................................... 23

Relationship of Performance on Three Subtests ............................................. 28

Discussion .......................................................................................................... 28

Psychometric Properties ................................................................................ 29

Theoretical Implications .................................................................................. 31

Limitations and Future Research Directions .................................................. 33

Conclusions and Future Research Directions ............................................... 35

References ......................................................................................................... 37

III. STUDY TWO: INVESTIGATION OF THREE INTEGRATED TASKS
OF LANGUAGE AND MEMORY ............................................................................ 42

Purpose of the Current Research Study .......................................................... 45

Methods ............................................................................................................ 46

Participants ....................................................................................................... 46
Table of Contents—Continued

CHAPTER

General Testing Procedures .............................................................. 48
Measures ......................................................................................... 49
Data Preparation and Analysis ......................................................... 52
Results ............................................................................................ 54
Descriptive Statistics and Differences Between Groups ..................... 54
Variance in Language Task Performance Accounted for by Age, Vocabulary Awareness, and Memory ....................................................... 55
Discussion ....................................................................................... 58
Strengths, Limitations, and Future Research Directions ...................... 61
Conclusions ..................................................................................... 62
References ....................................................................................... 63

IV. EVIDENCE FROM SCHOOL-AGE CHILDREN FOR ASSOCIATION AND DISSOCIATION OF MEMORY AND LANGUAGE .................. 66

Memory Defined ............................................................................... 67
Memory and Specific Language Impairment ........................................ 68
Necessary and Sufficient .................................................................. 71
Purpose ............................................................................................. 72
Methods ............................................................................................ 77
Participants ....................................................................................... 77
Table of Contents—Continued

CHAPTER

General Testing Procedures ................................................................. 78

Measures of Language Impairment, Memory, and Language ............... 79

Data Preparation and Analysis Plan .................................................... 83

Results ............................................................................................... 84

Results of the Classification ............................................................... 84

Results of Chi-square Analysis ........................................................... 88

Discussion .......................................................................................... 89

Association and/or Dissociation of Memory Skills
and Language Skills ........................................................................... 89

Memory Skills as Necessary and/or Sufficient for Language ............... 91

Consistency with Existing Literature and Theory .................................. 91

Strengths and Limitations ................................................................... 92

Clinical Implications ......................................................................... 93

Future Research and Recommendations ............................................ 94

References ......................................................................................... 95

V. CONCLUSION .................................................................................. 100

Central Purpose ................................................................................ 100

Summary of Findings .......................................................................... 102
Table of Contents—Continued

CHAPTER

Study One Summary ................................................................. 102

Study Two Summary ............................................................... 103

Study Three Summary ............................................................. 104

Discussion .................................................................................. 105

Conclusions ............................................................................... 106

Clinical Implications ................................................................. 107

Recommendations for Further Research ...................................... 108

References ................................................................................ 109

APPENDIX

A. HSIRB Approval Letters ......................................................... 110
LIST OF TABLES

2-1. Demographic Information of Participants (N = 187) ........................................ 17
2-2. Traits of Pilot Tasks ........................................................................................................ 21
2-3. Reliability Results for Three Tasks ................................................................................ 24
2-4. Means and Standard Deviations (SD) for Each Task by Grade ................................. 26
2-5. Games-Howell Multiple Comparisons of Differences in Means by Task and Grade Level ........................................................................................................................................ 27
2-6. Kendall's Tau Correlations Between the Three Tasks .............................................. 28
3-1. Criteria for Group Determination .................................................................................. 47
3-2. Demographics of Typical Language and Language Impaired Groups ......... 48
3-3. Descriptive Statistics for Measures as a Function of Dyad Designation...... 54
3-4. Results of t-Test Differences Between Group (TL and LI) Performance on Independent and Dependent Measures ................................................................................................................................. 54
3-5. Correlations Between Measures for Language and Memory Tasks ............ 55
3-6. Results of Multiple Regression Analysis ...................................................................... 57
4-1. Descriptive Statistics for the TL and LI Participants ................................................. 78
4-2. Range and Mean of Standard Scores on CELF-4 Core Subtests for the TL and LI Groups .................................................................................................................................................. 81
LIST OF FIGURES

2-1. Example of Following Directions Stimulus Item.............................................. 18
2-2. Histogram of Nonword Spelling........................................................................ 24
2-3. Histograms for Listening Comprehension and Following Directions............... 25
2-4. Means of Raw Scores by Grade Level on Three Language Tasks.................... 26
4-1. Data Distribution Predicted by Hypothesis 1.................................................. 73
4-2. Data Distribution Predicted by Hypothesis 2.................................................. 74
4-3. Data Distribution Predicted by Hypothesis 3.................................................. 75
4-4. Data Distribution Predicted by Hypothesis 4.................................................. 76
4-5. Percent of Participants in Each Category of High/Low Skills.......................... 85
4-6. Participants Scoring Above and Below Language Cut-Scores Based on Language Status............................................................... 86
4-7. Participants Scoring Above and Below Memory Cut-scores Based on Language Status................................................................. 87
4-8. Distribution of Participants by Language Ability into Four Performance Categories........................................................................ 88
CHAPTER I

INTRODUCTION

This dissertation comprises five chapters, starting with an introduction, followed by three separate yet related research studies, and culminating with a discussion. The unifying construct under investigation is the relationship of memory and language, specifically in school-age children.

Memory Definitions

The construct of memory has been studied by researchers from a variety of fields such as psychology, education, cognitive science, and speech-language pathology. Terminology describing memory varies across fields and authors, however; thus it is necessary to provide definitions that will be used for the purpose of this dissertation.

*Short-term memory* (STM) is used in these studies to describe the aspect of memory that is limited to a few seconds in duration. Assessing short-term memory span is often accomplished by requesting repetition of a series of digits, words, or nonwords (verbal domain) or repeating a pattern for touching blocks (visual domain). STM has a limited capacity referred to as span. Span is defined as the largest number of units (e.g., words, digits, spatial positions that can be encoded, retained, and repeated accurately).

*Working memory* (WM) is a more complex aspect of memory. It requires not only short-term memory for encoding stimulus material, but also other attention-related
processes and coordination with long-term memory to complete complex, usually multi-step tasks. WM is defined as the ability to store information temporarily while concurrently performing a processing task, such as completing a 2-step mental math problem (Baddeley, 1996). The cognitive systems involved in WM are distinct from those used for STM, which requires storage, but only minimal processing of information (Atkinson & Shiffron, 1968). Verbal STM tasks involve serial recall of numbers or words without any requirements for mental manipulation of these or other stimuli (e.g., Conrad & Hull, 1964). Working memory is typically assessed through a dual task. For example, in assessing the verbal domain of working memory, a sentence span task may be used whereby the person is asked to verify the truth of a sentence and then to recall the final words in a set of those sentences (Daneman & Carpenter, 1980). Completing steps for a recipe, following directions to assemble furniture without repeating or omitting a step, and reconciling a bank statement are all examples of ordinary tasks that require the use of working memory.

Researchers in fields such as cognitive psychology, cognitive neuroscience, and cognitive linguistics are interested in how the human brain manages the feat commonly referred to as WM. Andrade (2001) expanded the definition of WM by describing it as “a system that enables temporary storage of the intermediate products of cognition and supports transformations of those products” (p. 5). Miyake and Shah (1999), in their review of contemporary theories of WM, found consensus among current researchers that “working memory is those mechanisms or processes that are involved in control, regulation and active maintenance of task-relevant information in the service of complex cognition” (p. 450).
Capacity and Processing in Assessment

Complex cognitive tasks requiring verbal WM have relevance to language development and processing, especially in classroom settings. Determining a student’s ability to perform complex tasks that require the ability to hold information in memory while manipulating it, is important to speech-language pathologists and other special service providers, as well as general education teachers. Clinicians need to understand the influences of verbal WM for language performance and the degree to which verbal WM interacts with language ability to make it possible to process language in the context of inter-modality classroom activities, as well as in common language assessment tasks.

Tasks constructed to measure specific language functions should account for the impact of verbal WM needed to complete the tasks successfully. Assessing students’ abilities in a manner consistent with curriculum-relevant tasks may yield results more valid for classroom performance, but it would be helpful to be able to disentangle the influence of memory difficulties from more specific problems involving language ability. In order to accomplish this, assessment in children at-risk for language or academic development should include an overt measure of verbal WM.

The dual goals of this dissertation are to contribute to theoretical models regarding the interaction of memory with language in assessment and as an underlying factor in explaining language impairment, while answering practical questions about the best ways to interpret integrated assessment tasks that include both language and memory demands. The nuances of the interactions in theory and differential diagnosis gain further practical significance when considering implications for designing intervention strategies for children struggling either with language, and/or memory or academic skills.
Introduction to the Three Studies

The first study is a descriptive correlational investigation of performance by students in 2nd, 4th, 6th, and 9th grades on three pilot curriculum-related language assessment tasks, which placed varying demands on language processing and working memory. The three tasks were developed to evaluate different language skills, but shared an administration characteristic in that the stimuli all were presented auditorially, introducing a verbal memory component for each task. The sound/word level task required nonword spelling, the sentence/discourse level task addressed listening comprehension, and the procedural sentence level task required a graphomotoric response to auditory directions. Each task therefore, was expected to reflect the integration of verbal memory skills with a specific language skill.

The second study is a regression study. Changes were made in the design for study two, based on the findings from study one. These included changes in participant sampling methods, stimulus items, and administration procedures. Inclusion of new assessment tasks, that is language tasks with low memory demands and memory tasks with low language demands, provided the opportunity to examine the amount of unique contributions each might add to the shared variance in performance. Revised versions of the three pilot tasks from study one were used as dependent measures. Empirical data from students who met inclusion/exclusion criteria as having language impairment (LI) and typical language (TL) also allowed for exploration of patterns of variables that predict performance in these two groups.

In the final study, questions of association or dissociation of memory skills and language skills were explored. Four hypotheses were proffered to account for four
theoretically driven patterns of data distribution that all were considered to be within the realm of possibility. Performances by students were categorized based on 50th percentile cutoff-scores into high or low memory skills, and high or low language skills. Combining the two skill profiles (high or low) allowed for further categorizing into high memory/low language, high memory/high language, low memory/low language, or low memory/high language combinations. Additionally, by considering the quadrants in which participants fell who were already categorized as having TL and LI, it was possible to examine whether the evidence supported one of the hypotheses over others.

These three studies were all designed to address questions related to the relationship between memory skills and language skills. A better understanding of how these two skills might relate—in children with typically developing language and those with language impairment—can inform both theory and practice.

References


CHAPTER II

A PILOT STUDY OF THREE INTEGRATED TASKS OF LANGUAGE AND VERBAL MEMORY: PSYCHOMETRIC PROPERTIES AND CORRELATIONS

Successful classroom performance requires processing of information through multiple sensory modalities—often simultaneously—while screening out irrelevant information (e.g., auditory, visual, or autonomic) from other sources. Group learning contexts of classroom activities require that children spend more time listening than talking in school. Teachers rely on informal observation and academic products from classroom activities to gain impressions of students’ abilities and teaching/learning needs. When problems are noted, professionals such as speech-language pathologists may be called upon, at which time, they need individualized formal assessment tools and observational methods to help determine the nature of the difficulty.

This study examines three language assessment tasks that were being piloted for inclusion in a comprehensive assessment battery of curriculum-related language and literacy skills. The three tasks were designed to reflect different aspects of curriculum-related language and literacy processing abilities: following verbal directions, listening comprehension, and nonword spelling. The purpose of this study to explore correlations among them for insights into the associations between children’s language and the latent trait of verbal memory that plays a role in each of them.
Each of these tasks was designed to measure curriculum-related language ability as an overt trait. Following directions requires vocabulary and sentence-level comprehension, along with procedural knowledge. Listening comprehension necessitates semantic and syntactic understanding of discourse and the ability to relate language to underlying meanings. Nonword spelling is a word-level task requiring phonological, morphological, and orthographic knowledge of word structure. In each task, however, it is possible that a child who has the basic language skills to complete the task may make an error due to an inability to remember the auditory verbal stimulus. Clinicians need tools that can help them to differentiate a primary impairment in language skills from underlying, more general, deficits of memory.

Research has shown that memory problems run the risk of being undetected or misdiagnosed by teachers as attentional deficits or lower intelligence (Gathercole, Lamont, & Alloway, 2006). Alloway, Gathercole, Kirkwood, and Elliott (2009) found that 10% of a sample of over 3,000 children in the United Kingdom exhibited working memory skills that were poor enough to interfere significantly with learning. With nearly 50 million students in the United States (U.S. Department of Education, 2009), projecting 10% of children to be affected would mean that 5 million children could be expected to have deficits in working memory.

**Short-term Memory and Working Memory**

Part of the problem in sorting out the roles of language and memory when children have difficulty in school is coming to a common definition of what constitutes memory. As a construct, memory is a general cognitive ability that plays an important
role in much of language processing in school. It has been a focus of research across disciplines such as educational psychology, special education, and speech-language pathology as shown by articles in various journals including: Learning Disabilities: A Contemporary Journal (Masoura, 2006), Applied Psycholinguistics (Marton, 2006), Journal of Experimental Child Psychology (Towse, Hitch and Hutton, 2002), and the International Journal of Language and Communication Disorders (van Daal, Verhoeven, & van Balkom, 2009). Terms describing aspects of memory, however, may vary by researcher and discipline. Some terms are used interchangeably for closely related, but not identical, concepts. For example, the terms short-term memory, and immediate memory may be treated as synonymous terms for recent memory in some sources, but they may be clearly distinguishable in others.

For current purposes, short-term memory is defined as temporary storage of sensory input (lasting a few seconds), with a limited, definable, capacity, which is often referred to as span. Span is usually operationalized as the longest number of units that can be received, retained, and then repeated without variation. Short-term memory span can be assessed by requesting immediate recall and repetition of a series of digits, words, or nonwords.

Contrasting with the construct of short-term memory, working memory is defined as “those mechanisms or processes that are involved in control, regulation and active maintenance of task-relevant information in the service of complex cognition” (Miyake & Shaw, 1999, p. 450). Working memory includes short-term memory in addition to other attention-related processes and activation of long-term memory to plan and complete tasks. Working memory can involve verbal and nonverbal components. It is
required, for example, to remember the results from the first part of an arithmetic problem to carry-out the next step, such as a clerk making change.

Assessments of working memory have included tasks such as digit span backwards, in which the list of numbers in short-term memory must be manipulated mentally and repeated in reverse order of presentation. Other examples include mathematical tasks requiring mental arithmetic or language tasks requiring sentence comprehension followed by requests for verification of the truth value of complex statements, or requests to recall final words from a set of preceding sentences. Examples of sentence-level working memory measures can be found in the Reading Span Test (Daneman & Carpenter, 1980) and the listening recall task from the Automated Working Memory Assessment (Alloway, 2007).

_A Model of Working Memory_

Sorting out the roles of language and memory in curriculum-related language tasks can benefit from a coherent theory of working memory. A prominently cited model of working memory, which was proposed by Baddeley and Hitch (1974) and updated in more recent publications (e.g., Baddeley, 1996), has been called the “fractionated” or multimodal model. That is because it differentiated memory processes for different modes of input (e.g. auditory or visual), but still with a common central processor.

According to this model, in acoustic or speech-based, hence “verbal memory” tasks, phonological information is stored temporarily and must be rehearsed actively; otherwise decay of the memory traces occurs within two to three seconds. The rehearsal is thought to occur through subvocal articulation, which Baddeley and Hitch called the
“phonological loop.” Thus, this theory explicitly ties working memory to linguistic systems at the phonological (sound/word) level.

Evidence supports that working memory capacity more generally is a developmental skill that improves with age, reaching an adult asymptote at around age 14 (Gathercole, Pickering, Ambridge, & Wearing, 2004). Other studies have shown age to be a factor in accounting for improvements in performance on diverse memory span tasks during childhood (Kail, 1992; Kail & Park, 1994). Rate and span factors also interact. Kail (1991) found that individuals perform most cognitive processes more rapidly as they mature. He proposed that developmental improvements in span reflect a corresponding increase in processing speed, which, in turn, provides more opportunity to rehearse or refresh auditory stimuli, such as words, in the articulatory loop. More rapid rehearsal is associated with increasing maturity, Kail argued, accounting for increases in memory span with age. Such interactions add to the difficulty of teasing apart the roles of language and memory when children have difficulty processing the language of the curriculum.

*Working Memory and Language Skills Integrated in the Classroom*

Many forms of classroom communication require children to integrate auditory verbal input with visual spatial concepts. A prime example of this is the language of directions. Children’s inability to follow verbal directions may be a concern of educators. In formal classroom discourse, as much as 13-15% of what the teacher says is in the form of procedural instructional language (Sturm & Nelson, 1997). In order to follow a teacher’s directions, a student must understand the teacher’s words and syntactic
structures and retain their essence in verbal short-term or working memory while completing step-by-step procedures, often involving visual spatial representation of concepts processed through auditory verbal systems (e.g., completing art projects or turning to a particular page in a text book and completing a set of problems).

Word-level processing skills also are in high demand in curricular language tasks, particularly when literacy skills are being developed, including reading decoding and spelling single words to dictation. Assessment tasks for measuring word level skills often use nonword tasks to avoid the influence of knowledge of specific vocabulary from long-term memory. Novel words, called “nonwords” or pseudowords, bear resemblance to English words but are not real words. Although vocabulary knowledge may influence the processing of nonwords (Edwards, Beckman, & Munson, 2004), it is not the primary trait assessed in nonword reading and spelling tasks. Nonword spelling tasks are considered a measure of a student’s ability to process the phonological and orthographic structure of words (Holmes & Quinn, 2009). In addition, however, nonword repetition tasks assess the latent trait of phonological short-term memory (Gathercole, Tiffany, Briscoe, Thorn & ALSPAC Team, 2005).

Comprehension of new material through language is a primary goal and also a primary tool of formal education. Sturm and Nelson (1997) found that 12-16% of teacher talk is used to convey content. Comprehension is clearly a language processing skill, but reading and listening comprehension both place heavy demands on working memory as well. Reading and listening comprehension require more than short-term memory for a stream of incoming words (Daneman & Merikle, 1996); students must construct meaning actively from the semantic and syntactic relationships of the words within and across
sentence boundaries as they listen to spoken discourse or read written discourse. Working memory is required to support processing of current incoming language symbols while accessing recently stored information and drawing from language knowledge in long-term memory to support the construction of meaning (Ericsson & Kintsch, 1995).

*Working Memory and Language Impairment*

Studies have shown that working memory supports the normal development of reading and writing (Hoskyn & Swanson, 2003; Swanson & Berninger, 1995, 1996). The nature of the relationship between impaired language and limited working memory, however, is a matter of debate. Some view children’s impairments in short-term or working memory as somewhat, but not completely, separate from their language problems (Cowan, 1996; Hoffman & Gillam, 2004; Towse et al., 2002). Others view working memory deficits as significant contributors to language problems, possibly underpinning the language-learning difficulties of children with specific language impairment rather than reflecting them (Archibald & Gathercole, 2006; Leonard et al., 2007; Montgomery, 2002).

Archibald and Joanisse (2009) conducted a study of 88 children ranging in age from 5 to 9 years of age in which they found support for profiles of deficits. This study was unusual because, rather than starting with a group of children with language impairments and then measuring their memory, Archibald and Joanisse started with a large group of children who were identified in a widespread study on the basis of their deficits in memory. Thus, they were able to identify groups of children who showed a pattern of either specific working memory impairments, specific language impairments,
or combined. Some children’s deficits were specific to language or working memory, but a subset of children had deficits in both. Archibald and Joanisse concluded that working memory and language skills may have an additive, rather than a unidirectional or causal pathway.

Archibald and Joanisse (2009) interpreted their findings to mean that working memory problems do not always cause language problems and language problems do not always cause working memory problems. The presence of a working memory problem in the verbal domain—but not the visual domain—which co-occurred in children with the most severely impaired language skills, suggests that impaired verbal working memory constrained performance in language skills. In addition, the results from their study showed specific working memory deficits to be a discrete category of impairment for a subgroup of children who demonstrated language skills that were within functional limits. Further study is needed to determine if more complex language skills, such as narration and discourse comprehension, require storage and processing demands that can not be supported by the weaker working memory of students with a relatively specific deficit in memory, but with intact basic language skills.

Language Assessment and Verbal Memory

Language assessment tasks that place dual demands on language and memory abilities reflect the authentic demands of curricular discourse processing, but they present challenges of interpretation. Without further assessment, it may be impossible to know whether an error on a particular language task was due to inadequate skill in the primary language trait assessed or weakness in a general processing capacity such as verbal short-
term or working memory. For purposes of differential diagnosis and intervention planning, it would be helpful to know whether any problems are better explained by deficits in a student’s existing language skills, such as phonemic awareness, vocabulary knowledge, and syntactic ability or deficits in verbal short-term or working memory processes used to access that knowledge in the context of a particular task. This distinction can influence the selection of intervention targets and intervention approaches for such students (Alloway, 2009; Gillam & van Kleeck, 1996; Montgomery, Magimairaj, & Finney, 2009).

Purpose of the Current Research Study

The purpose of the current pilot study was to investigate associations among the three language tasks as a preliminary step to further exploration of children’s performance on curriculum-related tasks with varying language and verbal memory demands (short-term and working). The researchers hypothesized that a nonword spelling task, a listening comprehension task, and a following directions task—each with auditory presentation of stimuli—should show a significant correlation because of the shared reliance on verbal memory. However, because they also were testing distinct language constructs (phonological, orthographic, sentence discourse or procedural), it was further hypothesized that the correlations should only be moderate in strength (.30-.69). If the correlations among the three tasks are >.70, then concerns arise that the tasks may be measuring the same latent construct (verbal memory). This purpose was addressed with the following research question:
Research Question 1: Are there statistically significant correlations between three language assessment tasks presented auditorally that support a hypothesis of verbal memory as a latent trait?

Research Question 2: If statistically significant correlations between the three language tasks are found, what is the level of correlation (weak .10-.29, moderate .30-.49, strong ≥ .49)?

Methods

This study used secondary data collected during a field test of language assessment tools under development. This was a correlational study designed to investigate the strength of correlations among three pilot versions of curriculum-related language assessment tasks.

Participants

Parental consent and student assent following a protocol approved by a Human Subjects Institutional Review Board were obtained for 218 student participants in second, fourth, sixth, and ninth grades. The students were recruited from classrooms in two Midwestern school districts (one urban and one rural) whose administrators and teachers had agreed to participate in the study. All students were invited to participate (no exclusion criteria). Demographic characteristics of the sample are summarized in Table 2-1.
Table 2-1. Demographic Information of Participants (N = 187)

<table>
<thead>
<tr>
<th></th>
<th>Sex</th>
<th>Race/Ethnicity</th>
<th>Special Needs</th>
<th>FRL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>W</td>
<td>NW</td>
</tr>
<tr>
<td>Urban</td>
<td>64 (48)</td>
<td>70 (52)</td>
<td>63 (47)</td>
<td>69 (52)</td>
</tr>
<tr>
<td>Rural</td>
<td>40 (48)</td>
<td>44 (52)</td>
<td>77 (92)</td>
<td>3 (4)</td>
</tr>
</tbody>
</table>

W=white; NW=nonwhite; FRL=free or reduced lunch

Note: numbers do not always add to 218 due to missing data for FRL (n = 7) and Race (n = 6)

Nearly 63% of the students were from the urban school district (n = 134; 48 second graders, 19 fourth graders, 35 sixth graders, and 32 ninth graders). The remaining participants were students in a rural school district (37%, n = 84; 15 second graders, 20 fourth graders, 22 sixth graders, and 27 ninth graders). Slightly more participants were female (52%; n = 114) than male (48%; n = 104). Only 4% (n = 3) of the students from the rural school were categorized as nonwhite. Thirty students with special needs participated in this study, including nine who had more than one of the following diagnoses: hearing impairment (n = 1), speech and language impairment (n = 12), learning disability (n = 3), emotional-behavioral impairment (n = 12), autism spectrum disorder (n = 1), cognitive impairment (n = 3), and/or attention deficit hyperactive disorder (n = 13).

Measures

Three pilot tasks were being evaluated for possible inclusion in a comprehensive “Test of Integrated Language and Literacy Skills” (TILLS; Nelson, Helm-Estabrooks, Hotz, & Plante, 2007). Construct validity had been considered by the test authors when they developed the model for each of the subtests of the comprehensive TILLS. The
test model was designed to assess language abilities across all four communication modalities (listening, speaking, reading, and writing) using curriculum-relevant tasks at the sound/word and sentence/discourse levels that would require integrated language and literacy processing skills. An expert panel reviewed the test model and offered suggestions, some of which resulted in the addition of several nonword tasks, including nonword spelling. Content validity for each task also was addressed by designing items to reflect actual examples of classroom discourse (spoken and written). The goal was to sample a range of length and complexity that would tax students' abilities across the full range of the school-age years (6;0 through 18;11), but remain at a manageable length in terms of numbers of items.

Following Directions Task. The Following Directions task required students to attend to a series of verbal directions presented by a researcher standing at the front of the class, while students were looking at visual stimuli arranged spatially in a linear array on their individual response booklets (see Figure 2-1).

"Draw a line from the lightning bolt over the moon and into the sun. Draw a circle under the arrow. Go."

Figure 2-1. Example of Following Directions Stimulus Item
To complete the task successfully, students had to understand the component parts, remember them, and carry them out using a grapho-motor (paper and pencil) response. The 19 items were constructed to increase systematically in number of words, number of steps to complete, and number of critical units of information in a hierarchical arrangement. The same 19 items were used at all grade levels.

*Listening Comprehension Task.* The Yes/No/Maybe task was designed to assess language comprehension of both spoken and written texts. The data from the Listening Only version of the task were included in this study to provide a measure of verbal working memory combined with language comprehension for content language (in contrast to the procedural language required for the following directions task).

Students listened to five short "stories" that were read aloud by the examiner. The stem stories ranged from one sentence (with embedded clauses) to three sentences. The language of the stories was constructed to represent the complex syntactic and semantic relationships of curricular language. The examiner was not allowed to repeat the stimuli. Students then were asked to answer four questions about the events of the story. The required response was to circle *Yes*, *No* or *Maybe* from the set of choices numbered and printed on their response forms. Students were directed to circle "yes" if they *knew* the answer was yes, "no" if they *knew* the answer was no, but to circle "maybe" if the story did not give the enough information to know the answer. The correct choices were balanced across the items as follows: 6 Yes, 7 No, and 7 Maybe. The demonstration item used for this task follows:

*One morning Susan missed the school bus. She thought she would be late for school, but her mother got her there on time.*
a. Did Susan miss the bus? Y N M
b. Was Susan late for school? Y N M
c. Was it raining that morning? Y N M
d. Did her father take her to school? Y N M

This Yes/No/Maybe format is based on a comprehension assessment paradigm used by Levine (1996). By requiring judgment of information adequacy, this task also was similar to the veracity judgment paradigm used in some other assessments of working memory (e.g., Daneman & Carpenter, 1980).

**Nonword Spelling Task.** The nonword spelling task involved a request for students in the class to respond chorally to the researcher’s production of nonsense words, some of which incorporated real morphemes (e.g., “say rebifology”). Then they were directed to spell each word, which was repeated in the directions (e.g., “spell rebifology”). Students produced a paper and pencil response on numbered lines in their response booklets. They were told that these were not real words, but they were to use what they knew about real words to spell the pretend words they heard. This task included 13 items with increasing phonological length and complexity.

Nonword repetition has been used frequently as a measure of phonological short term memory (Gathercole & Baddeley, 1989; 1990). Nonword spelling adds requirements for alphabetic knowledge to demands on phonological awareness and memory. To perform the nonword spelling task, the student must be able to hold the novel sequence of phonemes in phonological short term memory (via the phonological loop) while accessing long-term memory for correct orthographic representations.
Table 2-2 summarizes the task demands for the three pilot subtests used in this study. Note that all three use auditory presentation of a series of stimuli, thus assessing in part, verbal working memory. The table also specifies the differing types of language units that the student must hold in memory to complete each task, as well as variations in length in terms of the number of units that must be remembered while processing.

Table 2-2. *Traits of Pilot Tasks*

<table>
<thead>
<tr>
<th></th>
<th>NW spelling</th>
<th>YNM-L</th>
<th>Following Directions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulus Mode</strong></td>
<td>Auditory</td>
<td>Auditory</td>
<td>Auditory</td>
</tr>
<tr>
<td><strong>Types of Language Units</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sound/Word</td>
<td></td>
<td>Sentence/Discourse (narrative)</td>
<td>Sentence/Discourse (procedural)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vocabulary/World knowledge,</td>
<td>Vocabulary knowledge,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Syntactic cohesion,</td>
<td>Phrase/clause relationships,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Veracity judgment</td>
<td>Graphic representation</td>
</tr>
<tr>
<td><strong>Language Processes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonological awareness,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sound-symbol association,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter formation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Length/span of Stimuli</strong></td>
<td>1-5 syllables</td>
<td>1-3 sentences</td>
<td>3-30 words</td>
</tr>
<tr>
<td>1-10 phonemes</td>
<td></td>
<td></td>
<td>1-4 actions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2-11 content units</td>
</tr>
<tr>
<td><strong>Response Mode</strong></td>
<td>Orthographic (writing letters)</td>
<td>Paper/Pencil (circling choice)</td>
<td>Paper/Pencil (e.g. drawing lines, circles, boxes)</td>
</tr>
</tbody>
</table>

*General Testing Procedures*

Assent and consent were obtained under a protocol approved by the Human Subjects Institutional Review Board. In addition to the three subtests investigated in this study (following directions task, listening comprehension task, and nonword spelling task), the field test included an original story writing task, two short structured writing tasks, and two other Yes-No-Maybe tasks that involved reading comprehension (not discussed in this report).
All tasks were group-administered to whole classes of students in three 1-hour sessions in the fall term (in November). Testing was repeated with the same students near the end of the spring term of the same academic year (in April). Following item analysis of the fall data, an improved version of the language comprehension task was used in the second administration. Thus, the inter-task comparison data reported in the current study were drawn from the spring testing session. Because the listening comprehension task was modified significantly, test-retest data were only available for the following directions and nonword spelling tasks. The language comprehension and nonword spelling tasks were administered on the second day of the three 90 min daily sessions along with a sentence-combining formal writing task. The following directions task was administered on the third day of testing in both fall and spring sessions, along with a graphic-organizer formal writing task and a classroom questionnaire.

Data Preparation and Analysis

Three trained research assistants scored the test protocols. A subset of the original protocols (20%) was rescored to compute interscorer reliability. Data entry was checked for accuracy by a second assistant, and any discrepancies were corrected by referring to the original source data. Raw data that were entered using an ACCESS database system were imported into SPSS version 15.0 (SPSS, 2006) for analysis. Prior to running statistical analyses of the group data, tests of normality were conducted. When data distributions violated the assumptions of normality, nonparametric tests were used.

Since the language tasks were constructed to measure language skills known to be developmental, differences in means by grades were examined. One-way analysis of
variance (ANOVA) was conducted to determine if age was significantly related to performance on each task.

The research question asked about the relationships among the three tasks and their implications for understanding the relationships of language processing and verbal memory. The pilot nature of this investigation and the lack of independent measures of verbal short-term and working memory limited the degree to which the presence of relationship between the overt language traits and verbal memory or any other latent construct could be analyzed. A preliminary step, however, was to conduct this correlational study of the relationships between performances on all three tasks. The rationale for these correlational analyses was that each of the tasks requires verbal memory to successfully complete due to the auditory presentation of the stimuli. Therefore, the three tasks should be correlated, but only moderately (.30-.49), as each task also required a significantly different language processing skill—phonological, orthographic, discourse, and procedural knowledge.

Results

*Psychometric Properties of the Measures*

Interscorer reliability (agreement between two scorers), internal consistency (usually measured with Cronbach’s alpha), and test-retest reliability were assessed for each task. The point-by-point results for the subsets of 20% of randomly selected tests were compared to compute interscorer reliability as a percentage of inter-scorer agreement. The results for all three forms of reliability analysis are shown in Table 2-3.
Table 2-3. **Reliability Results for Three Tasks**

<table>
<thead>
<tr>
<th>Task</th>
<th>Inter-scorer reliability</th>
<th>Cronbach’s Alpha</th>
<th>Test-Retest Correlation* (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Following Directions</td>
<td>97.6%</td>
<td>.81</td>
<td>.631**</td>
</tr>
<tr>
<td>YNM-listening</td>
<td>99.0%</td>
<td>.87</td>
<td>NA</td>
</tr>
<tr>
<td>Nonword Spelling</td>
<td>99.6%</td>
<td>.83</td>
<td>.789**</td>
</tr>
</tbody>
</table>

* Pearson product-moment correlation coefficient**p ≤ .001;
NA = not applicable, due to significant revision of task between test trials

The data were examined for normality. The distributions for nonword spelling, $D(170) = .147, p < .0005$, listening comprehension, $D(170) = .105, p < .0005$, and following directions, $D(170) = .212, p < .0005$ were all significantly non-normal. As the histogram in Figure 2-2 shows, the skewness is likely due to the floor effect resulting from the significantly lower performance of the second graders.

---

*Figure 2-2. Histogram of Nonword Spelling*
The histograms for the listening comprehension task and following directions task are shown in Figure 2-3. The skewness in following directions task is likely the result of a ceiling effect in the upper grades.

![Histograms for Listening Comprehension and Following Directions](image)

*Figure 2-3. Histograms for Listening Comprehension and Following Directions*

To determine if the skills needed to complete the three language tasks successfully are developmental, an analysis of the differences in mean scores by grade level was conducted. ANOVA was selected as it is a robust test when the assumption of normal distribution is violated. The descriptive statistics (means and standard deviations) are shown in Table 2-4 and also are represented graphically in Figure 2-4.
Table 2-4. *Means and Standard Deviations (SD) for Each Task by Grade*

<table>
<thead>
<tr>
<th>Grade</th>
<th>Following Directions (possible 19) Mean (SD)</th>
<th>Listening Comprehension (possible 20) Mean (SD)</th>
<th>Nonword Spelling (possible 13) Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second</td>
<td>15.19 (2.094)</td>
<td>10.68 (3.078)</td>
<td>3.24 (2.808)</td>
</tr>
<tr>
<td>Fourth</td>
<td>16.82 (2.236)</td>
<td>12.52 (3.298)</td>
<td>6.16 (3.418)</td>
</tr>
<tr>
<td>Sixth</td>
<td>17.76 (1.544)</td>
<td>13.96 (2.074)</td>
<td>7.86 (2.090)</td>
</tr>
<tr>
<td>Ninth</td>
<td>17.90 (1.359)</td>
<td>14.10 (2.697)</td>
<td>8.04 (2.489)</td>
</tr>
</tbody>
</table>

*Figure 2-4. Means of Raw Scores by Grade Level on Three Language Tasks*

To determine if the grade-level effect was statistically significant, an ANOVA was conducted with a Games-Howell post hoc. The ANOVA results showed a statistically significant main effect for grades for all three tasks: following directions $F(3) = 24.783, p = .001$; listening comprehension $F(3) = 16.856, p = .001$; nonword spelling $F(3) = 36.747, p = .001$.

The results of Games-Howell post hoc analyses are summarized in Table 2-5, which shows grade-by-grade comparisons. The Games-Howell was chosen because it is a
more powerful post-hoc procedure (than Dunnett’s Cor Tamhane’s T2) and also is considered accurate when using unequal sample sizes (Field, 2005). Statistically significant differences were found between the second graders and students at all other grade levels for all three tasks. The one exception was that fourth graders were not significantly different from second graders on the listening comprehension task.

Table 2-5. *Games-Howell Multiple Comparisons of Differences in Means by Task and Grade Level*

<table>
<thead>
<tr>
<th>Task</th>
<th>Grade</th>
<th>Grade</th>
<th>Mean Diff</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(A)</td>
<td>(B)</td>
<td>(B-A)</td>
<td>p</td>
</tr>
<tr>
<td>FD</td>
<td>2</td>
<td>4</td>
<td>1.64</td>
<td>.006**</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>2.58</td>
<td>.001**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2.71</td>
<td>.001**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.94</td>
<td>.154</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.08</td>
<td>.071</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.14</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>YNM-L</td>
<td>2</td>
<td>4</td>
<td>1.82</td>
<td>.058</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>3.27</td>
<td>.001**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>3.41</td>
<td>.001**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.44</td>
<td>.130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.59</td>
<td>.109</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.14</td>
<td>.991</td>
<td></td>
</tr>
<tr>
<td>NWS</td>
<td>2</td>
<td>4</td>
<td>2.92</td>
<td>.001**</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>4.62</td>
<td>.001**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>4.80</td>
<td>.001**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.70</td>
<td>.068</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1.88</td>
<td>.044*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.18</td>
<td>.979</td>
<td></td>
</tr>
</tbody>
</table>

FD = Following Direction task, YNM-L = listening comprehension task, NWS = nonword spelling; *p ≤ .05  ** p ≤ .01

These grade-level results support the sensitivity of the three tasks to ongoing skill maturation; however, the post hoc tests revealed that an apparent asymptote was reached
between the fourth and sixth grade on all of the tasks. That is, the second graders scored significantly lower than the sixth and ninth graders but not the fourth graders. Fourth, sixth and ninth graders were not significantly different from each other—with the exception of fourth and ninth graders' performance on the nonword spelling task, which did show a significant increase from fourth to ninth grade.

**Relationship of Performance on Three Subtests**

Theoretical questions about relationships among the three tasks were examined using correlation analysis. Kendall’s Tau was selected as a nonparametric measure of correlations because task distributions were not normal. Significant moderate (.30-.49) correlations were found among the three tasks as shown in Table 2-6.

| Table 2-6. *Kendall’s Tau Correlations Between the Three Tasks* |
|---------------------|---------------------|
| Grade               | Following Directions| Listening Comprehension |
| Nonword Spelling    | .336*               | .324*               |
| Listening Comprehension | .316*               |

* Correlation is significant at the 0.001 level (2-tailed).

**Discussion**

The following directions task, listening comprehension and nonword spelling tasks are being developed for possible inclusion in a larger language assessment tool. The psychometric properties of the tests are discussed, as are the theoretical implications for the integrated nature of language assessment tasks with auditory presentation of stimuli.
Psychometric Properties

Analysis of validity and reliability are important steps for developing a measurement tool. As a primary requirement, an assessment task must be scorable by different examiners with a high degree of reliability. The interscorer reliability for each of these three tasks was well above 95%, which is considered a high level of agreement.

Test-retest reliability reflects the tool’s ability to provide consistent results when the same participants are tested at two points in time (AERA/APA/NCME, 1999; Field, 2005). Both the following directions and nonword spelling tasks had statistically significant test-retest correlations above .60, which are considered large (above .50; Field, 2005). The listening comprehension task underwent significant revision between the two administrations after the item analysis results for the data collected in the fall revealed problems with the items (Brennan, 2006; Nelson, Plante, Brennan, Anderson, & Johnson, 2005). Thus test-retest reliability was not able to be analyzed for the listening comprehension task in this study.

Item analyses for the three tasks (conducted as part of the larger study) indicated a need for both harder and easier items to avoid ceiling and floor effects for the following directions task, and for additional easy items to avoid a floor effect for the nonword spelling task. These findings were used to modify the items for inclusion in the next stage of test development and validation. Multivariate analysis was conducted to determine if bias based on unintended variables was present in the task design (sex, race/ethnicity, urban/rural, or SES). No bias was found based on sex or race/ethnicity. Nonword spelling, however, was found to vary significantly with the urban or rural school district
variable (i.e., students in the urban district had more difficulty with the nonword spelling task). This could be the result of confounding with SES based on the finding that 80% (103/129) of the urban students qualified for free or reduced lunch compared with only 50% of the participants from the rural district (41/82). Alternatively, if the two variables were completely confounded, the results also should have been similar for the other two tasks. Other systematic differences might exist between the urban and rural groups that could explain this finding related to spelling, but not the other two. Race was completely confounded with geography in that all of the nonwhite students were in the urban district, yet the nonword spelling task did not differ by race/ethnicity.

SES (defined as receiving free and reduced lunch or not) was found to be a statistically significant variable in differentiating performance of students. Factors such as these have to be considered when designing tests that are meant to diagnose language disorder, not socio-economic difference. Prior studies have shown that children from low socioeconomic groups routinely perform more poorly on measures of expressive and receptive vocabulary (Engel, Santos, & Gathercole, 2008). Others studies show support that the social environment can be instrumental in vocabulary acquisition (Hoff, 2003; Hoff & Tian, 2005). In fact, children with low SES tend to score lower on virtually all standardized measures of language and school achievement. Performance has been shown to be less influenced by SES on nonword repetition tasks which are less dependent on acquired knowledge (e.g., Campbell, Dollaghan, Needleman, & Janosky, 1997). In this study, children in the low SES group scored significantly lower on all of the measures. Culling items that DIF showed to be particularly sensitive to SES may improve the performance of these subtests in this area, but designing formal assessments of
curriculum-related language skills that are not sensitive to SES remains a major 
challenge.

Theoretical Implications

The three tasks used in this study all were created to mirror authentic classroom 
demands. They also were being investigated in the hope that analysis might shed light on 
the relationship between traditional language assessment tasks and the contributions of 
verbal memory needed for successful completion. Children with verbal memory deficits 
might be identified for whom interaction with a particular language component skill (e.g., 
phonemic, graphemic, syntactic, semantic) has yet to be mastered, but would not be 
recognized as problematic until the integrated demands of a complex task make it 
apparent.

The reasoning behind correlational analysis of the current study was that if the 
tests were highly correlated (> .70), they would more likely reflect a primary influence of 
verbal memory as a common factor. On the other hand, if the tasks were only moderately 
correlated, or not correlated at all, then this would provide evidence that the tasks were 
measuring distinctly different factors, with only a portion of their variance influenced by 
verbal memory processing (or some other unidentified common factor). For the purposes 
of comprehensive assessment of curriculum-relevant language and literacy abilities, it is 
desirable for the three tasks to be measuring different constructs. There is no reason to 
put a student through multiple subtests if they are all measuring essentially the same 
thing. The moderate correlation values found in the current study point to a reasonable 
conclusion that these three subtests are measuring different constructs, justifying their
inclusion as separate subtests in a comprehensive test of integrated language skills.

Caution should be used in interpreting these results, however, as the lack of variation in scores in all three tasks among those in the upper grades could reduce or underrepresent the degree of association.

The integrated nature of these three tasks—requiring both specific language skills and short-term and working memory—makes parsing of influences in performance challenging. The goal to design assessment tasks that were more authentic meant that students were required to coordinate multiple component skills, including auditory processing, phonological awareness, phonological short-term memory, verbal working memory, vocabulary knowledge, syntactic knowledge, discourse knowledge, graphic and orthographic response capabilities, as well as verbal memory. It was impossible with the current study design, however, to tease apart the specific influences of multiple factors in influencing student performance within the three tasks. With the inclusion of relatively language-free measures of short-term and working memory skills, it may be possible in future studies to derive the amount of shared variance each contributes to these integrated tasks.

The following directions task was originally included to assist in identifying children who were experiencing particular difficulty with verbal working memory. The results of this study are interpreted as evidence that this task also requires integration of language and memory skills. The integrated nature of the task, although more authentic, made it difficult to discern how much language or memory demands (or both) influenced performance. What is needed to answer the question of roles for language and verbal memory skills in integrated tasks are relatively language-free measure of memory (such
as digit span forward and digit span reversed) as an independent measure, and a relatively memory-free measure of language such as vocabulary knowledge. Such comparisons are planned in future research.

Developmental influences were apparent for all three tasks. Results of one-way ANOVA, which showed a main effect for grade, appeared to support findings of prior research that have shown that working memory is correlated with developmental skills, such as reading skills in younger children (Bayliss, Jarrold, Baddeley & Leigh, 2005; Gathercole & Pickering, 2000). The lower performance of the second graders on the following directions task that was intended to be more heavily dependent upon verbal working memory was not too surprising given that spontaneous use of rehearsal techniques—a strategy for improving verbal memory skills—does not emerge until approximately seven years of age (Gathercole, 1998).

Limitations and Future Research Directions

This was a pilot study intended to gather preliminary data on three newly developed tasks to assess three types of curriculum-related language processing. Several limitations in the design of the study, however, should be noted. A primary limitation to this study was the lack of independent measures of verbal memory that placed relatively low demands on language processing. Without this measure, it was not possible to accurately assess the influence of verbal memory skills on performance of different levels (sound/word and sentence/discourse) of language processing tasks. Another problem was that the range of difficulty of the items may have been insufficient to assess abilities adequately at the upper end of the age scale.
This study lacked sufficient numbers of children with confirmed language disorders to evaluate relationships of these children’s performance to children with typically developing skills. This information will be necessary to further validate the sensitivity and specificity characteristics of the three tasks.

Another limitation was introduced by the group test administration format. Although it mirrored classroom direction-giving contexts, group administration made it difficult to control for certain problems, such as students’ copying the responses of peers, or beginning a response before a set of directions was finished. Also, on the following directions task, the students were allowed to see the response items while listening to the verbal directions, so it is possible that visual rehearsal was used to aid verbal working memory. The group data collection method also made it impossible to identify errors of temporal sequencing.

It was identified that these pilot tasks require further revision to include both easier and higher items to reduce possible floor and ceiling effects. This would allow for further examination of how the relationships of performance on these tasks change with maturation.

Questions remain with regard to the nature of the relationship between verbal memory and language that could not be answered with this study. That verbal memory and language processing are related is generally accepted, but to what degree they are associated within these particular language tasks has not been answered. In particular, the potential for bidirectional cause and effect needs clarification, but that would require a longitudinal research design. Improved neuroimaging techniques and collaborative efforts among scientists from psychology, neuroscience, linguistics, and speech-language
pathology also should combine to yield a better understanding of group and individual differences in language and information processing.

Identifying causal relationships would be required to inform selection of interventions for children with difficulty performing in the classroom. It is hoped that further research can aid in the development of practical classroom screening tools to identify deficits in component skills and integrated abilities so as to promote appropriate intervention contexts, materials, and methods for students to improve the opportunity for academic success.

Conclusions and Future Research Directions

The findings from this study led to specific recommendations for modifying the tasks. The following directions task and nonword spelling task will be changed for based on item response analysis to improve the hierarchical structuring. Administration procedures also will be modified to reduce ceiling effects. Administration changes should include eliminating repetition of auditory stimuli by the examiner and preventing visual or motoric rehearsal by covering up the visual stimuli and only allowing access to the student response form after completion of the verbal stimulus. It was hoped that increases in relational unit complexity and increases in foil items, as well as changes in the administration procedure, by covering the items during the directions, ceiling effects will be reduced. In the next phase of testing, individual administration will replace the group administration used in piloting. The changes in items and administration and inability to look onto a classmate’s paper are expected to place more stringent demands on auditory comprehension and retention of the complexity components.
In addition, future research that incorporates digit span forward and backwards tasks and nonword repetition may help to sort out influences of both attentional skills and working memory on student performance across tasks. Inclusion of a story-retelling task and two individually administered phonological knowledge tasks also could allow for analysis of relationships with components of the following directions task. Each of these tasks is thought to be highly dependent on a different aspect of verbal memory, and also activation of long-term memory (in the case of the story retelling task). Ideally, a delayed component of story-recall such as that found in the Pediatric Test of Brain Injury (Hotz, Helm-Estabrooks, Nelson, & Plante, 2010) could be incorporated, as there is evidence that some individuals can perform within normal ranges on immediate recall, but show marked decreases in performance after a filled-delay. This would have significant impact on a child’s ability to perform successfully in an academic environment and would require specific interventions.

Given the high-stakes nature of the educational environment, the creation of ecologically valid screening tools that can be administered quickly—perhaps to groups of children—by clinicians or teachers for identifying students at-risk for academic failure would be welcomed. Such tools would be useful as part of programs measuring students’ response to intervention for identifying students with special needs and for preventing longer term difficulties (e.g., Troia, 2005). A continuing challenge is to provide tools that are not only sensitive and specific in identifying children who may need help, but tools that can provide more specific profiling, so that the type of intervention needed is apparent.
This study begins to shed light on how language assessment tasks that are more authentic to the classroom are likely to require integration of more than one skill (such as verbal memory and language). Therefore, caution is required in interpretation to determine what role the latent skill—in this case verbal memory—has in influencing the presence of a language impairment. Further studies need to address the relationship between verbal memory and language. To what degree are the skills associated and is it possible to determine causality for children with language impairment and/or verbal memory impairments?

References


CHAPTER III

STUDY TWO: INVESTIGATION OF THREE INTEGRATED TASKS OF LANGUAGE AND MEMORY

This study is a follow-up investigation to a descriptive correlational study (study one in this dissertation) of performance on pilot versions of three curriculum-related language assessment tasks—nonword spelling, listening comprehension, and following directions—that were thought to place varying demands on language processing and memory skills. For the purposes of this paper, the term memory is limited to aspects of verbal memory, as this study does not include examination of the role of visual memory. The memory construct under investigation is based on measurements of both short-term verbal memory and verbal working memory. Short-term memory assessments focus on capacity (number of items immediately recalled), whereas working memory assessments incorporate an additional processing feature (multi-step requirement). In this study, the verbal units used were the names of numbers (digits).

Results from study one (a correlational study) were inconclusive in parsing the influence of memory on performance of each of the three language assessment tasks because verbal working memory with low language demands was not included in that study as an independent measure. Each of the three language-assessment tasks being investigated was designed to measure a specific curriculum-related language skill at either the sound/word level (nonword spelling) or sentence/discourse level (listening
comprehension and following directions). Although discrete language abilities may underlie tasks, each task requires integration of more than one skill, as is often required in authentic classroom activities.

Each task requires not only language specific skills, but also memory skills, for successful completion. Thus, if a child performs poorly on an integrated task, it might be difficult to determine which core skill—language or memory—is most associated with any difficulty or strengths in performance. The primary focus of the current study was to determine the association of memory demands with performance on: (1) a complex word-level task of nonword spelling, (2) a sentence/discourse level task that measures listening comprehension for content language, and (3) a sentence/discourse level task that measures listening comprehension for the procedural language of directions. The results should have implications about areas to target for intervention.

As noted previously, because study one did not include an independent measure of memory, it was impossible to answer the important theoretical question regarding the role of memory in language assessment tasks. Such a measure, which would place a relatively low demand on language processing, could be used as a predictor variable for performances on a nonword spelling task, a listening comprehension task, and a following directions task. Thus, the current study added a memory variable comprising scores from digit span forward (short-term memory) and digit span reversed (working memory). The design also incorporated a vocabulary awareness composite score that could serve as a language measure that would be relatively independent of memory demands. By adding these two measures to the research model, the goal was to gain
insight into the degree to which poor student performance on a language assessment task is due to a language-specific difficulty or to a memory factor.

Other methodological limitations in study one were addressed in the current study by modifying administration procedures, participant sampling procedures, and item construction. The group administration procedures from the pilot version created difficulty in drawing firm conclusions, as did the lack of sufficient numbers of children with confirmed language disorders to study unique factors in their response to the tasks. Items in the following directions task were increased in complexity to address the ceiling effect that limited the variance in performance by older children.

A better understanding of the relationship of language and memory in language assessment tasks would address questions about the influence of a more general processing ability such as memory on language performance and its possible central role in language disorders. Previous researchers have debated alternative positions about the role of memory in language disorders. Some have emphasized evidence showing that poorer memory skills are concomitant with language impairments (Cowan, 1996; Hoffman & Gillam, 2004, Nation, Adams, Bowyer-Crane, Snowling, 1999); whereas others have emphasized a role for memory impairments as contributors to poorer language skills, with a possible causal influence (Archibald & Gathercole, 2006; Leonard et al., 2007; Montgomery, 2002). Some theoretical accounts assign a central role to deficits of working memory or information processing capacity to explain specific language impairment (Archibald & Gathercole, 2006; Bishop, 1997; Gathercole & Baddeley 1990; Johnston, 1999; Leonard et al., 2007; Montgomery, 2002). Other research provides evidence that memory and language skills may be concomitant or

A practical issue arises when a child performs poorly on integrated tasks commonly found in language assessments. Identifying the source of the problem can be difficult. Without knowing the source, selecting effective interventions can be challenging. This is not to suggest that integrated tasks should be avoided in assessment. Integrated tasks represent authenticity and relationship to the curricular contexts that present some of the greatest challenges to school-age children with language disorders. For some children, it may be at the point of integration where problems arise (Nelson, 2010). However, it may be necessary to examine the integrated assessment task via discrete skills to better identify the relationships of the component skills. Recognizing the different types of processing required in an integrated task should help clinicians understand the multiple sources that may influence performance. This understanding can lead to interventions that address the discrete skills as needed for remediation, but then scaffold for adequate performance in authentic integrated tasks.

**Purpose of the Current Research Study**

The purpose of this study was to explore the relationship of verbal memory and language in three language subtests developed for inclusion in a comprehensive formal test of language and literacy skills in school-age children. Determining the role of language and memory in language assessment tasks should assist clinicians to interpret results of assessment activities and tailor intervention targets and strategies. Questions of interest included:
1) How much of the variance in three language tasks (nonword spelling, listening comprehension, and following directions) can be accounted for by vocabulary awareness and memory scores (digit span forward and reversed) when age is controlled?

2) Does the amount of variance that can be accounted for by age, vocabulary awareness and verbal memory differ between students with typical language (TL) and language impairment (LI)?

Methods

Participants

The data for this study came from a larger data set. The differential characteristic for inclusion in the two groups compared in the current study was language ability—typical language (TL) and language impairment (LI). Participants in this study included students with LI only if they met all of the inclusion criteria described in Table 3-1, including scoring < 85 on the core language subtests of the Clinical Evaluation of Language Fundamentals-Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2003). For the larger study, some students were identified as having language-learning disabilities who were receiving language or literacy intervention services but who scored ≥ 85 on the CELF-4. Their data were not included in the current investigation.

For inclusion in the TL group, participants needed to earn a standardized language test score ≥ 85 on the four core subtests (for their age) of the CELF-4 and to have passed a recent hearing screening. Additionally, teacher and parent report of adequate language/literacy development was documented with a parent checklist (confirmed by the
language clinician based on school records) that indicated that the child had never been identified with a speech-language problem or any other special education diagnosis (ADHD was allowed: \( n = 4 \) in the TL group, and \( n = 7 \) in the LI group).

Table 3-1. Criteria for Group Determination

<table>
<thead>
<tr>
<th>TL</th>
<th>LI</th>
</tr>
</thead>
<tbody>
<tr>
<td>• No deficits in speaking, reading, writing, or social skills</td>
<td>• Deficits in spoken or written language</td>
</tr>
<tr>
<td>• ( \geq 85 ) standard score on CELF-4 core composite</td>
<td>• (&lt; 85 ) standard score on CELF-4 core composite</td>
</tr>
<tr>
<td>• No special education diagnosis</td>
<td>• No other special education diagnosis</td>
</tr>
<tr>
<td>• Passed hearing screening</td>
<td>• Passed hearing screening</td>
</tr>
</tbody>
</table>

Inclusion criteria for the LI group were (a) a diagnosis of LI or LD by interdisciplinary team, confirmed by clinical judgment, (b) language test score on CELF-4 core 4 subtests \(<85\), and (c) teacher report supporting concerns about language/literacy development. Students were excluded if they had a diagnosis of intellectual disability, an identified genetic syndrome, emotional impairment/behavioral disorder, autism spectrum disorder, overt neurological or seizure disorder, or blindness/low vision.

Data for this study came from a larger data set that included students who did not meet all criteria for LI outlined in Table 3-2. That is, all of the students with LI selected for the current study scored more than one standard deviations below the mean on the core composite score for a comprehensive test of spoken language and exhibited no other disability (except, in some cases, ADHD). Therefore, this sample population could be considered comparable to children labeled in prior research as having specific language impairment (SLI).
Table 3-2. Demographics of Typical Language and Language Impaired Groups

<table>
<thead>
<tr>
<th></th>
<th>Typical Language Group</th>
<th>Language Impaired Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>9 (24)</td>
<td>9 (24)</td>
</tr>
<tr>
<td>3-5</td>
<td>12 (32)</td>
<td>12 (32)</td>
</tr>
<tr>
<td>6-8</td>
<td>11 (30)</td>
<td>11 (30)</td>
</tr>
<tr>
<td>9-12</td>
<td>5 (14)</td>
<td>5 (14)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>18 (49)</td>
<td>18 (49)</td>
</tr>
<tr>
<td>Female</td>
<td>19 (51)</td>
<td>19 (51)</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>18 (49)</td>
<td>17 (46)</td>
</tr>
<tr>
<td>African-American</td>
<td>5 (13)</td>
<td>7 (19)</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>14 (38)</td>
<td>13 (35)</td>
</tr>
<tr>
<td>Free/Reduced Lunch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>18 (49)</td>
<td>14 (38)</td>
</tr>
<tr>
<td>No</td>
<td>13 (35)</td>
<td>16 (43)</td>
</tr>
<tr>
<td>Missing Data</td>
<td>6 (16)</td>
<td>7 (19)</td>
</tr>
</tbody>
</table>

Participants were between the ages of 6 and 19 years of age. They were recruited from the Midwest, Northeast, Southwest and West regions of the United States. Stratified sampling was used to create two groups that were comparable in terms of chronological age (within 9 months), sex, race/ethnicity, and socio-economic status, but differed in the area of language status based on the CELF-4 (TL versus LI; see Table 3-2).

General Testing Procedures

Informed parental permission and student assent were obtained according to a protocol approved by a Human Subjects Institutional Review Board. Each student participated in two assessments: the CELF-4 core subtests and the Test of Integrated Language and Literacy Skills (TILLS, beta research version; Nelson, Helm-Estabrooks, Hotz, & Plante, 2007). Administration was completed during one to three individual
sessions with a licensed or certified speech-language pathologist. Subtest order was randomized to control for possible order effects and as part of the pre-standardization trial of the full set of 11 TILLS subtests.

Measures

Scores from multiple subtests from the TILLS provided the data for the study. All subtests used in this study met the Cronbach’s alpha threshold of .80 or higher and interscorer reliability of 95% or higher.

*(What Goes Together; 64 items).* For this assessment of vocabulary knowledge, students hear three words while looking at the three words printed on a page (e.g., *car, water, boat*). They must select two of the words and tell how they “go together” by applying analogical, categorical, and relational concepts to thinking and talking about words. Then the words are read again and the student is asked to pick two different words and explain how they go together in a different way. The printed words remain in view as students complete the task to reduce the load on their working memory. They must hold the two selected items in mind while formulating an explanation of how the words go together, however, and must remember what they said previously to offer a different pair and explanation for the second set. Students receive a score of 0, 1, or 2 depending on the adequacy of their explanations about how two different pairs of words go together, each in a different way. Thus, the maximum score that could be earned on this measure was 128 points.

*Social vocabulary and message formulation (Acting a Scene; 19 items).* This task assesses knowledge of social-emotional vocabulary and pragmatic skills. Students are
given a verbal description of a short scene based on a specific vocabulary word and then
are asked to say what the character might say and to demonstrate how the character
would say it. An example is the following:

Sample: Jasmine always *whines* when her parents won't let her have her way. Her mother won't let her buy candy in the grocery store. What do you think Jasmine would say?

Students' responses are scored separately for vocabulary awareness (of the highlighted verb phrase), syntactic formulation, and pragmatic fit to the situation. Only the vocabulary score was used in the current study. Students earned a vocabulary score of 2 for responses that demonstrated immediate understanding of the key verb phrase (e.g., *argues*, *brags*), some of which were idiomatic (e.g., *fishes for compliments*); they earned a score of 1 if they could demonstrate understanding of the term after hearing a definition if their first attempt showed lack of understanding; and they earned a score of 0 if they could not demonstrate understanding even after hearing the definition. Thus, the maximum score that could be earned on this measure was 38 points.

*Following directions task* (18 items). For the FD task, a student response form is placed in front of the participant. A cover sheet prevents the participant from previewing the linear array of response items for each item. Once the verbal direction is given, the cover sheet is removed from the item and the participant completes the task using a pencil to carry out the instructions. The examiner records information such as temporal sequencing and directional components of the student's response in the examiner manual. Scoring is dichotomous with a 1 for correct and a 0 if any part of the answer is wrong, thus maximum possible for this task is 18.
Nonword spelling (18 items). Stimuli are presented using a CD recording. The directions for this subtest are as follows:

*I am going to play a voice recording for you. The person on the recording will say a pretend word and then ask you to say it. You will only hear the word one time before you try to say it. Remember, this is not a real word, so you need to listen carefully to say it just like the person you hear. Let’s practice.* [Activate the digital player.] *Now, the person will say the word one more time, and you spell it here.* [Point to Student Book.] *Use what you know about spelling real words to help you spell these nonsense words.*

Sample: Nonsense Words—Imitation and Nonsense Words—Spelling
*Say bup. _____ Spell bup.*

The participant is asked to spell the nonword in a student response booklet. Scoring is dichotomous, with 1 point for correct responses (allowing for correct alternatives that could appear in real words) and 0 for words with any unacceptable variations for a maximum of 18 points.

Digit span forward (6 items). Participants listen to a span of digits presented auditorally by the clinician at a rate of one per second. The participants must then say the numbers in the same order. Two trials are presented at each level, starting with a sequence of two digits. The examiner records participant production verbatim. If both trials for a level are correct a score of 2 is given, if one is missed then the score is 1. If both items at a level are missed the score is 0, and the task is ended. Maximum possible score is 12.

Digit span reversed (6 items). This task is administered in the same manner as digit span forward; however, the numbers must be recalled in reverse order of presentation. The maximum possible score is 12.
Listening comprehension (20 items). Participants listen to several sentences read aloud by the examiner. Repetition by the examiner is not permitted. In their response book, the students have four questions to answer about the material they just heard. They respond with Yes, No or Maybe. Here is an example:

One morning Susan missed the school bus. She thought she would be late for school, but her mother got her there on time.

a. Did Susan miss the bus? Y N M
b. Was Susan late for school? Y N M
c. Was it raining that morning? Y N M
d. Did her father take her to school? Y N M

Scoring is dichotomous, either 1 for right or 0 for wrong. There are five groups of sentences, each followed by four questions, for a maximum of 20 possible points.

Composite scores. Composite scores used in the current study were calculated for vocabulary awareness and verbal memory. The vocabulary awareness composite was the total of points earned on the What Goes Together items and the vocabulary score from Acting a Scene. It was designed to capture a principal component of language ability, but with a low demand on verbal memory skills. The memory composite was the total points earned on digit span forward and reversed tasks. These two tasks were administered as measures of short-term and working memory skills with low demands on language.

Data Preparation and Analysis

Raw data were scored and verified and entered into an ACCESS database. Data were then exported into SPSS v.15.0 (SPSS, 2006) for further analysis. Variables were examined for skewness and kurtosis and found to meet the assumptions of normal distribution.
Pearson’s correlation analysis was used to consider simple relationships among the variables in pairwise comparisons. However, it was expected that some of the independent variables could be influencing each other; therefore, it would be necessary to control for those influences through the use of multiple regression analysis.

Multiple regression models can help to answer questions about more complex relationships by identifying the association of any one variable to another after controlling for all other variables. It also allows for prediction of the value of a dependent variable given values of multiple independent variables. The dependent variables of interest in this study were the three language assessment tasks: nonword spelling, listening comprehension, and following directions. The first predictor variable entered in the regression models was chronological age in months (entered as a continuous variable) as each of the dependent measures requires skills that are developmental in nature. The remaining predictor variables are a composite language score (vocabulary awareness) and a memory composite score, which was created by combining the digit span forward and digit span reversed task scores.

To answer the second research question about differences between groups, separate models were analyzed for TL and LI groups. Of particular interest was the amount of unique variance that might be accounted for by memory skills and any differences in the predictive value between the two groups (TL and LI) for these three language assessment tasks.
Results

Descriptive Statistics and Differences Between Groups

Means and standard deviations for each of the dependent outcomes and predictor variables are presented in Table 3. Independent sample t-tests (Table 4) confirmed that the performance of the TL group was statistically significantly better than performance of the LI group for all tasks, \( p < .0005 \).

Table 3-3. Descriptive Statistics for Measures as a Function of Dyad Designation

| Measure | TL | | | | | LI | | | | | |
|---------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|         | n  | Min| Max| M  | SD | n  | Min| Max| M  | SD | n  | Min| Max| M  | SD |
| NWS     | 37 | 0  | 17 | 10.92 | 4.13 | 37 | 0  | 14 | 5.30 | 4.12 |
| LC      | 37 | 8  | 19 | 13.54 | 2.56 | 37 | 5  | 14 | 10.73 | 2.24 |
| FD      | 37 | 3  | 17 | 11.59 | 3.20 | 37 | 2  | 14 | 7.46 | 2.87 |
| VA      | 37 | 47 | 100| 7.54  | 14.01| 37 | 10 | 85 | 50.92| 19.98|
| DF      | 37 | 6  | 12 | 8.92  | 1.89 | 37 | 4  | 12 | 7.22 | 1.77 |
| DR      | 37 | 2  | 11 | 5.47  | 1.79 | 37 | 1  | 9  | 3.86 | 1.72 |

TL = typical language group; LI = language impairment; NWS = nonword spelling; LC = listening comprehension; FD = following directions; VA = vocabulary awareness composite; DF = digits forward; DR = digits reversed

Table 3-4. Results of t-Test Differences Between Group (TL and LI) Performance on Independent and Dependent Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>t</th>
<th>p-value</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWS</td>
<td>5.86</td>
<td>.0001</td>
<td>1.38</td>
</tr>
<tr>
<td>LC</td>
<td>5.03</td>
<td>.0001</td>
<td>1.19</td>
</tr>
<tr>
<td>FD</td>
<td>5.85</td>
<td>.0001</td>
<td>1.38</td>
</tr>
<tr>
<td>VA</td>
<td>6.64</td>
<td>.0001</td>
<td>1.65</td>
</tr>
<tr>
<td>DF</td>
<td>4.00</td>
<td>.0001</td>
<td>.94</td>
</tr>
<tr>
<td>DR</td>
<td>3.91</td>
<td>.0001</td>
<td>.92</td>
</tr>
</tbody>
</table>

TL = typical language group; LI = language impairment group; NWS = nonword spelling; LC = listening comprehension; FD = following directions; VA = vocabulary awareness composite; DF = digits forward; DR = digits reversed
Variance in Language Task Performance Accounted for by Age, Vocabulary Awareness, and Memory

Pearson’s correlation analyses were used first to evaluate the linear relationships between the scores of each of the three dependent language measurements (nonword spelling, listening comprehension, and following directions), age, and the two composite score variables (vocabulary awareness and memory). Table 3-5 shows the results for each of the sample groups.

Table 3-5. Correlations Between Measures for Language and Memory Tasks

<table>
<thead>
<tr>
<th></th>
<th>NWS</th>
<th>LC</th>
<th>FD</th>
<th>AGE</th>
<th>VA</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NWS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC</td>
<td>.475**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FD</td>
<td>.327*</td>
<td>.282</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>.422**</td>
<td>.123</td>
<td>.241</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VA</td>
<td>.687**</td>
<td>.597**</td>
<td>.369*</td>
<td>.621**</td>
<td></td>
</tr>
<tr>
<td>MEM</td>
<td>.564**</td>
<td>.408*</td>
<td>.524**</td>
<td>.371*</td>
<td>.482**</td>
</tr>
<tr>
<td>LI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NWS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC</td>
<td>.301</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FD</td>
<td>.073</td>
<td>.296</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>.583**</td>
<td>.505**</td>
<td>.500**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VA</td>
<td>.414*</td>
<td>.369*</td>
<td>.488**</td>
<td>.627**</td>
<td></td>
</tr>
<tr>
<td>MEM</td>
<td>.570**</td>
<td>.040</td>
<td>.074</td>
<td>.364*</td>
<td>.268</td>
</tr>
</tbody>
</table>

*p < .05 and **p < .01 (two-tailed). TL = typical language group; LI = language impairment group; NWS = nonword spelling; LC = listening comprehension; FD = following directions; VA = vocabulary awareness composite; MEM = digits forward + digits reversed

For the participants in the TL group both the vocabulary composite and the memory composite were associated with all language task variables at a statistically significant level. Vocabulary had the strongest relationship with the nonword spelling task and the listening comprehension task (r = .687 and .597 respectively, p < 0.01). Memory was moderately correlated with the nonword spelling task and the following
directions tasks ($r = .564$ and $.524$ respectively, $p < 0.01$). Age was significantly correlated only with nonword spelling ($r = .422, p < 0.01$).

For the participants in the LI group, age was correlated more strongly with each of the three language task than either memory or vocabulary. Vocabulary was found to be correlated significantly with all three tasks at moderate levels ($r = .414, p < .05$ with nonword spelling; $r = .369, p < .05$ with listening comprehension; and $r = .488, p < .001$ with following directions). Memory was associated at a significant level only with the nonword spelling task ($r = .570, p < 0.01$).

To further investigate the associations between memory and language abilities, a series of regression analyses were conducted to evaluate the unique variance accounted for by each of the independent measures (age, vocabulary awareness and verbal memory skills) for each of the dependent measures (nonword spelling, listening comprehension and following directions).

Age was entered into each of the models first to account for developmental effects. Next the vocabulary composite was added as a core language attribute. Lastly, the verbal memory composite was added to the model to see what unique contributions it added after accounting for the other key components that were expected to influence performance.

Whereas age showed moderate associations in the bivariate correlation for the TL group between age and performance on the nonword spelling task, when the influence of the vocabulary and memory variables was controlled, age was no longer significant. Other notable changes in the TL group are that (1) vocabulary (but no longer memory) is the only significant variable accounting for variance in listening comprehension
Table 3-6. Results of Multiple Regression Analysis

<table>
<thead>
<tr>
<th>Model by Group</th>
<th>Nonword Spelling</th>
<th>Listening Group</th>
<th>Following Directions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SEB</td>
<td>(\beta)</td>
</tr>
<tr>
<td>TL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.126</td>
<td>2.547</td>
<td>12.315</td>
</tr>
<tr>
<td>Age</td>
<td>.049</td>
<td>.018</td>
<td>.422*</td>
</tr>
<tr>
<td>(R^2 = .178 (p = .009)**)</td>
<td>(R^2 = .015 (p = .468))</td>
<td>(R^2 = .058 (p = .151))</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.779</td>
<td>2.911</td>
<td>5.574</td>
</tr>
<tr>
<td>Age</td>
<td>-.001</td>
<td>.018</td>
<td>-.008</td>
</tr>
<tr>
<td>VA</td>
<td>.204</td>
<td>.047</td>
<td>.692**</td>
</tr>
<tr>
<td>(\Delta R^2 = .294 (p = .0001)**)</td>
<td>(\Delta R^2 = .440 (p = .0001)**)</td>
<td>(\Delta R^2 = .078 (p = .088))</td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-7.269</td>
<td>2.955</td>
<td>4.589</td>
</tr>
<tr>
<td>Age</td>
<td>-.005</td>
<td>.017</td>
<td>-.094</td>
</tr>
<tr>
<td>VA</td>
<td>.167</td>
<td>.047</td>
<td>.566**</td>
</tr>
<tr>
<td>Memory</td>
<td>.414</td>
<td>.181</td>
<td>.308*</td>
</tr>
<tr>
<td>(\Delta R^2 = .072 (p = .029)*)</td>
<td>(\Delta R^2 = .029 (p = .178))</td>
<td>(\Delta R^2 = .156 (p = .011)**)</td>
<td></td>
</tr>
<tr>
<td>LI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.116</td>
<td>2.284</td>
<td>6.285</td>
</tr>
<tr>
<td>Age</td>
<td>.068</td>
<td>.016</td>
<td>.583**</td>
</tr>
<tr>
<td>(R^2 = .340 (p = .0001)**)</td>
<td>(R^2 = .255 (p = .001)**)</td>
<td>(R^2 = .250 (p = .002)**)</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.146</td>
<td>2.311</td>
<td>6.267</td>
</tr>
<tr>
<td>Age</td>
<td>.062</td>
<td>.021</td>
<td>.533**</td>
</tr>
<tr>
<td>VA</td>
<td>.016</td>
<td>.037</td>
<td>.080</td>
</tr>
<tr>
<td>(\Delta R^2 = .004 (p = .658))</td>
<td>(\Delta R^2 = .004 (p = .654))</td>
<td>(\Delta R^2 = .051 (p = .126))</td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-7.823</td>
<td>2.392</td>
<td>7.103</td>
</tr>
<tr>
<td>Age</td>
<td>.047</td>
<td>.019</td>
<td>.401*</td>
</tr>
<tr>
<td>VA</td>
<td>.011</td>
<td>.033</td>
<td>.053</td>
</tr>
<tr>
<td>Memory</td>
<td>.550</td>
<td>.179</td>
<td>.410**</td>
</tr>
<tr>
<td>(\Delta R^2 = .145 (p = .004)**)</td>
<td>(\Delta R^2 = .025 (p = .288))</td>
<td>(\Delta R^2 = .016 (p = .381))</td>
<td></td>
</tr>
</tbody>
</table>

* \(p \leq .05\) and ** \(p \leq .005\)
performance (contributing 59% to shared variance), and (2) memory (but no longer vocabulary) is the only significant variable accounting for variance in following directions performance (contributing 20% to shared variance).

For the LI participants, age remained statistically significantly correlated with nonword spelling and listening comprehension measures, but was no longer significant for the following directions task. Age was the only variable contributing significantly to the shared variance (26%) for listening comprehension performance; age (16%) and memory (17%) contributed to the shared variance of nonword spelling.

Discussion

The changes in participants, test administration, and task inclusion from study one to study two allowed for further exploration of the role of memory in language assessment tasks. The three language tasks—nonword spelling, listening comprehension, and following directions—were designed not only to evaluate skills that could be negatively affected by language impairment, but to mirror authentic classroom tasks.

Results of this study support the earlier findings from study one that the three language tasks are evaluating different constructs. Bivariate correlations of the TL group showed only mild associations between the three tasks, ranging from a statistically nonsignificant .282 between listening comprehension and following directions, to .475 between nonword spelling and listening comprehension. For the LI group, bivariate correlations showed none of the tasks to be related in a statistically significantly way.

The nonword spelling task was expected to tap into both the phonological loop of verbal memory and to be supported (to a degree) by vocabulary awareness, even though
nonwords were used for the task. As vocabulary grows, one should better recognize the orthographic representations of the phonotactic patterns of English syllables, some of which are morphemes (e.g., \textit{-ing}). The nonword items on the nonword spelling task were constructed to follow English phonotactic and orthotactic rules and to incorporate some real English morphemes (Tattersall, 2010). Results showed that the groups of TL and LI students appeared to employ different processes for this task. For the TL group, vocabulary awareness did indeed play a significant role; yet it was not even statistically significant for the LI group. This may be related to students in the LI group having weaknesses in vocabulary awareness, so that vocabulary skills could fail to support their attempts at nonword spelling. Another explanation could be that, even with adequate vocabulary knowledge, they could have difficulty calling on that knowledge in integration with other abilities in tasks that can be integrated by students with TL. The memory variable was statistically significantly related to nonword spelling performance for both groups and was the biggest contributor for the LI group.

The listening comprehension task was expected to require both vocabulary awareness and memory, along with comprehension of syntax. Surprisingly, verbal memory was not found to play a statistically significant role for either the TL or LI group. Vocabulary awareness was the variable accounting for the most variance for the TL group, whereas age was the only statistically significant variable predicting performance on the listening comprehension task for the LI group.

The following directions task was originally conceptualized as a test primarily to assess verbal working memory, albeit in an integrated form with language skills playing an important role. Results from performance of the TL group, consistent with this
expectation, showed that the memory composite was most predictive of performance on this task ($\beta = .453, p = .011$). Age and vocabulary awareness did not statistically significantly predict performance on the following directions task for students with TL.

Particularly interesting was the lack of significance of the memory composite as a predictor for the following directions performance by the LI group. It may be that children with LI who have weaknesses in vocabulary may need to devote more processing capacity in order to comprehend and follow verbal directions, interfering with rehearsal via the phonological loop, so that recall of the directions is negatively impacted. However, it also could be that weaker memory skills fail to support attempts to encode the language elements. Given the lack of a statistically significant relationship between the memory composite and the following directions task for the LI group, it may be that combinations of deficits exist. Further analysis of individual differences might reveal some students as having vocabulary awareness that is adequate for some purposes, but not strong enough to support following directions, in combination with poor memory skills; others might have inadequate vocabulary awareness and memory skills that are adequate for some purposes, but not strong enough to support integrated processing. In group analyses, these patterns could cancel each other out and obscure individual differences. When an integrated task requires multiple skill sets, skills that might appear on discrete tasks to be adequate, may be exposed as inadequate to support performance on an integrated task.

Of particular interest in this study is the observation that not only did the LI group have different predictor variables for performance of the three language tasks from the TL group, but the pattern changed from one task to another. In other words, it was not
always age or vocabulary awareness or verbal memory that was most important. This supports a conclusion that students with LI differ from students with TL in their approach to the integrated language and memory processes that are required for the language of the curriculum.

It is important to note that brief formal assessment tasks such as these are best conceptualized as screening tasks because time constraints for a comprehensive test model dictate that performance in a particular area can only be sampled. Thus, tasks, which, on the surface, appear to be measuring language skills such as spelling or verbal direction following, when analyzed further, they may place a significant demand on verbal memory skills as well. In recognizing this dual demand, professionals who interpret performance on such tasks may be well-served to probe further in determining which factor or factors may be contributing to a particular child’s difficulties in the classroom.

Strengths, Limitations, and Future Research Directions

The stratified sampling method in the study design reduced much of the potential for confounding. However, there may still have been insufficient numbers across ages to draw firm conclusions that can be generalized to the broader population. The sample was particularly small at the youngest and oldest ends of the age spectrum. The racial/ethnic and socioeconomic diversity of participants, although a strength of the study for some purposes, might also have introduced variation that could obscure patterns.

The question of association or dissociation of verbal memory and language skills deserves further study. The question should be explored whether children with different
profiles of relative strength in language and memory skills can be found that show
dissociation, with greater strength in either direction. Other research has shown that
children with SLI have a profile of weak verbal memory and language skills (Dollaghan
& Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990;
Montgomery 1995, 2000), and, by definition, typically developing children would be
expected to have both adequate verbal memory and language skills. Are there children
who present with adequate language skills, but relatively weak verbal memory skills and
vice versa? Archibald & Joanisse (2009) have presented some preliminary evidence that
such patterns do exist.

In addition, analysis of other integrated tasks, such as nonword imitation and
reading, and reading comprehension, in terms of their demands on verbal memory skills
might continue to clarify the picture regarding the association of these skills. Factor
analysis might provide further evidence for the roles played by various language skills
such as phonemic awareness and semantic knowledge versus more general processing
skills such as short-term and working memory.

Conclusions

The present study examined the relationships of language skills and memory
skills and their influences on three language assessment tasks designed to mirror
authentic classroom tasks. Children with typically developing language performed better
on all tasks than did children with language impairments as expected. However, the
variables (age, vocabulary awareness or memory) that accounted for more of the variance
in performance on each task were different between the groups and by task. One pattern
of note is that for the LI group, the vocabulary variable was never associated at a statistically significant level with task performance, but age was for the nonword spelling and listening comprehension tasks. On the other hand, for the TL group, age was never associated at a statistically significant level for the nonword spelling and listening comprehension tasks, but vocabulary was and memory contributed significantly to the nonword spelling and following directions tasks.

Interpretation of performance on language assessment tasks can be complicated by the role of latent variables such as memory. It may be important, therefore, to assess verbal memory as an overt trait as well as a latent trait. Future investigations will focus on the integrated nature of other language assessment tasks as well as individual differences of performance that can elucidate profiles of skills that may be employed by children with and without language impairment. All of these efforts are designed to better identify children at-risk for academic failure and to aid in the proper targeting of integrated component abilities in contextualized intervention activities.

References


CHAPTER IV

EVIDENCE FROM SCHOOL-AGE CHILDREN FOR ASSOCIATION AND DISSOCIATION OF MEMORY AND LANGUAGE

This theory-oriented study addressed a question about the relationships between verbal memory skills and language skills. The central question was whether verbal memory skills are necessary and sufficient for language skills.

Existing theories of language acquisition vary in the role they assign to memory. For example, the linguistic innateness theory as posited by Noam Chomsky (1957, 1980) and Stephen Pinker (1994, 2002) purports that language is innate to humans and requires minimal environmental input and little to no direct teaching. Only the parts of the brain known to specialize in language functions such as Broca's area and Wernicke's area are discussed in much detail (and they are thought to work as a module or "language acquisition device" as coined by Chomsky), with little to no mention of other brain regions used for information processing tasks such as attention and memory.

In contrast to the linguistic/nativist theorists, cognitive connectionists such as Bates and MacWhinney (1987), and Tomasello (2001) posit that language is not innate, but must be learned through reliance on cognitive information processing systems in the brain (e.g., attention, perception, and memory) combined with pattern-recognition. These are the same systems that support other forms of learning (e.g. procedural, or conceptual), not just language acquisition. Cognitive connectionists propose that general processing skills such as verbal memory are necessary for language skills.
If this supposition is true, then there are implications for both assessment and intervention. Speech-language pathologists who are primarily responsible for assessing and treating language impairments may need to rethink assessment methods and interpretation of common language assessment tasks. Many common language assessments designed primarily to evaluate a language skill (e.g., nonword repetition, sentence recall, listening comprehension, and following directions), are integrated; therefore, they are assessing the language skill and also verbal memory as a latent trait. A better understanding of the influence of verbal memory skills in combination with language skills could lead to identification of intervention strategies better suited to a particular profile of impairments.

Memory Defined

The construct of memory can be defined in a variety of ways depending upon discipline, modality, or temporal duration. For the purposes of this study, memory refers to verbal (not visual) memory. Working memory and short-term memory are of primary interest. Tasks assessing both forms of verbal memory are combined in this study to create a memory composite variable.

Baddeley (1986) and Hitch (Baddeley & Hitch, 1974) defined the construct of working memory as having two distinct features. The first is a domain specific (visual or auditory) storage component. The storage component provides workspace for the second feature, which is the manipulation or processing of information that is required for many complex tasks. Working memory is in contrast with short-term (or immediate) memory, which requires information to be stored passively and then reproduced without need for
significant processing or transformation of the stimuli. An example of a short-term memory task would be serial recall of letters or numbers, such as repeating a phone number. A working memory task would include multiple step activities such as hearing a series of numbers, and mentally manipulating them in order to repeat them in reverse sequence, or mentally calculating the amount, for example, to leave for a tip at a restaurant.

Memory and Specific Language Impairment

Specific language impairment (SLI) is a diagnostic term that is used to describe children who score below age level on tests of language, demonstrate normal range nonverbal IQ and intact hearing, have no other comorbid conditions, but fail to develop language as expected (Leonard, 1998). Despite being viewed as having basically intact cognition, children with SLI have been found to exhibit cognitive processing limitations, including deficits in working memory (Archibald & Gathercole, 2006; Ellis Weismer, Evans, & Hesketh, 1999; Gathercole & Baddeley, 1990; Marton & Schwartz, 2003; Montogomery, 1995, 2000a), attention (Cowan, Nugent, Elliott, Ponomarev, & Saults, 1999; Martinussen & Tannock, 2006), and visual and visual-spatial processing (Hick, Botting, & Conti-Ramsden, 2005). Some even posit deficits in phonological short-term memory—a component of verbal working memory—as one of the core deficits underlying SLI (Conti-Ramsden & Durkin, 2007).

In addition to the capacity component of working memory, prior research points to children with SLI as having slower speed of processing than their typical language (TL) peers on both verbal and nonverbal tasks (Hayiou-Thomas, Bishop, & Plunkett,
2004; Hoffman & Gillam, 2004; Kail, 1994; Lahey, Edwards, & Munson, 2001; Leonard et al., 2007; Miller, Kail, Leonard, & Tomblin, 2001; Montgomery, 2000b; Montgomery & Windsor, 2007). Hayiou-Thomas, Bishop, and Plunkett (2004) found that typically developing children with intact linguistic systems can be induced to produce SLI-like patterns of performance when cognitive processing demands are increased by manipulating two variables independently. They placed greater demands on speed of processing by compressing speech signals, and they increased demands on memory load by adding redundant verbiage to lengthen sentences. Then they measured the children’s accuracy in judging the grammatical correctness of sentences. Each condition of processing stress caused children who could achieve near-perfect performance under unstressed conditions to exhibit patterns of performance that were similar to children with SLI. This provided support for the theory that, in some children, a profile of language difficulty characteristic of SLI may result from an underlying generalized processing deficit—such as verbal memory—that is not specific to language.

It also has been shown that children with SLI are not able to retain and repeat the same number of spoken language components as their typically developing peers (Dollaghan & Campbell, 1998; Ellis Weismer et al., 1999; Gathercole & Baddeley, 1990; Hoffman & Gillam, 2004; Montgomery, 2004). Phonological short-term memory (PSTM), which is typically measured with nonword repetition tasks, has been strongly associated with intact linguistic abilities in typically developing (TD) children (Baddeley, 2003; Gathercole & Baddeley, 1989, 1990). To complete a nonword repetition task successfully, one must accurately coordinate multiple phonological and memory processes. The perception of the phonemes must be accurate, then encoded and stored for
later retrieval and production. Using nonwords in spoken language repetition tasks reduces the ability to rely on learned vocabulary, making the task a more sensitive and less biased assessment of PSTM.

Previous studies have shown an association between a variety of memory measures and language skills such as weak phonemic short-term memory skills and nonword repetition (Archibald & Gathercole, 2006; Archibald & Joanisse, 2009; Bishop, North, & Donlan, 1996; Botting & Conti-Ramsden, 2001; Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Gathercole & Baddeley, 1990; Gillam & Van Kleek, 1996; Kamhi & Catts, 1986), impaired working memory and sentence repetition and comprehension (Archibald & Joanisse, 2009; Montgomery 1995, 2000a, 2006, 2009; Riches, Loucas, Baird, Charman & Simonoff, 2010), and impaired verbal short-term memory and working memory in specific language impairment (Archibald & Gathercole, 2006; Montgomery, 2000b). Researchers have found evidence that memory is an important skill for language acquisition and learning (Baddeley, Papagno, & Vallar, 1988; Daneman & Merikle, 1996; Gathercole, 2006; Gathercole & Baddeley, 1989; Papagno, Valentine & Baddeley, 1991; Papagno & Vallar, 1992), but questions remain about the relationships of language and memory, and particularly any causal relationships, across childhood, especially during the school-age years.

Questions about the association between memory and language impairment have implications for informing both theory and practice. A fundamental question when working with children with SLI is whether deficits are specific to language, or if memory skills are not only associated with, but contributing to the difficulty. Clinicians can benefit from increased understanding of the interaction of memory with language ability,
which may influence how students process language in the context of classroom activities that require inter-modality complex processing (e.g., such as note-taking, solving math word problems, or following verbal directions to conduct a science lab experiment).

Causal relationships between cognitive limitations (e.g., memory) and language performance are difficult to discern, however, because impairment in one domain could cause impairment in the other—in either direction (Conti-Rasmden & Durkin, 2007; Leonard et al., 2007). Limitations in memory could fail to support language development adequately. An alternative explanation to be considered, however, is the reciprocal case, in which language knowledge deficits might provide insufficient support for performance of tasks requiring memory (e.g., nonword repetition, Gillam & Van Kleek, 1996).

Necessary and Sufficient

This study was designed to identify patterns of association or dissociation between verbal short-term memory and verbal working memory skills with language skills in children with typically developing language and those with language impairment. These patterns can be utilized as support for one of four hypotheses which address the questions of whether memory seems to be a necessary and sufficient skill for supporting language.

The term necessary and sufficient is borrowed from the field of philosophy. In determining causal connections, one must determine the conditions that are necessary and sufficient for an event to occur. A necessary condition is one that if it were absent, the event cannot occur (Copi, 1961). For example, water is necessary for crops to grow, for without it, the crops will not survive. However, water is not a sufficient condition for
crops to grow. A sufficient condition for an event is “a circumstance in whose presence the event must occur” (Copi, p. 355). In the example of the crops, multiple necessary conditions are required for growth (e.g., appropriate range of soil nutrients, temperature, water, and light in combination with absence of disease and pests), and if all these conditions are present, then growth will occur. They must all be included in the sufficient condition.

The plausible hypotheses explored are (1) complete association between memory and language skills, (2) no association (the null hypothesis, or complete dissociation) between memory and language skills, (3) memory skills are necessary, but not sufficient for language skills, or (4) language skills are necessary, but not sufficient for memory skills.

Purpose

The purpose of this investigation was to explore the relationship between performance on low-language load verbal memory tasks (short-term and working memory) and low-memory load language tasks (vocabulary awareness) of two groups of similar children differing primarily by the presence or absence of language impairment. In a sample in which 50% of the children have typical language (TL) and 50% have language impairment (LI), it should be possible to identify patterns of relationships that would support one of the above four hypotheses about the relationship of memory and language. The four hypotheses are elaborated and illustrated with hypothetical figures as follows:
Hypothesis 1. Complete Association

According to this hypothesis, participants should be split equally into two columns representing memory and language skills that are either both high (above the cut-off scores) or both low (below the cut-off scores) depending on whether they were members of the LI or TL group based on a "gold standard" measure (in this case the CELF-4 core language score). This hypothesis posits a situation in which language is completely associated with memory, so that it is impossible to have high language if one has low memory, and impossible to have low language if one has high memory. Figure 4-1 illustrates how the data would look.

Figure 4-1. Data Distribution Predicted by Hypothesis 1

Hypothesis 2. No Association

According to this Hypothesis 2, participants should be equally split into four columns representing high-memory/low-language, high memory/high language, low
memory/low language, and low-memory/high language. In this situation, language and memory would be completely dissociated, so that performance on one would not be indicative of performance on another. The hypothesis of dissociation would represent the null hypothesis, and the data would disperse in the pattern seen in Figure 4-2.

![Figure 4-2. Data Distribution Predicted by Hypothesis 2](image)

**Hypothesis 3: Memory is Necessary, but not Sufficient for Language**

Hypothesis 3 suggests that memory skills are necessary for language skills (some association), but are not sufficient (some dissociation). In this situation it is likely that a majority of the participants would be in the columns for high memory/high language and low memory/low language, but a substantial minority would be in the column for high memory/low language and very few to none should be in the column for low memory/high language. Figure 4-3 illustrates the data profile for Hypothesis 3.
Hypothesis 4: Language is Necessary, but not Sufficient for Memory

Hypothesis 4 represents the reciprocal of Hypothesis 3 positing that language skills are necessary, but not sufficient for memory skills. Again a majority of the participants should be found in the high memory/high language and low memory/low language columns, but this time the substantial minority would be found in the low memory/high language column and very few to none in the high memory/low language column. Figure 4-4 illustrates the expected data profile for Hypothesis 4.

Research Questions:

1) Do participants with TL and LI distribute differently in each of four language-memory quadrants (high-memory/high-language, low-memory/high-language, low-memory/low-language, and high-memory/low-language) based on performance on TILLS memory and language tasks, indicating the presence of patterns of both association and dissociation?
2) Is memory necessary for typical performance on language assessment tasks? If so, it would be expected given our stratified sampling (50% TL, 50% LI) that almost all (accounting for measurement error) of the sample with TL should fall into the high memory/high language and almost all of the sample of LI should fall in the low memory/low language quadrants, indicating an association; and very few if anyone in either group should be found in the low memory/high language category.

3) Is memory sufficient for typical performance on language assessment tasks? If so, then there should be very few, if any found in the high memory/low language category.

Figure 4-4. Data Distribution Predicted by Hypothesis 4
Methods

Participants

The current study utilized secondary data gathered as part of a beta discrimination trial for the Test of Integrated Language & Literacy Skills (TILLS; Nelson, Helm-Estabrooks, Hotz, & Plante, 2007). A total of 118 school-age students participated in the TILLS study, which was conducted with the help of volunteer clinicians in six states--California, Illinois, Massachusetts, Michigan, Nevada, North Carolina, and Texas. From the total group of 118, a subset of 74 students was selected who met inclusion/exclusion criteria for the current study.

Two groups were created for the study that were closely matched for gender, race/ethnicity, socio-economic status, and chronological age (within 9 months). Each group included 37 students. All students were administered the core four subtests on the Clinical Evaluation of Language Fundamentals-4th Edition (CELF-4; Semel, Wiig, & Secord, 2003) as part of the larger study. The groups differed based on whether they met criteria for having typically developing language skills or met criteria for having a language impairment as determined by performance < 85 on the CELF-4 and whether they were receiving services for a language impairment by a speech-language pathologist. Inclusion criteria for each group required passing a hearing screening, no special education diagnosis (a diagnosis of ADD was permitted in either group, which was noted for 4 of the students in the TL group, and 7 of the students in the LI group). Some of the students (8 in the TL group, 9 in the LI group) were in homes where other languages were spoken, but research clinicians were asked to enroll students in pairs with similar
language histories and to document that children were as proficient in English as any other school language and that it was appropriate to test them in English.

The mean age for each group was 11 years, 6 months. In each group there were the same number of participants at each grade level. Table 1 summarizes the demographic data showing comparability of the two groups.

<table>
<thead>
<tr>
<th>Table 4-1. Descriptive Statistics for the TL and LI Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TL (n=37)</strong></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
</tr>
<tr>
<td>Caucasian</td>
</tr>
<tr>
<td>African-American</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
</tr>
<tr>
<td><strong>SES</strong></td>
</tr>
<tr>
<td>Free/reduced lunch</td>
</tr>
<tr>
<td>No free/reduced lunch</td>
</tr>
<tr>
<td>Missing data</td>
</tr>
<tr>
<td><strong>Grades</strong></td>
</tr>
<tr>
<td>Grades 1-2</td>
</tr>
<tr>
<td>Grades 3-5</td>
</tr>
<tr>
<td>Grades 6-8</td>
</tr>
<tr>
<td>Grades 9-12</td>
</tr>
</tbody>
</table>

General Testing Procedures

Informed assent and consent were obtained according to Human Subject Institutional Review Board protocol. Each student participated in two assessments: the CELF-4 core subtests and the TILLS beta research version. Administration was completed during one to three individual sessions with a state-licensed or certified
speech-language pathologist. Subtest order was randomized to control for possible order effects.

*Measures of Language Impairment, Memory, and Language*

*Measure of language impairment versus typical language.* The CELF-4 core language subtests appropriate for a participant’s age (5-8 years, 9-12 years, or 13-21 years) were used as the gold standard measure of language impairment. This measure was selected for this purpose based on evidence in the test’s examiners’ manual and in the literature (Spaulding, Plante, & Farinella, 2006) showing high levels of sensitivity and specificity for identifying children with language-learning disorder. At -1 standard deviation (SD) specificity is reported to be 1.00 and sensitivity .82; at -2 SD, specificity is .87 and sensitivity is .96 (Semel, Wiig, & Secord, 2003).

Participants in all age groups are tested using the Recalling Sentences and Formulated Sentences subtests. The objective in Recalling Sentences is to evaluate a participant’s ability to repeat verbatim, sentences of increasing length and complexity. Accurately doing so requires internalizing sentence structures for accurate recall of both meaning and structure, as required for following academic instruction and note taking. The inability to do so is discriminative between normal and disordered language development. The internal consistency reliability coefficient for this task is reported to be .92 (Semel, Wiig, & Secord, 2003).

The Formulated Sentences task provides the participant with a word of increasing complexity (e.g. *car, if, because*) and an illustration which imposes contextual constraints. The participant is requested to formulate a spoken sentence that is complete,
and semantically and grammatically correct. The internal consistency reliability coefficient for this task is .83 (Semel, Wiig, & Secord, 2003).

The 5-8 and 9-12 year olds also complete the Concepts and Following Directions subtest. Directions of increasing length and complexity are spoken while the child looks at a page with an array of linearly presented objects (e.g. ball, house, apple, shoe). The participant must interpret and recall names, characteristics, and order of objects to correctly identify them among the choices. The internal consistency reliability coefficient for this task is .89 (Semel, Wiig, & Secord, 2003).

Additionally, the youngest group (5-8 year olds) is administered the Word Structures subtest. This task evaluates the participant’s knowledge of morphology, pronoun usage and possessive relationships. The examiner speaks an open-ended sentence while pointing to a picture and records the participant’s response. For example:

\textit{This boy [point] is standing and this boy [point] \underline{______}. (is sitting).}

The internal consistency reliability coefficient for this task is .84 (Semel, Wiig, & Secord, 2003).

The 9-21 year olds also complete Word Classes (receptive and expressive scores combined to create a total). The objective of these tasks is to evaluate the student’s understanding of semantically related words and to express verbally the relationships. Three to four words are presented auditorally and the student must determine which two words go together (receptive score) and tell how (expressive score). Although similar to the What Goes Together subtest from TILLS, the CELF-4 version does not provide the words in print, thus students must rely on verbal memory to complete the task.
The oldest group, 13-21 year olds, also completed a Word Definitions subtest as part of the core language tests. The student hears a word and that word is then used in a sentence. The student is asked to define the word. The objective is to evaluate the student’s ability to derive meaning from words and to classify by relationships that are unique to the reference. The internal consistency reliability coefficient for this task is .87 (Semel, Wiig, & Secord, 2003). The range and mean CELF-4 scores for the TL and LI groups can be found in Table 4-2.

Table 4-2. Range and Mean of Standard Scores on CELF-4 Core Subtests for the TL and LI Groups

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL</td>
<td>85</td>
<td>134</td>
<td>108.03</td>
<td>12.982</td>
</tr>
<tr>
<td>LI</td>
<td>46</td>
<td>84</td>
<td>69.28</td>
<td>9.579</td>
</tr>
</tbody>
</table>

Memory composite measure with low language load. Two of the subtests of the beta research edition of the TILLS were used to constitute the memory composite variable for this analysis. On the Digit Span Forward subtest (digit recall), participants listen to the clinician present a sequence of digits at a rate of one per second. They must then say the sequence in serial order. Two trials are presented at each level, starting with a sequence of two digits. The examiner records participant production verbatim. If a participant misses both trials at a given level, the task is ended. On the Digit Span Reversed (backward digit recall) subtest, the procedures are the same as for digit recall, only the participant must recall the numbers in the reverse order of presentation. Digit
recall is considered a verbal short-term memory task, whereas backward digit recall is recognized as a verbal working memory task. Processing demands are low for digit recall, but substantial for backward digit recall. The two scores were combined on the basis of previous research, which found close relations on these measures in typically developing populations (Alloway et al., 2006).

**Language composite measure with low memory load.** The measure of language used in this study was a composite of scores from two subtests of the TILLS, both of which measured forms of vocabulary awareness. Vocabulary awareness was selected to capture a core aspect of language ability that involves awareness of how verbal symbols can be used to represent meanings. The two subtests used to assess this ability are called “What Goes Together” and “Acting a Scene.”

In the What Goes Together subtest on the TILLS, students hear three words while looking at the same three words printed on a page (to reduce the need for working memory). They must select two of the words and tell how they “go together” by applying analogical, categorical, and relational concepts to thinking and talking about words. Then the same words are read again and the student is asked to pick two different words and explain how they go together in a different way. A sample item reads as follows:

**Sample:** car, water, boat

Analysis of the beta trial data showed interscorer reliability of 96% and Cronbach’s alpha of .90.

The second component of the vocabulary awareness composite score came from the vocabulary score on the subtest called *Acting a Scene*. This task assesses knowledge
of social-emotional vocabulary and pragmatic skills. Students are given a verbal
description of a short scene based on a specific vocabulary word and then are asked to
say what the character might say and to demonstrate how the character would say it, such
as the following:

Sample: Jasmine always whines when her parents won’t let her have her way. 
Her mother won’t let her buy candy in the grocery store. What do you think 
Jasmine would say?

Interscorer reliability was 90.1% for the vocabulary score on this subtest, and the
Cronbach’s alpha for Acting a Scene was .99 in the TILLS beta trial.

Data Preparation and Analysis Plan

Raw data were scored by the administering clinicians, verified by the research
team, then entered into an ACCESS database. Data were then exported into SPSS v.15.0
(SPSS, 2006) for further analysis. Prior to running statistical analyses of the group data,
tests of normality were conducted and no violations of assumptions were found.

A dichotomous variable was then created for all participants indicating “high” or
“low” performance for both the memory variable and the language variable. To create the
dichotomous variable, data for the total sample were first split into four age groups to
account for maturation (Grades1-2, 3-5, 6-8, and 9-12). Then a univariate descriptive
analysis was used for each composite score to yield percentiles for the memory
composite score and vocabulary composite score separately. Separate cut scores were
created for each of the four age categories, by using the 50th percentile for the age
category as the cut-off. Those who scored above the cut-off were categorized as “high,”
and those who performed below the cut-off were categorized as “low.” This simple
method of calculating who was above and below the 50th percentile was considered appropriate because the sample was constituted of two groups matched for all key demographic characteristics except for language ability. By gold standard definition, half of the group in the study were known to have a language impairment, thus a normal distribution would not be expected.

After participants were categorized as high or low on both the memory and the language variables in relation to the two composite score cut-offs, it was possible to create a new variable, categorizing each of them into one of four profiles representing four quadrants of relationships of memory and vocabulary (a) high/low; (b) high/high; (c) low/low; or (d) low/high. All participants were thus assigned to one of the four profile categories nested within their prior inclusion in the TL or LI group.

Crosstabs were then used to calculate numbers and percentages for each of the four categorical profiles representing the four quadrants of a 2 x 2 table. Chi square analysis and odds ratios were used in analyzing the sample for patterns that would be more or less consistent with predicted patterns for each of the four hypotheses listed in the background section of the paper.

Results

Results of the Classification

When the data for the total group of 74 participants were classified and plotted into each of the categorical profiles (see Figure 4-5), the results showed a majority (n = 57, 77%), but not 100%, to fall within either the high memory/high language or low memory/low language quadrant. A portion of the participants (n = 13, 18%)
demonstrated high memory skills in combination with low language skills, and a small group (n = 4, 5%) had low memory skills in combination with high language skills.

![Bar chart showing percent of participants in each category of high/low skills.]

**Figure 4-5.** Percent of Participants in Each Category of High/Low Skills

The data also were analyzed for the two subgroups of students who were categorized based on the independent gold standard measure (CELF-4) as having either TL or LI. Figure 4-6 shows the performance of the two groups on the vocabulary composite measure.

If the results of the vocabulary composite cut score classification were consistent with the results of the CELF-4 (both methods of measuring language ability), this a priori step should have led half of the total group (those with TL) to fall in the right column, and the other half (those with LI) to fall in the left column. The results showed that 32 students with TL (86%) did fall in the high range for the vocabulary cut score and 5 students with TL (14%) fell in the low language range on the vocabulary composite,
Despite having scores above 85 on the CELF-4, the results showed that 31 students (84%) who were classified as having LI based on the CELF-4 earned vocabulary composite scores that fell in the low language range, as expected, but 6 students (16%) fell in the high range on the vocabulary composite even though they met the original criteria for LI by having CELF-4 scores <85.

The performance of the two groups on the memory composite measure are shown in Figure 4-7. When classifying the split group by the memory composite score only, a majority of the TL group (n = 33; 89%) fell in right side column representing high memory (above the cut-off score on the memory composite for the age group); only 4 students with TL (n = 4; 11%) were classified as having low memory. The LI group performed less consistently, with 38% (n = 14) performing above the cut-off score, and
62% (n = 23) performing below the cut-off score on memory. Figure 4-8 shows the actual
distribution of TL and LI into the four categories of performance.

A majority of the LI group (n = 31; 84%) is on the left half of the quadrant, which
represents low language (below cut-off score on the vocabulary composite for the age
group). The LI group is split with 30% (n = 11) having high memory/low language, and
54% (n = 20) with low memory/low language. A majority of the TL group (n = 32, 86%)
is on the right half of the quadrant which represents high language. Only 16% (n = 6) of
the LI group is in the high language quadrants. Of those with high language, 89% (n =
34) also had high memory scores.
Results of Chi-square Analysis

An inferential statistical analysis of the split data (by the two ability groups, TL and LI) was conducted. This involved performing a chi-square analysis in a 2 x 2 arrangement to test the hypotheses of no association, which was essentially Hypothesis 2. The results showed that the null hypothesis (i.e., no association) could be rejected, in favor of a hypothesis of significant association between memory skills and language performance as categorized by TL or LI ($\chi^2(3) = 21.051, p < .0005$). Odds ratios
indicated that a child with low memory is 13.66 times more likely to have low language, than a child who has high memory; and conversely, a child with high memory is 13.52 times more likely to have high language, than a child who has low memory.

Discussion

The purpose of this study was to look for patterns of association or dissociation between memory skill and language skills, and to identify if memory skills are necessary and/or sufficient for typical language development. The memory tasks were selected for their low language-load and included one task of short-term memory (digit span forward) and one task of working memory (digit span reversed). Of interest was the association of memory with language skills, and whether patterns of high and low skills in both of these abilities considered together could inform the ongoing discussion in the literature about the nature of the relationships of memory and language.

Association and/or Dissociation of Memory Skills and Language Skills

In considerations of these results related to each of the four hypotheses, the following conclusions can be drawn:

Evidence for Hypothesis 1 (Complete Association). The data suggested an association between memory skills and language skills as evidenced by the majority (77%) of the sample falling into the high-memory/high-language or low memory/low language categories. Chi-square analysis found the association to be statistically significant. However, the substantial minority (23%) with mixed or dissociated skills precluded the acceptance of the strong hypothesis of association, as would have been
represented by the Hypothesis 1 profile. The conclusion is that Hypothesis 1 is not supported by the evidence.

Evidence for Hypothesis 2 (No Association). Clearly the data did not fit the Hypothesis 2 profile that would have been expected if there was no association between memory skills and language skills. The statistically significant chi-square results warranted the rejection of Hypothesis 2, which represented the null hypothesis.

Evidence for Hypothesis 3 (Memory Skills are Necessary, but Not Sufficient for Language Skills). The pattern of the data and the results of the chi-square analysis indicated an association between memory skills and language skills. This hypothesis predicted that for those children for whom the skills were not associated, they would show a pattern of high memory/low language skill, not high language/low memory skills. The interpretation would be that memory skills are necessary for language skills (the majority of the group showing association), but are not sufficient (1% with high memory/low language). Conversely, based on this pattern, it would be highly improbable to have high language skills concurrent with low memory skills. This Hypothesis is supported by the evidence and is considered to have the strongest empirical support.

Evidence for Hypothesis 4 (Language Skills are Necessary, but Not Sufficient for Memory Skills). As previously mentioned, the findings did support the complete association between language skills and memory skills as predicted by the first part of this hypothesis. However, for this hypothesis to be correct, no participants should fall into the high memory/low language category. This was not the case as 18% did demonstrate this pattern. Thus, it can be concluded that Hypothesis 4 is not supported by the evidence.
Memory Skills as Necessary and/or Sufficient for Language

These results can be interpreted as evidence for Hypothesis 3 that memory skills are necessary for good language skills, but not sufficient. The fact that the LI group presented with a substantial minority (30%) who had high memory scores but low language, indicates that memory is not sufficient for language. The odds ratios also indicated that children with low memory scores were nearly 14 times more likely to have low language scores, than children who had high memory scores. Conversely, children with low memory scores were only 0.07 times more likely to have high language scores, than children who had high memory scores.

Consistency with Existing Literature and Theory

The findings supporting Hypothesis 3 suggest that children with LI are not a homogenous group in regard to memory skills. Some children’s language impairments are likely due to specific rather discrete impairments in language skills (e.g., deficits in phonemic knowledge or weaker morphological skills or orthographic ability) and not an underlying memory processing skills deficit as evidenced by the LI participants with high memory skills, but low language skills. For others, evidence points to weaker memory skills as contributing to, or perhaps underpinning, the language impairments (the low memory/low language category). These findings support earlier works that found children with LI to have memory skills that are significantly lower than their typically developing peers (Archibald & Gathercole, 2006; Gathercole & Baddeley, 1990; Montgomery, 2000a).
Perhaps going one step further, these findings hint at causality. These findings support the theory that memory skills are necessary, but not sufficient for language skills. It may be that for certain individuals, difficulties become apparent only when tasks require integration of language and memory skills. These findings are in contrast with those from previous studies that have suggested that children with SLI are unable to succeed on short-term memory tasks such as the Children’s Nonword Repetition Test (Gathercole & Baddeley, 1996) because of less efficient lexical abilities to mediate the necessary processing requirements (Archibald & Gathercole, 2006).

Strengths and Limitations

Although attempts were made to gather this sample from various regions around the United States, generalization may be limited because the sample was neither random nor large. The stratified sampling did strengthen the results as the two groups were nearly identical for age, gender, race/ethnicity and socio-economic status.

In determining the two groups, the CELF-4 core score performances were used. This test has wide-spread accepted use clinically. However, several of the core tasks are reliant on memory for verbal input to complete successfully. An example is the Word Classes subtest which required the child to hold stimulus words in mind, while identifying relationships. As such, children who scored < 85 on the CELF-4, and who were therefore categorized as LI, may actually be unintentionally confounding the results of performance on our memory and language tasks. If it was the memory feature of the tasks on the CELF-4 that actually was responsible for the lower score, then more children in the LI group might be representing “poor language” when in actuality they should be
recognized for poorer memory abilities. The "gold standard" test might be overrepresenting language impairment, and underidentifying memory impairment.

The memory skills chosen for this experiment were verbal in nature due to the auditory stimuli, even though they were selected to have low language demands. No independent measures were used to document nonverbal memory or cognition. Although vocabulary awareness can be justified as a relatively memory-free core language skill, there is more to language and more to memory than either of these proxy measures could capture. Caution must be used in interpreting these results as our study was purposefully simplified in an attempt to try to isolate the relationship between memory and language skills.

In the same vein, the simple method of determining cut scores for high and low based on dividing them at the 50th percentile might be perceived as problematic because of the similarity of the scores immediately before and after the cut-off. However, this was deemed appropriate given the stratified sampling and non-normal distribution. If anything, it is probably a conservative approach to dividing the participants into equal groups.

Clinical Implications

The children in this study who were categorized as LI based on the results of the CELF-4 yielded two distinct profiles of performance on memory tasks. Nearly 38% of the LI participants had memory performances above the cut-off score for memory. Only one child was classified as TL and performed below the cut-off score on the memory composite. Given that nearly all the children categorized as TL had memory scores above
the cut-off, it is logical to conclude that consideration should be given to interventions for children with LI that addresses memory skill strategies that support language learning and utilization. This is not to say that addressing memory skills in isolation would yield improvements in language.

The conclusion that can be drawn from these data is that memory skills are necessary for but not sufficient for language. For academic success, a clinical implication is that clinicians should provide children with opportunities to remediate and compensate both language and memory deficits, using authentic integrated tasks that can be scaffolded by clinicians, teachers and parents.

**Future Research and Recommendations**

Of additional interest would be determining in greater detail the language skill and cognitive skill differences (not just memory, but for example, executive skill functions) between the group of low language performers with and without good memory. Continuing with the work of Alloway et al. (2009) who began by screening children for memory deficits, studies that then explored language performance—the opposite approach to this study—might provide further evidence in support for the hypothesis of memory being necessary, but not sufficient for language.

In order to truly identify causality, longitudinal research is needed that could incorporate structured equation modeling. Intervention studies with multiple measures over time also could measure the influence of targeting memory remediation in the context of authentic language tasks.
References


CHAPTER V

CONCLUSION

The purpose of this final chapter is to review the central purpose of the dissertation and discuss the integrated findings. This section includes summation of the study findings, limitations, implications, and recommendations for future research.

Central Purpose

The central purpose of this dissertation was to investigate the relationship between verbal memory skills and language skills. The genesis of this inquiry centered on a practical question: “If a child misses a question related to information he or she has heard, how do we know if the problem stems from lack of comprehension (language) or lack of remembering (memory)?” This question has at its root, the recognition that many tasks used in language assessment, academic assessment, and classroom activities are integrated with respect to language and verbal memory skills. For classroom teachers, special educators, and speech-language pathologists working in schools, a better understanding of the roles played by verbal memory and language can assist with interpretation of student performance—a necessary first step in identifying why a child may be struggling.

The three studies in this dissertation approached the question of the nature of the relationship between verbal memory and language skills using three different methods—simple correlations, regression modeling, and hypothesis testing. The first study tested
whole classes of children looking for a simple correlational relationship between three language tasks that were expected to have a differing language construct—sound/word level, semantic/discourse level and procedural discourse level—but a similar integration of verbal memory as a latent trait because of the oral administration of the stimuli. The second and third studies made use of a different population represented by equal numbers of children with and without language impairment and tested individually.

The second study made use of the same three language tasks as study one. With the inclusion of additional composite scores representing “verbal memory” and a basic measure of “language” (vocabulary awareness), however, it was possible to use regression to look for linear relationships that might predict performance on the three language tasks based on verbal memory skills and vocabulary awareness skills.

The final study examined the relationship from a theoretical perspective. The creation of 50th percentile cut-scores on the verbal memory composite and the language composite for each of the four age groups were utilized to categorize participants into “high” and “low” groups on each variable. Theoretically derived distributions were plotted in a two by two matrix that represented four theoretically plausible hypotheses for explaining the association of verbal memory and language. Distributions of the empirical data were then compared to the theoretical models to begin to look for causal relationships using a necessary and sufficient model.
Summary of Findings

Study One Summary

Study one (chapter II) examined the relationship between three language assessment tasks—nonword spelling (NWS), listening comprehension (LC), and following directions (FD). The tasks were being developed for inclusion in a larger comprehensive evaluation of language and literacy skills. Administration was to whole class groups, not to individuals. Each task required integration of a discrete language skill with verbal memory skills, because of the auditory presentation of the stimuli.

As part of test development, the psychometric properties of the three tasks were tested. No bias was found based on sex or race ethnicity, but some evidence was found for bias based on socio-economic status. Test item analyses indicated a need for both harder and easier items for the following direction task to avoid floor and ceiling effects, and easier items for nonword spelling to avoid floor effects. In addition to changes in item structure, suggestions were given for changes in sampling and administration procedures to include individual administration, and stratified sampling to include two groups of children—one with typically developing language (TL), and one representing language impairment (LI).

To investigate the theoretical implications, correlation analysis was utilized. It was hypothesized that if the correlations between tasks were high (> .80), then the primary construct being measured was likely to be verbal memory, the latent trait they all shared—and not the discrete language skills that differed by task. Results indicated a developmental relationship between grade and performance in all tasks. This was
expected and desirable because there is evidence that these curriculum-related language skills improve with grade level. Correlations were statistically significantly moderately correlated among all three tasks. This was interpreted as evidence for the three tasks representing separate language constructs, but the influence of memory skills was inconclusive. Additional studies were warranted, which included direct measures of verbal short-term and working memory tasks with low language loads. Their inclusion would make it possible to parse the influence of these verbal memory abilities on performance of integrated language tasks.

Study Two Summary

Study two was designed to address the recommendations from study one. The same three language tasks—nonword spelling, listening comprehension, and following directions were used, but changes were made to items and other aspects of tasks to improve the psychometric properties of the measures. The study population differed from study one as well. Participants were recruited with the help of speech-language pathologists in Michigan, Illinois, Texas, Nevada, California and Massachusetts. Seventy-four students recruited as 37 pairs were categorized either as TL or LI based primarily on their performance on the CELF-4 (scores ≤ 85 = TL, >85 = LI). These two groups were comparable in age, sex, race/ethnicity and socio-economic status.

The primary purpose of study two was to explore the amounts of variance on three language tasks that could be accounted for by performance on a composite verbal memory measure with low language demands (digit span forward and reversed), above
the contributions of age and language knowledge. Language knowledge was represented by a composite score from vocabulary awareness measures.

The results confirmed that children with LI are processing the language assessment tasks both with more difficulty and also differently from their TL peers. Regression analysis showed that for each of the tasks, a different variable or variables accounted for more of the shared variance. For nonword spelling it was vocabulary and memory that accounted for most of the variance in the TL group; whereas for the LI group, vocabulary did not play a statistically significant role, but age did. For the LI group, only age was predictive of listening comprehension and none of the three predictor variables was statistically associated with performance on following directions. In contrast, for the TL group, age was never significant, but for listening comprehension, vocabulary contributed significantly to the variance, and for following directions memory accounted for most of the shared variance.

*Study Three Summary*

Study three utilized the same data as study two. The focus of this study was on the association or dissociation between memory skills and language skills. The central question was whether verbal memory was necessary and sufficient for language skills. Cut-off scores for “high” and “low” performance were created for both the memory composite (digit span forward and reversed) and vocabulary awareness composite measure by creating a dichotomous variable based on the 50th percentile. Next, the students were categorized based on their “high-low” profiles into one of four quadrants with language represented by the x axis and verbal memory by the y axis. Students were
also grouped based on whether they met a priori criteria as having language impairment or not.

Results showed that nearly all of the TL students not only had sufficient language skills, but also had sufficient memory skills. Only one TL student had a profile of low memory, but high language. The LI group was very different. Nearly half the group had sufficient memory, but insufficient language, and the other half had “low” profiles in both memory and language. These findings were interpreted as showing that memory skills are necessary, but not sufficient for language skills because almost no children had low memory and high language, but children with low language did not all have low memory.

Discussion

Although it was well-established in the literature that students with LI may be likely to have impaired verbal short-term or working memory skills as well, little research has been conducted to illuminate the integrated nature of many common language assessment tasks. Results of studies one and two helped illustrate the importance of recognizing the influence of the latent variable of verbal memory that can influence language performance during assessment.

Study two results helped illustrate the difference in processing patterns between children with typically developing language and those with language impairment. Not only were different variables more predictive of performance on language tasks between groups, but the patterns changed depending on the task. Of particular interest was the lack of predictive ability of verbal memory skills on a procedural direction following task for the LI group, whereas verbal memory skills was the most significant predictor for the TL
group. This raises more questions as to the nature of the interaction of these two skill areas for children struggling with language. Can the answer be simply that weak language skills that impede comprehension prevent the use of memory skills for carrying out more complex tasks (e.g., story recall, direction-following, reading for learning)? More research is needed.

Results of study three seem to support a different conclusion, which is that weaker memory skills are a potential cause of language impairments for some children, and not just a co-occurring disorder. Alloway and colleagues (2009) provided some impetus for the idea of profiling children based on memory skills in combination with a domain specific skill such as language. The results from study three showed three profiles of children with language impairment: (a) those with complete association of language and memory skills (both above or both below cut-offs), (b) some with above cut-off memory skills, but below cut-off language skills—but perhaps most tellingly—(c) rarely any with language skills above cut-off, and memory skills below cut-off. This was interpreted as evidence that memory skills are necessary for language skills, but are not by themselves sufficient. Longitudinal studies are needed in order to examine causality more definitively.

Conclusions

These three studies were conducted with the hope of adding to the discussion regarding the relationship between verbal memory skills and language skills in school-age children. Several conclusions follow:

• Common language assessment tasks (nonword spelling, listening comprehension and following directions) may be assessing specific types of
language skills, but they are also influenced by the latent skill of verbal memory.

- Children with typically developing language appear to be using verbal memory skills and language skills in different ways to complete language assessment tasks. The pattern of usage for verbal memory skills and language skills may change by task.

- It appears that verbal memory skills are necessary for children to perform within normal limits on language assessment tasks when the stimuli are presented auditorally.

- It is possible to profile children with language impairment by assessing verbal memory skills with low language loads (e.g., digit span forward and digit span reversed tasks) to gain a better understanding of possible causal factors. For some, the memory skills may not just be co-occurring with a language impairment, but may actually be underpinning the deficit.

Clinical Implications

The notable but complex relationships between common language assessment tasks and verbal memory skills observed in these three studies have practical implications for those working to assess or treat children with potential language impairment. Recognizing the influence of verbal memory skills can aid in interpreting assessment results and identifying potential treatments. It may be necessary to assess verbal memory skills explicitly to fully understand the nature of a particular child’s difficulties. Best methods for doing so need to be explored.
For some children, intervention may need to include targeted work on improving verbal memory skills. This may include teaching of specific strategies such as chunking information, creating associations, recognizing patterns, or using paraphrasing. It is hoped that those providing intervention will recognize the need for any treatment to be scaffolded into authentic tasks. Working to increase memory span via drill and repetition (i.e. to recall a series of digits, words or sentences), is not likely to yield improvements in real-world activities. Only by working with a child to identify strategies that help in the performance of integrated tasks, and then working to generalize use of those strategies outside of the treatment session, are meaningful outcomes expected. Studies are needed to identify best practice methods for providing intervention.

Recommendations for Further Research

To better understand the relationship between the variables of language skills and verbal memory skills, longitudinal studies are needed. This could provide the information necessary to use methods such as structured equation modeling.

Future research might look at patterns of performance on other commonly used language assessments or authentic classroom tasks. Identifying more specific processing differences in activities such as reading to learn, retelling stories, and expository writing may add meaningful insights.

Focusing on executive skill functions and their role in task performance could also help explain important differences between the child with typically developing language and one with language impairment. As with verbal memory skills,
understanding the specific role of cognitive abilities with language acquisition and utilization has both theoretical and clinical applications.

Intervention studies to determine which techniques and strategies prove most beneficial in closing the gap of performance for children with language impairments is critical. Better profiling of characteristics (cognitive, linguistic and personality/motivation) may provide key elements for the selection of treatment that is best-suited for a particular child. The results of such studies may in turn lead to refinement of assessment tools and theories regarding the nature of language impairment.

References

Appendix A

HSIRB Approval Letters
To: Nickola Nelson, Principal Investigator  
Elaine DeRoover, Co-Principal Investigator  
Adelia Van Meter, Co-Principal Investigator  
Andrea Quast, Student Investigator for honors thesis  
Michele Anderson, Student Investigator  
Barbara Johnson, Student Investigator  
Andrea Brennan, Student Investigator  
Sally Anderson, Student Investigator  
Brandi Harveth, Student Investigator

Re: HSIRB Project Number: 04-10-16

Date: December 15, 2005

This letter will serve as confirmation that the change to your research project “Language Assessment: Formal and Informal” requested in your memo dated 12/14/2005 (collaborating investigator added) has been approved by the Human Subjects Institutional Review Board.

The conditions and the duration of this approval are specified in the Policies of Western Michigan University.

Please note that you may only conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: November 9, 2006
Date: September 3, 2008

To: Nickola Nelson, Principal Investigator
    Amy Curtis, Co-Principal Investigator
    Barbara Johnson, Student Investigator
    Michele Anderson, Student Investigator
    Dawn Anderson, Student Investigator
    Catherine Whitsel, Student Investigator
    Patricia Tattersall, Student Investigator

From: Amy Naugle, Ph.D., Chair

Re: HSIRB Project Number: 06-12-23

This letter will serve as confirmation that the change to your research project “TILLS-Beta Trial” requested in your memo dated September 2, 2008 (status of Michele Anderson changed from student investigator to student investigator for dissertation) has been approved by the Human Subjects Institutional Review Board.

The conditions and the duration of this approval are specified in the Policies of Western Michigan University.

Please note that you may only conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: January 26, 2009